Intelligent Hybrid Multi-Stage Feature Selection and Assessment for 5G Base Station Antenna Health Effect Detection

(5G 基地局アンテナの健康影響検出のための知的ハイブリッド多段特徴選択とその評価)

山梨大学大学院

医工農学総合教育部

博士課程学位論文

修了年月 2024 年 3 月

氏名 TASNEEM BINTI SOFRI

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENT

Alhamdulillah, first and foremost, I would like to thank Allah Almighty for granting His blessing upon me to be on this postgraduate journey.

I express my sincere gratitude to Assoc. Prof. Ir. Ts. Dr. Hasliza binti A Rahim @ Samsuddin for her precious supervision, support, and guidance throughout my academic journey. Her expertise played a crucial role in shaping my experimental methods and critiquing my research performance. I deeply admire her wisdom and dedication to education and research, making her a significant role model for both my career and personality development. Additionally, I appreciate her enthusiastic efforts in seeking a grant and scholarship to support my studies. I extend my heartfelt gratitude to Prof. Dr. Nishizaki Hiromitsu for his generous support and invaluable assistance during my studies in Japan. His warm welcome and kindness were evident during my time at the Acoustics, Linguistics, Picture, and Signal (ALPS) Processing Laboratory, situated in the Department of Mechatronics Engineering, University of Yamanashi. I also appreciate his valuable opinions regarding my PhD progress and dissertation.

I extend my gratitude to my supportive co-supervisors, Ts. Dr. Allan Melvin Andrew and Assoc. Prof. Ts. Dr. Soh Ping Jack, for their mentorship and guidance. I would like to convey my deepest thanks to Ts. Dr. Allan Melvin for his assistance from the inception of my exploration into machine learning research. His wealth of knowledge and expertise significantly contributed to the development of my research. He also provided crucial support when I encountered challenges in my research, helping me find effective solutions. Assoc. Prof. Soh Ping Jack has been helpful in enhancing my publication skills, offering valuable assistance in refining, and advancing my work. His esteemed guidance in journal submissions has undoubtedly played a fundamental role in shaping me into a proficient researcher.

I want to also express gratitude to Mr Mohd Hafizi Omar and Mrs. Khatijahhusna Abd Rani for their unforgettable assistance in evaluating my data collection experiment and analyzing the results. Without their support, I would have faced challenges in progressing with my experiment assessments.

Very special thanks to Assoc. Prof. Dr. Latifah Munirah Kamarudin for getting me accepted into the Ph.D. Dual Degree Program between Universiti Malaysia Perlis (UniMAP) and University of Yamanashi under the supervision of Prof. Dr. Nishizaki Hiromitsu with the scholarship from Japan Student Services Organization (JASSO) Student Exchange Support Program with University of Yamanashi for a period of 1 year. I would like to express my sincere gratitude to Prof. Xiaoyang Mao for providing the opportunity to participate in this program. The memorable experience of spending a year in Japan has not only enriched my academic studies but has also broadened my life knowledge.

I am deeply grateful to the University of Yamanashi’s staff, including Miss Mizoguchi, Miss Nambu, Miss Ishikawa, Mr Aida, Miss Uematsu, Miss Ozawa and Miss Miko, for their unwavering support and care during my time in Japan. I truly appreciate their kindness in ensuring my overall well-being. They generously shared insights into Japanese culture and introduced me to unique places, enriching my experience.

I would like to thank to the Malaysian Communications and Multimedia Commission (MCMC) and UniMAP-Private Matching Fund (UniPRIMA) for the precious opportunity and financial assistance granted to me in pursuing the highest degree of education in Faculty of Electronic Engineering Technology, UniMAP. I extend my appreciation to MCMC with grant number 2020/01/001 and UniPRIMA, Ref. No. UniMAP/RMC/9001-00702 (UniMAP Fund)/ 9030-00006 (FILPAL (M) Sdn Bhd Fund) for making this research possible. The Malaysian Communications and Multimedia Commission (MCMC) approved the use data generated from the UniMAP research grant funded by MCMC (grant number: 2020/01/001) for this publication.

I express heartfelt gratitude to my beloved parents, Sofri Yahya and Hasmiza Mohd Amin, as well as my family members, Muadz, Syafaf and Huda for their unwavering moral support. I sincerely thank them for their prayers and encouragement throughout my postgraduate journey. Their steadfast backing has been a source of inspiration and motivation.

I would like to extend special thanks to Miss Nurul Fatihah, Miss Ain Najihah, and Miss Kushalya for their assistance in assessing my experiment data collection. Additionally, I express my gratitude to the Advanced Communication Engineering (ACE) Laboratory at UniMAP for providing an excellent facility that greatly facilitated my work as a postgraduate student. A special acknowledgment goes to Sadia, my friend from ACE, for dedicatedly staying late at the laboratory to meet deadlines and complete tasks together.

I extend my sincere gratitude to my friends from ALPS Laboratory – Hakimi, Husaini, Maoling, Heng, CC, Miss Jang, Mr. Leow, Mr. Kitagawa, Mr. Yaw, Mr. Soya, Mr. Hotta, and other unnamed friends – for their cooperation and support throughout my research study. Additionally, I express my thanks to my second family in Japan, friends from Yamanashi University Malaysian Society (YUMS) – Ermi, Arnidah, Syazwani, Hayati, Iffah, Haqeem, Faris, Suit Mun, Nazrul, Aniq, Syahmin, Aqil, Syahmi, Anas and Huzairi– for making my daily life in Japan exceptionally convenient. Your support has been truly meaningful.

I extend my apologies to those not explicitly mentioned in this acknowledgment, particularly those contributing behind the scenes without being named. Your assistance is deeply appreciated, and only God can truly repay your kindness.

ABSTRACT

Fifth Generation (5G) technology able to support more terminals (device density up to one million per square kilometer) with much higher data rates (peak rate up to 20 Gbps), extremely low latency (not more than 1 millisecond) and very high reliability (99.999%). However, in Malaysia, the accelerating of the telecommunication towers buildings raises concern among residents about possible health effects of the radiation coming from those structures in the past few years including the 5G base station exposure. The first main problem of this research work is all studies on the effect of Radio Frequency Electromagnetic Field (RF-EMF) are related to experiments that are either studies on animals or short-term studies in human subjects limited to the effects of GSM900/GSM1800/UMTS/4G Mobile Phones (MP), GSM900/GSM1800/UMTS BS, Digital European Cordless Telecommunications (DECT) and Wireless Fidelity (Wi-Fi) exposures, without considering the effects of 5G 700 MHz, 5G 3.5 GHz or 5G 28 GHz BS signal. The second main problem is limited integration hybrid feature selection method design for physiological parameters and cognitive performance studies on 5G base station antenna health effects. The objective of this research is to investigate the effects of 5G 700MHz, 3.5 GHz and 28 GHz BS antenna fields exposures and Sham on physiological parameters and cognitive performance of adults in the double blinded condition on Electromagnetic Hypersensitivity (EHS) subject and non-EHS subject, to design the hybridized Multi-Stage Feature Selection (MSFS) and hybrid feature for 5G BS antenna health effect detection based on the physiological parameters and cognitive performance of adults dataset and lastly to validate the performance of the proposed MSFS hybrid feature dataset using supervised machine learning in terms of machine learning classification accuracy, precision, f1-score, sensitivity, and specificity. The outcomes from this research are verification of the hypotheses that the effects of 5G 700MHz, 3.5 GHz and 28 GHz BS antenna fields exposures and Sham on the physiological and cognitive parameters of the subjects are statistically significant or not and also evaluation of the hybridized MSFS and supervised ML for 5G BS antenna health effect detection classification based on the assessed parameters in order to reduce misclassification in the classification. Based on the p-value ($p>0.05$) result analysis, the findings from the assessment indicated that there are no statistically significant effects from short-term 5G radiation exposure from adults in terms of cognitive function and physiological parameters. The initial application of the prepared dataset involves utilizing classification algorithms such as K-nearest neighbours (KNN), Support Vector Machine (SVM), Ensemble Method, Naïve Bayes, and Probabilistic Neural Network (PNN) without implementing any feature selection approach. However, the outcomes of this approach indicate suboptimal results, with accuracy levels falling below 50% for both the classification of subjects and exposure classification. This suggests that the models, when applied to the dataset in its entirety without feature selection, do not perform satisfactorily in accurately classifying subjects and exposure scenarios. Following the initial suboptimal results, the dataset undergoes normalization using 20 different normalization methods. A thorough statistical examination is conducted on the newly normalized datasets to identify the top three normalization techniques. Subsequently, an in-depth analysis of the data properties is undertaken to extract features that are most conducive to

accurately classifying subjects and exposure scenarios. Machine learning algorithms are then applied to these datasets, and the algorithm that demonstrates the highest accuracy is selected for further consideration. This methodological sequence aims to improve classification outcomes by strategically normalizing the data and selecting features that enhance the effectiveness of the machine learning algorithm. The proposed technique obtained a high average accuracy of 99.5% good performance of the SVM machine learning algorithm.

CHAPTER 1 : INTRODUCTION

1.1 Research Background

Nowadays, wireless communication devices and connected objects are part of our personal and professional daily life through the widespread usage of local, wide-area, and mobile networks that connect computers. In 2018, the number of mobile devices and connections reached 8.8 billion (Cisco, 2020). High-speed wireless communication is anticipated to be primarily enabled by 5G technology. Additionally, research indicates that by 2023, there will be 13.1 billion mobile devices connected to 4G and 5G networks worldwide. Furthermore, it is predicted that each user would own 3.6 devices per person due to an upsurge in smartphone use, Machine-to-Machine (M2M) connections, and etc. (Regrain et al., 2020). The International Telecommunications Union (ITU) anticipated that the number of connected devices on the Internet is projected to reach 50 billion from 2025 onwards (ITU, 2022).

Compared to existing mobile networks, 5G will be able to support much more terminals (device density up to 1 million per square kilometre) with much higher data rates (peak rate up to 20 Gbps), extremely low latency of not more than 1 milliseconds (ms) and very high reliability (99.999%). In this way, 5G will ensure high Quality of Service for users, and enable highly reliable massive communication between devices (Pawlak et al., 2019). These developments indicate that large parts of the global population are now exposed to Radio Frequency Electromagnetic Field (RF-EMF), and the exposure is expected to increase in the coming years. Concern has been raised regarding public health consequences from RF-EMF and it is therefore crucial to perform a health risk assessment to support decision-makers and to inform the general public (Röösli et al., 2021; Verbeek et al., 2021; Bosch-Capblanch et al., 2022).

In this research, the term intelligent is to intelligently analyse the assessment data used from the use of Artificial Intelligence (AI) and machine learning. Machine learning studies generally differ from traditional research in two ways. The first is a focus on prediction (explanatory power of the model) rather than inference (hypothesis testing) (Chekroud et al., 2021). As AI is increasingly applied to biomedical research, there is an urgent need to start using technology to identify the possible impact of RF-EMF on any 5G frequency band exposure from the base station on the cognitive performance and physiological parameters (heart rate, blood pressure, body temperature) of adults in epidemiology studies while using domain knowledge, with Multi-Stage Feature Selection (MSFS) techniques. Technology is essential to compensate for the lack of experiences and required training in handling data collected from assessment investigations. This applies specifically to the assessment of 5G base station antenna exposure on physiological parameters and cognitive performance in adults. Implementing technology in this context aims to ease the burden on future researchers. However, developing these applications involves transforming the data from the original format to a format that is understandable by the system. It also involves using suitable machine learning algorithms appropriate for the problem to be solved. This work discusses about a framework that is developed to a hybrid feature strategy a prediction model tested on new, unseen datasets to observe the possible classification for the feature of exposure and subject in this epidemiology study.

1.2 Problem Statements

i. Limited Studies in the Assessment of 5G Base Station Antenna Exposure on Physiological Parameters and Cognitive Performance of Adults

Previous research on the effects of RF-EMF radiation on human cognitive performance, physiological parameters, and measurements has primarily focused on wireless communication devices such as base station antennas operating in different frequency bands (Eltiti et al., 2007; Regel et al., 2006; Ridgewell et al., 2007; Wallace et al., 2010; Malek et al., 2015; Andrianome et al. 2017; van Moorselaar et al. 2017; Bogers et al., 2018), mobile phones (Tahvanainen et al., 2004; Oftedal et al., 2007; Cinel et al., 2008; Eltiti et al. 2009; Kwon et al., 2012; Choi et al., 2014, Huang et al. 2022), Wireless Fidelity (Wi-Fi) (Andrianome et al., 2017), Terrestrial Trunked Radio (TETRA) base station (Wallace et al., 2010) and portable TETRA handsets (Sauter et al., 2015). Most available studies on the cognitive performance and physiological parameters of human are limited to the effects of Global System for Mobile Communication (GSM) 900 MHz, GSM 1800 MHz, Universal Mobile Telecommunications System (UMTS), 4G mobile phones, GSM 900 MHz /GSM 1800 MHz/UMTS base station, Digital European Cordless Telecommunications (DECT) and Wi-Fi exposures, without considering the effects of 5G 700 MHz, 3.5 GHz, or 28 GHz base station signal. Based on an exhaustive literature review, there are no published investigations on evaluating any 5G frequency band exposure from the base station on the cognitive performance, and physiological parameters (heart rate, blood pressure, body temperature) of adults. The information gathered from such studies will provide important insights into the significance of 5G base station signal exposure on humans.

ii. Single-Stage Feature Selection Method for Physiological Parameters and Cognitive Performance Studies on 5G Base Station Antenna Health Effects

While traditional research approaches focused on p-values for specific coefficients in a model, prediction studies focus on the overall explanatory power of the model, often in terms of R2, balanced accuracy, or Area Under the Receiver Operating Characteristic Curve (AUC). Predictive studies require a keen focus on validation approaches, to examine whether the model is learning patterns that are substantive and consistent from one dataset to another, or whether the model has simply learned idiosyncrasies of the initial training data. (Chekroud et al., 2021). Usual single-stage feature selection methods may overlook the intricate relationships between features. They often fail to capture the nuanced interactions that contribute to the overall predictive performance. The methodological approach for selecting relevant features from physiological parameters and cognitive performance data in the context of studying the potential health effects of 5G base station antennas are limited. Previous studies primarily focusing on statistical analysis in order to discover the changes of health effect from the RF-EMF exposure. The biological effects of RF-EMF exposure on human health remain unclear due to inconsistent and contradictory findings of various studies. Halgamuge (2020) mentioned significant of feature selection refers to the process of identifying the most significant characteristics or variables that provide the best predictive capability in modelling data as this is one of the key ideas in machine learning, which tremendously impacts the model or classifier performance. Previous researchers depicted the use of conventional feature selection method, basically, by using a single-stage feature selection method. In the single-stage feature selection method, the important features are extracted from the raw data, and the extracted data is further filtered to select only important and useful features (Halim et al., 2022; Elkhouly et al., 2023). The exploration and exploitation of the data will be insufficient during the feature selection as the features are reduced at the initial stage. As a result, only some redundant features are selected, and some useful features are lost due to poor data management (Tran et al., 2018; Vijayasarveswari et al., 2020; Alwohaibi et al., 2021).

1.3 Objectives

1. To assess the effects of 5G 700MHz, 3.5 GHz and 28 GHz base station antenna fields exposures and Sham on physiological parameters and cognitive performance of adults in the double blinded condition.

2. To design the hybridized MSFS framework and hybrid feature for 5G base station antenna health effect detection based on the physiological parameters and cognitive performance ability of adults' dataset.

3. To validate the performance of the proposed MSFS hybrid feature dataset for 5G base station antenna health effect detection using supervised machine learning in terms of machine learning classification accuracy, precision, f1-score, sensitivity, and specificity.

1.4 Scope of Research

The first scope of this research focuses on the experimental investigation of the effects of 5G 700 MHz, 3.5 GHz and 28 GHz base station antenna fields exposures and Sham (No Exposure) antenna exposure on cognitive performance and physiological parameters of adults. The aim of this study's scope is to test the hypothesis concerning whether a relationship exists between 5G base station antenna exposure and cognitive performance and physiological changes. The changes in cognitive functions and physiological parameters of subjects during the 5G base station antenna exposure (including Sham) will be determined by applying statistical techniques. The effects of different types of bands for 5G base station antennas exposures also were examined on physiological parameters of subjects.

The second scope of this research focuses on designing method of selecting features for hybridized MSFS for 5G base station antenna health effect detection classification that is based on multi-stage of data processing, feature selection and model optimization with the proposed parameters including cognitive performance and physiological. The proposed hybridized MSFS method have four stages. The first stage consists of data pre-processing and data normalization methods. The second stage consist of feature selection methods, while third stage and fourth stage consist of feature fusion and feature extraction, respectively. The selection of data normalization methods and features are done by computing the p-value and F-value in the first stage. The raw data samples go through these stages to identify the best data normalization techniques, the best feature extraction methods, and the optimum features to be hybridized (fused). The features are fused together using feature fusion technique. This newly hybrid feature dataset will be used for 5G base station antenna health detection framework. The supervised machine learning classifier algorithm testing from the determine the accurate and appropriate classifier prediction for this input dataset involved. Hence, the performance of proposed classifier accuracy, precision, f1-score, sensitivity, and specificity can be validated. This part also highlights a focused approach in the methodological design, with an emphasis on integrating biomedical data and a notable shift towards leveraging machine learning techniques.

1.5 Main Contributions

1. New scientific evaluation of assessment 5G 700MHz, 3.5 GHz and 28 GHz base station antenna fields exposures effects on physiological parameters (body temperature, blood pressure and heart rate) and cognitive performance of adults.
2. Evaluation relationships to understand how varying levels of exposure to different frequencies (700MHz, 3.5 GHz, and 28 GHz) correlate with physiological and cognitive changes.
3. Investigate the influence of 5G fields exposures and Sham (no exposure) on cognitive functions performance of adult subjects.
4. Advanced pre-processing technique for analysing the impact of weak radiofrequency radiation on human subjects.

5. New technique of selecting features using hybridized MSFS and formulation of hybrid feature for 5G base station antenna health effect detection classification based on the assessed parameters.
6. Development of a novel system for the classification of health effects in 5G base station antennas.

1.6 Organization of the Thesis

This thesis's contents are organized as follows: Chapter 2 is the literature review where spectrum band for each 5G signal, overview of RF-EMF exposure in 5G to human, overview on the effect of 5G BS signal exposure on cognitive and physiological performance are discussed. Then machine learning technique, machine learning algorithm and MSFS methods are introduced. Lastly in this chapter, the findings comparison of previous studies with respect to physiological parameters and cognitive parameters and overview of the machine learning and MSFS method are compared to highlight the research gap in the last section. Followed by chapter 3 describing the methodology used in the research where assessment of 5G base station antenna design setup and measurements are described as well as its output parameters. The preprocessing step before applying to machine learning and the algorithm that is used to build the classifier using supervised machine learning, is explained. In chapter 4, the results of the performance of the proposed MSFS hybrid feature dataset for 5G base station antenna health effect detection using supervised machine learning in terms of machine learning classification accuracy, precision, f1-score, sensitivity, and specificity are demonstrated, explained, and discussed. The machine learning model is validated using the collected data. Finally, future work and recommendations are presented in chapter 5.

CHAPTER 2 : LITERATURE REVIEW

2.1 Introduction

This chapter covers an introduction to basic compulsory concepts of 5G RF-EMF exposure and human exposure to RF-EMF along with the parameters that are studied, including cognitive performance and physiological parameters. A thorough literature review to describe the assessment of 5G BS antenna exposure on Malaysian adult subjects with Normal attributes and Sensitive in order to know the changes of the assessed parameters before, during or after the exposure (including Sham) are presented. The findings of these studies are summarized. It includes some of the feature selection method and related studies on the hybridized MSFS techniques for 5G BS antenna health effect detection classification based on the assessed parameters. Next, machine learning techniques are introduced, then the mechanism of machine learning and how to improve machine learning model prediction by data normalisation techniques are explained. The findings of these studies are summarized. Next, the overview techniques of MSFS are discussed. From this literature review, a conclusion is drawn to highlight the research gap of this study. Lastly, a summary for the work is included.

2.2 Spectrum Bands for 5G

Band selection is based on global 5G trends (commercial launches) and ecosystem maturity and these bands are 3.5 GHz band and 26/28 GHz where this 3.5 GHz mid-band spectrum (<6 GHz) is emerging as the core band for 5G globally due to the technical characteristic that offers an optimal balance of high capacity (amount of traffic supported) and coverage (the distance of the signal travelled). Meanwhile, 26/28 GHz band is being proposed due to the availability of large contiguous bandwidth. These high band spectrums (>6 GHz) are important to provide extremely high data speeds and ultra-reliable services in 5G. To meet the national 5G needs, two spectrum bands have been identified as the priority 1 spectrum for 5G deployments (MCMC, 2019):

1. 3.5 GHz band (range of 3.3 GHz to 4.2 GHz)

2. 26/28 GHz (range of 24.25 GHz to 29.5 GHz)

Spectrum Management & Allocation Working Group (SWG) has conducted a study in the 3.5 GHz band for potential spectrum allocation for 5G. The 3.5 GHz band or popularly known as C-band is heavily utilised for Fixed Satellite Service (FSS) in Malaysia. The Malaysian satellite operator and other FSS operators currently operate in the 3.4 - 4.2 GHz band. Figure 2.1 shows the existing allocation of 3.5 GHz band in Malaysia along with the proposed allocation for 5G. Based on the analysis, an initial total of 100 MHz (limited to indoor use) and 400 MHz (general use) are proposed to be allocated for 5G deployments, with the following arrangement on the C-Band frequencies, refer to Figure 2.2. The range of 3.3 - 3.4 GHz is for 5G indoor use only, the range of 3.4 - 3.8 GHz is for 5G and the range of 3.8 - 4.2 GHz is for satellite.

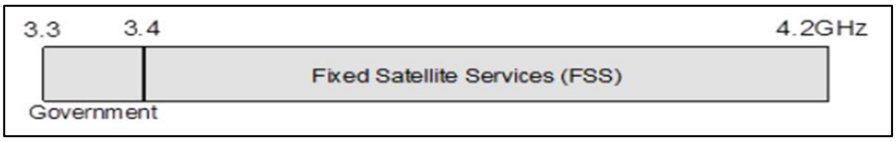


Figure 2.1 Current usage in 3.5 GHz Band
Source: MCMC (2019)

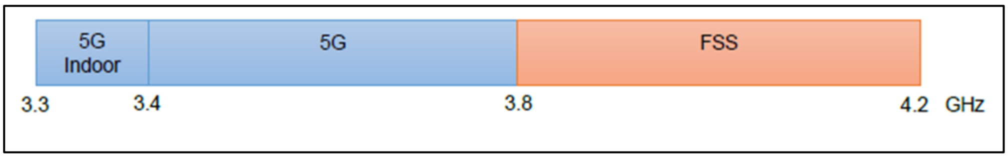


Figure 2.2 Proposed frequency arrangements for C-Band
Source: MCMC (2019)

The 26/28 GHz band are the spectrum bands being considered internationally for 5G and ranges from 24.25 – 29.5 GHz. This includes the 26 GHz 3GPP band, n258 (24.25 – 27.5 GHz) and the 28 GHz, n257 (26.5 GHz to 29.5 GHz). Although the 26/28 GHz band is not suitable for ubiquitous 5G nationwide coverage and has relatively poor outdoor to indoor penetration, it is more suitable for outdoor hotspot, in-building coverage, and Fixed Wireless Access (FWA) with outdoor Customer-Premises Equipment (CPE). The 26/28 GHz is important in the overall 5G ecosystem as it will address specific 5G use case requirements and demands (MCMC, 2019). In most use cases, the 26/28 GHz band is a complementary “layer” to 3.5 GHz and other lower bands to deliver 5G services. The band is seen to have huge potential due to its characteristics to provide very high capacity, speeds, low latency and easier to manage interference compared to mid band spectrum. The current usage of 26 GHz and 28 GHz bands in Malaysia is shown in the Figure 2.3.

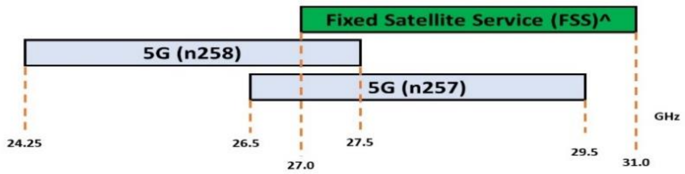


Figure 2.3 Existing FSS in the 24.25 GHz to 31 GHz
Source: MCMC (2019)

Three band for 5G in Malaysia as proposed which are 700 MHz, 3.5 GHz and 28 GHz. 5G 700 MHz band is to secure pervasive national coverage. It is likely to be deployed from the traditional BS tower. Only modest data capacity can be supported. The 3.5 GHz band sits between the current Wi-Fi bands at 2.4 GHz and 5 GHz that are widely deployed in homes, offices, and public places. 3.5 GHz is the ‘good spot’ for achieving the best capacity over the largest areas for the lowest cost and has wide international support. The mass deployment of small low power BS in towns and cities will most likely use this band. 28 GHz of high frequency (mmWaves) supports the largest capacity but at the highest cost of coverage. Research engineers see a potential for 28 GHz to be used for a data capacity lift in the limited number of locations where the 3.5 GHz frequency maxes out over the next 10 years. Another use may be as a low power advanced manufacturing broadband access point (Industry 4.0). Such examples of relatively short distance applications only need relatively low power levels.

2.2.1 5G New Radio Grid Structure

Like 4G Long-Term Evaluation (LTE), 5G New Radio (NR) supports both Frequency Division Duplexing (FDD) and Time Division Duplexing (TDD) and signals are modulated by using Orthogonal Frequency Division Multiplexing (OFDM) with a cyclic prefix. Moreover, 5G NR also uses a grid structure consisting of subcarriers in the frequency domain and OFDM symbols in the time domain. The basic granularity of the 5G NR resource grid (i.e., in frequency and time) is the Resource Element (RE), which spans one OFDM symbol in time and one subcarrier in frequency. In the frequency domain, the grid structure is further organized in Resource Blocks (RBs), with each RB consisting of twelve contiguous subcarriers. The total number of RBs available for data transmission depends on the channel bandwidth (up to 100 MHz for sub-6 GHz signals) and the numerology or Sub-Carrier Spacing (SCS), which is 15 kHz, 30 kHz, or 60 kHz for sub-6 GHz signals. This contrasts with LTE, where the SCS is fixed at 15 kHz and the bandwidth at up to 20 MHz. In the time domain, the structure is organized in frames. A 5G NR radio frame is 10 ms long and consists of ten subframes of each 1 ms. A subframe is further divided into slots, which each comprise 14 (in the case of a normal cyclic prefix) or 12 OFDM symbols (in the case of an extended cyclic prefix). The number of slots and the duration of a symbol depend on the SCS. For example, in the case of an SCS of 30 kHz, a subframe consists of two slots and the symbol duration is 35.68 µs. Analogous to an RB in the frequency domain, a slot is the basic transmission unit in the time domain (Aerts et al., 2019).

The Synchronization Signals (SS) and Physical Broadcast Channel (PBCH) collectively constitute the SS/PBCH block, often referred to as the SS block or SSB. The SS/PBCH block, which comprises the constant-power signal components of 5G NR, spans four OFDM symbols in the time domain and 240 contiguous subcarriers, or 12 RBs, in the frequency domain and with indication of the minimum (B BW_{SSB}, min) and maximum bandwidth (B BW_{SSB}, max) as shown in Figure 2.4. As opposed to the constituting signal equivalents in LTE, in 5G NR the SSB is not fixed to the centre frequency of the radio channel, but instead its position (denoted by SSREF) is determined by the Global Synchronization Raster Channel (GSCN) value, which fixes it on a discrete raster. Furthermore, whereas in 4G LTE the synchronization signals are transmitted over the entire cell, 5G NR systems can apply beamforming, in which case the base station repeatedly transmits the SSB in a number of predefined directions or beams in an SS burst or SS burst set (consecutive SS bursts). The SS burst set is transmitted at regular time intervals, which can be 5, 10, 20, 40, 80 or 160 ms (with the default being 20 ms), within the span of one half-frame (5 ms) (Aerts et al., 2019).

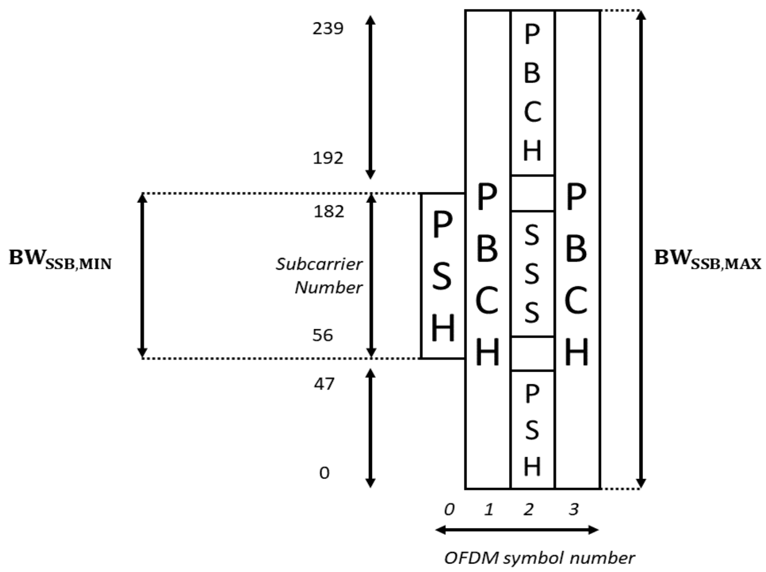


Figure 2.4 Structure of the SS/PBCH block in time and frequency.
Source: Aerts et al. (2019)

In contrast to 2G to 4G mobile technologies such as GSM, UMTS, and LTE, the 5G NR technology will make use of a huge span of Radio Frequencies (RF), split in two broad ranges: one spanning from 410 MHz to 7.125 GHz (`sub-6 GHz'), and the other from 24.25 GHz to 52.6 GHz (`mmWaves'). Furthermore, one of the main technological advances introduced or enhanced in 5G NR will be the widespread use of Massive Multiple-Input Multiple-Output (MaMIMO), in which many antenna elements (up to hundreds) can be used to narrow and steer the transmit beam in order to optimize the signal at the receiver device (Aerts et al., 2019). Figure 2.5 shows the concept of MaMIMO system employing ample of transmit antennas. It was found in that adding more antennas at the base station is always beneficial even with very noisy channel estimation can be reduced out of cell interference because the base station can recover information even with a low Signal-to-noise-ratio (SNR) once it has sufficiently many antennas. This motivates the concept of using a very large number of transmit antennas, where the number of antenna elements can be at least an order of magnitude more than the current cellular systems. MaMIMO systems have the potential to revolutionize cellular deployments by accommodating a large number of users in the same time-frequency slot to boost the network capacity and by increasing the range of transmission with improved power efficiency (C. X. Wang et al., 2016).

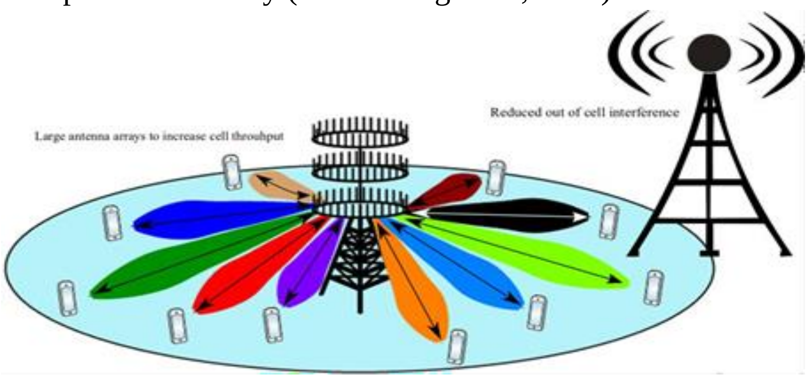


Figure 2.5 Massive MIMO Concept for Large Antenna Array
Source: C. X. Wang et al. (2016)

2.2.2 Exposure Recommendation

Technical studies on RF-EMFs with frequencies range from 100 kHz to 300 GHz have progressively increasing in a diversity of applications in medicine, industry, household items, safety and security management and, most significantly, telecommunications since the 1950s (Masrakin et al., 2019; Pophof et al., 2021). Several international organizations have been formed to address these challenges. Examples of them include the International Commission on Non-Ionizing Radiation Protection (ICNIRP, 2020), the Institute of Electrical and Electronics Engineers (IEEE, 2019), the National Radiological Protection Board (NRPB) of the United Kingdom (NRPB, 2004) and the Australian Radiation Protection and Nuclear Safety Agency (ARPANSA, 2017). In order to ease people's concerns about the consequences of radio frequency radiation, these organizations released guidelines and standards. Exposure to RF-EMF has been a significant issue for telecommunications organizations in the previous 10 years to preserve safety.

The Federal Communications Commission (FCC) of the United States (US) and the ICNIRP established recommendations for the maximum quantity of EMF radiation that may be administered to a person's body. It is significant that the FCC's Specific Absorption Rate (SAR) recommendation is averaged over 1 gram (g) of tissue, whereas the ICNIRP's is averaged over 10g. Recommendation by the FCC looks to be more cautious, whereas 2-3 times energy absorption as the ICNIRP permits. In addition, the US Food and Drug Administration (FDA) states that current knowledge of the negative effects of EMF emissions on human health is insufficient to determine whether exposure to the emissions is safe or not, and that more research is needed to fill in the gaps in the literature on human health safety in wireless systems use (United States Government Accountability Office, 2012). Meanwhile, the International Agency for Research on Cancer (IARC) of the World Health Organization (WHO) categorizes EMF exposure as probable carcinogen (FDA, 2018).

The ICNIRP exposure recommendations establish a maximum power density of 10 W/m^2 for the general population between 10 GHz and 300 GHz, measured as an average across any 20 cm^2 of exposed area. In addition, the spatial maximum power density averaged over any 1 cm^2 shall not exceed 200 W/m^2. The uncontrolled power density exposure limit for FCC between 6 GHz to 100 GHz is also 10 W/m^2, which in general is to be considered as a spatial peak value (FCC, 1997; Robert F. et al., 1997). However, spatial peak is not a well-defined quantity, and the answer achieved will be dependent on the method used to measure exposure. Averaging across the probe dimensions will be achieved for measurements, and a suitable sample density will be required for calculations. Between 3 GHz to 100 GHz, the IEEE general public basic restriction on power density is 10 W/m^2 (Andrianome et al., 2017). In the frequency range between 3 GHz to 30 GHz, the power density is to be spatially averaged over any contiguous area corresponding to $100\lambda^2$ where λ is the free space wavelength of the RF field. In addition, IEEE also species maximum spatial peak power densities of 18.56f 0.699 W/m^2 at frequencies between 3 GHz and 30 GHz, where f shall be taken as the frequency in GHz. The IEEE specifies neither average area nor spatial sampling density for the spatial peak power density limitations. The spatial peak power density was measured with a minimum spatial sampling density of four samples per wavelength in this study (Thors et al., 2016).

A summary of the RF-EMF exposure limits, S_{lim}, is provided in Table 2.1. For convenience, the spatially averaged power densities are for all RF exposure limits determined assuming square-shaped averaging areas (Thors et al., 2016). S_{lim}, as defined by ICNIRP, FCC, IEEE. The parentheses behind the power density limits indicate the applicable average area. Absence of averaging area implies spatial peak power density.

Table 2.1 General Public Basic Restrictions on Power Density

	ICNIRP	**FCC**	**IEEE**
***f* (GHz)**	10 - 300	6 -100	3 - 30
S_{lim} (W/m^2)	10 (20 cm^2)	10 (1 cm^2)	10 (100 λ^2)
	200 (1 cm^2)		$18.56f^{0.699}$

2.3 Overview of RF-EMF Exposure to Human

The potential adverse health effects associated with mobile phones overall, including their base stations, have sparked considerable public concern. The accelerating of the telecommunication towers buildings raises a concern among residents about possible health effects of the radiation coming from those structures in the past few years. Many demonstrations and complaints have been made against the construction of telecommunication tower in their residential area. Concerns regarding the harmful effects of radio frequencies on human health might potentially be a stumbling block to broad 5G infrastructure deployment. The mmWaves spectrum will be utilized to build a dense network of small picocells, resulting in the installation of ample new radio transmitters.

As a result, activists are concerned that 5G will expose the population to new sources of dangerous radio frequency radiation (Pawlak et al., 2019). On 27 August 2020, UK Engineering & Technology (E&T) Magazine reported that the UK government has published a guide for the public about 5G networks due to the increase of 5G conspiracies on social media platforms, including theories that COVID-19 pandemic could be linked to the new networks in some way. The misinformation spread quickly and led to numerous accounts of people vandalizing 5G masts over this concern (The Engineering & Technology (E&T) Magazine, 2020). In Malaysia, Sani & Tay (2018) found that the level of awareness concerning likely health hazards for residents near the mast was significantly different in demographics and type of resident. Currently, a similar debate rages over 5G technology, with a sizable number of individuals persuaded that 5G poses a serious threat to human health (Nyberg & Hardell, 2017). As a result, the terms "5G" and "risks" are frequently used interchangeably, negatively impacting public impression of 5G. However, it is critical to continue to investigate potential health impacts associated with realistic exposure (i.e., below maximum limits) to 5G devices. Clearly, there is a need for further study to be done in order to correctly build exposure-aware cellular networks for 5G and beyond systems (Chiaraviglio et al., 2021).

RF-EMF exposure potentially affects the human body, including ‘heating' of the skin. The temperature of a skin's outer surface is usually between 30 and 35ºC. The pain detection threshold temperature for human skin is around 43ºC (Robert F. et al., 1997) and any temperature surpassing it can cause a long-term injury. Heating is a significant influence since it can result in cell damage and protein induction. High-frequency EMF is also known to influence the sweat glands. (which may serve as helical antennas), peripheral nerves, the eyes and the testes, and may have indirect effects on many organs in the body (Markov, 2019). 5G wireless is projected to deliver considerably greater data speeds than previous generation wireless networks to meet the newest skyrocketing bandwidth demand. The necessity of a very high data rate in 5G necessitates an increase in signal power received at the user's end, possibly increasing the electromagnetic radiation inflicted on the user in the vicinity (Nyberg & Hardell, 2017; Nasim & Kim, 2019). Furthermore, three technical features of 5G are identified, which can potentially increase human EMF exposure further. Firstly, 5G aims to operate at greater frequencies (e.g., 28, 60, and 70 GHz) in addition to the existing lower-frequency bands for cellular communications. However, as the frequency of EMFs increases, so does the rate of signal absorption into human skin. Second, the variation in cell size between mmWaves 5G, 4G, and 3.5G is a key motivating factor in investigating the level of human EMF exposure prior to the deployment of 4G, the 3GPP released 3.9G. There will be a greater number of transmitters in operation in the vicinity of the community due to the use of small cells in mmWaves 5G. These base stations service smaller geographic regions and are consequently closer to human users, resulting in an increased risk of EMF exposure.

Among the three technologies, 5G communication systems feature the smallest cell diameter (200 m) in forming a small-cell network with an Inter-Site Distance (ISD) of 100 m. This distance is also the maximum distance between a user and a mmWaves 5G base station. Third, in 5G, directed beams will be used as a solution to faster signal power attenuation due to operating in high frequency bands. It is important to note that the major reason for utilizing a multiple-antenna system is to enhance antenna gain. Due to the larger concentration of electromagnetic radiation, an EMF has a better chance of penetrating deeper into a human body (Cinel et al., 2008). Most previous research have focused solely on the uplink, with little attention paid to EMF emissions generated by base stations in a 5G network. Figure 2.6 illustrates the geometric difference between two directions of communication which are uplink and downlink in 5G. The uplink in 5G is described as the allocation of power resources among users via User Equipment (UE), i.e., mobile phones. The power resource centralization inside the base stations is the downlink in 5G. Due to the changes in coverage exposure area in mmWaves 5G, the downlink may also pose a hazard to human health (Seungmo Kim & Nasim, 2020).

The changes adopted by 5G mmWave can be summarized as follows:

i. Increased carrier frequency operation

ii. Reduction in cell size (resulting in an increase in the number of base stations)

iii. Greater EMF energy concentration in an antenna beam

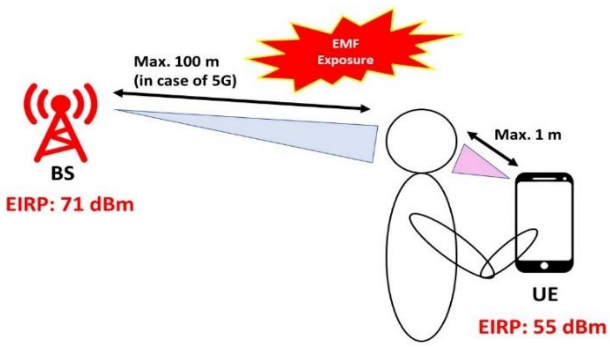


Figure 2.6 Comparison of uplink and downlink in 5G
Source: Seungmo Kim & Nasim (2020)

2.4 Overview on the Effect of Base Station Signal Exposure on Cognitive Performance and Physiological Parameters

Prior studies related employing full text analysis and a more extensive examination of relevant papers was conducted, with studies chosen based on the following inclusion criteria:

- Blind condition (single or double blind)/randomized/balance study with a cross-over design;
- Physiological parameter (blood pressure, body temperature or heart rate);
- Cognitive performance as experimental approach;
- Radiofrequency range related to base station and mobile phone technologies;
- Radiofrequency range related to 5G technologies;

As a result, 22 studies (9 for cognitive performance and 13 for physiological) were selected for inclusion in this review. A summary of the selected studies categorized on cognitive performance and physiological parameters is shown in Table 2.2 and Table 2.4. In the following sections, the main results reported the experimental protocols, materials and methods of each selected study, and parameters which will be compared and discussed in the final section includes the likes of volunteers' inclusion criteria and physiological measures, SAR, RF-EMF, exposure period, etc. This section explains the findings from the literature, followed by an analysis and discussion of the previous findings.

2.4.1 Cognitive Performance Studies

Competent processing of information is of great importance in daily life. This is achieved through neurobehavioral performance, which also is known as cognitive performance. Cognitive performance has been the focus of the research related to exposures to RF-EMF, and the goal has been to determine whether any changes occur in human cognitive function resulting from such exposures. These observations commonly are conducted by applying cognitive tasks with different complexities that facilitate in detection of any intellectual disturbances caused by the RF-EMF exposure. Preece et al. (1999) defined the basis of cognition as knowing or perceiving that involves perception and sensation. To a further extent, cognition consists of mental processes that include multiple tasks dealing with acquiring, storing, retrieving, and manipulating information, thereby making the cognitive function a very crucial feature in supporting all of people's daily activities. For about 120 years, the accepted figures for mean Simple Reaction Times (SRT) for college age individuals have been about 190 ms for light stimuli and about 160 ms for sound stimuli. Many researchers have confirmed that reaction to sound is faster than reaction to light, with mean auditory reaction times being 140-160 ms and visual reaction times being 180-200 ms. Perhaps this is because an auditory stimulus only takes 8-10 ms to reach the brain but a visual stimulus takes 20-40 ms. Reaction time to touch is intermediate, at 155 msec. Differences in reaction time between these types of stimuli persist whether the subject is asked to make a simple response or a complex response. The time between a stimulus and a response is known as reaction time, and it refers to the time it takes to recognize a situation, choose a reply, and initiate action by activating certain muscles (Radák, 2018). SRT tests are among the most basic measures of processing speed, in which individuals simply reply as quickly as possible to the presence of a stimulus (Woods et al., 2015). Slow reaction is defined as the reaction which takes longer time to complete. Figure 2.7 shows that cognitive function can be classified into four distinct features, i.e., motor control, attention, memory, and episodic secondary memory.

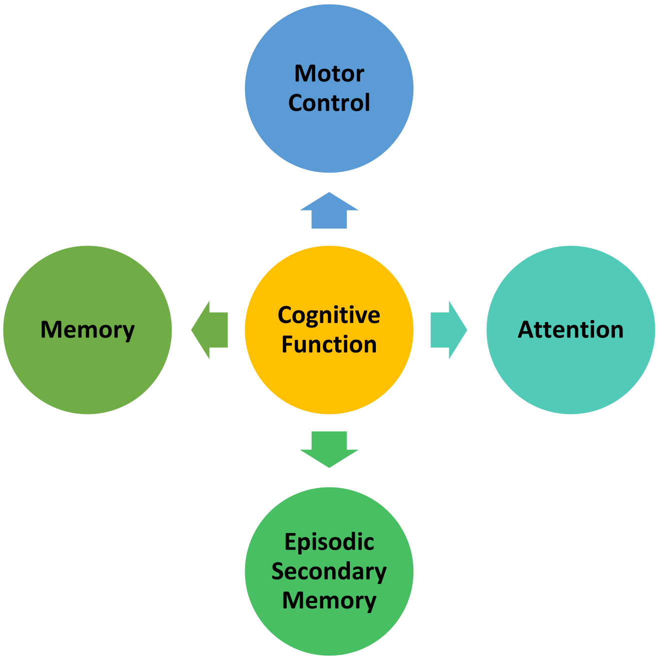


Figure 2.7 Four Main Features of Cognitive Function

Healthy adult volunteers who described feeling a range of symptoms such as headaches in the proximity of RF sources were studied in terms of cognitive function. Many studies have examined how mobile phone radiation affects cognitive performance using behavioral metrics, including response speed and accuracy in a variety of tasks as shown in Table 2.2. Preece et al. (1999) tested the short-term and long-term memory, simple and choice response time, and sustained attention on 36 participants, yielding a total of 15 dependent variables. Using a single-blind, counterbalanced, randomized cross-over method, volunteers were exposed (or Sham exposed) to continuous or pulsed 915

MHz GSM-type transmissions for about 30 minutes. In the Choice Response Time (CRT) task, there was a statistically significant reduction in reaction time when exposed to the continuous signal. The impact was not followed by a decrease in response accuracy, indicating that it was not a speed-accuracy trade-off. SRT were unchanged, and word, number, and image recall, as well as spatial memory, were constant. Exposure to a pulsed GSM signal had no significant impact.

Koivisto et al. (2000) evaluated at 48 individuals and used a variety of cognitive tests. Volunteers were exposed or sham-exposed to a 902 MHz GSM signal using a single blind counter-balanced crossover setup. In basic reaction time and vigilance activities, Koivisto et al. (2000) reported slower reaction times. Furthermore, during exposure, the time required to complete a mental arithmetic subtraction assignment was reduced. However, the effect of exposure on a choice reaction time task similar to Preece et al., (1999) was investigated was far from significant. However, a second research used a similar experimental design to evaluate the effects of GSM RF on the execution of a task with varying working memory demand, Koivisto et al. (2000) and colleagues reported a statistically significant reduction of reaction time when the memory load was particularly demanding. However, a similar group's attempt to validate and expand the findings of both investigations was failed (Koivisto et al., 2000).

Curcio et al. (2004) studied a small number of volunteers (N = 20) using various cognitive tests in a double-blind counterbalanced crossover design. The experimental procedure of this study is similar with the one reported in Eltiti et al. (2007). The subjects were tested on four cognitive tasks, i.e., an acoustic SRT task, a visual search task, an arithmetic descending subtraction task, and an acoustic CRT task. The results indicated that both SRT and CRT tests were reduced during exposure to a 902 MHz GSM signal than Sham exposure. However, an attempt by the same research group to replicate and confirm the finding was not successful, as no significant effects in the same SRT task was observed the CRT test was not performed by Curcio et al. (2005). Other studies also have failed to detect significant effects of mobile phone and base station signal exposure on cognitive performance of human. Cinel et al. (2008) replicated the effect of GSM mobile phone on attention and memory functions presented earlier in Koivisto et al. (2000) with a larger sample size (N=168). Nevertheless, the effect of exposure on any of the six cognitive tests parallel to those performed by Koivisto et al. (2000) was far from significant.

Sauter et al. (2015) found no evidence of a detrimental impact of a short-term EMF of a TETRA hand-held transmitter on cognitive performance in healthy young males, according to the researchers. Computer tests on three distinct elements of attention (i.e. divided attention, selective attention, vigilance) and working memory were used to assess cognitive functions. Recently, Vecsei et al. (2018) observed the short-term RF-EMF exposure from 3G and 4G mobile phone. Similarly, the Stroop test revealed that these signals had no effect on the cognitive functioning of executive function measurements, processing speed, or selective attention.

Psychology Experiment Building Language (PEBL) is a free, open-source software system that allows researchers and clinicians to design, run, and share behavioral tests. At its core, PEBL is a programming language and interpreter/compiler designed to make experiment writing easy. It is cross-platform, written in C++, and relies on a Flex/Bison parser to interpret programming code that controls stimulus presentation, response collection, and data recording. It is designed to be an open system, and is licensed under the GNU Public License 2.0, as shown in Figure 2.8.

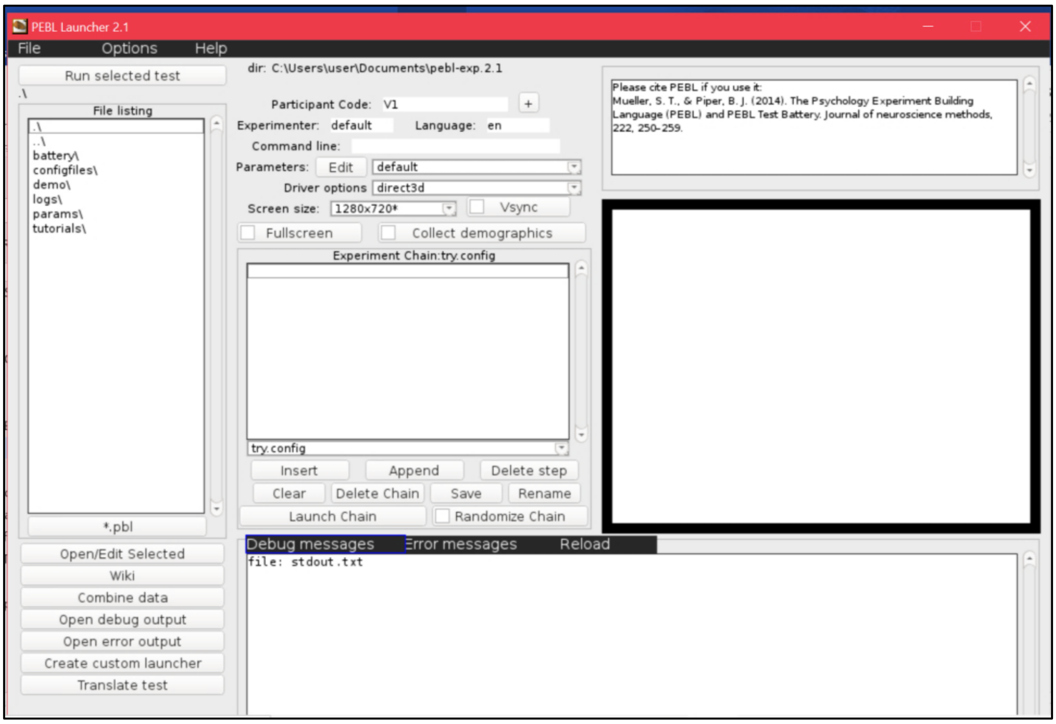


Figure 2.8 PEBL Software

Cognitive functioning is measured using four comprehensive test battery of well-developed and commonly used tests that specifically addressed the following core executive functions: working memory, inhibitory control and cognitive flexibility and the higher-order executive function response planning and problem-solving. The four comprehensive test battery of cognitive function are selected based on recent peer-reviewed journals in Stöckel et al. (2017) and Vecsei et al. (2018).

Table 2.2 Studies on the Effects of EMF Exposure on Cognitive Performance

Study	Exposure Type	Design	Subject	Exclusion Criteria	No and period of exposure assessments	Exp. time (min)	Exposure Setup	SAR	E-field strength	Crossover	Room	Measurements	Results
Koivisto et al. (2000)	902 MHz; 30 min L	Single-blind, counterbalanced, pseudorandom	Healthy subjects: 24M, 24F (age range 18-34, mean age 23.2 years)	No neurological diseases	Single session	30	GSM MP mounted on the subject's head; earphone positioned on left ear with 4cm apart from antenna	NR	-	(ON, OFF)	NR	n-back (0-3)	RT ↓ to targets (3-back Task)
Koivisto et al. (2000)	902 MHz; 60 min R	Single-blind, counter-balanced, randomized	Healthy volunteers: 24M, 24F (age range 18-49, mean age 26 years)	NR	2 sessions with 1 day interval	60	GSM MP mounted on the subject's head; earphone positioned on left ear with 4cm apart from antenna	NR	-	(ON, OFF)	NR	SRT, CRT, SUB, VER, VIG, etc (12 tasks)	SRT ↓; VIG ↓SUB ↓; VIG accuracy ↑
Curcio et al. (2004)	MP GSM 902.40 MHz; 45 min L (peak power of 2 W, equivalent to an average power of 0.25 W)	Double-blind, counter-balanced	Healthy subjects: 10M,10F (age range 22-31 years, mean age 26.4 ±2.86 years)	No MP/ No neurological & psychiatric history/ No medication/ No drug intake/No sleep complaints	3 sessions with interval of ≥48 h between session	45	Helmet (antenna oriented to temporo-parietal areas & microphone oriented towards the mouth), 1.5 cm from the left ear, 2nd MP (off) on the right side of the headBSL: only helmet ON/OFF: helmet & MP	Max value: 0.5 W/kg	-	(BSL, ON, OFF)	NR	SRT, CRT, VS, SUB	SRT ↓ (POST); CRT ↓ (POST)

Table 2.2 Studies on the Effects of EMF Exposure on Cognitive Performance

Study	Exposure Type	Design	Subject	Exclusion Criteria	No and period of exposure assessments	Exp. time (min)	Exposure Setup	SAR	E-field strength	Crossover	Room	Measurements	Results
Regel et al. (2006)	BS UMTS 2140; 45 min, 2 m	Double-blind, randomized	Healthy subjects: 33 EHS (14M 19F; 84 controls (41M 43F) (age range 20- 60 years, mean ± SD=37.7 ± 10.9), BMI 19–30 kg/m^2	No pacemakers, hearing aids, artificial cochlea, drugs/No smoking/No chronic diseases/No pregnancy/ No head injuries, neurologic, psychiatric/ No sleep disturbances /Average alcohol, caffeinated/ No shift workers/No long-haul flights (> 3 h time zone difference) within the last month	3 experimental sessions at 1-week intervals scheduled at the same time of day (~ ± 2 h)	45	The antenna at 1.5 m height and 2 m distance from the subjects, targeting the left side of the body from behind, with a field incidence angle of 25° with respect to the ear-to-ear vertical plane	0 V/m, 1 V/m, or 10 V/m		(0, 1, 10 V/m)	One-side-open chamber shielded with RF radiation absorbers	SRT, CRT, N-back and Visual Selective Attention	None

Table 2.2 Studies on the Effects of EMF Exposure on Cognitive Performance

Study	Exposure Type	Design	Subject	Exclusion Criteria	No and period of exposure assessments	Exp. time (min)	Exposure Setup	SAR	E-field strength	Crossover	Room	Measurements	Results
Curcio et al. (2008)	MP GSM 902.40 MHz; 45 min R (peak power of 2 W (equivalent to an average power of 0.25 W)	Double blind, counter-balanced	Healthy Subjects: 12F,12M (age range 19-36, mean age 28.17±4.78 years)	Regular sleep cycle/No coffee/ No alcohol/ No MP	Weekly interval, conducted 2 sessions between 9.00 & 11.30 a.m.	45	Helmet (antenna oriented to temporo-parietal areas & microphone oriented towards the mouth), 1.5 cm from the left ear, 2nd MP (off) on the right side of the head	Max value: 0.5 W/kg (absolute uncertainty within 20%)	-	(ON, OFF) x (BL, 15, 30, 45 min)	Shielded, sound proof & temperature-controlled room	SRT, sequential finger tapping	None
Cinel et al. (2008)	GSM 888 MHz and CW; Exp 1:45 min L/R; Exp 2:40 min L/R	Double-blind, randomized, counterbalanced	EXP 1: Healthy subjects: 116F 44M (avg age 22.2 years) EXP2: Healthy subjects; 112F 52M (age range: 18-42, avg age 23 years, SD=5)	No MP for at least 1 h before each session	EXP1 & EXP 2: Weekly interval, conducted 2 sessions	EXP 1: 45 EXP 2: 40	MP was fixed on a 'cage/cap' that was mounted on the head	1.4 W/kg (±30%) (SAR average for CW & GSM) 11.2 W/kg (peak of SAR for GSM)	-	(ON, OFF)	NR	Exp 1: n-back (2-3), VIG; Exp 2: Stroop, VS, Sternberg	None

Table 2.2 Studies on the Effects of EMF Exposure on Cognitive Performance

Study	Exposure Type	Design	Subject	Exclusion Criteria	No and period of exposure assessments	Exp. time (min)	Exposure Setup	SAR	E-field strength	Crossover	Room	Measurements	Results
Eltiti et al. (2009)	BS GSM 900 + 1800; UMTS 2020; 50 min, 5 m	Double-blind, counter-balanced, randomized	44 EHS (M=46.1, SD=13.3); 44 control (M=46.1, SD=13.2)	NR	4 sessions, weekly interval at approximately the same time of day (±3 h)	50	GSM signal (combining both 900 & 1800 MHz) & UMTS signal (2020 MHz) over the area where the participant was seated	-	10 mW/m^2	(GSM, UMTS, OFF)	Shielded room with high shielding effectiveness	DSST, DS, mental arithmetic	None
Malek et al. (2015)	BS GSM 945 MHz, 1840 MHz; UMTS 2140 MHZ; 2 m	Single-blind, counter-balanced, randomized	100 EHS; 100 non-EHS	No shift worker	4 sessions	50	BS antenna (Kathrein 800 10046/GSM900/ GMS1800/UMTS) is placed at 1.5 m from the ground & 2 m distance from the subjects	-	1 V/m	(GSM900, GSM1800, UMTS, OFF)	RF shielded room, lined using microwave absorbing sheets	Paired Associates Learning, RT, Rapid Visual Processing, Spatial Span	None
Sauter et al. (2015)	TETRA hand-held transmitter 385 MHz; 2 h 30 min L	Double-blind, balanced, randomized	Healthy subjects: 30M (age range 20–30 years, mean ±SD: 25.4±2.6 years)	No sleep disorder/ Non-smoker/ No drugs & medication/ No implantation	Intervals of 2 weeks, 9 daytime assessment in the afternoon at a fixed	150 /day	Cushioned light weight antenna on the left side of their heads. Each exposure condition was applied at the left side of the head three times	(1) TETRA low level (maxSAR10g=1.5 W/kg) (2) TETRA high level	-	(TETRA 1.5 W/kg, TETRA 6.0 W/kg, OFF) (UMTS/LTE, OFF)	Shielded room with low background field	Test for Attentional Performance (Divided attention, VIG), Vienna Test system (Selective	None

Table 2.2 Studies on the Effects of EMF Exposure on Cognitive Performance

Study	Exposure Type	Design	Subject	Exclusion Criteria	No and period of exposure assessments	Exp. time (min)	Exposure Setup	SAR	E-field strength	Crossover	Room	Measurements	Results
				s & tattoos on head	time frame			(maxSAR10g=6.0 W/kg)				attention) & n-back	
Vecsei et al. (2018)	MP UMTS WCDMA 1947; LTE 1750; 20 min L	Double-blind, counter-balanced, randomized	UMTS: Healthy subjects 20F 14M (aged 20 ± 3 years) LTE: 13F 13 M (aged 21 ± 3 years)	No smoking/ No alcohol/ No coffee/ Moderate MP use	2 sessions with 1 week interval between session at 8 am - 6 pm	20	Patch antenna mounted in a position mimicking the normal use of an MP: the centre of the patch antenna was near the exit of the ear canal, above the tragus, at 7 mm	1.8 W/kg	-	(UMTS/ LTE/OFF)	Dimly lit room	Stroop test (executive function, processing speed, selective attention)	None

BSL/BL – Baseline, BS-Base Station, CRT – Choice Reaction Time, DSST - Digit Symbol Substitution Task, DS – Digital Span, EXP – Experiment, F – Females, GSM - Global System for Communication, h – hours, L – Left, LTE - Long-Term Evaluation, Max- Maximum; min – minutes, MP – Mobile Phone, M – Males, NR – Not Reported, POST – Post Exposure, R – Right, SAR - Specific Absorption Rate; SD – Standard Deviation, SRT – Simple Reaction Time, SUB – Subtraction Time, TETRA - Terrestrial Trunked Radio, UMTS - Universal Mobile Telecommunications System, W -Watts; VER – Verification Time, VIG – Vigilance, VS – Visual Search, ↑ - increased, ↓ - decreased.

2.4.2 Physiological Parameters (Heart Rate, Blood Pressure and Body Temperature)

It is hypothesized that the subjects' blood pressures would increase if they experienced increased body temperatures and heart rates. Blood pressure is a measure of the force of blood pushing against the walls of arteries and it is characterized by two values, i.e., Systolic Blood Pressure (SYS) and Diastolic Blood Pressure (DIA). SYS refers to the pressure inside the artery when the heart contracts and pumps blood through the body; DIA is the pressure inside the artery when the heart is relaxed and filled with blood. Hypertension has been identified as a main contributing risk factor for coronary heart disease (heart attack), stroke (brain attack) and renal failure, and this is supported by strong studies that have established direct relationship between blood pressure and cardiovascular disease. The Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure (JNC7) defined several stages of blood pressure that lead to hypertension, as shown in Table 2.3. The prehypertension stage that is defined in intervening levels, i.e., 120 to 139 and 80 to 89 millimeters of mercury (mmHg) refers to a group that has higher health risk.

Table 2.3 Classification of Normal Blood Pressure and Hypertension

BP Classification	**BPS, mm Hg***	**BPD, mm Hg***
Normal	<120	<80
Prehypertension	120-139	80-89
Hypertension	≥140	≥90

**Classification determined by the higher BP category. The term 'mm Hg' (millimeters of mercury) is a unit used for blood pressure.*

The WHO, the International Society of Hypertension, and the European Society of Hypertension/European Society of Cardiology have recommended the classification of normal blood pressure and hypertension (Pickering et al., 2005). They also refer to

<120/<80 as optimal, 120 to 129/80 to 84 as normal, and 130 to 139/85 to 89 as high normal. If blood pressure increases beyond the normal range due to the exposure to 5G radiation, there would be an association with an increasing hypertension risk, and this must be further confirmed by medical experts.

A few studies have testified that mobile phone (Tahvanainen et al., 2004; Oftedal et al., 2007; Cinel et al., 2008) and base station exposures (Eltiti et al., 2007; Malek et al., 2015) did not induce physiological effects on blood pressure and heart rate. When both self-reported EHS and Non-EHS groups were exposed to 3G Wide Code Division Multiple Access (WCDMA) mobile phones in Kwon et al. (2012). Kwon et al. (2021) discovered no significant physiological changes in heart rate, heart rate variability, and respiration rate. Moreover, WCDMA was found to not affect the heart rate, respiration rate, heart rate variability, or subjective symptoms in adults, according to Choi et al. (2014). Malek et al. (2015) also reported that base station transmissions had no significant short-term impacts on heart rate, blood pressure, or body temperature. Andrianome et al. (2017) observed the effects of continuous RF-EMF exposure from signals emulating GSM 900 MHz, GSM 1800 MHz, DECT and Wi-Fi 2.45 GHz signals. Similarly, it was discovered that these signals had no effect on the EHS participants' autonomic nervous system, which included blood pressure and heart rate variability. Masrakin et al. (2019) observed that wearing textile antennas operating at 2.45GHz had no effect on adults' blood pressure, heart rate, or body temperature. In a more recent study, Huang et al. (2022) discovered that physiological parameters Blood Pressure (BP), Heart rate variability (HRV) and Heart Rate (HR) at base station GSM 900 MHz, GSM 1800 MHz, 2100 MHz no changes. Table 2.4 presents the physiological parameters findings in the previous studies.

Table 2.4 Findings Comparison of Previous Studies with Respect to Physiological Parameters

Study	Exposure Type	Design	Subject	Exclusion Criteria	No and period of exposure assessments	Exposure time (min)	Exposure Setup	SAR	E-field strength	Crossover	Room	Measurements	Results
Thavanainen et al. (2004)	MP; GSM 900 and 1800; 35 min	Double-blind, randomized, placebo-controlled	Healthy subjects: 16F 18M (avg age of 38.8 years, SD=10.3, mean BMI of 23.5 (SD 2.2) kg/m^2	NR	2 sessions, weekly interval at between 1 pm - 3 pm	35	Dual band MP & a physically identical but inactive MP were located on a plastic head helmet. RF field recording antenna placed around 20 cm from the active MP	900 MHz: 1.58 W/kg 1800MHz: 0.70 W/kg	-	(GSM 900, GSM 1800, OFF)	EMF shielded laboratory	Physiological parameters (BP and HR)	No effect on BP and HR
Regel et al. (2006)	BS UMTS 2140; 45 min, 2 m	Double-blind, randomized	33 EHS (14M 19F; 84 control (41M 43F) Age range 20-60 years, mean ± SD 37.7 ± 10.9, BMI 19–30 kg/ m^2	No pacemakers, hearing aids, artificial cochlea, drugs/No smoking/No chronic diseases/No pregnancy/ No head injuries, neurologic, psychiatric/ No sleep disturbances /Average alcohol, caffeinated/ No shift workers/Nol long-haul flights (> 3	3 experimental sessions at 1-week intervals scheduled at the same time of day (~ ± 2 h)	45	The antenna at 1.5 m height and 2 m distance from the subjects, targeting the left side of the body from behind, with a field incidence angle of 25° with respect to the ear-to-ear vertical plane	-	0 V/m, 1 V/m, or 10 V/m	(0, 1, 10 V/m)	One-side-open chamber shielded with RF radiation absorbers	5 subjective well-being symptoms	No effect on subjective symptoms

Table 2.4 Findings Comparison of Previous Studies with Respect to Physiological Parameters

Study	Exposure Type	Design	Subject	Exclusion Criteria	No and period of exposure assessments	Exposure time (min)	Exposure Setup	SAR	E-field strength	Crossover	Room	Measurements	Results
				h time zone difference) within the last month									
Oftedal et al. (2007)	MP GSM 900 MHz; 30 min	Double-blind, randomized, Sham-controlled	Healthy subjects: 12M 5F (age range 20-58 years, mean=39)	No MP/No other serious health conditions/ No frequent headache/	Max takes 4 session, ≤ 2 days interval between sessions	30	Antenna mounted symmetrically at the sides of the subject's head. Wooden bars restricted the sideways movements of the head. Antenna positioned 8.5 cm from the head	Spatial peak SAR_{1g}: 1.0 W/kg SAR_{10g}: 0.8 W/kg	-	(ON, OFF)	Control room next to the shielded exposure room	4 subjective well-being symptoms Physiological parameters (BP and HR)	No effect on subjective symptom, BP, and HR
Cinel et al. (2008)	GSM 888 MHz and CW; 40 min	Double-blind, randomized, counterbalanced	Adults (116 M 330F, age range 18 to 42 years, mean=23, SD=4.4)	No MP	Weekly interval, conducted 2 sessions	40	MP attached to a cap that was then positioned on participant's head	1.4 W/kg (±30%) (SAR average for CW and GSM) 11.2 W/kg (peak of SAR for GSM)	-	(ON, OFF)	NR	5 subjective well-being symptoms	No consistent effect on subjective symptoms
Eltiti et al. (2009)	BS GSM 900 + 1800; UMTS 2020; 50 min, 5 m	Double-blind, counter-balanced, randomized	44 EHS (M=46.1, SD=13.3); 44 control (M=46.1, SD=13.2)	NR	4 sessions, weekly interval at approximately the same time of day (±3 h)	50	GSM signal (combining both 900 & 1800 MHz) & UMTS signal (2020 MHz) over the area where the participant was seated	-	10 mW/m^2	(GSM, UMTS, OFF)	BS GSM 900 + 1800; UMTS 2020; 50 min, 5 m	6 VAS subjective well-being symptoms, 57 EHS symptoms, EMF Perception, Physiological parameters (BP and HR)	No effect on subjective symptoms, EMF perception, BP, HR

Table 2.4 Findings Comparison of Previous Studies with Respect to Physiological Parameters

Study	Exposure Type	Design	Subject	Exclusion Criteria	No and period of exposure assessments	Exposure time (min)	Exposure Setup	SAR	E-field strength	Crossover	Room	Measurements	Results
Wallace et al. (2010)	BS TETRA 420; 50 min, 5 m	Double-blind, randomized, counter-balanced	48 EHS (61%F, mean ± SD = 42 ± 16; range, 18–73); 132 controls (51%F, mean ± SD =41 ± 19; range, 18–80)	No brain injury, diagnosis of epilepsy, claustrophobia, diagnosis, treatment for a mental disease/ No pacemaker/ No physical impairment or illness/ No medication	3 sessions (inc open provocation session) with interval ≤ 1 week apart & same time of day	50	Participants seated 4.95 m from antenna of the BS & use TETRA signal release 1 [specification 390 392-2; European Telecommunications Standards Institute (ETSI)	271 μW/kg	10 mW/m^2	(ON, OFF)	Screened semi-anechoic chamber	6 VAS subjective well-being symptoms, 57 EHS symptoms, EMF Perception, Physiological parameters (BP and HR)	No effect on subjective symptoms, EMF perception, BP, and HR (double-blind) Have effects on subjective symptoms (exposure is known)
Kwon et al. (2012)	MP 3G WCDMA 1950; 64 min	Double-blind, counter-balanced, randomized	17 EHS 8M 9F (mean=30.1 SD=±7.6); 20 control 11M 9F (mean=29.4 SD=±5.2)	No caffeine/ No Smoke/ No exercise/ Enough sleep	2 sessions with 1-10 days interval between sessions	64	Dummy MP containing a WCDMA module within a headset placed on the head	1.57 W/kg	-	(ON, OFF)	Laboratory and other electrical devices were unplugged except for instruments	8 subjective well-being symptoms Physiological parameter (HR)	No effect on subjective symptoms and HR for EHS, Non-EHS subjects
Choi et al. (2014)	MP 3G WCDMA 1950; 64 min	Double-blind, randomized	26 adults 13M 13F (mean=28.4 SD=±5.1);	No caffeine/ No smoker/ No exercise before day experiment	2 sessions with 1-10 days interval between sessions	64	Dummy MP containing a WCDMA module within a headset placed on the head	1.57 W/kg	6.9 V/m	(ON, OFF)	Laboratory & other electrical devices	8 subjective well-being symptoms Physiological parameter (HR)	No effect on subjective symptoms and HR

Table 2.4 Findings Comparison of Previous Studies with Respect to Physiological Parameters

Study	Exposure Type	Design	Subject	Exclusion Criteria	No and period of exposure assessments	Exposure time (min)	Exposure Setup	SAR	E-field strength	Crossover	Room	Measurements	Results
			26 teenagers 13M 13F (mean=15.3 SD=±0.7)								were unplugged except for instruments		
Malek et al. (2015)	BS GSM 945 MHz, 1840 MHz; UMTS 2140 MHZ; 2 m	Single-blind, counter-balanced, randomized	100 EHS; 100 non-EHS	No shift worker	4 sessions	50	BS antenna (Kathrein 800 10046/GSM900/GMS1800/UMTS) is placed at 1.5 m from the ground floor & at 2 m from the subjects	-	1 V/m	(GSM900, GSM1800, UMTS, OFF)	RF shielded room, lined using microwave absorbing sheets	Physiological parameters (BT, BP, and HR)	No effect on physiological parameters (BT, BP and HR)
Andrianome et al. (2017)	BS GSM 900, GSM 1800, DECT & Wi-Fi 2.45 GHz; 5 min	Double-blind, counter-balanced	10 EHS (8F 2M age range 35-63 years, mean age: 48 ± 10) ; 25 non-EHS, mean age: 46 ± 10	No alcohol/ No coffee for the 24 hours prior to & during the study; None EHS participants were on medication	2 session intervals of ≤ 1 week	5	No external EMF sources were allowed, the exposure consisted of a series of EMF signals emitted from a generator & a horn antenna	-	1 V/m	(GSM900, GSM1800, DECT, Wi-Fi, OFF)	Shielded chamber	Autonomic nervous system that includes BP and HR	No effect on physiological BP and HRV

Table 2.4 Findings Comparison of Previous Studies with Respect to Physiological Parameters

Study	Exposure Type	Design	Subject	Exclusion Criteria	No and period of exposure assessments	Exposure time (min)	Exposure Setup	SAR	E-field strength	Crossover	Room	Measurements	Results
van Mooselaar et al. (2017)	BS GSM 900, GSM 1800, UMTS, DECT & Wi-Fi 2.45 GHz; 150 min	Double-blind, randomized, controlled	42 EHS (32F 10M, mean age 55 years, range 29–78)	Inability to complete the administered questionnaires, communicate with the study assistant/ self-reported symptoms exceeded 15 min	Testing group then follow up at 2 months interval	150	2 custom-made mobile exposure units. Different types of non-ionizing EMF can be generated a) radiofrequency EMF (“RF-unit”): GSM 900, GSM 1800, cordless MP (“DECT phone”, 1880–1900 MHz), UMTS & Wi-Fi; and b) extremely low-frequency magnetic fields (“ELF-unit”)	-	Max: 6 V/m (average exposure levels at the upper body level)	GSM 900, GSM 1800, UMTS, DECT, Wi-Fi	Home and another location where they felt comfortable	EMF Perception, symptoms	No effect on EMF perception but have effects on EHS symptoms
Boger et al. (2018)	BS GSM 900, GSM 1800, UMTS, DECT & Wi-Fi 2.45 GHz; 6 h	NR	7 EHS (4F 3M)	No applicants with knowledge on their personal EMF exposure/No depression, anxiety disorder, burnout, psychosis, chronic fatigue syndrome, fibromyalgia	4 sessions with intervals of 6 h prior to filling out the diaries in the morning, afternoon & evening	360	EME-SPY 121 exposimeters (Satimo, Cortaboeuf, France) worn at the hip in a camera bag.	-	2.5 V/m	GSM 900, GSM 1800, UMTS, DECT, Wi-Fi	At home inside, at home outside, at work or educational institution, elsewhere, travelling	EMF Perception, EHS subjective symptoms	Have effects on EHS symptoms

Table 2.4 Findings Comparison of Previous Studies with Respect to Physiological Parameters

Study	Exposure Type	Design	Subject	Exclusion Criteria	No and period of exposure assessments	Exposure time (min)	Exposure Setup	SAR	E-field strength	Crossover	Room	Measurements	Results
Masrakin et al. (2019)	Wearable textile antenna 2.45 GHz; 50 min	Single-blind, counter-balanced, randomized	Healthy subjects 10M 10F (age range= 23-31 years, avg age 25 years & SD = 2.4, body weights 19 -26 kg/m^2)	No artificial cochlea, hearing aids, pacemakers/ No smoking/No alcohol/ No caffeinated drinks/ No psychiatric disease/ No drug in the previous 6 months/ No long-haul flight for >3 h of different time zones/No shift workers/ Matched menstrual cycle	2 sessions	50	Mounted Tx onto the upper right arm. Both the Tx & Rx antennas were both vertically oriented (with the radiator placed on the top section) when mounted on the subject's body. Rx was mounted on the left chest of the subjects	For 10g SAR TM: (2.88 W/kg) 10g SAR TP: 0.35 W/kg)	-	(ON, OFF)	RF-shielded room	10 subjective well-being symptoms, Physiological parameters (BT, BP & HR)	No effect on subjective symptoms and physiological parameters (BT, BP and HR)
Huang et al., (2022)	BS; GSM 900, 1800 GSM, 2100 MHz	Double	58 EHS adults 92 Non-EHS adults	No cancer, claustrophobia, pregnancy, coronary heart disease, or psychologic	2	30	2G base stations of 900 MHz GSM and 1800 MHz GSM as well as 3G base stations of 800 MHz GSM and 2100 MHz GSM. Peak power of each band was set at 0.25 W/m2 for an average combined power of 1 W/ m^2	0.25 mW/m^2	-	GSM 900, GSM 1800, UMTS, DECT, Wi-Fi	Anechoic laboratory	Physiological parameters (BP, HRV and HR)	None

Table 2.4 Findings Comparison of Previous Studies with Respect to Physiological Parameters

Study	Exposure Type	Design	Subject	Exclusion Criteria	No and period of exposure assessments	Exposure time (min)	Exposure Setup	SAR	E-field strength	Crossover	Room	Measurements	Results
				al disorders. No possessed catastrophic illness certification issued by the National Health Insurance or had pacemakers.									

AVG –Average, BP - Blood Pressure, BS-Base Station, BT – Body Temperature, CW – Continuous Wave, EHS– Electromagnetic Hypersensitivity, ELF – Electromagnetic Low Frequency, F – Females, H – Hours, HR – Heart Rate, HRV – Heart Rate Variability, LTE - Long-Term Evaluation ; M – Males, Min – Minutes, MP-Mobile Phone, TETRA - Terrestrial Trunked Radio, UMTS - Universal Mobile Telecommunications System, NR – Not Reported, Rx – Received Antenna, SD – Standard Deviation, Tm – Textile Monopole Antenna, Tp – Textile Patch Antenna, Tx – Transmitted Antenna, VAS – Visual Analogue Scale.

2.5 Machine Learning Techniques

Machine learning is one of the fields of artificial intelligence. It uses different statistical methods so that the computers able to "learn" from data, without being programmed to perform certain tasks (Awad & Khanna, 2015). This means that the computer has learned from experience as humans and animals learn. The computer uses computational and statistical methods to learn information from data. The selected machine learning technique depends on the type of data available. Data that is annotated or tagged by domain experts, is called labelled data. Supervised learning uses the input data with known responses (labels) to train a model to generate predicted responses for new data, meanwhile unsupervised learning detects data patterns or structures to draw inferences based on input data that is not labelled (MathWorks, 2016). Third technique is semi-supervised learning, an integration between supervised and unsupervised learning. It includes a data type that is mostly unlabelled data. The labelled data is used to train a model where this model is used to predict labels to the unlabelled data. Then the original labelled data and the newly labelled are used to enhance the model design (Lee, 2013). Lastly, Reinforcement learning is a machine learning technique that uses trial and error to find the relation between the outcome and sequence of trials that lead to a successful outcome (IBM, 2019).

Rather than discussing predictors or covariates for an outcome or dependent variable, the machine learning literature refers to features that may be used to classify outputs or targets. Further, supervised learning to predict a categorical outcome is referred to as classification in the machine learning literature (Tammy Jiang et al., 2020; Olsen et al., 2020). It is optimal to characterize supervised learning and unsupervised learning according to class definitions. In supervised learning, classes are identifiable, and class boundaries are clearly delineated in the provided (training) dataset. Learning takes place using these classes, or class labels, leading to its designation as a classification method (Suthaharan, 2015). The machine learning classifiers cover a variety of parametric and non-parametric classification algorithms such as Logistic regression, boosted and bagged decision trees, including AdaBoost, LogitBoost, GentleBoost, RobustBoost, Naïve Bayes classification, K-Nearest Neighbours (KNN), Discriminant analysis (Linear and Quadratic), SVM binary or multiclass classification. Machine learning classifiers encompass a diverse array of both parametric and non-parametric classification algorithms for examples, logistic regression, boosted and bagged decision trees (such as AdaBoost, LogitBoost, GentleBoost, RobustBoost), Naïve Bayes classification, KNN, Discriminant analysis (both Linear and Quadratic), and SVM for binary or multiclass classification. The Classification Learner tool in MATLAB 2022 proves valuable for performing various tasks like interactive data exploration, feature selection, specifying cross-validation schemes, model training, and results evaluation (Ciaburro, 2017). While supervised machine learning techniques exhibit promise in controlled experiments with standardized imaging protocols, their efficacy may decline when applied to new images acquired under slightly different conditions. These techniques assume that both training and test datasets are random samples drawn from the same distribution (de Bruijne, 2016; Książek et al., 2019). The supervised learning approach, a prominent method, involves training sets of data with labelled classes to create models, known as classifiers, based on features or attributes (Asmita Singh et al., 2017).

2.5.1 Machine Learning Mechanisms for Prediction Models and Feature Selection Techniques Using Data from Weak Radiofrequency Radiation Effect on Human

In data science, description, prediction, and causal inference are three key tasks that address different aspects of understanding and analysing data (Hernán et al., 2019). Description entails summarizing and presenting the primary characteristics of a dataset, aiding in the identification of patterns, trends, and overall insights within the data. Prediction, on the other hand, revolves around forecasting or estimating future outcomes based on historical data patterns.

Machine learning starts with access and explore data in which it is explored data availability, how to collect a convenient sample. If the sample is an accurate representation, it will lead to building a good model with high prediction accuracy. Then, data preparation which starts by checking if there are any outliers which are defined as data points that lie outside the rest of the data. Outliers should be checked if they could be safely removed, or they indicate some features that worth studying. Missing values should be checked if they can be ignored or substituted by approximated values. Inconsistent and redundant values should be removed (Simplilearn, 2020).

Features are the extracted data properties to capture the main patterns in data. These features contribute most to the accuracy of the prediction variable. Features can describe data statistically as; mean, median, and standard deviation. It can be related to signal frequency contents such as signal bandwidth, fundamental frequency component, and power spectrum. In step 3, features are extracted from data by analysing data with a deep understanding of its features. This the most important step in machine learning as it converts data from numeric values to valuable information. Then, a predictive model is designed and developed in step 5. This process starts with dividing the data into two parts; 1. To train the model and 2. To validate data. The validation set is used to detect the accuracy of the developed model. If the data set is large; Holdout Validation is recommended and can be done by holding out a certain percentage to one of the two parts of data. Cross-Validation is appropriate for small data set as it maximizes the data used to train the model. Next, machine learning algorithm is chosen by iterating with various algorithms to achieve the highest prediction accuracy; the optimized machine learning algorithm is identified by trial and error. The most used classification algorithms are Discriminant Analysis (DA), Naïve Bayes Classifier (NBC), Decision Trees (DT), SVM, Nearest Neighbour Classifiers (NNC) and Ensemble Classifiers (EC) (MathWorks, 2019). For step 6, the model is optimized. where the performance of the model is improved by tuning model parameters. In machine learning, some parameters have a high impact on the results and vice versa. There are several techniques to perform parameter tunings such as Grid Search and Random Search (Filion, 2019). Model can be improved by adding more data to train the model, aiming for a better accuracy. This can be done until a point of overfitting is reached where the error increases and model accuracy decrease. Model can be enhanced by adding more features into it or if it is noticed signs of overfitting, a feature reduction technique can be used such as Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA) (MathWorks, 2019). If some data patterns are highly important to be correctly classified, a cost function can be introduced that sets high penalties to undesired misclassifications. The final step is to deploy the model into production system by integrating to the desktops, IT systems. At this stage, it is needed to create a function that takes raw data from sensors and reform the data as input to the model design (MathWorks, 2019).

The research using data from weak radiofrequency radiation effect on human described, focusing on machine learning mechanisms for prediction models and feature selection techniques using data from weak radiofrequency radiation effects on humans, is indeed new and holds potential for the broader field of bioelectromagnetics research. This approach can bridge the gap between traditional bioelectromagnetics research and advanced data analytics, offering new insights and methodologies. The prospect of applying machine learning to epidemiological studies involving human populations is particularly significant. Machine learning techniques can potentially identify patterns and associations in large datasets, helping researchers uncover links between weak radiofrequency radiation exposure and health outcomes in real-world scenarios (Halgamuge & Davis, 2019).

2.5.2 Supervised Machine Learning Classification Algorithms

In the field of machine learning, various algorithms are used to classify data according to the required response. machine learning algorithms can be classified into two significant methods: supervised machine learning, unsupervised machine learning and Regression Analysis. Classification and regression methods are known as supervised machine learning, while clustering and association methods are known as unsupervised learning. An approach that lies in between supervised and unsupervised machine learning method is called semi-supervised learning. The most practical applications utilize supervised machine learning algorithms (classification algorithms) for prediction.

Supervised machine learning takes a known set of input variables, x (the training set), the known responses to the data or output variable (Y), and an algorithm that learns the mapping function or trains a model from the input to the output variables, $Y = f(X)$. In this method, all the data are labelled, and the algorithms attempt to figure out how to predict the output from the input data. Thus, the mapping function can be approximated adequately. With this, a classifier (machine learning algorithm) can predict the output variables (Y) for that for new input data (x). Since the outcome of the training data are known, it is known as supervised machine learning technique. The algorithm iteratively makes predictions on the training data and learning ends when the algorithm delivers a satisfactory level of performance (Halgamuge, 2020).

2.6 Normalization Techniques in Machine Learning

Normalizing data is the first stage before applying any machine learning technique as it has very important role on the performance of the model (Jain et al., 2018; Shahriyari, 2019). Many studies have shown the importance of data normalization for different data sets on the accuracy of the model results (Shahriyari, 2019; Raju et al., 2020). Data normalization is considered to be very crucial at the stage of data pre-processing as it improves the predictive accuracy (Raju et al., 2020). If there are high inconsistency and magnitude mismatch problems in the data, hybrid normalization could be implemented which means two or more normalization techniques can be used (Zhou et al., 2020). When the selection of normalization techniques is a function of data characteristics, this is called dynamic but if the technique is fixed for all classification problems, it is called static. If the data complexity is high, dynamic selection of normalization technique can be implemented based on data characteristics (Jain et al., 2018). They used the most popular normalization two techniques; min–max and z-score to study the effect of proper normalization on models' prediction accuracy. In Liang et al. (2018), the two data sets used in gene expression data analysis were normalized by scaling to 0–1. In Shahriyari, (2019), three normalization techniques have been used which are scaling, vector normalization and z-score to discover the gene expression that predicts the survival in patients with colon tumour. In Isaksson et al. (2020), the authors used four normalization techniques: histogram normalization, generalized scale normalization, generalized ball-scale normalization and custom normalization to check the impact on the accuracy of radiomic machine learning models. Andrew et al. (2016) used five normalization techniques: Logarithmic Sum Squared Voltage value (RLSSV), Relative Logarithmic Voltage value (RLV), Relative Sum Squared Voltage (RSSV), Relative Voltage value (RV), and Fractional Voltage Change value (FVC) to build a classifier for early fire detection. In Vijayasarveswari et al. (2020), the authors used the five normalization techniques implemented in Andrew et al., (2016) and another four normalization techniques which were Decimal Scaling (DS), Z-score (ZS), Min-Max (MM) and Mean & Standard Deviation (MSD) for early breast cancer size prediction. In (Raju et al., 2020), the authors used seven normalization technique: Min-Max Scaler, Standard Scaler, Robust Scaler, Quantile Transformer, Scale, power Transform and Max-Abs Scaler to conclude that all the used data normalization techniques enhanced the accuracy of three machine learning classifiers; Radial Kernel SVM, sigmoid SVM and KNN. The most relevant normalization techniques to this research are listed in Appendix B.

2.7 Overview of the Machine Learning Algorithms for Bioelectromagnetics

Previous research involved utilizing full text analysis and conducting a thorough review of pertinent literature. The selection of studies was determined by adhering to the following inclusion criteria:

- RF-EMF exposure assessment;
- Bioelectromagnetics;
- Machine learning;
- Prediction models;
- Health effect;
- Human adults;

As a result, few studies were selected for inclusion in this review. A summary of the selected studies categorized on machine learning classifier performance as displayed in Table 2.5. In the following sections, the main results reported the experimental dataset, feature selection technique, classifier, validation measurement, prediction output and results of each selected study, and limitation which will be compared and discussed in the final section. This section explains the findings from the literature, followed by an analysis and discussion of the previous findings.

The initial investigation by Halgamuge and Davis (2019) utilized a relatively small dataset comprising 8 attributes related to RF-EMF and 169 experimental case studies. They employed KNN and Random Forest (RF) algorithms to assess the impact of RF-EMF on plants, yielding a notable prediction accuracy of 91.17% with the KNN machine learning model. In a subsequent study conducted by Halgamuge (2020), it was found that the Random Forest classifier exhibited superior classification performance in comparison to other models. The study mentions a low sample size of 169 reported experimental case studies, limiting the statistical power and generalizability of the findings. A larger and more diverse dataset would strengthen the reliability of the machine learning models and enhance the robustness of the conclusions. The raw data extracted from

Table 2.5 Findings Comparison for Prediction Models and Feature Selection Techniques Using Data from Weak Radiofrequency Radiation Effect

Reference	Dataset	Feature Selection Technique/ Pre-Processing Techniques	Features Selected	Classifier/Machine Learning Approach	Validation Measurement	Best Classifier Performance	Prediction Output	Result	Presence of Hybrid Dataset	Limitations
(Halgamuge & Davis, 2019)	8 attributes of RF-EMF and 169 experimental case studies	Presence	Plant type, frequency, SAR, power flux density, electric field strengh	KNN and RF	Accuracy	KNN	Plant sensitivity and prediction of RF-EMF effects on plants	kNN with accuracy of 91.17%	No	Low sample size of 169 reported experimental case studies.
(Halgamuge, 2020)	1127 human and animal cells response to RF-EMF	Domain knowledge or expert knowledge, PCA and Chi-squared feature selection method	frequency, SAR, exposure time, and SAR×exposure time (impact of accumulated SAR within the exposure period)	Random Forest, Bagging, J48, SVM (Linear Kernel), Jrip, Decision Table, BayesNet, NB, LR	Accuracy, error rate, precision, sensitivity or recall, specificity, area under the ROC Curve, and precision-recall	Random Forest	To know whether or not the non-thermal low power RF-EMF's impact on the cellular response was observable (presence or absence)	RF shows AUC = 0.903	No	-Generalization to In-Vitro Experiments -Low sample size (1127 reported experimental case studies)

(Anushikha Singh et al., 2020)	Microscopic image of brain of Drosophila	Presence	Microscopic segmented brain image of drosophila	SVM, NB, RF, and ANN	Accuracy, true positive, true negative, false positive, false negative, f1-score, recall and precision	SVM	Segmentation of brain image area	Accuracy: 94.66 % using SVM classifier for classification of EMF exposed/without exposure drosophila	No	The findings may not capture the complexity of the human brain and the potential neurological effects of EMF on higher-order cognitive functions.

Support Vector Machine (SVM), Probabilistic Neural Network (PNN), Naïve Bayes (NB), Principal Component Analysis (PCA), K-nearest Neighbor (KNN), Random Forest (RF), Artificial Neural Network (ANN) different articles may exhibit heterogeneity in terms of experimental protocols, plant types, and RF-EMF exposure conditions. This variability can introduce noise and confounding factors, making it challenging to draw clear conclusions about consistent biological impacts. The research aims to predict cellular responses without in-vitro laboratory experiments, which might raise concerns about the generalization of the findings to real-world scenarios. In the investigation conducted by Anushikha Singh et al. (2020), an image dataset of Drosophila brains was employed. The study employed various machine learning models, including SVM, Naïve Bayes (NB), Random Forest (RF), and Artificial Neural Network (ANN). Notably, the SVM classifier achieved a high accuracy rate of 94.66% in the classification of the brain images. Drosophila melanogaster is a simpler organism compared to humans, and the study focuses on brain morphology changes. The findings may not capture the complexity of the human brain and the potential neurological effects of EMF on higher-order cognitive functions and the study may not fully replicate the real-world exposure conditions that humans experience with mobile phones and cell towers. Addressing these limitations would contribute to a more nuanced and comprehensive understanding of the potential effects of EMF on health effect from radiation based on the study's findings.

To establish computational simplicity and potential applicability in embedded applications, it is essential to explore the most suitable features and classification algorithm (Andrew et al., 2016). Another widely employed technique for enhancing data quality and improving the accuracy of machine learning models is the utilization of MSFS classifiers. This method has been applied across diverse applications to boost classifier accuracy (Elkhouly et al., 2023). MSFS involves both data pre-processing (managing numerical features, addressing missing values, and handling outliers) and data processing (encompassing tasks like feature normalization, dimension reduction, selection classifiers, extraction, and fusion) (Halim et al., 2022). The proposed MSFS method offers a solution to address the limitations of single-stage feature selection methods, as highlighted by Vijayasarveswari et al. (2020). This approach proves to be a comprehensive strategy for mitigating the drawbacks associated with traditional feature selection methods.

The proposed MSFS is computationally efficient because it only requires simple multiplication operations in the alternative optimization process. The data was compare it with several representative multi-label feature selection methods on each single-view feature space or the concatenated multi-view feature space to validate the effectiveness of MSFS which are Convex Semi-Supervised Multi-Label Feature Selection (CSFS), Sub-Feature Uncovering With Sparsity (SFUS), Multi-Label Informed Feature Selection (MIFS), Manifold Regularized Discriminative Feature Selection (MDFS), Block-Row Sparse Multi-View Multi-Label Learning (F2L21F) and Multi-Label Sparse Feature Selection (MLSFS). With different numbers of selected features, MSFS mostly outperforms other methods on each single-view feature space, which indicates that MSFS is effective in finding more discriminative features for different views by jointly exploiting the complementary information provided by multiple view features and the correlations among multiple class labels. Halgamuge (2020) mentioned that they proposed feature selection technique using domain knowledge or expert knowledge, PCA technique, and the Chi-squared feature selection method to know whether or not the non-thermal low power RF-EMF’s impact on the cellular response was observable (presence or absence). Their evaluation measures of knowing the performance of classifiers was by using the confusion matrix in order to avoid accuracy inconsistency. And from input of confusion matrix, they analysed by calculating on accuracy (PCC—Percent Correct Classification), error rate (RMSE), precision, p is the percentage of predictive items which are correct, p = TP = (TP + FP), sensitivity or recall (true positive rate), TP = (TP + FN), 1− specificity (false positive rate, FP/(FP + TN), area under the ROC Curve, and precision-recall (PRC Area). Six key features extracted from the analysis is specie, frequency of weak RF-EMF, SAR, exposure time, SAR×exposure time, cellular response (presence or absence) while removed the other features. They used 10 classification algorithms to make the best predictions for the given dataset and selected top seven (Random Forest, Bagging, J48, Decision Table, BayesNet, kNN, and JRip) show classification algorithm that were performed in terms of Area under the ROC Curve and accuracy. Outputs are estimated using the k-fold cross-validation method and the features selected for confusion matrix (4x4) analysis included are frequency, SAR, exposure time, and SAR×exposure time (impact of accumulated SAR within the exposure period). It is reported that using robust predicting methods to identify the impact has become increasingly more critical. Strong correlations were observed between SAR and exposure time of weak RF-EMF, while an insignificant relationship was observed between frequency and exposure time. Halgamuge (2020)‘s study confirmed that supervised machine learning is a viable strategy for discovering features best characterizing the RF-EMF exposure scenarios. Essentially, in this manner, it is critical to examine model viability in a specific data set. More research in this space is crucial to learn whether and how some RF-EMF features (e.g., frequency of weak RF-EMF, SAR, exposure time) influence the prediction of reactions in living organisms. Future applications in public health and occupational and environmental epidemiology should utilize machine learning algorithms as shown in Figure 2.9.

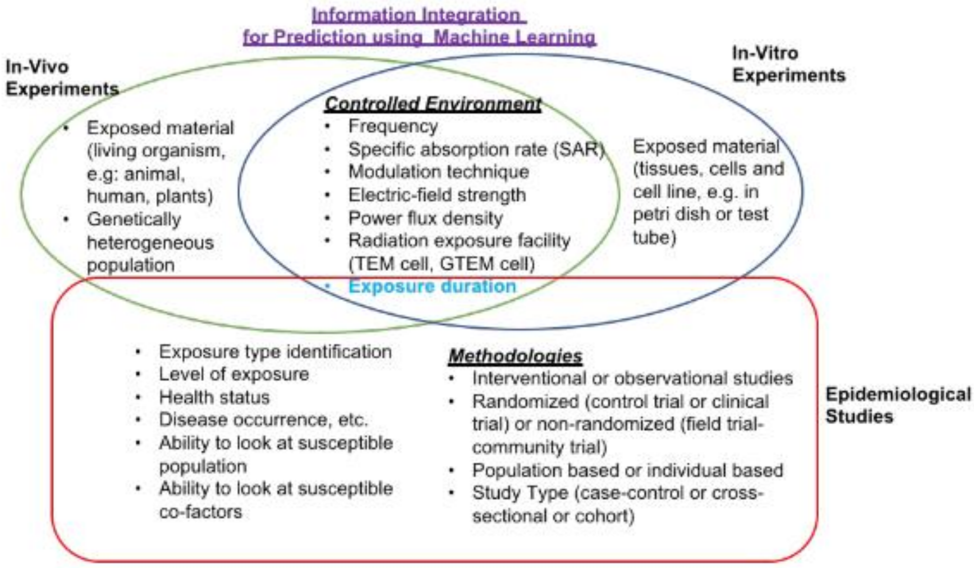


Figure 2.9 Potential features, attributes, or variables of bioelectromagnetic experiments (in-vitro, in-vivo, and epidemiological studies) that could be utilized in machine learning algorithms (M. N. Halgamuge, 2020)

2.8 Research Gap

Many studies have been conducted to investigate the effects of RF-EMF exposure from the communication devices and infrastructures on the human health. Most available studies on the cognitive performance, physiological parameters, and well-being of human are limited to the effects of GSM900/GSM1800/UMTS/4G mobile phone, GSM900/GSM1800/UMTS base station, Digital European Cordless Telecommunications (DECT) and Wi-Fi exposures, without considering the effects of 5G 700 MHz, 3.5 GHz or 28 GHz base station signal. However, to the best of our knowledge, none of previous reviews focused on the health effects resulting from exposure from the

5G MP and BS antennas from 700 MHz to 30 GHz on the cognitive performance and human physiological parameters of adults. Moreover, a recent study by Seungmo Kim & Nasim, (2020) highlighted that 5G radiation at 28 GHz may represent a hazard to human health, making such assessment urgently needed. This is significant in determining whether 5G technology is indeed safe for human. The information gathered from such studies will provide important insights into the significance of 5G BS signal exposure on humans as highlighted in the most recent state-of-the-art study, (Seungmo Kim & Nasim, 2020), the authors reported that 5G exposure at 28 GHz may threaten human health. The search for publications indicated no human epidemiology studies by 5G and potential health effects (Karipidis et al., 2021) nor at the RF-EMF frequencies higher than 2500 MHz. The results of the current review demonstrate no consistent relationship between the character of RF-EMF effects and parameters of exposure by different generations of telecommunication technology (Hinrikus et al., 2022). Thus, further studies regarding human health impacts caused by 5G exposure are urgently needed.

Datasets typically have high dimensions and consist of various types of features. The nature of features can differ between datasets, and the presence of diverse feature types may necessitate the use of distinct feature selection techniques (R. Zhang et al., 2018, Alwohaibi et al., 2021). Further, the correlation between each feature and the target variable (class) can be perceived by computing the correlation coefficient between each feature and the class in feature selction method (Xue et al., 2016). The conventional method to validate whether there is an effect of RF-EMF on human health is by conducting manual analysis using the statistical technique analyses (Koivisto et al., 2000; Curcio et al., 2004; Regel et al., 2006; Eltiti et al., 2007; Oftedal et al., 2007; Curcio et al., 2008; Cinel, et al., 2008; Kleinlogel et al., 2008; Eltiti et al., 2009; Trunk et al., 2015; Malek et al., 2015; Andrianome et al., 2017; van Moorselaar et al., 2017; Vecsei et al., 2018; J. Wallace et al., 2020) through recent peer-reviewed articles. The statistical analyses are performed by comparing the assessed parameters under no exposure and exposure of the RF-EMF signal. The *p*-values are calculated using statistical technique analyses such as Analysis of Variance (ANOVA), independent t-test, Pearson Chi-Square, and Wilcoxon signed rank tests. If there was significant difference between the values of the investigated parameters under no exposure and exposure, the *p*-values are less than 0.05 ($p<0.05$). This indicates that there is an RF-EMF effect on the investigated parameters. If $p>0.05$, this indicates that no RF-EMF effect on the investigated parameters. However, these manual statistical technique analyses led to time consuming and with implementing machine learning in the bioelectromagnetic research can attempts to discover the undiscovered pattern in data as well as aims to address users to make intelligent judgements from their research outcomes. Furthermore, machine learning advances the use of prediction tools to support future health checks (ex-vivo) and enables researchers to see how environmental factors may affect a final decision (M. N. Halgamuge, 2020).

Feature selection is a challenging task due mainly to the large search space (Xue et al., 2016) and the appropriate feature selection method is crucially needed for investigation on the proposed dataset involved as mentioned by Halgamuge (2020) reported that more study intends on investigation robust predicting techniques for identifying the impact of RF-EMF on bioelectromagnetics datasets. With the use of reliable prediction techniques to identify the effect of 5G exposure (700 MHz to 30 GHz) on human health and cognition using supervised machine learning RF-EMF, this study aims to present the merit of utilizing machine learning algorithms (supervised learning, i.e., prediction) to develop higher accuracy classifiers for predicting the potential impact of weak RF-EMF on human to discover features in occupational and environmental epidemiology and public health studies. To the best of our knowledge, the studies to classify hybrid dataset for data from short-term 5G base station exposure on the cognitive performance and physiological parameters of adults are limited trials and it needs further research, especially, the ones that applies intelligent solutions. None of the previous reviews focused on the health effects resulting from the exposure from the 5G mobile phone and base station antennas from 700 MHz to 30 GHz on the cognitive performance and the human physiological parameters utilizing machine learning algorithms, especially the supervised learning in the scope of prediction model with result to develop high accuracy classifiers for predicting the potential impact of RF-EMF exposure on human in experimental studies. By solving the problem with novel manner, an AI algorithm is utilized to analyze collected data systematically and make reasonable conclusions, making the whole process automatic (Wei. Y.et al., 2020).

CHAPTER 3 : EVALUATION OF 5G BASE STATION ANTENNA DESIGN SETUP AND MACHINE LEARNING APPROACH

3.1 Introduction

This chapter presents the methodology for describes the design of assessment 5G BS antenna as well as technical steps to design 5G BS antenna health effect detection based on supervised machine learning. Firstly, the specification and technical measurement needed for 5G BS antenna assessment design is introduced and the selection of subjects and ethical approval for assessment. Then, the steps to improve the classifier accuracy by applying supervised machine learning are presented. This section illustrates the evaluation, demonstration, and statistical outcomes to validate the performance of the proposed MSFS hybrid feature dataset using supervised machine learning in terms of machine learning classification accuracy, precision, f1-score, sensitivity, and specificity.

3.2 Research Methodology Design

This section introduces the research methodology design which includes three phases stages namely Phase 1: 5G BS Antenna Exposure Assessment on the 5G base station antenna for the inputs of the dataset to be used in this work consists of sixty healthy subjects (30 EHS and 30 Non-EHS) who participated in the 5G base station antenna RF-EMF effect study and completed all four 5G base station signal exposures (Sham, 700 MHz, 3.5 GHz and 28 GHz) during pre-exposure, exposure, and post-exposure. Two types of data including the physiological measurements of the individuals in terms of body temperature, SYS, DIA and heart rate, as well as four cognitive performance outcomes. In Phase 2 of the MSFS Process for 5G Base Station Antenna Health Effect Detection, the second step in data processing involved labelling the raw data. Afterward, the MSFS process included the application of a filtering method for data reduction. Each dataset was then transformed into 20 individual normalized datasets, and the top three normalization methods were selected based on the results of the paired t-test and F-test. The subsequent step involved feature fusion to create a hybrid dataset. Feature extraction was then performed using the PCA method, followed by feature selection through correlation analysis. The final step was testing classifiers for each dataset to determine the most suitable machine learning model. In the final phase, Phase 3: Validate Performance of Proposed Classifier using Supervised Machine Learning, machine learning models were created for each dataset. Following that, the performance of the proposed MSFS hybrid feature dataset was validated using supervised machine learning. This validation considered machine learning classification metrics such as accuracy, precision, f1-score, sensitivity, and specificity. The main framework of the research is illustrated in Figure 3.1.

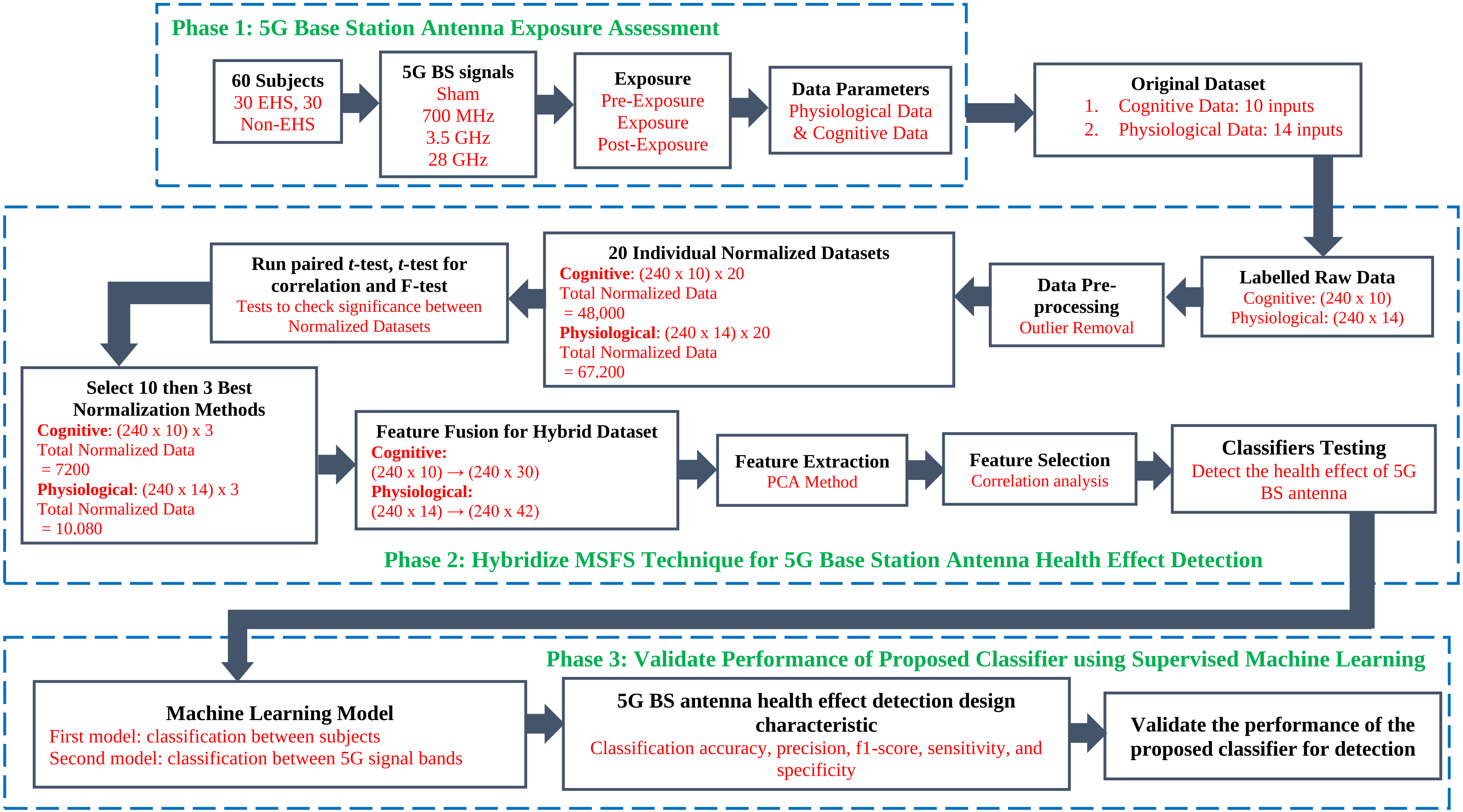


Figure 3.1 Research Methodology Framework Block Diagram.

3.3 Subject Recruitment

In this exploratory study, adult subjects recruited in various ways such as through local advertising or newspaper, Facebook, the website of UniMAP, Faculty of Electronic Engineering & Technology, Advanced Communication Engineering Centre of Excellence (ACE), word of mouth and via other participants. Recruitment continued until all participants complete the measurement protocol. From all applications, participants were initially selected on the basis of EHS questionnaire during a face-to-face, telephone interview or online questionnaire via Google Form. The biographical questions obtained information on age, gender, race, marital status, employment, number of hours per week the individual worked, volunteered, or spend studying. All potential volunteers were encouraged to make registration in the Google Form. The selected participants are notified via WhatsApp.

3.3.1 Subject Declaration

The EHS represents the group of subjects which members have attributed complaints suspected to 5G exposure (IEI-EMF). On the contrary, the Non-EHS group denotes the reference group – subjects without any complaints (Non-IEI-EMF category). To determine whether a subject is an EHS individual or not, all subjects are required to answer a survey prior to the experiments. A selection of questions was generated to assist the respondent connect the dots between their symptoms and their exposure to different electrical appliances that emit EMFs. A 57-item of self-declared hypersensitive people attributed their EHS to a defined source of RF-EMF. This definition of the self-declared

EHS symptoms of participants are based on well-established EHS studies by (Eltiti et al., 2007; Wallace et al., 2010; Malek et al., 2015). The EHS subject sorting for classification is shown in Figure 3.3. On a scale from 0 (not at all) to 4 (a great deal), participants were asked to rate symptoms experiencing for each symptom. The selected EHS symptoms from each candidate are map to 8 cluster which are allergy-related, cardiorespiratory, locomotor, neurovegetative, skin, auditory, headache, and cold-related. The classification of each subject group for the total symptoms scores are as tabulated in Table 3.1.

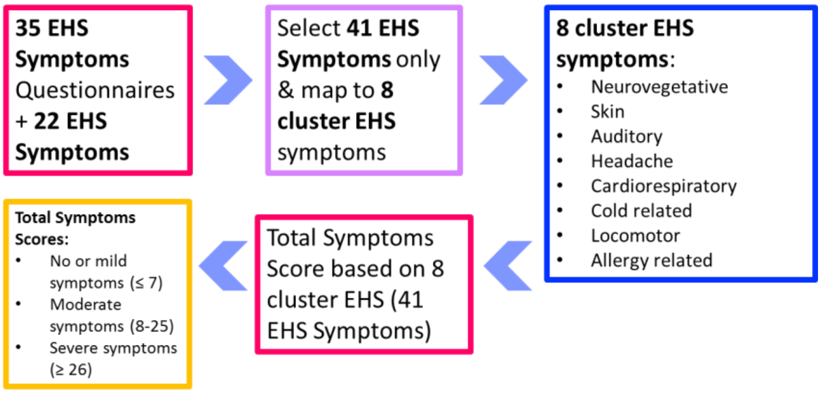


Figure 3.2 Methodology of the EHS Subjects Classification (Eltiti et al., 2007)

Table 3.1 The EHS Subjects Classification (Eltiti et al., 2007)

Symptoms Score	**Symptoms**	**Subject Group**
≤ 7	No or mild symptoms	Non-EHS
8-25	Moderate symptoms	Non-EHS
≥ 26	Severe symptoms	EHS

3.3.2 Involved Subject

A total of 148 subjects volunteered, as shown in Table 3.2. 54% female and 46% male are all registered volunteers, while Malay is the primary race of respondents, led by Chinese, Indian, and others. The highest rank for the age of respondents is 22 and even for the full-time schooling population of employment status. The average age for the control group was 22.83 years with a standard deviation, σ= 3.37, while the EHS group had an average age of 23.5 years with standard deviation σ= 5.16. Females made up most of both the EHS (65%) and the Non-EHS (51%) groups.

Table 3.2 Demographic data for the subject recruitments.

		EHS	**Non-EHS**
Age		23.5±5.16	22.83±3.37
Gender	Male: Female	16:30	50:52
Race	Malay	38	82
	Chinese	1	6
	Indian	5	8
	Others	2	6
Employment	Not Schooling	0	3
	Unemployed	11	16
	Part Time Schooling	0	2
	Self Employed	0	3
	Full Time Schooling	30	74
	Training	0	0
	Employed	5	4

The minimum age of the subjects is 18 years, and the maximum age is 41 years old. This complies to the definition of the WHO defines an adult as a person who is older than 19 years old, unless national legislation specifies an earlier age restriction (without defined maximum age limit of adults) In Malaysia, an adult is defined as a person of 18 years old and above. The average age for the Non-EHS group is 22.83 years, Standard Deviation (SD) of 3.37 whereas the EHS group had an average age of 23.5 years, SD of 5.16. However, upon screening, some subjects were disqualified due to reasons such as health conditions issues, unable to be on site throughout the duration of experiments, being underage, or has employment commitments. The final experiments were performed on 60 adults and the specifics are outlined in Table 3.3.

Table 3.3 Subject Involved Specifications

Subject	Male	Female	Average Age	SD Age	Average height	Average weight
Non-EHS	19	11	22.3 years old	2.3	166.33 cm	63.89 kg
EHS	11	19	22.6 years old	5.3	161.60 cm	59.61 kg

3.3.3 Exclusion Criteria

The main exclusion criteria to reduce biased and inaccurate results in determining the effects of 5G BS exposure which includes subjects fitted with pacemakers, using hearing aids and artificial cochlear, chronic illness polymorbidity, a history of head trauma, or neurological or mental conditions such as depression, phobia, exhaustion, and psychosis, fibromyalgia, or syndrome of protracted weariness (Eltiti et al., 2007) and issues with sleep. In addition, persons who regularly consumed psychotic drugs in the preceding six months, over 10 times a week consumed alcohol, or consumed caffeinated beverages with an average daily caffeine intake of more than 450 mg (e.g., three cups of coffee) were disqualified from the study. Finally, subjects who worked shifts during the month before the trial and those who took a long-haul journey with a time zone difference of more than 3 hours were also eliminated. These exclusion criteria are important in this study to reduce biased and inaccurate results in determining the effects of 5G BASE STATION exposure.

3.3.4 Sample Size of Study Population

To investigate the influence of 5G field exposures and Sham on well-being, physiological parameters, and cognitive performance between EHS and Non-EHS subjects, a critical decision that needs to be made is on the choice of sample size. In this study, each participant will be randomized to the exposure (700 MHz, 3.5 GHz, and 28 GHz) including Sham. This type of design is called a repeated measure design. The number of participants required in this experiment is determined by using the G-power statistical technique for the power analysis calculation (Malek et al., 2015; Masrakin et al., 2019). As a result, a group of subjects are required to have 24 subjects in order to assess the significance of the model by using the G*Power software, a medium is assumed, 0.06 partial eta squared where the effect size = 0.25 with significance level 0.05 and statistical power 0.80. To increase this significance, the sample size is enlarged to 30 samples for each group, with a total sample size of 60. A partial eta squared of 0.02 (where "exposures" accounted for nearly none of the EHS and Non-EHS group) is assumed, and power 0.82 with significance model level at $\alpha = 0.05$ as shown in Figure 3.3. Thus, 60 sampling size is sufficient to provide significant statistical power to evaluate the effects of 5G base station exposure to the general population. The adult subjects for this study are classified into two groups. Group A denotes the group of subjects that have previously reported to experience complaints and have attributed these complaints to 5G exposure (i.e., sensitive category); Group B denotes the reference group, namely a group of subjects without any complaints (i.e., non-sensitive category).

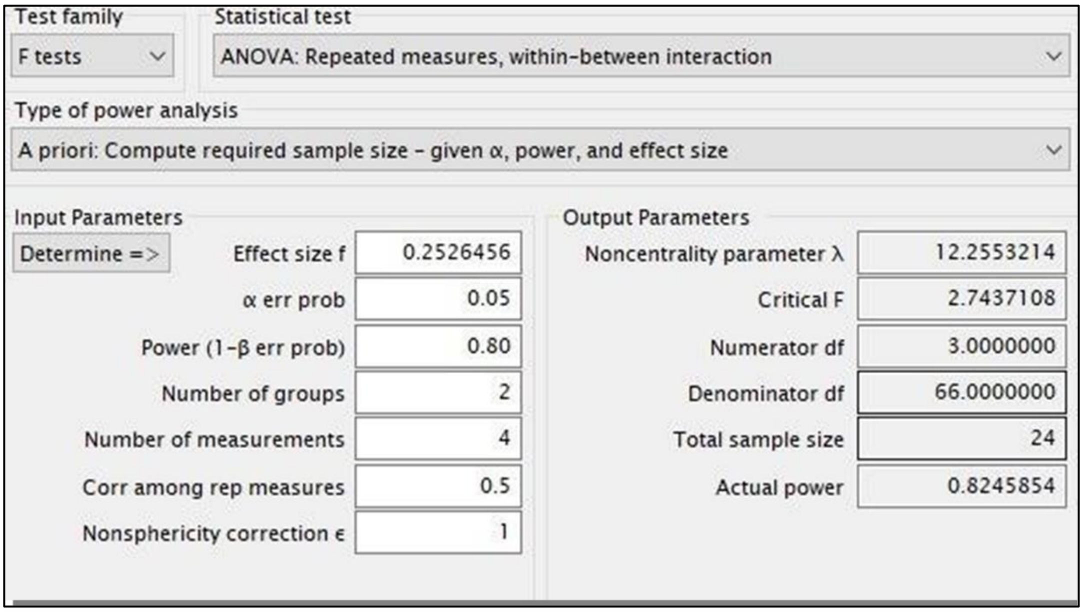


Figure 3.3 Analysis of G-power

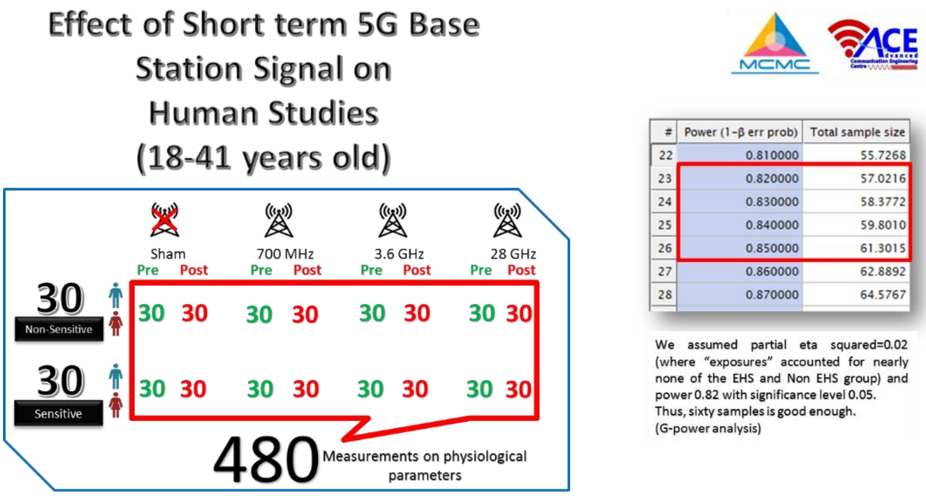


Figure 3.4 A maximum total of 30 adult subjects each for two groups (IEI-EMF and non-IEI-EMF categories)

3.3.5 Ethical Approval from UniMAP

Each participant was informed of the study's objectives before the experiment, and those who agreed to participate voluntarily provided signed informed consent. The procedures used comply with the standards set out by the Universiti Malaysia Perlis (UniMAP) Ethical Committee (Reference no: UniMAP/PTNC(P&) I/100-1()) as shown in Appendix A.

3.4 Assessment of 5G BS Antenna Design Setup

This part of the research focused on assessing technical aspects related to 5G signal bands emitted at 700 MHz, 3.5 GHz, and 28 GHz. The assessment included measuring the modulated signal and setting up the E-field, using appropriate tools for signal analysis. The E-field setup was designed to evaluate electromagnetic field exposure, ensuring adherence to safety standards as per ICNIRP recommendation. The assessment procedure strictly followed standardized methods for signal band measurement, with monitoring and adjustments to equipment from the researchers as necessary. Data parameters involved reporting the strength of 5G signal bands and the levels of electromagnetic field exposure. The evaluation included analysing collected data for cognitive and physiological impacts.

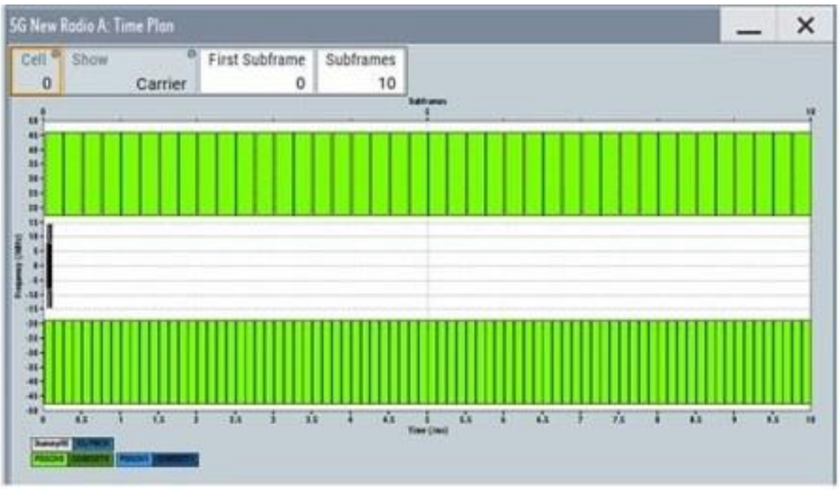
3.4.1 5G NR Modulated Signal

One of the key elements of 5G is the use of Cyclic-Prefix Orthogonal Frequency Division Multiplex (CP-OFDM) and Discrete Frequency Transform Spread Orthogonal

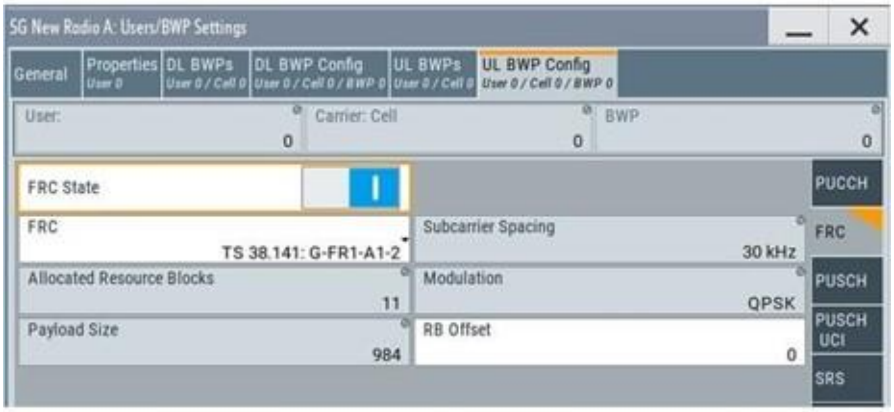
Frequency Division Multiplex (DFT-S-OFDM) as the signal bearer. OFDM is used in a number of other of systems from WLAN, WiMAX to broadcast technologies including DVB and DAB. OFDM has many advantages including its robustness to multipath fading and interference. In addition to this, even though, it may appear to be a particularly complicated form of modulation, it lends itself to digital signal processing techniques. In view of its advantages, the use of ODFM is natural choices for the new 5G cellular standard. A detailed grid structure of 5G NR has been described in Section 2.2.1.

In this research work, a new modulation of 5G NR, 5G NR SMBV-K444 software will be utilized in producing the modulated 5G NR for the investigated frequency band 700 MHz, 3.5 GHz and 28 GHz. The 5G NR software option (-K444) simplifies uplink and downlink 5G NR signal configuration. It supports all waveforms, channel bandwidths, modulation schemes and numerology options specified in the standards. The intuitive GUI allows configuring these and many other parameters, such as multiple bandwidth parts or MIMO precoding, directly on the instrument. Figure 3.5 shows a screenshot of 5G NR Transmit signal in R&S 5G NR SMBV-K444 software.

(a)



(b)



(c)

Figure 3.5 Screenshot of 5G NR Transmit signal in R&S 5G NR SMBV-K444 software.

3.4.2 EMF Measurement

To represent the exposure by from the existing base stations in the environment, two key parameters of the 5G signals are the E-field strength, E in equations (3.1) (Fernandes, 2017; Masrakin et al., 2019) and Power Density, S (3.2) as per recommendation in (Fernandes, 2017; Pawlak et al., 2019) and ICNIRP (ICNIRP, 2020), are listed in Table 3.4 and research study by Wali et al., 2022 the investigation focused on determining the peak exposure levels emanating from a 5G mm-Wave base station. The findings revealed that the highest recorded exposure from the 29.5 GHz base station was 5.71 V/m, while the maximum average exposure amounted to 2.02 V/m. In this area, the field has the character of a plane wave, the vectors of E and the Magnetic Field, H are perpendicular to each other, and the Power Density, S is related to the Electric Field Strength, E by the free space impedance Z_o = 120π (Ω). P_r is the power received by the RF-EMF probe antenna, in watts, η=120π Ω is the intrinsic impedance of the air. $A_{eff} = \left(\frac{\lambda^2}{4\pi}\right)$ is the effective area of the probe antenna, in square meter, λ is the wavelength of the radio source, in meter. Maximum value of E-Field strength is recorded to compare the compliances level of human exposure to radiation.

$$E\left({}^{V}/_{m}\right) = \sqrt{\frac{\eta \times P_{r;watt}}{A_{eff}}} \tag{3.1}$$

$$P_{r;watt} = \frac{E^2 A_{eff}}{\eta}$$

$$S = \frac{E^2}{Z_0} = \frac{E^2}{120\pi} \tag{3.2}$$

Table 3.4 Restriction for EMF exposure – ICNIRP Guidelines

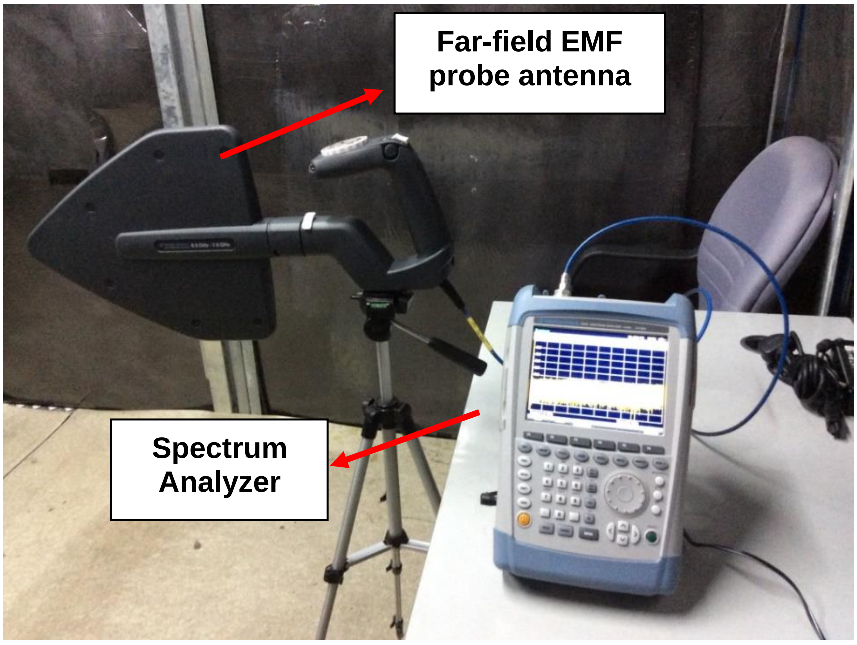
Frequency, f (MHz)	**E (V/m)**	**S (W/m²)**
>30 - 400	28	2
>400 - 2000	1.375 x $\sqrt{f}$	f/200
>2000 - 300000	61	10

The evaluation of compliance with the permissible EMF exposure in the vicinity of a base station is achieved by measuring the levels of EMF produced by the base station-like antenna in the RF shielded room. This is to ensure that the E-field strength of 1 V/m and

2 V/m with power densities of 2.652 mW/m^2 and 10.61 mW/m^2, respectively can be obtained based on (Malek et al., 2015) which is 10 times higher than the real 2G/3G base station antennas exposure in Malaysia. Measurements was performed in the far field area and carried out in a RF-shielded room anechoic chamber with the same standard procedure of measurement as published in (Masrakin et al., 2019) . Far field distance measurement setup EMF probe and spectrum analyser as listed in Table 3.5. The instruments used in this measurement were the Rohde & Schwarz (R&S) Signal Generator (SG) model SMBV100A, the R&S Handheld Spectrum Analyzer (SA) model FHS4 (9 kHz – 3.6 GHz), the R&S Handheld SA model FSV, base-station antenna and far-field EMF probe antenna model R&S HE300 Antenna module 4067.6458.00, as shown in Figure 3.6 (a). The directional antenna was placed approximately 1.5 m from ground in order to record the measurement (Ismail et al., 2009; [ITU-T, K.61], 2014; Maccartney et al., 2015). It represents the head position of average adults which resembles the experimental procedure in (Regel et al., 2006; Malek et al., 2015) as shown in Figure 3.6 (b), Figure 3.6 (c) , Figure 3.6 (d) and Figure 3.6 (e). The measurement time of six minutes was carried out as in the standard recommended by ICNIRP and IEEE (ICNIRP, 2020; IEEE, 2019) to ensure the data acquired is accurate. In this assessment, three type of signal exposures were emitted which were 5G 700 MHz, 5G 3.5 GHZ and 5G 28 GHz.

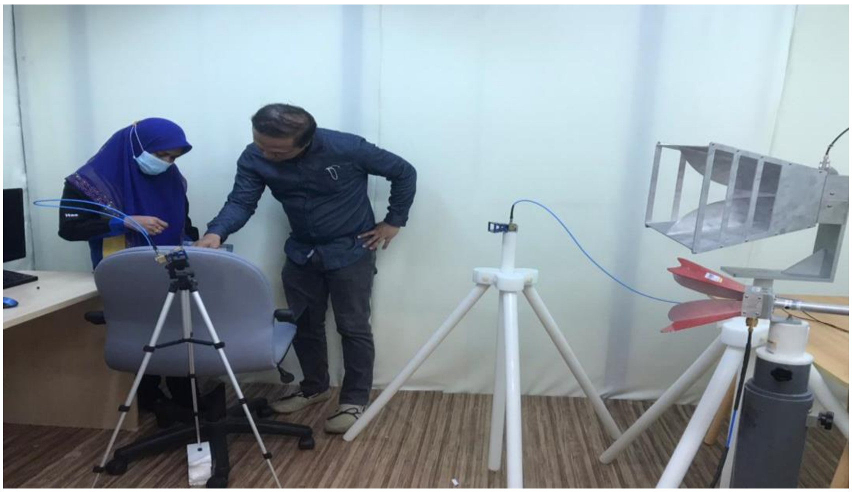
Table 3.5 Type of Probes and Instrument Use in the Measurement

Probe Type and Instrument	Frequency Range
Far-field EMF probe antenna model R&S HE300 Antenna module 4067.6458.00	500 MHz – 7.5 GHz
R&S Handheld Spectrum Analyzer (SA) model FHS4	9 kHz – 3.6 GHz
R&S Signal Analyzer model FSV	10 Hz - 30 GHz

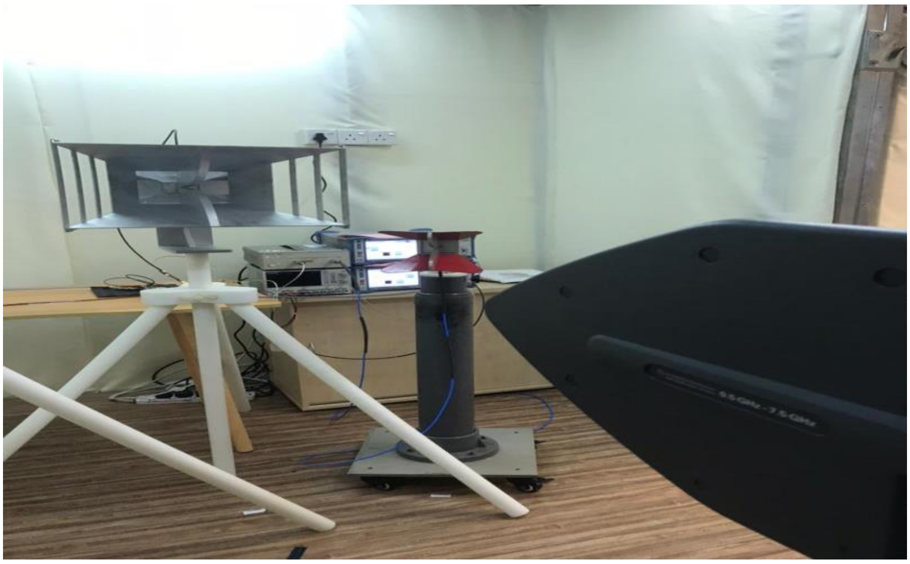


(a)

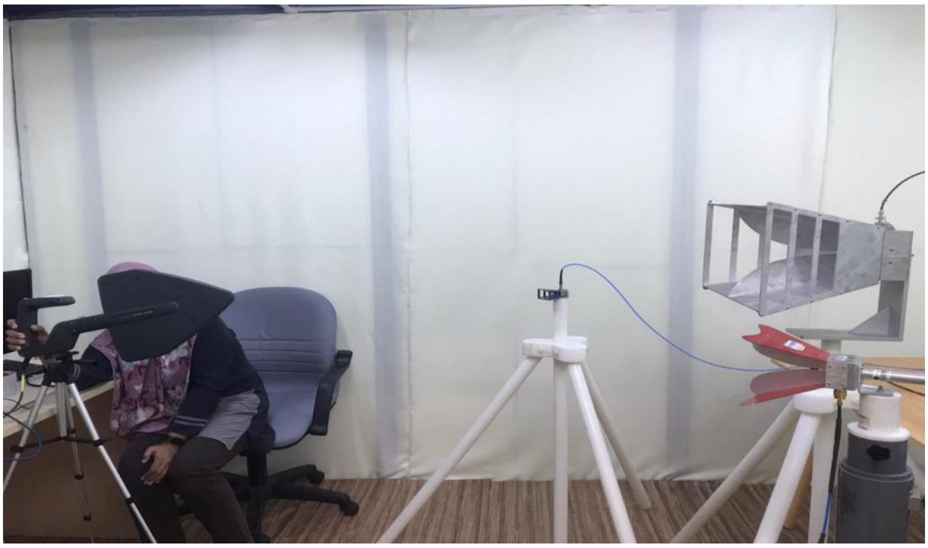
(b)



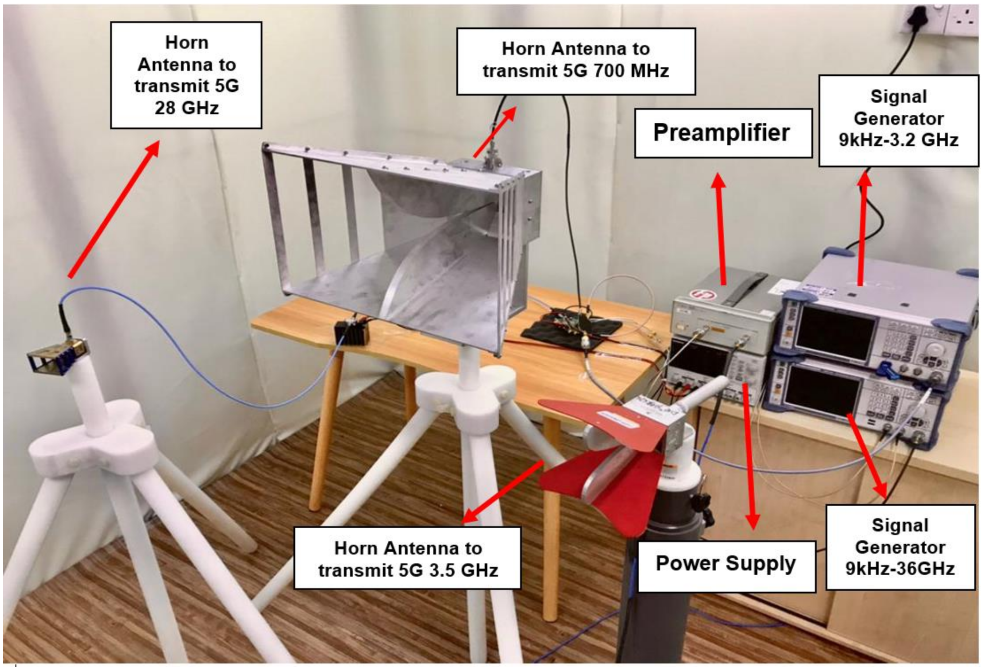
(c)



(d)



(e)



(f)

Figure 3.6 Far field distance measurement setup (a) EMF probe and spectrum analyzer, (b) Each 5G frequency spectrum between subject and exposure, (c) Measurement on E-field strength for 5G 28 GHz, (d) Measurement on E-field for 5G 700 MHz and 3.5 GHz, (e) Measurement on E-field for 5G 700 MHz and 3.5 GHz and 5G 28 GHz (f) Complete setup on E-field for 5G 700 MHz, 5G 3.5 GHz and 5G 28 GHz.

To assess the effects of 5G exposure, two different carrier 5G modulated signals are generated at 700 MHz, 3.5 GHz and 28 GHz to cater to multiple bands implementation (and possible carrier aggregation). The low band is considered for coverage and high band for speed and data capacity. They were being set up as follows:

- A-Infomw LB-660-NF Broadband Horn Antenna 0.6 GHz-6 GHz is used as the transmitting antenna at 700 MHz.

- ETS-Lindgren 3117 Double-Ridged Waveguide Horn Antenna 1 GHz - 18 GHz is used as the transmitting antenna at 3.5 GHz.
- RT-RF HA-1840GA1-NF Double Ridged Broadband Waveguide Horn Antenna 18GHz-40 GHz is used as the transmitting antenna at 28 GHz.

This setup designed for comprehensive testing and measurement of 5G base station E-field setup. A handheld spectrum analyzer (Rohde & Schwarz FHS4) and a far-field EMF probe (R&S HE300 module 4067.6458.00) is used to measure the E-field strength, which is set at 1 V/m identical to that used in (A. P. M. Zwamborn et al., 2003; Regel et al., 2006; Eltiti et al., 2007; Malek et al., 2015). 5G 700 MHz signal transmission involved an RF splitter (Passion Radio) operating between 0 and 5000MHz Attenuation 6 dB with SMA was connected between the 700 MHz antenna and the signal generator which obtained its input from an upconverter (Analog Devices ADMV1013044718C) to generate the 5G 700 MHz. For the connection of the 5G 3.5 GHz signal, a different signal generator of the same model and different operating limits (Rohde & Schwarz SMBV100A, 9 kHz to 6 GHz) is used. The generated 5G NR signal is amplified to a 28 dBm output level using a preamplifier (Agilent 8449B) prior to connecting to the 3.5 GHz antenna (ETS-Lindgren 3117 antenna). The 1 V/m level is measured on the same handheld spectrum analyser and EMF probe antenna as similar reading of base station used in the experiments study by (Regel et al., 2006; Eltiti et al., 2009; Malek et al., 2015;). The 5G 28 GHz signal can be generated at the output but due to the low power level, an active amplifier (Mini-Circuit ZVE-403-K+, 26000-40000 MHz, S N705502043) is used to increase the power level up to -10 dBm. Finally, the amplifier’s output is connected to the 28 GHz antenna to transmit exposure signal. Spectrum analyser (Rohde & Schwarz FSV 30 GHz) used to measure the output power. Next, E-field strength value of 5G 28 GHz is calculated manually from the received signal of R&S Signal Analyzer model FSV. To obtain electric field strength at 1 V/m from 5G 700 MHz and 5G 3.5 GHz, the transmitted RF power is set at 17.06 dBm and -5.5 dBm in the SG, respectively. The transmitted RF power for 5G 28 GHz is set at 10 dBm. The complete setup of the RF 5G base station antenna exposure is illustrated in Figure 3.7.

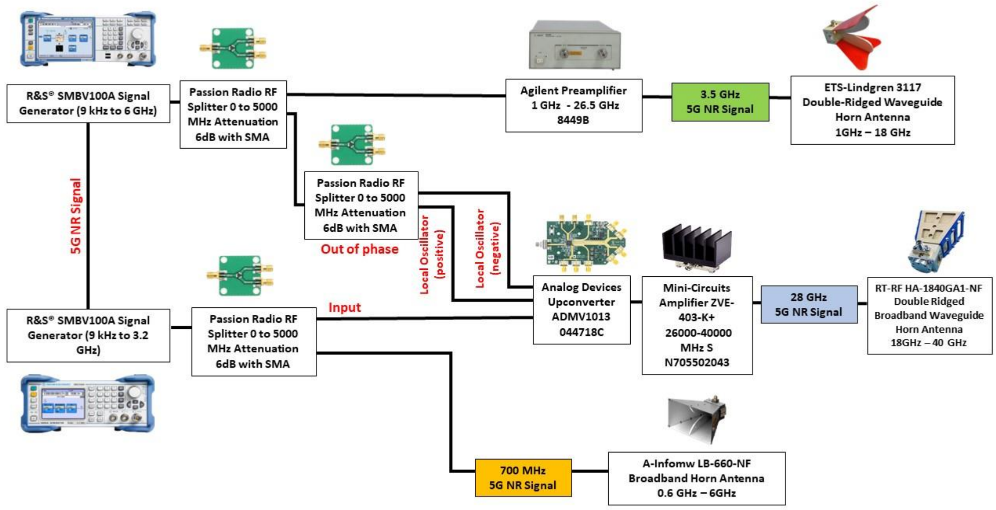


Figure 3.7 Schematic diagram of 5G base station exposure E-field setup.

During the exposure session, the experiment will be conducted under counterbalanced randomised double-blind conditions in a randomised crossover design. Hence, during the double-blind tests neither the adult subjects nor the experimenters will not be notified of which exposure was being generated. The equipment setting for the E-Field at the 5G base station antenna will be handled by a second experimenter. To maintain a double-blind study design, the sequence of sham and RF exposures 5G 700 MHz, 5G 3.5GHz, and 5G 28 GHz assigned to each participant was disclosed solely to the second experimenter responsible for overseeing the exposure (R Huber et al., 2002; Oftedal et al., 2007). In order to prevent bias in the study findings, the adult test subjects and the second researcher are not aware or recognize the type of exposure that occurs during the double-blind tests. Another researcher will remotely control the signal generator which was invisible to the other researcher and the subjects present. In order to enforce double-blind experimental circumstances, this can be done by either activating or signal output being disabled (according to Real or Sham exposure, respectively) Table 3.6 shows the measurement setting for the E-field during exposure for each equipment setup.

Table 3.6 The Measurement Equipment Setting for E-Field During Exposure

<table>
<tr><th>The Devices</th><th>700 MHz</th><th>3.5 GHz</th><th>28 GHz</th></tr>
<tr><td rowspan="2">R&S® SMBV100A Signal Generator (9 kHz to 3.2 GHz)
</td><td>Level:
17.06 dBm</td><td rowspan="2">RF OFF</td><td>Level:
10 dBm</td></tr>
<tr><td>Frequency:
700 MHz</td><td>Frequency:
2.5 GHz</td></tr>
<tr><td rowspan="2">R&S® SMBV100A Signal Generator (9 kHz to 36 GHz)
</td><td rowspan="2">RF OFF</td><td>Level:
-10dBm</td><td>Level:
10 dBm</td></tr>
<tr><td>Frequency:
3.5 GHz</td><td>Frequency:
4.25 GHz</td></tr>
<tr><td>Agilent Preamplifier model 8449B 1-26.5 GHz
</td><td>OFF</td><td>ON</td><td>OFF</td></tr>
</table>

GW INSTEK DC Power Supply model GPD-4303S 	OFF	OFF	ON

The far-field distance between antenna signal exposure and subject are determined that depends on the characteristics of the antenna and the frequency of the signal. The far-field distance, also known as the Fraunhofer distance, is the point at which the electromagnetic waves emitted by the antenna become predominantly plane waves, and the wavefronts are approximately parallel. The far field distance, R required to generate a base station-like signal is calculated by Equation (3.3), where D = largest dimension of the source of the radiation and λ is the wavelength corresponding to the appropriate frequency.

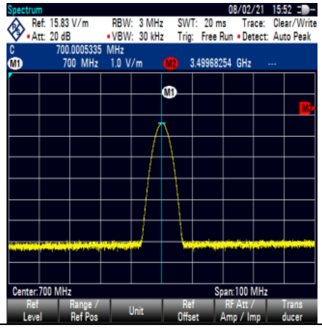
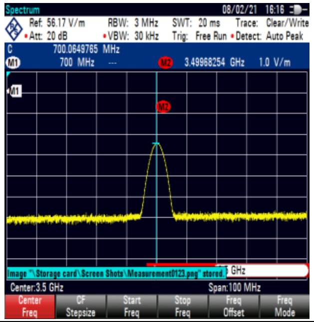
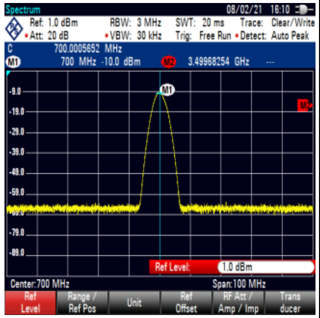
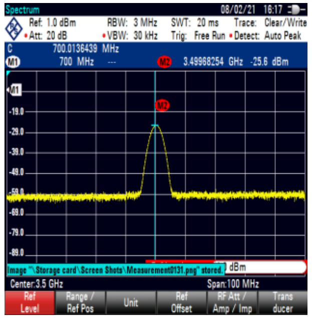
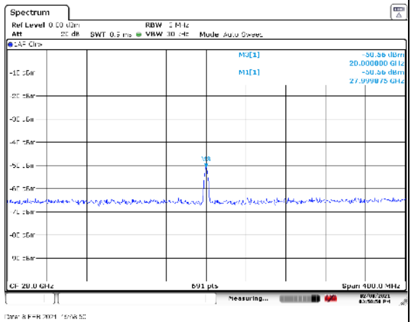
$$R > \frac{2D^2}{\lambda} \tag{3.3}$$

The calculation of R utilizes the information from the datasheets of the respective antennas (A-Infomw-NF Broadband Horn Antenna LB-660 Datasheet; ETS-Lindgren’s Model 3117 Double-Ridged Waveguide Datasheet; RT-RF HA-1840GA1-NF Double Ridged Broadband Waveguide Horn Antenna Datasheet), as outlined in Table 3.7. The E-field strength in V/m was measured as shown in Table 3.8. The E-field strength measured in this assessment and the exposure limit set by the ICNIRP guidelines were calculated, revealing a percentage difference of less than 5%. This indicated that the measured E-field strength was significantly within the ICNIRP guideline, affirming that the measured value was below the recommended exposure limit according to ICNIRP guidelines.

Table 3.7 R in meter for each transmitted antenna.

Antenna	**R (m)**
5G 700 MHz: A-Infomw LB-660-NF Broadband Horn Antenna 0.6 GHz-6 GHz (A-Infomw-NF Broadband Horn Antenna LB-660 Datasheet)	D = 0.435m $\lambda = 0.429$ $R > \frac{2(0.435)^2}{0.429} > 0.879$
5G 3.5 GHz: ETS-Lindgren 3117 Double-Ridged Waveguide Horn Antenna 1 GHz - 18 GHz (ETS-Lindgren's Model 3117 Datasheet)	D = 0.33m $\lambda = 0.0857$ $R > \frac{2(0.33)^2}{0.0857} > 2.541$
5G 28 GHz: RT-RF HA-1840GA1-NF Double Ridged Broadband Waveguide Horn Antenna 18GHz-40 GHz (RT-RF HA-1840GA1-NF Double Ridged Broadband Waveguide Horn Antenna Datasheet)	D= 0.038m $\lambda = 0.0107$ $R > \frac{2(0.038)^2}{0.0107} > 0.27$

Table 3.8 Comparison of the measured electric field with the exposure limit recommend by ICNIRP for 5G 700 MHz, 5G 3.5 GHz and 5G 28 GHz exposure.

	Electric Field (V/m)			**Power Density (W/m^2)**		
	5G 700 MHz	**5G 3.5 GHz**	**5G 28 GHz**	**5G 700 MHz**	**5G 3.5 GHz**	**5G 28 GHz**
This research value	1 	1 	0.64	2.652×10^{-3}	2.652×10^{-3}	1.086×10^{-3}
Exposure limit for general public	$1.375 \times \sqrt{f(MHz)}$ $= 1.375 \times \sqrt{700 \times 10^6}$ $= 36.379$	61	61	f(MHz)/200 =(700)/200 =3.5	10	10
Comparison with exposure limit (%)	2.75	1.639	1.049	0.076	0.027	0.011
Signal Received (dBm)	-10 dBm 	-25.6 dBm 	-50.56 dBm 			

3.4.3 Exposure Assessment Setup

This measurement of assessment was carried out in the High Voltage Laboratory, Faculty of Electrical Engineering Technology, Universiti Malaysia Perlis (UniMAP), Arau, Perlis, Malaysia. Note that testing methods, antenna base station measurement setup, and experimental exposure procedures are all conducted in the same RF-shielded room. The RF shielded room is built using iron plates, with an overall dimension of 3.7 m (length) x 2.4 m (width) x 2.47 m (height) as shows in Appendix C (Masrakin et al., 2019). The inner walls are covered by microwave absorbing sheets to absorb reflected RF signal, providing a controlled environment to limit exposure to RF-fields to those generated within the room. Along with the 700 MHz, 3.5 GHz, and 28 GHz antennas, this area is outfitted with a flat screen monitor on a wooden table, plastic armed chair, and other items. The RF-shielded room's shielding effectiveness is calculated using the free space measurement method based on Standard No. IEEE 299(1), with a 40dB shielding effectiveness for the tested frequency range (Malek et al., 2015; Masrakin et al., 2019). This parameter is important in determining the level of signal blockage to meet the site criteria. The shielded room accommodates both the experimenters and the subjects. Exposures are conducted using counterbalanced randomized double-blind conditions. The randomized controlled trials design is implemented in which the subjects are randomly assigned to in order to minimizes biases and strengthen the case for causation.

Exposure for each single session will last for 60 minutes. All subjects are first briefed and trained in an office room (outside the shielded room) prior to being escorted to the exposure room. During pre-exposure session, cognitive test training will be performed, where subjects will be explained about the cognitive function tests when the

5G radiation is turned ‘off’. The research assistants additionally emphasize the subjects that no one will be exposed to EMFs throughout the training, ensuring that there is no bias about whether they are aware of the sort of exposure being emitted. Cognitive changes performance tests recorded during exposure session. Physiological changes of adults will be monitored before, during and after the exposure session were recorded using a blood pressure wrist measurement device Omron Automatic Blood Pressure Monitor HEM-7320 and Omron Forehead Thermometer MC-720 that complied with the WHO’s classification. The four field conditions—Sham, 5G 700 MHz, 5G 3.5 GHz, and 5G 28 GHz—are implemented on several days that are at least a week apart. To rule out any potential carryover effects, four people will be evaluated each day at the same time of day (about ±2 hours) (Regel et al., 2006; Curcio et al., 2008; Choi et al., 2014; Sauter et al., 2015). The experiments also will be scheduled twice a week for 2 to 3 weeks (with a minimum gap of three days after each session) also for the same purpose. The exposure schedule for the assessment is as shown in Table 3.9 while Figure 3.8 illustrates the top view of the 5G exposure (5G 700 MHz, 5G 3.5 GHz and 5G 28 GHz) positioned inside an RF-shielded room with the horn antenna towards the subject. Figure 3.9 shows the complete flowchart of the proposed assessment of 5G base station antenna exposure. In this flowchart, the inputs consist of subjects, crucial for acquiring data in the assessment. The assessment encompasses both Non EHS and EHS subjects. Physiological and cognitive data are measured at distinct conditions: before, during, and after exposure. The objective is to gather comprehensive physiological and cognitive data to facilitate the subsequent machine learning process.

Table 3.9 The Exposure Schedule.

Section	**Pre-Exposure**			**Exposure**		**Post-Exposure**	
Location	Outside RF shielded room			Inside RF shielded room		Outside RF shielded room	
Activity	Registration	Trial/Practice Session	Physiological Test	Cognitive Test	Physiological Test	Physiological Test	Cool Down Time

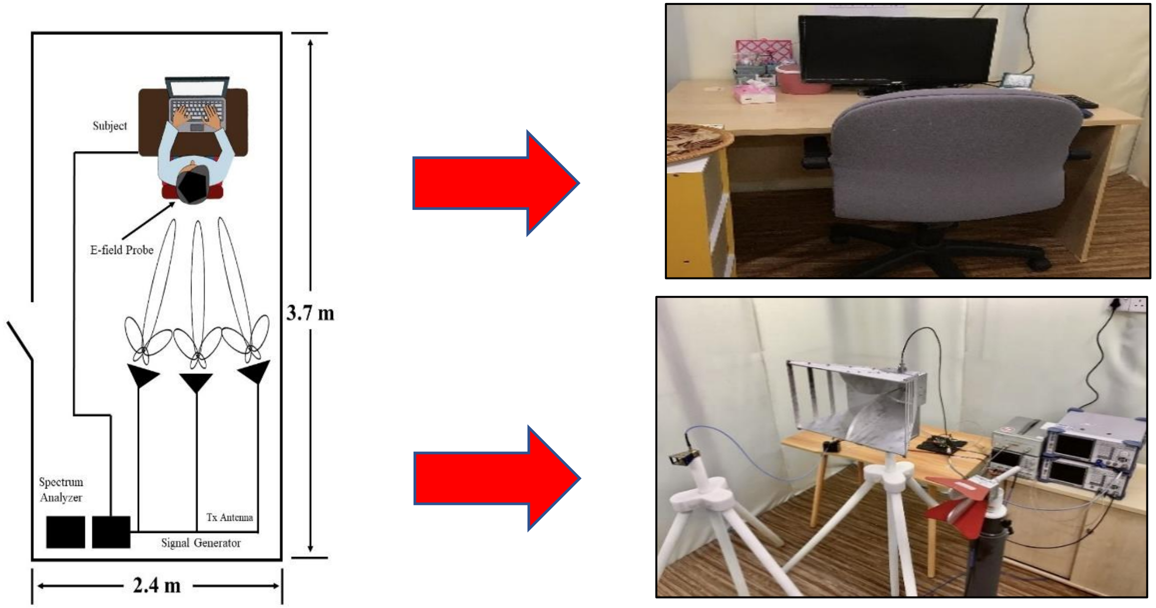
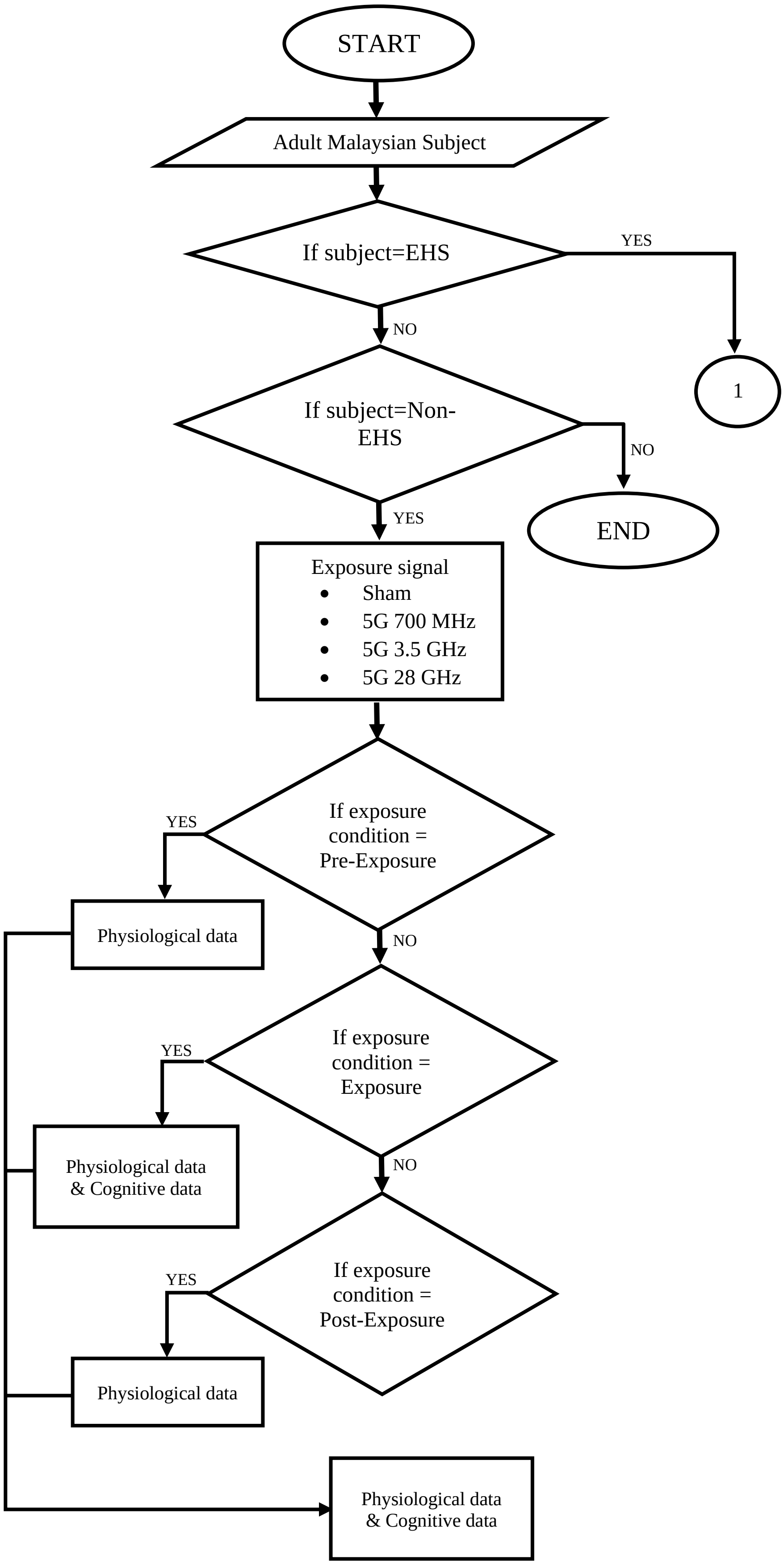


Figure 3.8 View of the 5G exposure setup from above in the RF-protected space, displaying the three horn antenna pointing in the direction of the subject.


START
Adult Malaysian Subject
If subject=EHS
YES
NO
1
If subject=Non-EHS
NO
YES
END
Exposure signal
Sham
5G 700 MHz
5G 3.5 GHz
5G 28 GHz
If exposure condition = Pre-Exposure
YES
NO
Physiological data
If exposure condition = Exposure
YES
NO
Physiological data & Cognitive data
If exposure condition = Post-Exposure
YES
Physiological data
Physiological data & Cognitive data

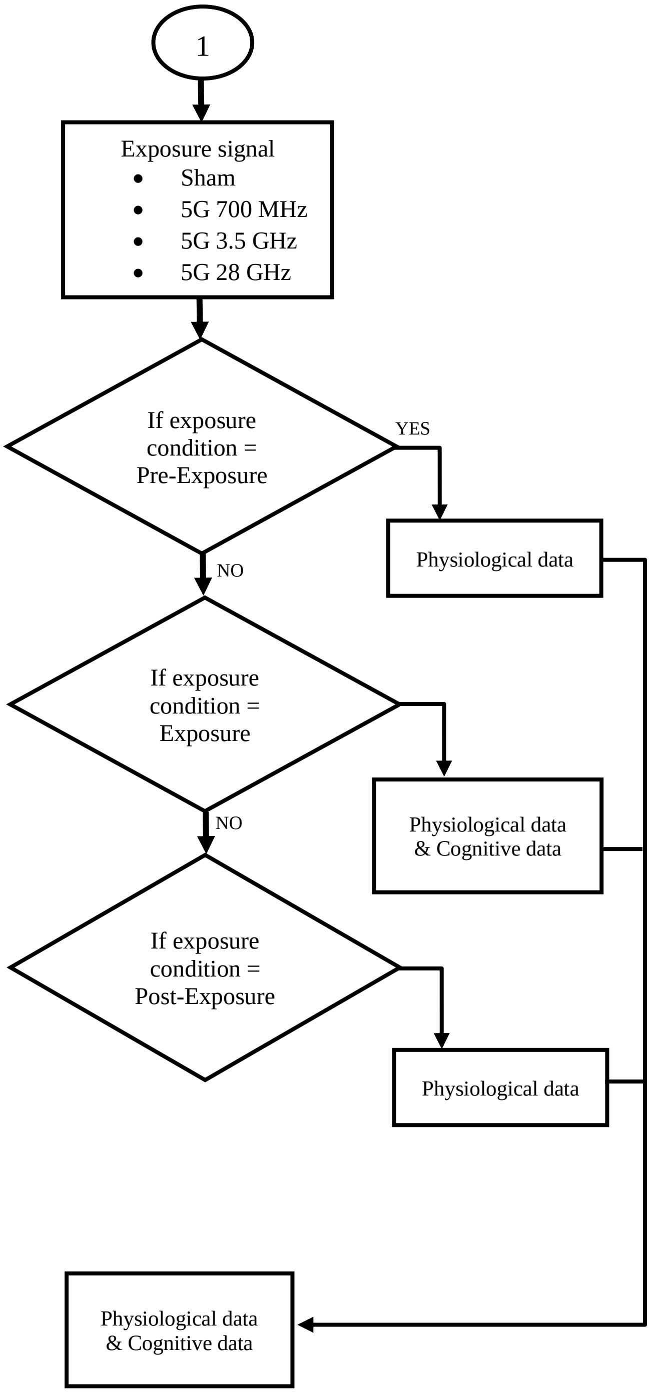


Figure 3.9 Flowchart of the Proposed Assessment of 5G base station Antenna Exposure

3.4.4 Output Parameters

There are two output parameters in which are considered for this research and these parameters will be investigated based on quantitative and qualitive methods throughout this study: Cognitive performance and Physiological Parameters.

The training for cognitive test activity (32 minutes) will be done prior to the first exposure session in each experiment session. During the training activity, the subjects will be explained on the cognitive function test for training reasons only. It is stressed that during this activity none of the subjects have been exposed to electromagnetic fields. The subjects were informed on the absence of 5G fields. The subject will be given a briefing on the procedures and tests involved within the 60 minutes' exposure period. The cognitive test randomized counterbalanced table is stated at Table 3.10 for Normal Subjects and Table 3.11 for EHS Subjects.

Table 3.10 Cognitive Schedule for Non-EHS Subjects

ADULT NORMAL (NS1 – NS30) (n = 30)				
Subject	**Exposure Condition**			
NS01	**DS**	**FT**	**BCST**	**TOL**
NS02	**DS**	**FT**	**TOL**	**BCST**
NS03	**DS**	**BCST**	**FT**	**TOL**
NS04	**DS**	**BCST**	**TOL**	**FT**
NS05	**DS**	**TOL**	**FT**	**BCST**
NS06	**DS**	**TOL**	**BCST**	**FT**
NS07	**FT**	**DS**	**BCST**	**TOL**
NS08	**FT**	**DS**	**TOL**	**BCST**
NS09	**FT**	**BCST**	**DS**	**TOL**
NS10	**FT**	**BCST**	**TOL**	**DS**
NS11	**FT**	**TOL**	**DS**	**BCST**
NS12	**FT**	**TOL**	**BCST**	**DS**
NS13	**BCST**	**DS**	**FT**	**TOL**
NS14	**BCST**	**DS**	**TOL**	**FT**
NS15	**BCST**	**FT**	**DS**	**TOL**
NS16	**DS**	**FT**	**BCST**	**TOL**
NS17	**DS**	**FT**	**TOL**	**BCST**

NS18	DS	BCST	FT	TOL
NS19	DS	BCST	TOL	FT
NS20	DS	TOL	FT	BCST
NS21	DS	TOL	BCST	FT
NS22	FT	DS	BCST	TOL
NS23	FT	DS	TOL	BCST
NS24	FT	BCST	DS	TOL
NS25	FT	BCST	TOL	DS
NS26	FT	TOL	DS	BCST
NS27	FT	TOL	BCST	DS
NS28	BCST	DS	FT	TOL
NS29	BCST	DS	TOL	FT
NS30	BCST	FT	DS	TOL

DS = Backward Digit Span (DS)
FT = Flanker Task (FT)
BCST = Berg's Card Sorting Task (BCST)
TOL = Tower of London (TOL)

Table 3.11 Cognitive Schedule for EHS Subjects

ADULT SENSITIVE (SS1 – SS30) (n = 30)				
Subject	**Exposure Condition**			
SS01	DS	FT	BCST	TOL
SS02	DS	FT	TOL	BCST
SS03	DS	BCST	FT	TOL
SS04	DS	BCST	TOL	FT
SS05	DS	TOL	FT	BCST
SS06	DS	TOL	BCST	FT
SS07	FT	DS	BCST	TOL
SS08	FT	DS	TOL	BCST
SS09	FT	BCST	DS	TOL
SS10	FT	BCST	TOL	DS
SS11	FT	TOL	DS	BCST
SS12	FT	TOL	BCST	DS
SS13	BCST	DS	FT	TOL
SS14	BCST	DS	TOL	FT
SS15	BCST	FT	DS	TOL
SS16	DS	FT	BCST	TOL
SS17	DS	FT	TOL	BCST
SS18	DS	BCST	FT	TOL
SS19	DS	BCST	TOL	FT
SS20	DS	TOL	FT	BCST
SS21	DS	TOL	BCST	FT
SS22	FT	DS	BCST	TOL
SS23	FT	DS	TOL	BCST
SS24	FT	BCST	DS	TOL
SS25	FT	BCST	TOL	DS
SS26	FT	TOL	DS	BCST
SS27	FT	TOL	BCST	DS

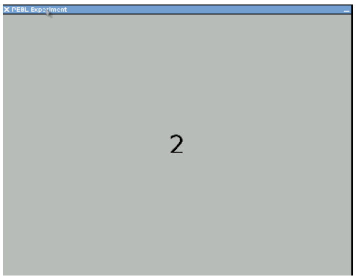
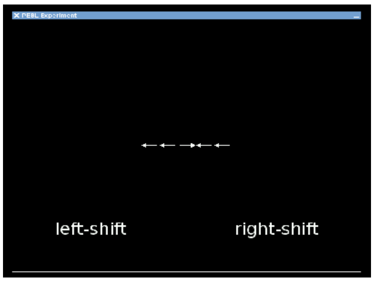
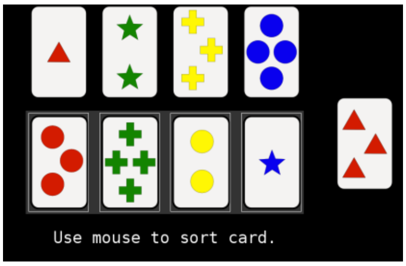
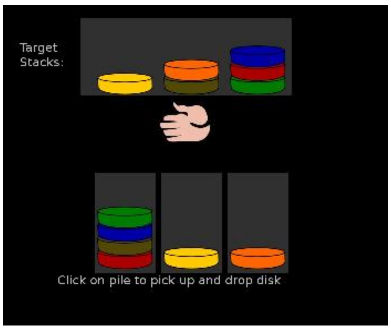
SS28	BCST	DS	FT	TOL
SS29	BCST	DS	TOL	FT
SS30	BCST	FT	DS	TOL

DS = Backward Digit Span (DS)
FT = Flanker Task (FT)
BCST = Berg's Card Sorting Task (BCST)
TOL = Tower of London (TOL)

In this research, any changes in cognitive performance when adults are exposed to 5G exposure were noted in the post-statistical analysis. This is to determine if there is any significant effect of 5G signal exposure on attention and memory function. The cognitive performance of subjects is evaluated with computerized tests via the established Psychology Experiment Building Language (PEBL), a free psychology software package. Based on (Vecsei et al., 2018a), (Stöckel et al., 2017), the four thorough, well-designed, and widely utilized tests that particularly addressed the fundamental executive functions were chosen which include higher-order executive function, working memory, inhibition, flexibility of thought, and reaction preparation and problem-solving (Mueller & Piper, 2014).

Working memory for verbal assessment is measured using the Backward Digit Span Task (DSPAN). Each trial begins with the presentation of a series of single numbers between 0 and 9 in the center of the presented screen and the stimulus interval of 1000 milliseconds. Then, the subject is requested to enter the displayed number in reversal order. Each span-length was assessed twice and ranged from three to ten digits. Prior to each trial, the sequence length is disclosed, and each trial's results are disclosed. If the subject was able to organize the order of the single number correctly in at least one of the two trials for a particular length of span, one digit from the prior trials will be increased. When a person fails to replicate the right sequence on both trials, the test is deemed invalid. The Flanker task is used to measure attention and inhibitory control. Each trial begins with a fixation cross presentation for 500 milliseconds, followed by an 800-millisecond horizontal array of five arrows. The flanking arrows in the congruent condition all point in the same direction as the target arrow (→→→→→), while those in the incongruent condition all point in the opposite direction (←←→←←) and neutral condition (−−→−−). The participant is instructed to concentrate on the direction of the center arrow and quickly press the computer keyboard using the right or left shift key button and ignoring the flanking arrows. The participants are informed of the outcomes following each experiment. Every study had a complete randomization. Accurate information is considered as a covariate while RT is also considered. This is since changing one variable frequently affects the other variables as well, for example, to the speed-accuracy trade-off. Thus, reading value of RT residuals mean in which adjusted for accuracy (RT-acc) are used to assess task-specific processing speed, whereas the score for mean RT interference residuals (RT-accinterference) are important to identify inhibitory control of attention. The Berg's Card Sorting Task (BCST) measures a person's capacity for set shifting and hence, their capacity for cognitive flexibility (Gläscher et al., 2012). The primary need of the assignment is to arrange the stimulus cards the four piles other cards according to their same characteristics in terms of color, shape, and quantity of symbols for every task. The classification rule is kept a secret from the participants, but after each try, they receive feedback (either "correct" or "incorrect") on whether the card was categorized correctly or incorrectly based on the current rule. The categorization rule changes when the cards are successfully sorted ten times in a row. Until all 128 cards are sorted, testing is still in progress. The task’s important outcomes are the number of perseverative errors, or the number of mistakes made when the individual applied the identical rule as in the trial preceding. This variable gives a broad indication of a person's capacity for flexible rule change (or rule abandonment). Contrarily, Non-Perseverative mistakes assess a person's capacity to adhere to a set of rules. The Tower of London task (TOL) is used to evaluate participants' capacity for problem-solving and response planning. A stack of discs must be rearranged by participants in the task such that it matches the given arrangement of the prepared discs. Subjects are instructed to move one disc at a time and not to add more discs to a full pile. Participants are further asked to attempt to complete the job in as few steps as feasible. The 12 tasks in the standard stimulus set are built around 3 discs and a few different pile heights (1, 2, 3). The desired result evaluates the number of trials with flawless answers (TOLpercent success, or trials solved in the smallest move), as well as the time the subject takes to make the first move on every problem (TOLfirstmove). Cognitive functioning components using PEBL test and the outcome measure are tabulated in Table 3.12.

Table 3.12 Cognitive functioning components using PEBL test and the outcome measure.

Cognitive Function Component	**PEBL Test**	**Cognitive Outcome Measure**
Working memory	Backward Digit Span Task 	1. $\text{Dspan}_{\text{backward}}$
Inhibitory control of attention (selective attention) and processing speed	Flanker Task 	1. RT-acc (controlled for accuracy) 2. $\text{RT-acc}_{\text{interference}}$ (congruent minus incongruent conditions)
Cognitive flexibility – shifting abilities	Berg's Card Sorting Task 	1. Correct, % 2. Perseverative error, % 3. Non-perseverative error%
Response planning and problem-solving abilities	Tower of London Task 	1. Success, % (trials solved in the minimum number of moves)result 2. First move time, seconds (the time needed until first

move for each problem)

The physiological changes, focusing on three vital parameters, i.e., body temperature, blood pressure, and heart rate also were recorded. The blood pressure and heart rate of each subject were recorded using a blood pressure wrist measurement device Omron Automatic Blood Pressure Monitor HEM-7320 and Omron Forehead Thermometer MC-720 as shown in Figure 3.10 (a) and Figure 3.10 (b) that complied with the WHO’s classification.

(a) (b)

Figure 3.10 (a) Omron Automatic Blood Pressure Monitor HEM-7320 and (b) Omron Forehead Thermometer MC-720

The aim of this investigation was to measure the possible effects of 5G base station exposure on the physiological parameters and cognitive performance of the subjects as illustrated in Table 3.13. The first dataset consists of Body Temperature, Blood Pressure and Pulse that were recorded before (Pre-Exposure), during (Exposure) and after (Post-Exposure) the assessment of 5G exposure. The Body Temperature, blood pressure and the pulse were recorded in Celsius (°C), millimeters of mercury (mmHg), and beats per minute (BPM) respectively. The second dataset was measured during exposure of 5G only. It consists of cognitive function component data parameters, which, were computed from the Psychology Experiment Building Language (PEBL) tests of Backward Digit Span Task (DSPAN) and Flanker Task, with outcome (Controlled for Accuracy (RT-A). Next, Berg's Card Sorting Task has three measured outcomes of Correct Percentage (C%), Percentage of Perseverative error (PE) and Percentage of Non-perseverative error (NPE). Lastly, the cognitive task named Tower of London Task has two outcomes, which are the Percentage of Success (S %) and the time needed until first move for each problem (FM). The physiological dataset involves 12 columns of normalized data parameters but for the analysis, the data is divided into each parameter based on the physiological parameter,

Table 3.13 Dataset parameter for physiological and cognitive.

<table>
<tr><th colspan="9">Cognitive Dataset</th></tr>
<tr><th>Type of Subject</th><th>Type of Exposure</th><th>Backward Digit Span Task (DSPAN)</th><th>Flanker task</th><th colspan="3">Berg's Card Sorting Task (BCST)</th><th colspan="2">Tower of London task (TOL)</th></tr>
<tr><td>EHS & Non EHS</td><td>Sham, 5G 700 MHz, 5G 3.5 GHz & 28 GHz</td><td>DSPAN Data</td><td>RT-A & RT-Ai Data</td><td>C % Data</td><td>PE Data</td><td>NPE Data</td><td>S% Data</td><td>FM Data</td></tr>
</table>

<table>
<tr><th colspan="6">Physiological Dataset</th></tr>
<tr><th>Type of Subject</th><th>Type of Exposure</th><th>Body Temperature</th><th>Systolic Blood Pressure</th><th>Diastolic Blood Pressure</th><th>Heart Rate</th></tr>
<tr><td rowspan="2">EHS & Non EHS</td><td rowspan="2">Sham, 5G 700 MHz, 5G 3.5 GHz & 28 GHz</td><td colspan="4">Pre-Exposure
During Exposure
Post-Exposure</td></tr>
<tr><td>PreBT, ExpBT & PostBT Data</td><td>PreSYS, ExpSYS & PostSYS Data</td><td>PreDia, ExpDIA & PostDIA Data</td><td>PreHR, ExpHR & PostHR Data</td></tr>
</table>

which are Body Temperature recorded before 5G exposure (PreBT), the Body Temperature recorded during 5G exposure (ExpBT), the Body Temperature recorded after 5G exposure (PostBT), the Diastolic Blood Pressure recorded before 5G exposure (PreDIA), the Diastolic Blood Pressure recorded during 5G exposure (ExpDIA), the Diastolic Blood Pressure recorded after 5G exposure (PostDIA), the Systolic Blood Pressure recorded before 5G exposure (PreSYS), the Systolic Blood Pressure during 5G exposure (ExpSYS), the Systolic Blood Pressure recorded after 5G exposure (PostSYS), the Heart Rate recorded before 5G exposure (PreHR), the heart rate recorded during 5G exposure (ExpHR) and the heart rate recorded after 5G exposure (PostHR).

3.4.5 Statistical Analysis

In order to quantitatively summarize the research findings when testing a hypothesis, the p-value method is used. The value under the null hypothesis of no impact or difference is the probability of experiencing an outcome that is equally likely as or more unlikely than what was seen. The purpose of these hypotheses is to determine whether the experiments were valid and whether the radiation from the 5G base station antenna significantly alters or influences the parameters being examined. P-value is only known following the observation of a result. Rejecting one hypothesis and accepting the other are the two possible outcomes of the hypothesis test. The area in the crucial zone is where the importance level must be determined. The region to the right or left of the test statistics contains the p-value. If the test statistic is inside the crucial zone, the p-value is less than the level of significance. As a result, the null hypothesis—that there is no impact or difference—can be rejected, and it is concluded that there is an effect. In this study, there are two hypotheses that must be proven, explained as follows.

- First, the null hypothesis: There is no statistically significant difference with respect to any of the adult subjective complaints and tests on cognitive function performance, well-being conditions and physiological parameters as recorded during sham exposure, relative to standardized 700 MHz, 3.5 GHz and 28 GHz 5G field exposures. This means that there is no effect of 5G field exposure for all investigated frequency in both groups, EHS and Non-EHS groups on cognitive function performance, well-being conditions and physiological parameters.
- Second, alternative hypothesis: The data analysis shows that there is a statistically significant difference between one or more adult subjective complaints and tests on cognitive function performance, well-being conditions and physiological parameters as recorded during sham exposure, relative to standardized 700 MHz, 3.5 GHz and 28 GHz 5G field exposures. This means that there is an effect of 5G field exposure for all investigated frequency in both groups, EHS and Non-EHS groups on cognitive function performance, well-being conditions and physiological parameters.

The effects of the 5G exposures on the physiological parameters, cognitive performance and well-being parameter are studied via the three sets of analyses performed, as follows.

- The aim of the first analysis of four physiological parameters is to determine whether there is a significant difference between the four different signals in affecting the average body temperature, Blood Pressure Systolic (SYS), Blood Pressure Diastolic (DIA), and heart rate under two exposure conditions, i.e., pre-exposure and post-exposure as in (Malek et al., 2015; Masrakin et al., 2019). The independent sample t-test is applied to analyze the probability effects between pre-exposure and post-exposure of the Sham, 700 MHz, 3.5 GHz and 28 GHz 5G signals on the physiological parameters. Besides that, any significant different values of physiological parameters between Signal (Sham/700 MHz/3.5 GHz/28GHz) and Group (EHS and Non-EHS) will indicate that there are indeed effects of 5G exposure towards physiological parameters of adults.

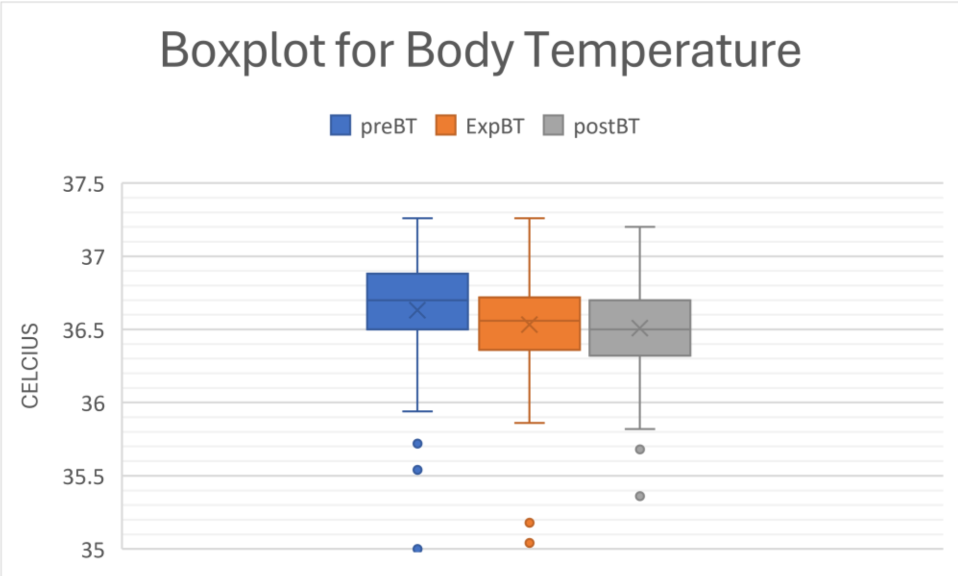
- The second analysis is performed using Analysis of Variance (ANOVA) repeated measure to determine whether if any four measured cognitive performance components is affected by the 700 MHz or 3.5 GHz or 28 GHz for both EHS and Non-EHS subjects. Any significant different values of cognitive outcome measure parameters between Signal (Sham/700 MHz/3.5 GHz/28 GHz) and Group (EHS and Non-EHS) will indicate that there are indeed effects of 5G exposure towards the cognitive performance of adults.

3.5 Multi-Stage Feature Selection (MSFS) and Machine Learning

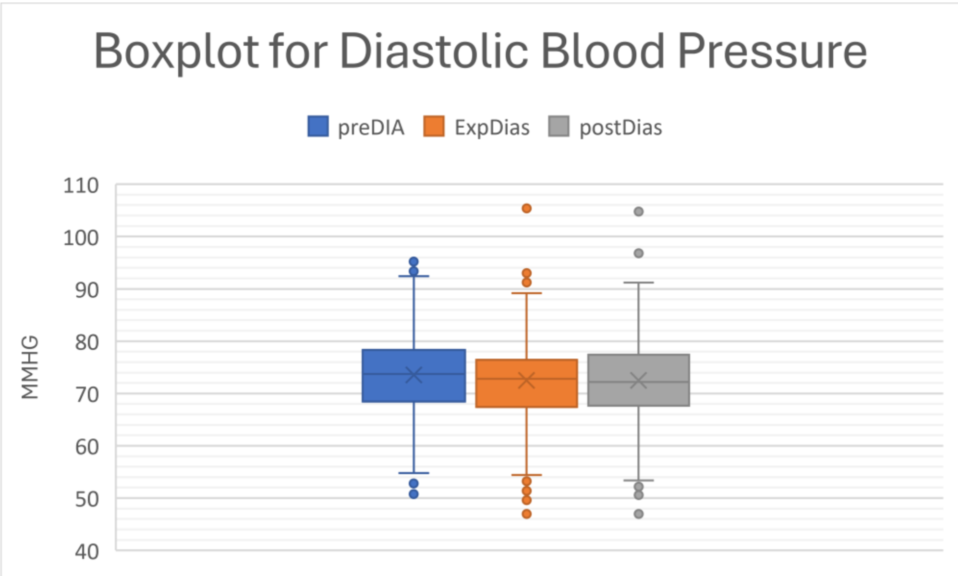
For this part of the research, MSFS refer to an extended or optimized version technique that incorporates elements or techniques from different approaches of selecting features using formulation of hybrid data feature for 5G base station antenna health effect detection classification to improve the overall effectiveness. MSFS addresses limitations associated with traditional feature selection methods by combining their strengths and enhancing adaptability, interaction modeling, and context-awareness. The modification enhancing and adapting the feature selection process of MSFS based on insights gained from analyzing data related to 5G base station exposure. The performance of KNN, SVM, Ensemble Methods and PNN classifiers have been validated, and the evaluation includes metrics such as classification accuracy, precision, f1-score, sensitivity, and specificity. These metrics provide a comprehensive assessment of how well the classifier, integrated with the MSFS, performs in terms of its ability to accurately classify data.

3.5.1 Data Preparation

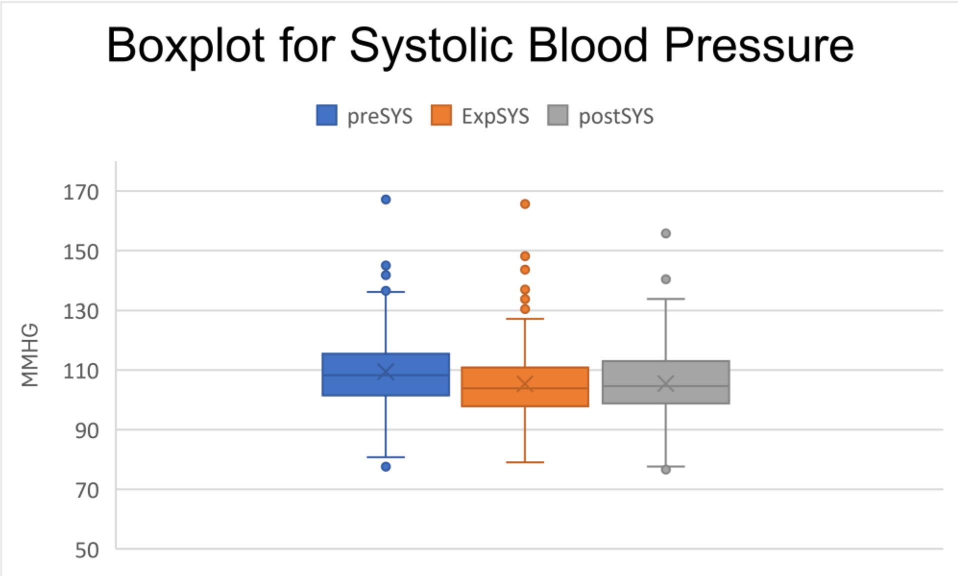
The initial processing steps before applying to machine learning is to understand nature of data. The supervised machine learning approach with data that focus on the methods that are designed to predict or classify an outcome of interest (Jiang et al., 2020) as data can be of various forms, such as structured, semi-structured, or unstructured (Sarker, 2021). The dataset involved for physiological and cognitive from this analytical epidemiological study is categorical outcomes as well as numerical parameter as illustrated in Figure 3.11 (a) until Figure 3.11 (g). Categorical outcomes and referred to as classification in the machine learning literature (Jiang et al., 2020). Categorical outcomes refer to situations where the target variable or response variable falls into distinct categories or classes, observations are predicted to belong to the most commonly occurring class/category in a node. The features involved for cognitive consists of subject data, exposure data and cognitive data. Subject data category includes whether that subject is EHS or Non-EHS within the healthy and normal adult age. The exposure data refers to the 5G signal (700 MHz, 3.5 GHz and 28 GHz) emitted includes Sham as one of the factors influencing cognitive performance. The cognitive data category directly measures cognitive abilities, providing insights into the actual cognitive functioning of individuals in the study. By combining these three types of data, the connections between subject characteristics, environmental exposures, physiological and cognitive outcomes can be explored. The analysis aims to evaluate which factors related to subjects or exposures are linked to variations in cognitive performance. In the case of the physiological dataset, subject data and exposure data share the same features. Subject data helps put physiological responses into context by considering individual differences that might influence how the body reacts to exposures. The physiological data, recorded before, during, and after exposure, includes parameters such as body temperature, blood pressure, and heart rate. Descriptions of the selected six features (attributes or variables) as tabulated in Table 3.14 of the analysis.


Boxplot for Body Temperature
preBT
ExpBT
postBT
CELCIUS
37.5
37
36.5
36
35.5
35

(a)

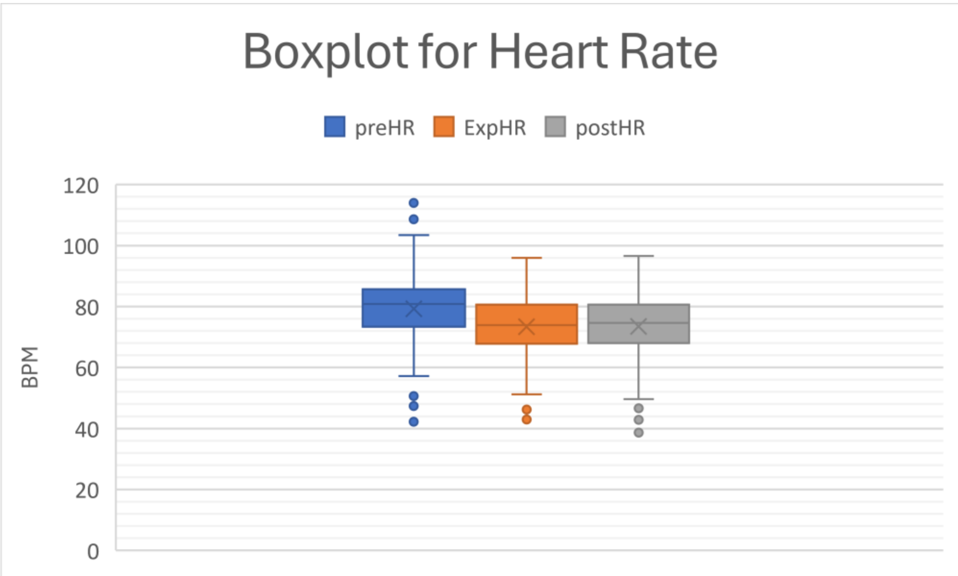

Boxplot for Diastolic Blood Pressure
preDIA
ExpDias
postDias
MMHG
110
100
90
80
70
60
50
40

(b)



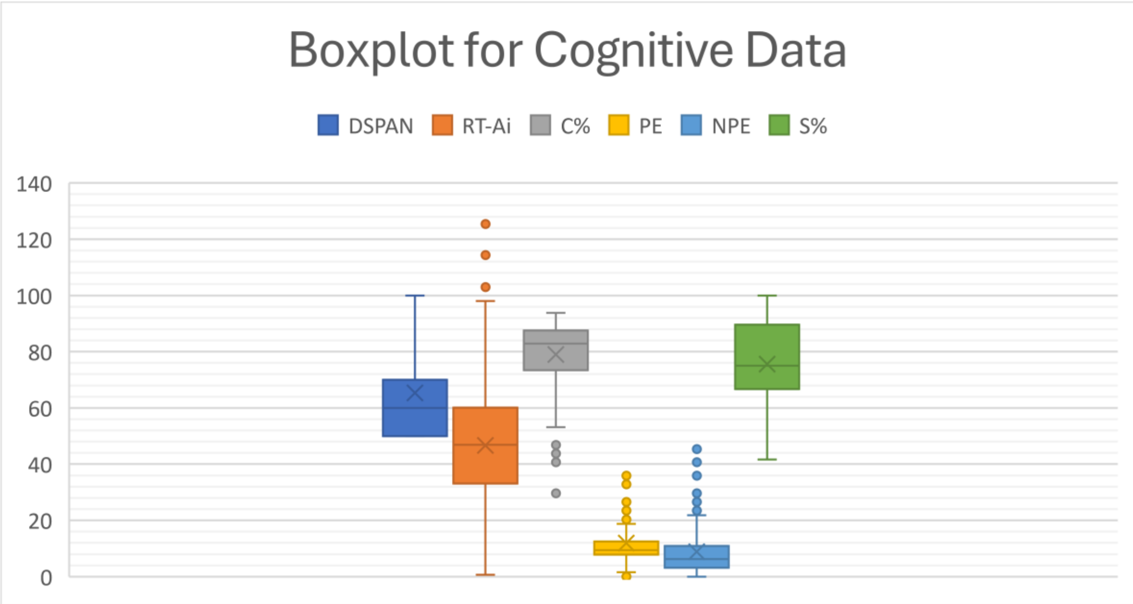
Boxplot for Systolic Blood Pressure
preSYS
ExpSYS
postSYS
MMHG
170
150
130
110
90
70
50

(c)

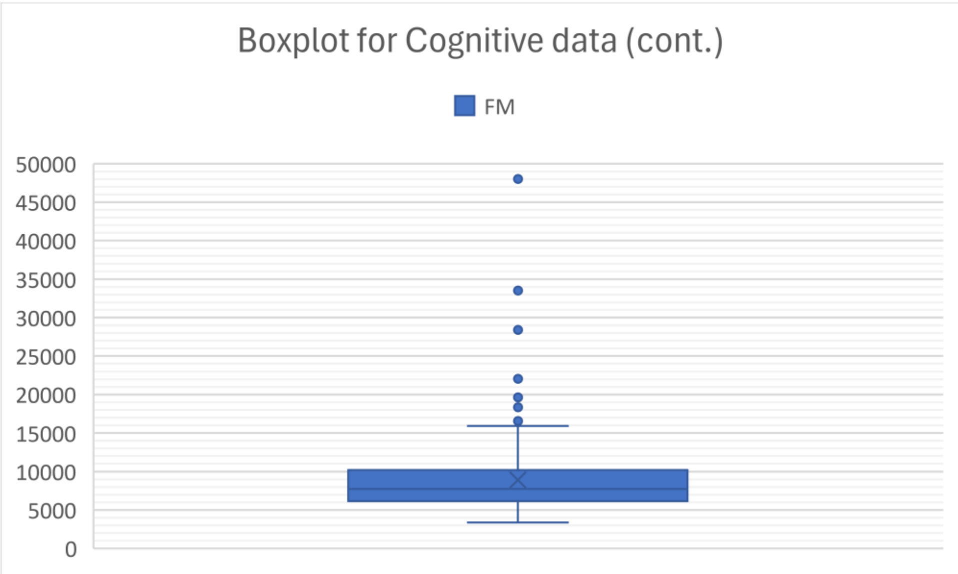


Boxplot for Heart Rate
preHR
ExpHR
postHR
BPM
120
100
80
60
40
20
0

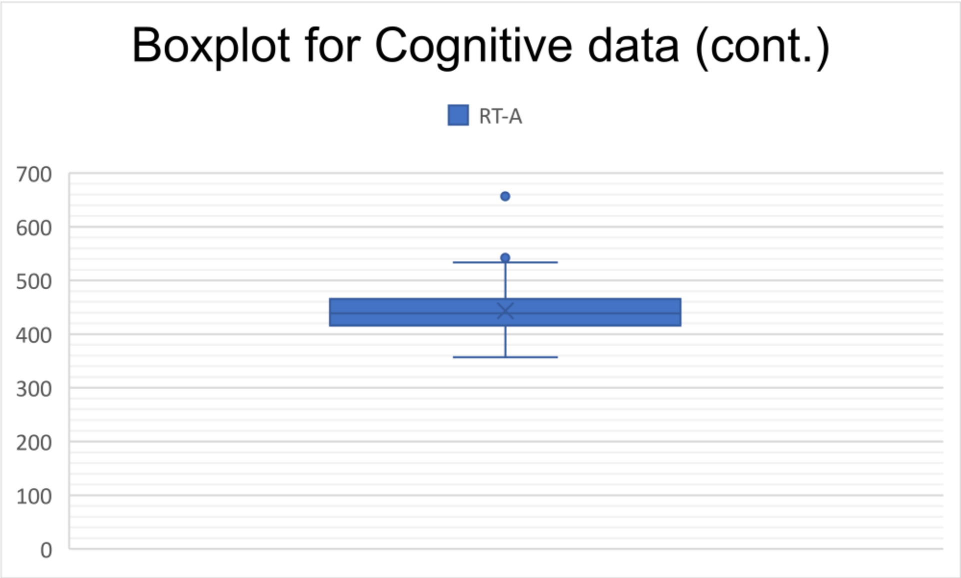
(d)


Boxplot for Cognitive Data
DSPAN
RT-Ai
C%
PE
NPE
S%
140
120
100
80
60
40
20
0

(e)


Boxplot for Cognitive data (cont.)
FM
50000
45000
40000
35000
30000
25000
20000
15000
10000
5000
0

(f)



(g)

Figure 3.11 (a) The boxplot for dataset collected from pre-exposure, exposure, and post-exposure for body temperature physiological category, (b) The boxplot for dataset collected from pre-exposure, exposure, and post-exposure for diastolic blood pressure physiological category (c) The boxplot for dataset collected from pre-exposure, exposure, and post-exposure for systolic blood pressure physiological category, (d) The boxplot for dataset collected from pre-exposure, exposure, and post-exposure for physiological category, (e) The boxplot for dataset collected from cognitive data category for DSPAN, RT-Ai, C%, PE, NPE and S%, (f) The boxplot for dataset collected from cognitive data category for FM and lastly (f) The boxplot for dataset collected from cognitive data category for RT-A.

Table 3.14 Descriptions of the selected six features (attributes or variables) of the analysis.

Features	**Type**	**Feature Type**	**Description (Domain)**
Subject (EHS, Non EHS)	Nominal	Input	EHS represents the group of subjects which members have attributed complaints suspected to 5G exposure (IEI-EMF). On the contrary, the Non-EHS group denotes the reference group – subjects without any complaints (Non-IEI-EMF category).
Frequency of 5G Signal Exposure	Numeric	Input	Sham, 700MHz, 3.5GHz and 28GHz
Cognitive data	Numeric	Input	DSPAN Data, RT-A Data, C% Data, PE Data, NPE Data, S% Data, FM Data
Physiological data	Numeric	Input	PreBT, ExpBT, PostBT, PreSYS, ExpSYS, PostSYS, PreDia, ExpDIA, PostDIA, PreP, ExpP & PostP Data
Subject	Binary	Output	EHS or Non-EHS subject
Exposure	Binary	Output	Sham, 700MHz, 3.5GHz or 28GHz

3.5.2 Initial Data Processing without Feature Selection

The raw data underwent an initial processing step without employing any feature selection technique. This implies that all features or attributes present in the dataset were utilized as inputs for the classifier without prior filtering or dimensionality reduction. The classifier involved are KNN, SVM, Ensemble Method and PNN. The classifier received the raw data directly, and predictions or classifications were made based on this unaltered dataset. Subsequently, the outcomes of the classifier predictions were assessed or evaluated.

3.5.3 Multi-Stage Feature Selection

MSFS is a technique as shown in Figure 3.12 used in machine learning and data processing to enhance the performance of models by selecting the most relevant features from the input data.

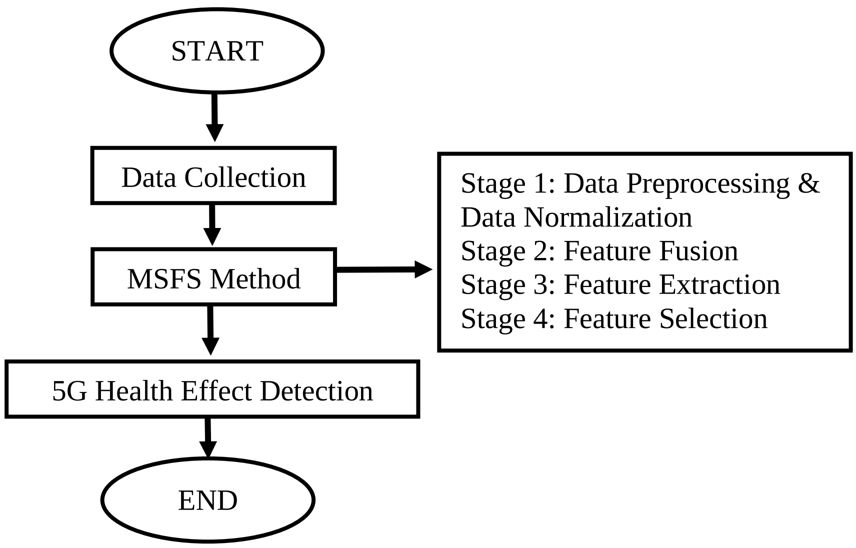


Figure 3.12 MSFS Technique

For MSFS, the first stage consists of data pre-processing and data normalization methods. Data normalization was performed after removing outliers to avoid their influence on the scaling process. The second stage consists of feature fusion method, while the third stage and the fourth stage consist of feature extraction and feature selection, respectively.

3.5.4 Data Pre-Processing

The primary goal of stage 1 is to clean and prepare the raw data for further analysis in subsequent stages, through the removal of outliers after calculating their Interquartile Range (IQR) process involves identifying and eliminating data points that deviate significantly from the expected range of values within the dataset. First, eliminate outliers to address potential distortions in subsequent normalization processes. Afterward, apply average imputation to rectify out-of-range values, ensuring a more streamlined and effective data preparation.

3.5.5 Data Normalization

Data normalization is the process of scaling and centering the features of the dataset. Normalized datasets are to ensure consistency and remove biases that could impact feature selection. This step is important to create a model with good accuracy (Elkhouly et al., 2023). In this study, 20 normalization techniques as shown in Appendix B are applied to the data which are; Z-score Normalization (ZS), Linear Scaling (LS), Binary Normalization (BNN), Bipolar Normalization (BPN), Min-Max Scaling (MMS), t-score Normalization

(TS), Decimal Inverse Logarithmic Scaled Normalization (DILSN), Relative Mean Normalization (RMN), Relative Standard Deviation Normalization (RSDN), Variation Normalization (VN), Robust Normalization (RN), Relative Interquartile Normalization (RIN), Differential Moment Normalization (DMN), Absolute Percentage Error Normalization formula 1 (APE1), Absolute Percentage Error Normalization formula 2 (APE2), Arctan APE formula 1 (ARCAPE 1), Arctan APE formula 2 (ARCAPE 2), Gaussian Normalization (GN), Relative Sum Squared Value (RSSV), and Relative Logarithmic Sum Squared Value (RLSSV). The analysis was conducted by choosing the ideal normalization techniques or approaches to analyze the data prior to the machine learning methodology. The descriptive statistical analysis of the physiological and cognitive datasets was performed in order to determine the *p*-value and *F*-value for each normalization method for each dataset. The second process is applying different statistical analysis to remove any normalized method that leads to data redundancy or change in the relative levels which will cause a change in the shape of the signal and input data. This process had to pass through 3 different statistical analysis; paired *t* -test for mean difference, *t*–test for correlation significance then *F*–test for data variability. Then, in pairs, data normalization methods are compared using paired *t*-test at 0.05 level of significance, to test the hypothesis:

H0: There is no significant difference in the performance of data normalization methods when compared in pairs.

H1:. There is a significant difference in the performance of at least one pair of data normalization methods.

In simpler terms, the null hypothesis suggests that any observed differences in the performance of data normalization methods are due to random chance, while the alternative hypothesis implies that there is a genuine and significant difference in the performance of at least one pair of methods. During the paired t-test, if the p-value is less than 0.05, one may reject the null hypothesis in favor of the alternative hypothesis, indicating that there is sufficient evidence to suggest a significant difference in the performance of at least one pair of data normalization methods. If the p-value is greater than 0.05, one would fail to reject the null hypothesis, suggesting that observed differences are likely due to random variability.

The *t*-test statistic value calculated using Equation (3.3), which produced *p*-value then a decision should be made according the decision rule; if *p*-value < α, reject *H0*.

$$t_{STAT} = \frac{\overline{D} - \mu_D}{\frac{S_D}{\sqrt{m}}} \tag{3.3}$$

This should be repeated for two different normalization methods at a time Out of the 20 normalization techniques studied, only the best 10 datasets with the best statistical readings are selected to be combined. The data is pre-selected to form a normalized dataset after has been normalized and statistically studied. Sequence of statistical analysis is illustrated with the flowchart in Figure 3.13.

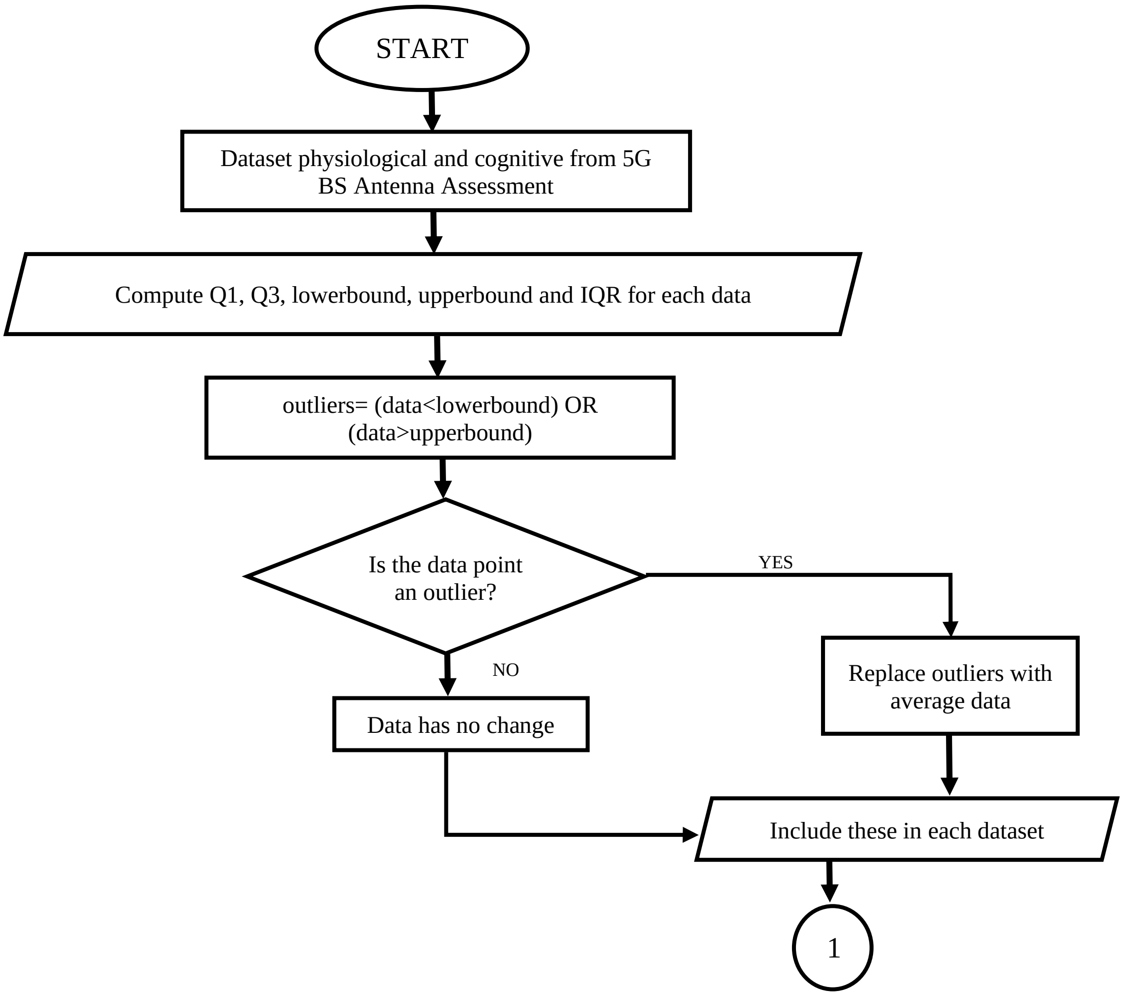
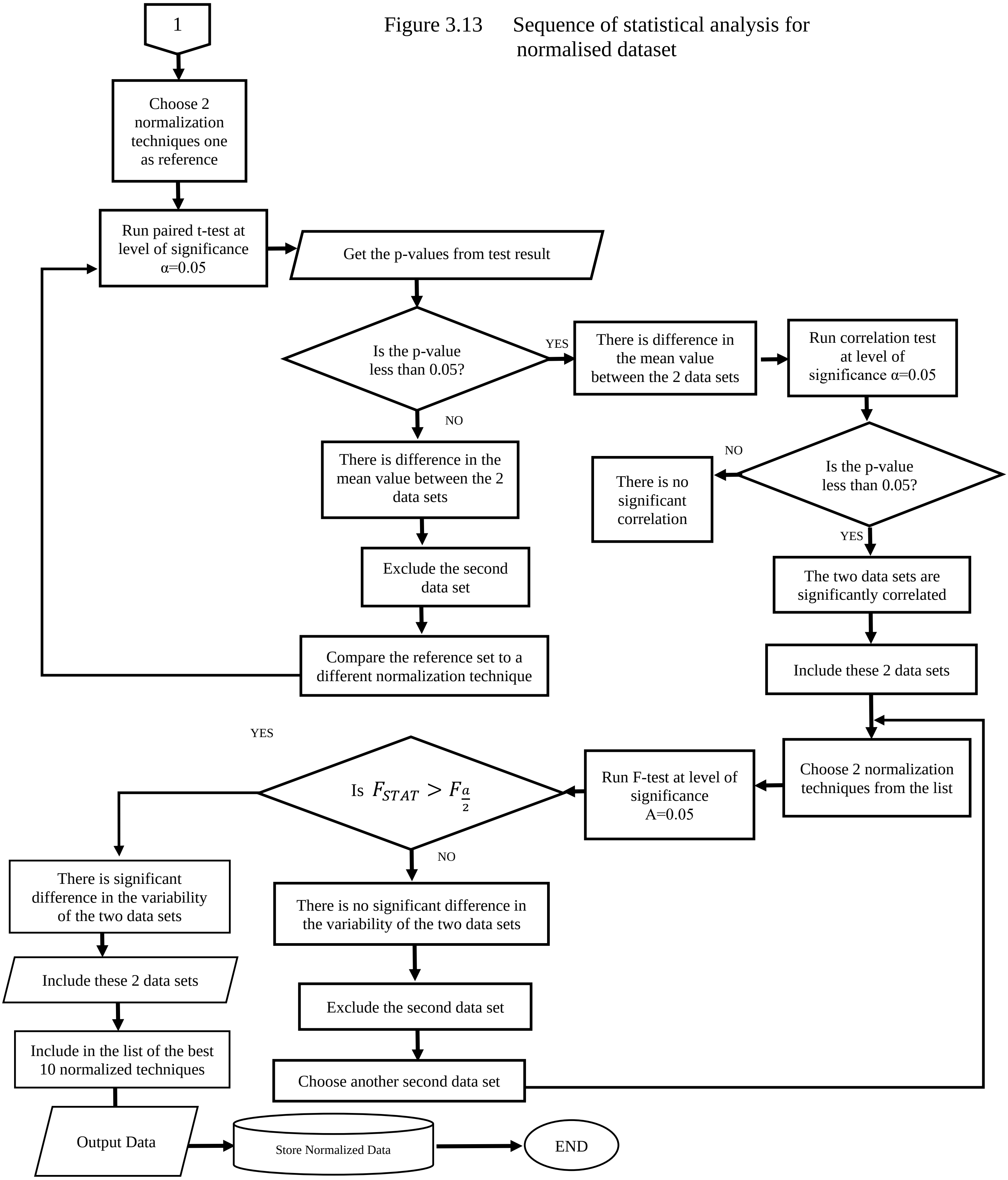

START
Dataset physiological and cognitive from 5G BS Antenna Assessment
Compute Q1, Q3, lowerbound, upperbound and IQR for each data
outliers= (data<lowerbound) OR (data>upperbound)
Is the data point an outlier?
YES
NO
Data has no change
Replace outliers with average data
Include these in each dataset
1

Figure 3.13 Sequence of statistical analysis for normalised dataset



3.5.6 Feature Fusion

The integration of feature fusion technique is merging multiple sets of features into hybrid feature dataset which are their exposure data, type of subject and the cognitive and physiological dataset. Feature fusion involves combining information from multiple sets of features into a unified dataset to generate a hybrid feature dataset that incorporates the most valuable information from each contributing set of features. Each feature set as detailed below:

Exposure Data: Information related to the exposure of subjects which are Sham, 5G 700 MHz, 5G 3.5 GHz and 28 GHz.

Type of Subject: Categorization of subjects based on EHS and Non EHS

Cognitive and Physiological Dataset: Data related to cognitive functions and physiological measurements.

Concatenation feature fusion technique involves combining different sets of features by joining them along a common axis as shows on Figure 3.14. The flow technique for feature fusion is illustrated in Figure 3.15. The integrated hybrid feature dataset is the output result from the MSFS method in this study.

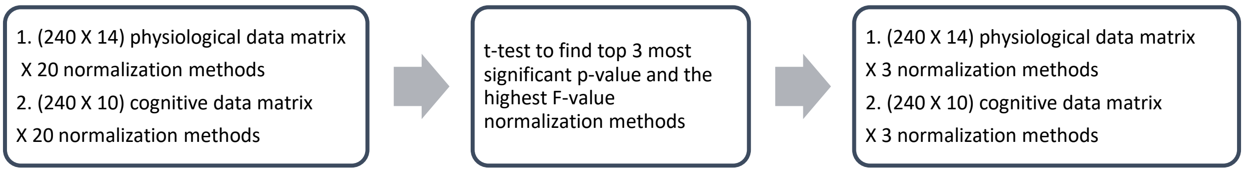


Figure 3.14 Modification process to the data matrix due to normalization technique.

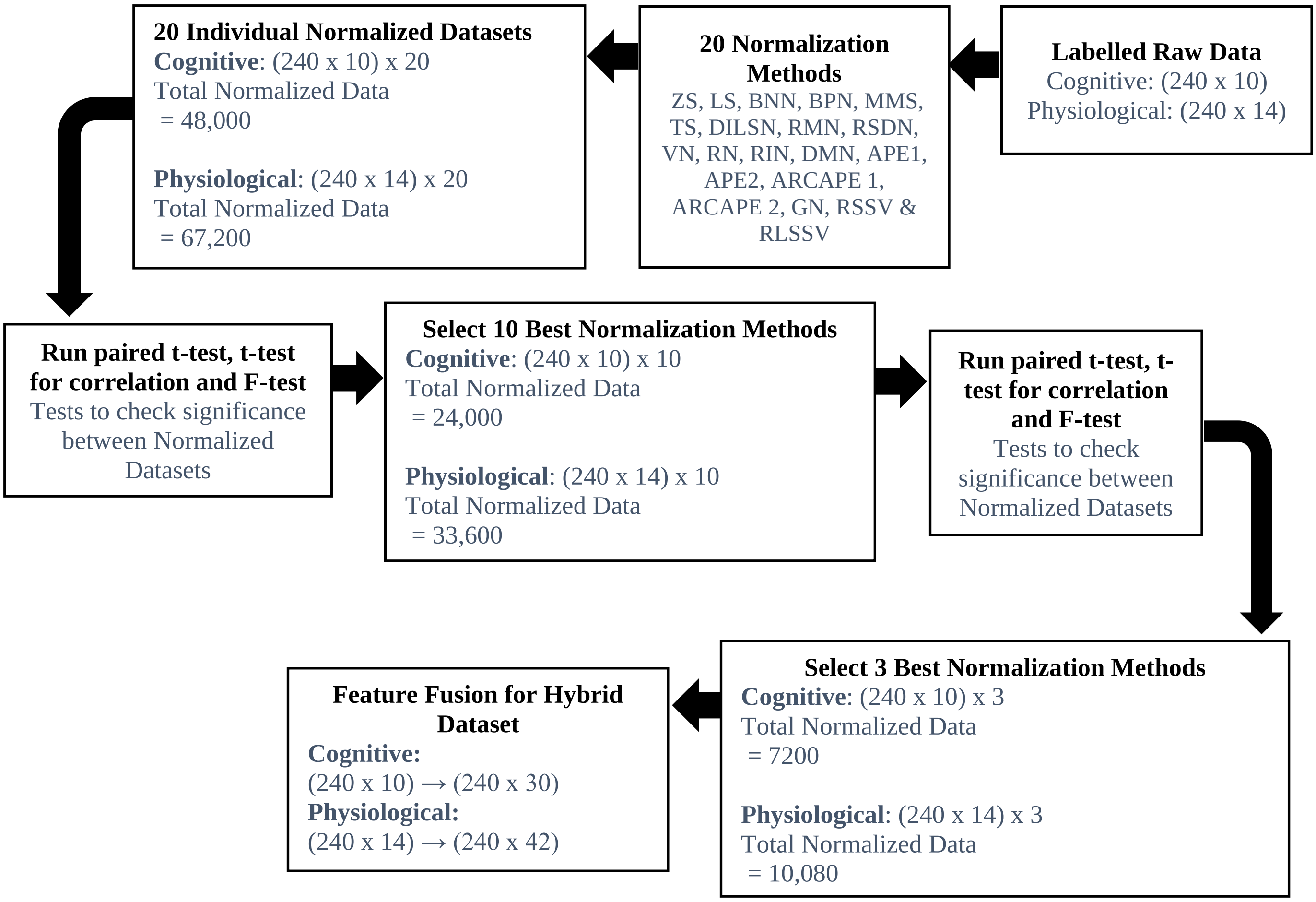


Figure 3.15 Flow Technique for Feature Fusion

3.5.7 Feature Extraction

The exploration and exploitation of the data will be insufficient during the feature selection as the features are reduced at the initial stage. As a result, only some redundant features are selected, and some useful features are lost due to poor data management. The proposed multi-stage approach consists of feature engineering within natural language processing, signal reconstruction, feature selection, feature extraction, improved learning techniques for resampling and cross-validation, and the configuration of hyperparameters. Conventional single stage feature selection has the drawback of possibly selecting data after eliminating useful data during feature extraction stage after the data has undergo first phase of MSFS and subjected to three distinct data normalization techniques. PCA for the feature extraction stage in a subset of the hybrid dataset represents a comprehensive and consolidated view of relevant information from exposure data, subject type, and cognitive and physiological datasets. PCA works by transforming the original features into a new set of uncorrelated variables called principal components, ordered by their variance. The first few principal components capture most of the variability in the data, allowing for a reduced-dimensional representation. PCA application to this hybrid dataset to reduce its dimensionality while retaining as much of the original variability as possible. The process illustrated in Figure 3.17 outlines the sequential steps of PCA. It commences by standardizing the features, ensuring uniform contributions by subtracting the mean and dividing by the standard deviation. Following this, the covariance matrix of the standardized features is calculated, and the eigenvectors and eigenvalues of this matrix are obtained. In the subsequent stage of principal component selection, eigenvalues are sorted in descending order, and the top-k eigenvectors are chosen to constitute the principal components. Lastly, the original data is projected onto these selected principal components to derive the reduced-dimensional representation.

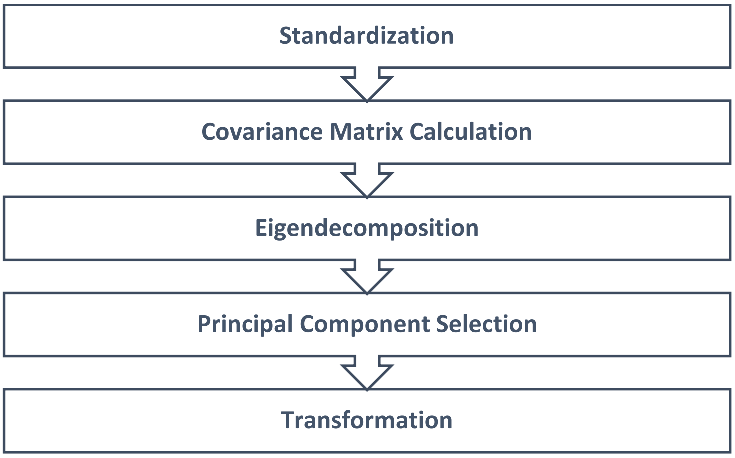


Figure 3.16 Flow Technique for Feature Extraction

3.5.8 MSFS Process

In the context of machine learning, the MSFS method is a systematic approach to choosing and refining features used in a predictive model. This method involves multiple steps or stages to carefully curate the feature set, ensuring that the model is trained on the most relevant and informative variables. The proposed hybridized MSFS method consists of four stages. The first stage consists of data normalization methods and data pre-selection. The second stage consist of feature extraction methods, while third stage and fourth stage consist of feature selection and feature fusion, respectively. The selection of data normalization methods and features are done by computing the p-value and F-value in the first stage. The raw data samples go through these stages in order to identify the best data normalization techniques, the best feature extraction methods and the optimum features to be hybridized (fused). The features from stage 3 are fused together using feature fusion technique in stage 4. This newly hybrid feature dataset will be used for 5G base station antenna health detection framework.

3.5.9 Classification Analysis

The classification method that classifies the data into two categories, e.g., whether or not the subject category and exposure from each data parameter was observable. Independent variables, such as the presence of 5G signal and their frequency, subject, cognitive data, and physiological data on human. A principal assumption of machine learning is that the training data is the representation of the distribution from which test data (future data) will be picked. The data are independent and distributed identically.

which remains an assumption of this study. The analysis is performed using MATLAB (MathWorks Inc) R2022. The KNN, SVM, Ensemble Method, Naive Bayes and PNN classifier were used in conjunction with the selected feature outcome from the MSFS. After the feature selection, the selected subset of features is used as input to all the classifiers. KNN classifies data points based on the majority class of their k-nearest neighbours. The distance metric used for defining "closeness" might be influenced by the selected features. The mathematical equation in Equation (3.4) is used to calculate the distance between data points and determine the nearest neighbours. KNN is a type of instance-based learning or lazy learning, where the model does not explicitly learn a function from the training data but memorizes the entire training dataset instead. The most common distance metric is Euclidean distance. For the hybrid datasets outcome from the MSFS given dataset, two data points A=$(a_1, a_2, \dots, a_n)$ and B=$(b_1, b, \dots, b_n)$ in which represent the data features involved. The Euclidean distance (d) between these two points is calculated as shown in (3.4). n is the number of dimensions (features) in the dataset and $a_i - b_1$ are the values of the i-th feature for points A and B, respectively.

$$d(A,B) = \sqrt{\sum_{i=1}^{n} (a_i - b_1)^2} \tag{3.4}$$

This formula computes the straight-line distance between two points in an n-dimensional space. In the context of KNN, this distance metric is used to find the nearest neighbours of a given data point. Once the distances between the query point and all other points in the dataset are calculated, the K-nearest neighbours are determined by selecting the K points with the smallest distances. The most common class label (for classification tasks) or the average label (for regression tasks) among these K neighbours is then assigned to the query point. SVM classifier aims to find a hyperplane that best separates different classes. The choice of features can significantly impact the position and orientation of this hyperplane. In the case of a linear SVM, the mathematical formulation involves the use of vectors and a weight vector, along with a bias term as the equation for a linear SVM in Equation (3.5). For a two-class classification problem, where x is the input feature vector, w is the weight vector, x is the input feature vector, b is the bias term, · denotes the dot product between w and x, sign (·) is the sign function, which returns +1 for positive values, -1 for negative values, and 0 for zero.

$$f(x) = sign(w \cdot x + b) \tag{3.5}$$

The decision boundary is determined by the hyperplane $w \cdot x + b = 0$, and the sign of f(x) indicates the predicted class (either +1 or -1). For training the SVM, the goal is to find the optimal w and b that maximize the margin between the two classes while minimizing classification errors. This optimization problem involves the use of Lagrange multipliers and leads to the formulation of a dual optimization problem. For a non-linear SVM, the kernel trick is applied to map the input features into a higher-dimensional space. The equation is then expressed in terms of the transformed features, allowing SVMs to learn non-linear decision boundaries. In a Random Forest, individual decision trees are trained on random subsets of the data and features. The final prediction is made by aggregating the predictions of all individual trees. The ensemble prediction $\hat{y}_{(ensemble)}$ as in Equation (3.6) is obtained through a majority vote for classification tasks, where $\hat{y}_i$ is the prediction of the i-th decision tree.

$$\hat{y}_{(ensemble)} = mode(\hat{y}_1, \hat{y}_2, \dots, \hat{y}_N) \tag{3.6}$$

The Naïve Bayes algorithm is based on Bayes' theorem and the naïve assumption of feature independence given the class. In the context of a binary classification problem with classes C_1 and C_2 as outlines in Equation (3.7) and features $X = (x_1, x_2, \dots, x_n)$ the Naïve Bayes algorithm calculates the posterior probability of each class given the features.

$$P(C_k|X) = \frac{P(X|C_k) \cdot P(C_k)}{P(X)} \tag{3.7}$$

$P(C_k|X)$ is the posterior probability of class C_k given the features.

$P(X|C_k)$ is the likelihood of the features given class C_k. The naïve assumption is that the features are conditionally independent given the class, so this is often calculated as the product of individual feature probabilities:

$P(X|C_k) = P(x_1|C_k) \cdot P(x_2|C_k) \dots . P(x_n|C_k)$

$P(C_k)$ is the prior probability of class C_k representing the probability of observing class C_k without considering the features.

$P(X) = \sum_{i=1}^{K} P(X|C_i) \cdot P(C_i)$

Classification task of PNN prior probability $P(C_k)$ in Equation (3.8) for each class is estimated based on the proportion of training samples for each class C_k and N is the total number of training samples.

$$P(C_k) = \frac{N_k}{N} \tag{3.8}$$

The model will be trained and tested using k-fold cross validation method. The average classification accuracy for 50 trials are taken and the average classification accuracy will be determined. A good classification accuracy, sensitivity and specificity of the model should be more than 90%, which indicates that 90 out of 100 trials yield correct classification result.

3.5.10 Classifier Performance Validation

Evaluating the performance of all classifiers involved using evaluation metrics to assess how well the classifier is performing in terms of classifying data points with 50 repetitions. The primary metrics which result from the confusion matrix are used to evaluate the classifier's performance include accuracy, precision, f1-score, sensitivity, and specificity respectively (Y. Wang et al., 2021). The equations involved to obtain the evaluation metrics is the classified outputs are compared with the actual test targets and from the confusion matrix computed. Equation (3.9) calculates the accuracy of the model, representing the correctness of the model's classification of data into their respective classes. Equation (3.10) measures precision, also known as positive predictive value, which indicates how well the model identifies positive cases accurately. Equation (3.11) quantifies recall, which measures the proportion of correctly predicted positive instances (true positives) out of all actual positive instances. It evaluates the model's ability to capture positive cases effectively. Equation (3.12) computes the f1-score, which provides a balanced measure of the model's performance by considering both precision and recall. It combines these metrics to assess overall performance. Equation (3.13) represents specificity, also known as true negative rate, which measures how well the model accurately identifies negative cases. Lastly, equation (3.14) corresponds to sensitivity, which is synonymous with recall. It measures the proportion of correctly predicted positive instances (true positives) out of all actual positive instances. Collectively, these equations offer a comprehensive set of metrics to evaluate and analyse the performance of a binary classification model across different aspects, including accuracy, precision, recall, f1-score, specificity, and sensitivity. Calculating the accuracy, specificity and sensitivity are important as to have a successful 5G base station antenna health effect detection and to reduce misclassification in the classification. There are possible chances of high misclassification to happen where misclassification happens when there is health effect of 5G base station antenna, but not detected by the system, or no health effect available but the classifier detects health effects from the five output parameters. Having such possibility will affect the overall efficiency of the system, and thus, must be eliminated or reduced. The assessment of a classifier’s prediction performance is critical to get the decision on its suitability as irrelevant or less essential features can severely affect model performance. Essentially, in this manner, it is critical to examine model viability in a specific data set. The classifier was compared for prediction performance utilizing seven measures to get the choice on its suitability, utilizing precision, accuracy, recall and f1-score.

$$\text{accuracy} = \frac{\text{True}_{\text{positive}} + \text{True}_{\text{negative}}}{\text{True}_{\text{positive}} + \text{False}_{\text{positive}} + \text{True}_{\text{negative}} + \text{False}_{\text{positive}}} \tag{3.9}$$

$$\text{precision} = \frac{\text{True}_{\text{positive}}}{\text{True}_{\text{positive}} + \text{False}_{\text{positive}}} \tag{3.10}$$

$$\text{recall} = \frac{\text{True}_{\text{positive}}}{\text{True}_{\text{positive}} + \text{False}_{\text{negative}}} \tag{3.11}$$

$$\text{f1} - \text{score} = \frac{2 \text{ x precision x recall}}{\text{precision} + \text{recall}} \tag{3.12}$$

$$\text{specificity} = \frac{\text{True}_{\text{negative}}}{\text{True}_{\text{negative}} + \text{False}_{\text{ positive}}} \tag{3.13}$$

$$\text{sensitivity} = \frac{\text{True}_{\text{positive}}}{\text{True}_{\text{positive}} + \text{False}_{\text{ negative}}} \tag{3.14}$$

In the final step of the classifier, a Graphical User Interface (GUI) is designed, as depicted in Figure 3.17. The GUI serves as an interactive platform for users, and MATLAB is employed to create this interface. The purpose is to allow users to input their own data and obtain classification results related to specific parameters, namely, either the subject or the exposure.

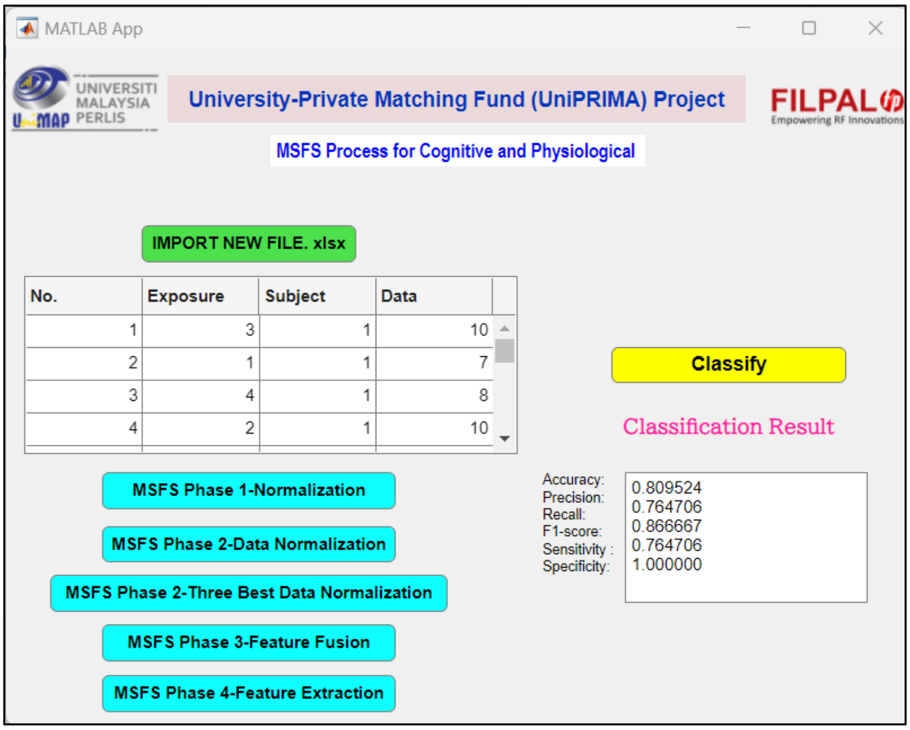


Figure 3.17 GUI Design

3.6 Summary

This research aims in assessing the effect of 5G base station antenna exposure on the adults' health and designing 5G base station antenna health detection based on the framework. The dataset from investigation of the effects of 5G 700MHz, 3.5 GHz and 28 GHz base station antenna fields exposures and Sham on physiological parameters (body temperature, blood pressure and heart rate) and cognitive performance of adults in the double blinded condition will become the inputs for hybridized MSFS using supervised machine learning, then the output from this classifier will be used for 5G base station antenna health effect detection classification based on the proposed parameters and the performance of the MSFS hybrid dataset will be validated in terms of machine learning classification accuracy, precision, f1-score, sensitivity, and specificity. In this research several computational tools are used like; SPSS to develop polynomial representation to the data and analyse it statistically; Excel for data representation and some statistical analysis, MATLAB 2022Rb is used in many parts of the research and it is fundamental tool to build the required classifier and enhance it.

CHAPTER 4 : RESULTS & DISCUSSION

4.1 Introduction

In this chapter the objectives' results are presented and discussed. It starts with Section 4.2, where the results of the assessment outcomes from the 5G base station antenna health effect and their statistical analysis for physiological and cognitive datasets. Followed by Section 4.3, in which different classification methods are applied to the raw datasets to assess the proposed machine learning model. The normalization and their statistical analysis process starts with Section 4.4 where the used data is prepared and pre-processed. Lastly, in Section 4.41 and Section 4.42 discuss on the results of classifier performance for each dataset.

4.2 Assessment Outcomes

The physiological and cognitive parameters outcomes are attained from 60 subjects (30 EHS and 30 Non-EHS) who participate in the experiment and have completed all four 5G base station signal exposures. The analysis is performed to evaluate whether the objectives of this research work are achieved or not.

4.2.1 Statistical Data Analysis for Physiological

The difference in relation to any subjective well-being complaints, cognitive function, and physiological parameters recorded before, during, or after Sham and the 5G signal exposures is assessed using the double-blind, randomized, counterbalanced, and cross-over experiment design. According to 5G specifications, the signals are planned to broadcast at 700 MHz, 3.5 GHz, and 28 GHz. Continuous variables are summed up using descriptive statistics, which include mean, standard deviation, minimum and maximum values. Discrete variables are encapsulated by frequency and percentage. The results of four physiological parameters between Sham and real 5G exposure signals are presented in Table 4.1. The statistical analysis confirms that there is no significant effect ($p>0.05$) of body temperature, SYS, DIA, and heart rate between Sham and 5G signal exposures (5G 700 MHz, 5G 3.5 GHz, and 5G 28 GHz) before and after exposure conditions. This means that the body temperature, SYS, DIA, and heart rate of EHS and Non-EHS adults are almost the same before and after they are exposed to the 5G signal exposures or not exposed. In addition, statistical analysis findings illustrated that there is no statistically significant difference ($p>0.05$) between EHS and Non-EHS in either the pre-exposure or post-exposure circumstances for any of the physiological parameters that were assessed as shown in Table 4.2. This is determined from the independent sample t-test between EHS and Non-EHS groups, which indicates no significant difference ($p>0.05$) between EHS and Non-EHS for all measured physiological parameters in EHS and Non-EHS subjects.

Table 4.1 Descriptive statistics are used to investigate the significant difference between Sham, 700 MHz, 3.5 GHz, 28 GHz signals during each pre-exposure and post-exposure of physiological parameters.

		Sham				**700 MHz**				**3.5 GHz**				**28 GHz**				
		EHS		**Non-EHS**		**EHS**		**Non-EHS**		**EHS**		**Non-EHS**		**EHS**		**Non-EHS**		
		M	**SD**	**M**	**SD**	**M**	**SD**	**M**	**SD**	**M**	**SD**	**M**	**SD**	**M**	**SD**	**M**	**SD**	**p-value**
Body Temperature	PRE	36.60	0.06	36.75	0.05	36.70	0.05	36.65	0.05	36.66	0.05	36.66	0.06	36.65	0.04	36.62	0.05	0.82
	POST	36.54	0.05	36.54	0.04	36.51	0.05	36.50	0.06	36.54	0.05	36.51	0.04	36.49	0.04	36.47	0.05	0.64
SYS	PRE	107.13	2.09	113.67	3.29	108.51	2.17	108.69	2.09	108.06	2.05	110.14	2.35	108.99	1.86	109.98	2.01	0.88
	POST	103.55	1.67	108.43	2.78	103.55	2.29	106.43	1.80	105.17	1.94	106.14	2.51	104.20	1.83	106.57	2.10	0.97
DIA	PRE	73.32	1.27	74.00	1.75	73.87	1.39	70.73	1.48	72.85	1.42	73.57	1.64	73.45	1.12	76.48	1.81	0.35
	POST	72.30	1.47	72.79	1.56	72.29	1.37	73.09	1.78	72.64	1.52	72.03	1.91	73.19	1.21	71.78	1.51	0.99
Heart Rate	PRE	79.23	1.59	80.84	2.86	78.09	1.51	78.50	2.53	79.89	1.33	79.69	2.49	78.94	1.65	79.07	2.48	0.84
	POST	74.19	1.42	73.36	2.17	74.74	1.50	72.37	2.20	74.75	1.45	72.67	2.12	74.61	1.40	71.35	2.05	0.97

Note: (*) significant at $p < 0.05$.

Table 4.2 The significant difference between the groups (EHS and Non-EHS) during pre-exposure and post-exposure for physiological parameters is examined using descriptive statistics.

		EHS		Non-EHS		
		M	**SD**	**M**	**SD**	**p-value**
Body Temperature	PRE	36.65	0.03	36.67	0.03	0.58
	POST	36.52	0.02	36.51	0.02	0.57
SYS	PRE	108.17	1.01	110.62	1.24	0.13
	POST	104.12	0.96	106.89	1.15	0.07
DIA	PRE	73.37	0.64	73.70	0.85	0.76
	POST	72.61	0.69	72.42	0.84	0.87
Heart Rate	PRE	79.04	0.75	79.53	1.28	0.74
	POST	74.57	0.71	72.44	1.06	0.09

4.2.2 Statistical Data Analysis for Cognitive Performance

The results of the four cognitive tests between Sham and real 5G RF exposure signals are tabulated in Table 5. The analysis shows that there is no significant effect ($p>0.05$) of learning for DSPAN, BCST, TOL and Flanker Task between EHS and Non-EHS subjects. Similarly, there is no significant difference ($p>0.05$) of learning for DSPAN, BCST, TOL, and Flanker Task between Sham and 5G signal exposures. Thus, there is no difference between working memory, attention, shifting, and problem solving of EHS and Non-EHS adults for both Sham and 5G signal exposures. However, only one parameter, the Perseverative Error %, showed significant difference ($p>0.05$) between EHS and Non-EHS subjects when there is no exposure. The capacity to recognize when what one is doing is ineffective and to make the necessary adjustments to accommodate novel circumstances is known as perseverative error. It is described as a series of mistakes when the participant applies the same rule (as in the prior trial), and it acts as a broad indicator of someone's ability to accept a new rule or give up on an existing rule. Despite its p>0.05, this error is observed to be small between EHS and Non-EHS subjects when there is no exposure, with an error percentage difference of about 4.3%. On the contrary, minor differences of perseverative error % are observed when the EHS and Non-EHS subjects are exposed to all 5G signals, up to a maximum of 2.8 % (for 28 GHz) is insignificant. These findings reveal that both subject groups' brain shifting skills are unaffected by 5G signal exposure. Most notably, the overall findings imply that 5G base station exposure has no effect on brain functions in information processing such as working memory, attention, shifting, and problem solving.

Table 4.3 The statistical results for cognitive performance.

		DSPAN		**Flanker Task**				**BCST**						**TOL**			
		DSPAN		**RT-Acc (ms)**		**RT-Acc inteference**		**Correct %**		**Perseverative error %**		**Non-perseverative error**		**TOLpercent success %**		**TOLfirst move**	
		M	**SD**	**M**	**SD**	**M**	**SD**	**M**	**SD**	**M**	**SD**	**M**	**SD**	**M**	**SD**	**M**	**SD**
Sham	EHS	6.50	0.29	447.33	6.31	48.59	4.35	76.62	2.08	14.48	1.70	10.08	1.37	76.39	2.77	8022.40	768.14
	Non-EHS	6.93	0.28	436.76	5.46	46.67	3.49	77.87	2.43	10.11	0.92	10.68	1.94	78.06	2.45	10023.75	1302.44
700 MHz	EHS	6.33	0.26	447.03	8.02	46.39	3.92	80.63	1.80	12.14	0.94	7.24	1.19	73.89	2.73	9552.03	1454.29
	Non-EHS	6.67	0.30	441.68	6.44	46.05	2.71	79.69	2.23	11.10	1.19	7.76	1.40	74.72	3.52	8918.35	527.45
3.5 GHz	EHS	6.27	0.26	445.13	7.02	45.44	3.90	78.18	2.09	12.45	1.06	9.09	1.37	74.44	3.01	8889.03	694.54
	Non-EHS	6.67	0.28	446.21	9.34	42.31	2.73	78.65	1.82	11.62	1.45	9.58	1.58	79.72	2.87	8615.79	512.45
28 GHz	EHS	6.37	0.23	448.66	7.49	49.82	4.23	77.29	2.08	12.97	1.35	10.63	1.76	71.39	2.76	8107.87	544.00
	Non-EHS	6.53	0.25	434.59	5.64	47.99	3.66	82.87	1.44	10.21	0.90	5.84	0.81	76.39	2.12	9126.87	917.57
p-value	(Group)	0.08		0.146		0.48		0.27		**0.01***		0.45		0.11		0.41	

(Signal)	0.78	0.85	0.57	0.49	0.92	0.25	0.27	0.89

Note: (*) significant at $p < 0.05$.
p-value (Group): to investigate the significance difference of each cognitive test between Group (EHS and Non-EHS)
p-value (Signal): to investigate the significance difference of each cognitive test between Signal (Sham, 700 MHz, 3.5 GHz, 28 GHz)

4.2.3 Evaluation of Performance for the Assessment

The results of the statistical analysis revealed that there were no substantial differences observed between exposure to Sham and various 5G radiation frequencies, specifically at 700 MHz, 3.5 GHz, and 28 GHz (P > 0.05). This lack of significant difference was consistent across all conditions, both before and after exposure, for all the physiological parameters and cognitive functions that were measured. The comparison of these findings with earlier studies is presented in Table 4.4, providing a comprehensive overview of how the proposed assessment aligns with the outcomes of previous research. This analysis contributes to the understanding of the impact of 5G radiation exposure on physiological parameters and cognitive function, suggesting that, based on the statistical analysis conducted, the measured variables did not significantly differ between Sham exposure and exposure to the specified 5G radiation frequencies across the tested conditions.

Table 4.4 Evaluation of performance in terms of cognitive function and physiological indicators using the suggested technique in comparison to earlier research.

Study	Exposure Type	Blind Design	Subject	Exposure assessment period	Exposure duration (min)	SAR	E-field strength	Crossover	Assessment location	Measurements	Outcomes
Koivisto et al. (2000)	Mobile phone; GSM 902 MHz; 60 min R	Single	48 adults	2 sessions, separated by 1 day	60	NR	-	(ON, OFF)	NR	Cognitive parameters (SRT, CRT, SUB, VER, VIG, etc (12 tasks))	SRT ↓; VIG ↓SUB ↓; VIG accuracy ↑
Koivisto et al. (2000)	902 MHz; 30 min L	Single	48 adults	Single session	30	NR	-	(ON, OFF)	NR	Cognitive parameters (n-back (0-3))	RT ↓ to targets (3-back Task)
Thavanainen et al. (2004)	Mobile phone; GSM 900 and 1800; 35 min	Double	34 adults	2 sessions, separated by 1 week	35	900MHz: 1.58 W/kg 1800MHz: 0.70 W/kg	-	(GSM 900, GSM 1800, OFF)	EMF shielded laboratory	Physiological parameters (BP and HR)	No effect on BP and HR
Curcio et al. (2004)	Mobile phone; GSM 902.40 MHz; 45 min L	Double	20 adults	3 sessions, separated by more than 2 days	45	Max: 0.5 W/kg	-	(BSL, ON, OFF)	NR	Cognitive parameters (SRT, CRT, VS, SUB)	SRT ↓ (POST); CRT ↓ (POST)
Regel et al. (2006)	BS; UMTS 2140; 45 min, 2 m	Double	33 EHS adults 84 Non-EHS	3 sessions, separated by 1 week	45	-	0 V/m, 1 V/m, or 10 V/m	(0, 1, 10 V/m)	Open chamber	5 subjective well-being symptoms Cognitive function (SRT, CRT, SUB, VER, VIG, etc (12 tasks))	No effect on subjective symptoms and cognitive performance
Oftedal et al. (2007)	Mobile phone; GSM 902.4; 30 min	Double	17 adults	4 sessions, separated by 2 days	30	Spatial peak SAR_{1g}: 1.0 W/kg SAR_{10g} : 0.8 W/kg	-	(ON, OFF)	Control room	4 subjective well-being symptoms Physiological parameters (BP and HR)	No effect on subjective symptoms, BP and HR

Study	Exposure Type	Blind Design	Subject	Exposure assessment period	Exposure duration (min)	SAR	E-field strength	Crossover	Assessment location	Measurements	Outcomes
Eltiti et al. (2007)	BS; GSM 900 + 1800; UMTS 2020; 50 min, 5 m	Double	44 EHS adults, 114 Non-EHS adults	4 sessions, separated by 1 week	50	-	10 mW/m^2	(GSM, UMTS, OFF)	Shielded room and high shielding effectiveness	6 VAS subjective well-being symptoms EMF Perception Physiological parameters (BP and HR)	No effect on subjective symptoms, EMF perception, BP, HR
Cinel et al. (2008)	Mobile Phone; GSM 888 MHz and CW; EXP 1:45 min L/R; EXP 2:40 min L/R	Double	EXP 1: 446 adults EXP 2: 164 adults	2 sessions, separated by 1 week	40	1.4 W/kg (±30%) (SAR average for CW and GSM) 11.2 W/kg (peak of SAR for GSM)	-	(ON, OFF)	NR	5 subjective well-being symptoms	No consistent effect on subjective symptoms
Eltiti et al. (2009)	Mobile phone; UMTS WCDMA 1947; LTE 1750; 20 min L	Double	UMTS: 34 adults LTE: 26 adults	2 sessions, separated by 1 week	20	1.8 W/kg	-	(UMTS/LTE, OFF)	Dimly lit room	Stroop test	None
Wallace et al. (2010)	BS; TETRA 420; 50 min, 5 m	Double	48 EHS adults 132 Non-EHS adults	3 sessions, separated by 1 week	50	271 μW/kg	10 mW/m^2	(ON, OFF)	Screened semi-anechoic chamber	6 VAS subjective well-being symptoms EMF Perception Physiological parameters (BP and HR)	No effect on subjective symptoms, EMF perception, BP and HR (double-blind) Have effects on subjective symptoms (exposure is known)
Kwon et al. (2012)	Mobile Phone; 3G WCDMA 1950; 64 min	Double	17 EHS adults 20 Non-EHS adults	2 sessions, separated by 1-10 days	64	1.57 W/kg	-	(ON, OFF)	Except for instruments, every electrical equipment in the lab was disconnected.	8 subjective well-being symptoms Physiological parameter (HR)	No effect on subjective symptoms and HR for EHS, Non-EHS subjects

Study	Exposure Type	Blind Design	Subject	Exposure assessment period	Exposure duration (min)	SAR	E-field strength	Crossover	Assessment location	Measurements	Outcomes
Choi et al. (2014)	Mobile Phone; 3G WCDMA 1950; 64 min	Double	26 adults 26 teenagers	2 sessions, separated by 1-10 days	64	1.57 W/kg	6.9 V/m	(ON, OFF)	Except for instruments, every electrical equipment in the lab was disconnected.	8 subjective well-being symptoms	No effect on subjective symptoms and HR
										Physiological parameter (HR)	
Malek et al. (2015)	BS; GSM 945 MHz, 1840 MHz; UMTS 2140 MHZ; 2 m	Single	100 EHS; 100 Non-EHS	4 sessions	50	-	1 V/m	(GSM 900, GSM1800, UMTS, OFF)	RF shielded room, lined using microwave absorbing sheets	Physiological parameters (BT, BP and HR)	No effect on physiological parameters (BT, BP and HR) and cognitive parameter
										Cognitive parameter (paired associates learning, rt, rapid visual processing, spatial span)	
Sauter et al. (2015)	TETRA hand-held transmitter 385 MHz; 2 h 30 min L	Double	30 adults	9 sessions, separated by 2 weeks	150 /day	(1) TETRA low level (max SAR 10g=1.5 W/kg) (2) TETRA high level (max SAR 10g=6 W/kg)	-	(TETRA 1.5 W/kg, TETRA 6.0 W/kg, OFF) (UMTS/LTE, OFF)	Shielded room with low background field	Test for Attentional Performance (Divided attention, VIG), Vienna Test system (Selective attention) and n-back	None
Andrianome et al. (2017)	BS; GSM 900, GSM 1800, DECT and Wi-Fi 2.45 GHz. 5 min (for each signal)	Double	10 EHS adults 25 Non-EHS adults	2 sessions, separated by more than 1 week	5	-	1 V/m	(GSM 900, GSM 1800, DECT, Wi-Fi, OFF)	Shielded chamber	Physiological parameters (BP and HR)	No effect on physiological BP and HRV
van Moorselaar et al. (2017)	BS; GSM 900, GSM 1800, UMTS, DECT and Wi-Fi 2.45 GHz; 150 min	Double	42 EHS adults	Testing group then follow up at 2 months interval	150	-	Max: 6 V/m (average exposure levels at the upper body level)	GSM 900, GSM 1800, UMTS, DECT, Wi-Fi	Home and other comfortable location	EMF Perception symptoms	No effect on EMF perception but have effects on EHS symptoms
Bogers et al., (2018)		NR	7 EHS adults	4 sessions with intervals of 6 h	360	-	2.5 V/m		Inside and outside home, at work or	EMF Perception	Have effects on EHS symptoms

Study	Exposure Type	Blind Design	Subject	Exposure assessment period	Exposure duration (min)	SAR	E-field strength	Crossover	Assessment location	Measurements	Outcomes
	BS; GSM 900, GSM 1800, UMTS, DECT and Wi-Fi 2.45 GHz; 6 h							GSM 900, GSM 1800, UMTS, DECT, Wi-Fi	educational institution,	EHS subjective symptoms	
Masrakin et al. (2019)	Wearable textile antenna 2.45 GHz; 50 min	Single	20 adults	2 sessions	50	For 10g SAR TM: (2.88 W/kg) 10g SAR TP: 0.35 W/kg)	-	(ON, OFF)	RF-shielded room	10 subjective well-being symptoms Physiological parameters (BT, BP, and HR)	No effect on subjective symptoms and physiological parameters (BT, BP and HR)
Huang, PC et al. (2022)	BS; GSM 900, 1800 GSM, 2100 MHz	Double	58 EHS adults 92 Non-EHS adults	2 sessions	30	-	0.25 mW/m^2	GSM 900, GSM 1800, UMTS, DECT, Wi-Fi	Anechoic laboratory	Physiological parameters (BP, HRV and HR)	None
Proposed work	BS; 5G 700 MHz, 3.5 GHz, 28 GHz; 60 min	Double	30 EHS adults 30 Non-EHS adults	4 sessions with at least a gap of 3 days after each session	60	-	1 V/m (700 MHz and 3.5 GHz), 0.64 V/m (28 GHz)	(5G/OFF)	RF-Shielded Room	23 subjective well-being symptoms Physiological parameters (BT, BP and HR) Cognitive function (DSPAN, Flanker task, BCST and TOL)	No effect on physiological parameters (BT, BP and HR) and cognitive parameters

BCST – Berg's Card Sorting Task, BT – Body Temperature, BP – Blood Pressure, BSL – Baseline, CW – Continuous Wave, CRT – Choice Reaction Time, DSPAN-Backward Digit Span Task, EHS– Electromagnetic Hypersensitivity, H – Hours, HR – Heart Rate, HRV – Heart Rate Variability, L – Left, Min – Minutes, NR – Not Reported, Rx – Received Antenna, POST – Post Exposure, TM – Textile Monopole Antenna, TOL- Tower of London Task, TP – Textile Patch Antenna, Tx – Transmitted Antenna, , min – minutes, RT – Reaction Time, R – Right, SRT – Simple Reaction Time, SUB – Subtraction Time, VAS – Visual Analogue Scale, VER – Verification Time, VIG – Vigilance

4.3 Classification of Initial Data Processing without Feature Selection

Physiological data obtained from the outcomes of the MSFS yielded classification results when subjected to KNN, SVM, Ensemble Method, Naïve Bayes, and PNN classifier algorithms through exposure classification using MATLAB 2022b, as detailed in Table 4.5. In the classification context, accuracy, a widely used metric, gauges the overall correctness of a model's predictions. This metric is computed by dividing the number of correctly predicted instances by the total instances in the dataset. The Fine KNN model stands out with an accuracy of 26.104%, marking it as the most accurate among the models listed. However, it's noteworthy that if the majority class dominates the dataset and the baseline accuracy hovers around 19.868%, the model may not contribute significantly to improvement. Despite achieving the highest accuracy reading of 52.760% in subject classification, this result falls short of meeting the research objective, which aims for an accuracy exceeding 90%. Similar result of accuracy can be seen with cognitive dataset as outlined in Table 4.6. Achieving an accuracy goal of more than 90% appears to be challenging based on the provided accuracy values for the mentioned classifiers. The highest accuracy among the listed models is 24.688% (Ensemble - Subspace KNN), which is considerably below the target of more than 90%. Reaching high accuracy often involves a combination of selecting appropriate features, tuning model parameters, and possibly exploring more sophisticated modelling techniques.

Table 4.5 Classification accuracy for both exposure and subject assessed using the initial data processing without feature selection for physiological data.

	Exposure Classification (%)	Subject Classification (%)
Fine KNN	26.104	52.760
Gaussian Naïve Bayes	25.111	50.000
Quadratic Discriminant	25.069	50.000
Ensemble (Subspace KNN)	24.858	50.212
Cosine KNN	24.691	49.715
Coarse Tree	24.639	50.844
Fine Tree	24.611	51.024
Weighted KNN	24.542	52.677
Ensemble (Boosted Trees)	24.434	51.392
Ensemble (RUSBoosted Trees)	24.434	51.299
Medium Tree	24.392	51.392
Ensemble (Bagged Trees)	23.993	52.281
Ensemble (Subspace Discriminant)	23.722	48.639
Coarse Gaussian SVM	23.712	48.413
Linear Discriminant	23.701	48.663
Medium KNN	23.698	49.486
Cubic KNN	23.698	49.483
Kernel Naive Bayes	23.538	50.688
Logistic Regression Kernel	22.764	49.031
SVM Kernel	22.691	48.767
Medium Gaussian SVM	22.625	47.257
Fine Gaussian SVM	22.451	48.917
PNN	22.222	47.701
Coarse KNN	19.868	47.538

Table 4.6 Classification accuracy for both exposure and subject assessed using the initial data processing without feature selection for physiological data.

	Exposure Classification (%)	Subject Classification (%)
Fine KNN	23.281	51.563
Gaussian Naïve Bayes	22.917	47.031
Quadratic Discriminant	22.188	45.885
Ensemble (Subspace KNN)	24.688	52.135
Cosine KNN	20.781	49.635
Coarse Tree	24.427	50.365
Fine Tree	22.188	51.927
Weighted KNN	22.083	50.521
Ensemble (Boosted Trees)	24.375	52.760

Ensemble (RUSBoosted Trees)	24.531	51.563
Medium Tree	24.375	50.990
Ensemble (Bagged Trees)	21.615	50.573
Ensemble (Subspace Discriminant)	18.802	45.469
Coarse Gaussian SVM	18.906	46.042
Linear Discriminant	18.958	45.469
Medium KNN	20.677	49.115
Cubic KNN	20.677	49.167
Kernel Naive Bayes	22.240	47.708
Logistic Regression Kernel	21.354	45.469
SVM Kernel	21.719	52.396
Medium Gaussian SVM	16.927	42.188
Fine Gaussian SVM	13.438	36.146
PNN	23.111	49.325
Coarse KNN	14.896	43.281

4.4 Data Normalization and Normalized Data Statistical Analysis

Normalization methods are generally used to reduce the impact of differences in scale and units across variables, and to ensure that variables are comparable in a statistical analysis. The normalization method is a tool for creating clean data sets from the raw data as well as for improving machine learning performance. The choice of normalization method depends on the nature of the data and the specific statistical analysis being performed and was published in our paper (Sofri et al., 2023). The p-values and the F-values are typically used in the hypothesis testing and the Analysis of Variance (ANOVA), respectively to assess the statistical significance of differences among the groups or the variables. To choose the best normalization method with the p-value and the F-value, the data distribution, the sample size, the research questions, and the outliers' factors should be considered. With these factors to be considered during the analysis, the data pre-processing based on MSFS starts with the data normalization using 20 different techniques after pre-processing phase. The data normalization method is crucial in this study as datasets consist of different sets of collected discrete data.

Once the features are extracted from the data, in the third stage, the 10 extracted feature datasets from normalization method pass through 3 different statistical analysis; paired *t*-test for mean difference, *t*–test for correlation significance then *F*–test for data variability. This analysis will be run using IBM SPSS Statistics 24.0 and Microsoft Excel. Through the analysis, the three best features are selected based on the best *p*-value and *F*-value, *p*-value less than 0.05 and highest *F*-value. If the data matrices did not meet the first selection criterion of ($p<0.05$), the data matrices will be rejected. Then, the second selection criterion (highest *F*-value) will be checked for the remaining data matrices and selected. The statistical analysis result of *p*-value, *F*-Value and the best Normalization Method (NM) for each data set is as shown in Table 4.7 to Table 4.14.

The normalization techniques with *p*-values higher than 0.05 are rejected, and for those that are accepted, a second analysis will be performed to get the highest *F*-value. Therefore, all features were selected and rearranged accordingly from the highest f-value to the lowest, as shown in Table 4.7 to Table 4.14. The top 10 normalization methods for body temperature datasets for PreBT are MMS, LS, ARCAPE 2, RMN, APE 2, ARCAPE 1, BPN, RN, DILSN and VN, while for ExpBT are GS, RSSV, BPN, MMS, LS, BNN, APE 1, APE 2, ARCAPE 1 and ARCAPE 2. ARCAPE 2, APE 2, MMS, LS, RMN, ARCAPE 1, BNN, GS, VN, and DILSN are the top 10 normalization methods selected for the PostBT dataset. This table illustrates a few normalizations with similar potential of the three suitable (PreBT, ExpBT and PostBT) datasets which are MMS, LS, ARCAPE 2, APE 2, and ARCAPE 1. For the PreBT dataset, it seemed that MMS and LS were paired of normalization with the desired results, whereas VN and DILSN for both PreBT and PostBT are the two lowest ranking among all. The *F*-value for the ExpBT was significantly higher value of $2.53\text{x}10^{10}$. For systolic body temperature datasets in Table 4.8 for PreSYS, the top 10 normalization methods are LS, MMS, RMN, APE 2, BNN, ARCAPE 2, ARCAPE 1, APE 2, RSSV, and RLSSV, whereas for ExpSYS, the top 10 normalization methods are ARCAPE 1, APE 1, APE 2, RMN, MMS, LS, BNN, ARCAPE 2, and RLLSV. The 10 best normalization techniques chosen for the PostSYS dataset are APE 1, MMS, ARCAPE 1, BNN, ARCAPE 2, RMN, LS, BPN, RSSV, and RLSSV. The present study shows that the normalization techniques RSSV and RLSSV are appropriate for the systolic blood pressure data in PreSYS, ExpSYS, and PostSYS and a few normalizations of the three relevant (PreSYS, ExpSYS, and PostSYS) datasets, LS, MMS, BNN, ARCAPE 2, APE 1 ARCAPE 1, RSSV, revealing comparable capabilities. Similar to Table 4.9, Table 4.10 and Table 4.11 display RSSV and RLSSV as the two best normalization methods for diastolic blood pressure datasets for PreDIA and ExpDIA. It also lists the 10 best normalization methods for all datasets concerned. ARCAPE 2, APE 2, RMN, ARCAPE 1, LS, MMS, APE 1, BNN, RSSV, and RLSSV are the most effective techniques for PreDIA. 10 best normalization techniques used for the PostDIA dataset are ARCAPE 1, BNN, APE1, MMS, LS, BPN, ARCAPE 2, RMN, RSSV, and RLSSV. The three appropriate (PreDIA, ExpDIA, and PostDIA) datasets. ARCAPE 2, RMN, ARCAPE 1, LS, MMS, APE 1, BNN, RSSV, and RLSSV, are shown in this table along with a few normalizations that have similar potential. The pulse dataset from Table 4.12, which was obtained from physiological datasets, demonstrates that only APE 2 is present for the PreHR, whereas LS, MMS, RMN, ARCAPE 1, APE 1, BNN, ARCAPE 2, RSSV, AND RLSSV are the normalization methods that are suitable for the three characteristics presented. For the cognitive dataset as tabulated for Table 4.13 until Table 4.14 as mentioned before only collected during the

Exposure Time in comparison between each cognitive task with different result output. For DSPAN task, *F*-value with the highest pair is DILSN and RSSV with 51.061 followed by ARCAPE 2, APE 2, RMN, BNN, ARCAPE 1, APE 1, LS, and MMS. As indicated in Table Table 4.14, both outcomes tasks for the Flanker task use normalization techniques of a similar type, namely RSSV, GS, APE 2, RMN, APE 1, ARCAPE 2, MMS, LS, and BPN, with the exception of BNN from the RT-A dataset. PE and NPE were observed to have less than 10 best normalization methods for the BCST task in Table 4.13,, considering the fact that their *p*-values are not considered valid for the research methodology. RSSV, GS, BNN, MMS, and LS are the normalization techniques selected for this work. Lastly, the TOL task of the cognitive dataset in Table 4.14 for the percentage of perseverative error and percentage of non-perseverative error has similar normalization method which are RLSSV, RSSV, GS, ARCAPE 1, BNN and MMS and the highest F-value can be seen here at $3.77\text{x}10^{12}$.

Table 4.7 The statistical analysis result of p-value, F -value and the best Normalization Method (NM) for body temperature dataset.

No.	BODY TEMPERATURE											
	PreBT				ExpBT				PostBT			
	NM	*p*-value	*F*-value	10 Best NM	NM	*p*-value	*F*-value	10 Best NM	NM	*p*-value	*F*-value	10 Best NM
1	RSSV	0	0	MMS	LS	0	1.562	GS	RSSV	0	0	ARCAPE 2
2	RLSSV	0			BNN	0			RLSSV	0		
3	DILSN	0	0.321	LS	BPN	0	3.240	RSSV	ARCAPE 2	0	1.151	APE 2
4	VN	0.001			MMS	0			APE2	0		
5	ARCAPE 2	0.002	0.988	ARCAPE 2	TS	0.001	0.126	BPN	VN	0	0.532	MMS
6	RMN	0.002			DMN	0.001			DILSN	0		
7	APE2	0.002	0.903	RMN	VN	0.001	0.661	MMS	RMN	0	0.997	LS
8	ARCAPE 1	0.002			DILSN	0.001			ARCAPE 1	0		
9	APE1	0.002	0.060	APE 2	APE1	0.001	1.008	LS	APE1	0	0.006	RMN
10	BNN	0.007			APE2	0.013			RN	0.013		
11	MMS	0.009	0.999	ARCAPE 1	ARCAPE 1	0.014	1.008	BNN	BNN	0.015	0.905	ARCAPE 1
12	LS	0.009			ARCAPE 2	0.017			GS	0.016		
13	BPN	0.017	0.532	BPN	GS	0.017	2.53e10	APE 1	MMS	0.019	1	BNN
14	RN	0.023			RSSV	0.019			LS	0.019		
15	GN	0.032	NM not accepted	RN	RLSSV	0.031	NM not accepted	APE 2	BPN	0.034	NM not accepted	GS
16	ZS	0.127			RMN	0.116			RIN	0.094		
17	RSDN	0.127		DILSN	ZS	0.127		ARCAPE 1	RSDN	0.127		VN
18	RIN	0.174			RSDN	0.127			ZS	0.127		
19	TS	0.985		VN	RIN	0.985		ARCAPE 2	TS	0.985		DILSN
20	DMN	4.393			RN	2.778			DMN	2.5045		

PreBT – Body Temperature recorded before 5G exposure, ExpBT - Body Temperature recorded during 5G exposure, PostBT - Body Temperature recorded after 5G exposure, NM – Normalization Method, ZS - Z-score Normalization, LS - Linear Scaling, BNN - Binary Normalization, BPN - Bipolar Normalization, MMS - Min-Max Scaling Normalization, TS - t-score Normalization, DILSN - Decimal Inverse Logarithmic Scaled Normalization, RMN - Relative Mean Normalization, RSDN - Relative Standard Deviation Normalization, VN - Variation Normalization, RN - Robust Normalization, RIN - Relative Interquartile Normalization, DMN - Differential Moment Normalization, APE 1 - Absolute Percentage Error Normalization formula 1, APE 2 - Absolute Percentage Error Normalization formula 2, ARCAPE 1 - Arctan APE formula 1, ARCAPE 2 - Arctan APE formula 2, GN - Gaussian Normalization, RSSV - Relative Sum Squared Value, and RLSSV - Relative Logarithmic Sum Squared Value.

Table 4.8 The statistical analysis result of p-value, F -value and the best Normalization Method (NM) for systolic blood pressure dataset.

No.	SYSTOLIC BLOOD PRESSURE											
	PreSYS				ExpSYS				PostSYS			
	NM	*p*-value	*F*-value	10 Best NM	NM	*p*-value	*F*-value	10 Best NM	NM	*p*-value	*F*-value	10 Best NM
1	RSSV	0	0	LS	RSSV	0	0	ARCAPE 1	RSSV	0	0	APE 1
2	RLSSV	0			RLSSV	0.001			RLSSV	0.001		
3	ARCAPE 1	0.014	0.970	MMS	ARCAPE 1	0.014	1.033	APE 1	ARCAPE 2	0.002	0.007	MMS
4	APE1	0.014			APE1	0.014			RMN	0.015		
5	BNN	0.014	0.991	RMN	BNN	0.014	0.990	APE 2	ARCAPE 1	0.015	0.831	ARCAPE 1
6	ARCAPE 2	0.014			ARCAPE 2	0.014			BNN	0.015		
7	RMN	0.014	0.999	APE2	APE2	0.015	0.999	RMN	APE1	0.015	0.547	BNN
8	APE2	0.014			RMN	0.015			MMS	0.018		
9	LS	0.018	0.999	BNN	MMS	0.018	0.999	MMS	LS	0.018	3.240	ARCAPE 2
10	MMS	0.018			LS	0.018			BPN	0.033		
11	BPN	0.032	NM not accepted	ARCAPE 2	BPN	0.032	NM not accepted	LS	RIN	0.111	NM not accepted	RMN
12	RIN	0.114			RIN	0.120			APE2	0.124		
13	ZS	0.127		ARCAPE 1	RSDN	0.127		BNN	VN	0.124		LS
14	RSDN	0.127			ZS	0.127			ZS	0.127		
15	VN	0.179		APE 1	VN	0.175		ARCAPE 2	RSDN	0.127		BPN
16	TS	0.985			TS	0.985			TS	0.985		
17	RN	21.882		RSSV	RN	19.683		RSSV	RN	10.314		RSSV
18	GS	35.450			GS	31.792			GS	15.962		
19	DMN	362.5641		RLSSV	DMN	344.990		RLSSV	DMN	155.304		RLSSV
20	DILSN	4.307E7			DILSN	3.050E7			DILSN	2508.969		

PreSYS – Systolic Blood Pressure recorded before 5G exposure, ExpSYS - Systolic Blood Pressure recorded during 5G exposure, PostSYS - Systolic Blood Pressure recorded after 5G exposure, NM – Normalization Method, ZS - Z-score Normalization, LS - Linear Scaling, BNN - Binary Normalization, BPN - Bipolar Normalization, MMS - Min-Max Scaling Normalization, TS - t-score Normalization, DILSN - Decimal Inverse Logarithmic Scaled Normalization, RMN - Relative Mean Normalization, RSDN - Relative Standard Deviation Normalization, VN - Variation Normalization, RN - Robust Normalization, RIN - Relative Interquartile

Normalization, DMN - Differential Moment Normalization, APE 1 - Absolute Percentage Error Normalization formula 1, APE 2 - Absolute Percentage Error Normalization formula 2, ARCAPE 1 - Arctan APE formula 1, ARCAPE 2 - Arctan APE formula 2, GN - Gaussian Normalization, RSSV - Relative Sum Squared Value, and RLSSV - Relative Logarithmic Sum Squared Value.

Table 4.9 The statistical analysis result of p-value, F -value and the best Normalization Method (NM) for diastolic blood pressure dataset.

No.	DIASTOLIC BLOOD PRESSURE											
	PreDIA				ExpDIA				PostDIA			
	NM	***p*-value**	***F*-value**	**10 Best NM**	**NM**	***p*-value**	***F*-value**	**10 Best NM**	**NM**	***p*-value**	***F*-value**	**10 Best NM**
1	RSSV	0	0	ARCAPE 2	RSSV	0	0	ARCAPE 1	RSSV	0	0	ARCAPE 1
2	RLSSV	0.001			RLSSV	0.001			RLSSV	0.001		
3	ARCAPE 2	0.014	1.026	APE 2	ARCAPE 2	0.014	0.980	APE 2	ARCAPE 2	0.002	0.016	BNN
4	APE2	0.014			BNN	0.014			RMN	0.015		
5	RMN	0.014	1.004	RMN	ARCAPE 1	0.014	1.003	MMS	ARCAPE 1	0.015	1.002	APE1
6	ARCAPE 1	0.014			APE2	0.014			BNN	0.015		
7	APE1	0.014	0.589	ARCAPE 1	RMN	0.014	0.964	LS	APE1	0.015	0.665	MMS
8	BNN	0.019			APE1	0.014			MMS	0.018		
9	LS	0.024	0.999	LS	MMS	0.018	0.999	ARCAPE 2	LS	0.018	0.308	LS
10	MMS	0.024			LS	0.018			BPN	0.033		
11	BPN	0.042	NM not accepted	MMS	BPN	0.032	NM not accepted	BNN	RIN	0.111	NM not accepted	BPN
12	RIN	0.108			RIN	0.116			APE2	0.124		
13	VN	0.117		APE 1	VN	0.118		RMN	VN	0.124		ARCAPE 2
14	ZS	0.127			ZS	0.127			ZS	0.127		
15	RSDN	0.127		BNN	RSDN	0.127		APE 1	RSDN	0.127		RMN
16	TS	0.985			TS	0.985			TS	0.985		
17	RN	10.15		RSSV	RN	9.4		RSSV	RN	10.314		RSSV
18	GS	18.41			GS	15.513			GS	15.962		
19	DMN	153.889		RLSSV	DMN	152.342		RLSSV	DMN	155.304		RLSSV
20	DILSN	423.155			DILSN	2.854E3			DILSN	2.518E3		

PreDIA – Diastolic Blood Pressure recorded before 5G exposure, ExpDIA - Diastolic Blood Pressure recorded during 5G exposure, PostDIA - Diastolic Blood Pressure recorded after 5G exposure, NM – Normalization Method, ZS - Z-score Normalization, LS - Linear Scaling, BNN - Binary Normalization, BPN - Bipolar Normalization, MMS - Min-Max Scaling Normalization, TS - t-score Normalization, DILSN - Decimal Inverse Logarithmic Scaled Normalization, RMN - Relative Mean Normalization, RSDN - Relative Standard Deviation Normalization, VN - Variation Normalization, RN - Robust Normalization, RIN - Relative Interquartile

Normalization, DMN - Differential Moment Normalization, APE 1 - Absolute Percentage Error Normalization formula 1, APE 2 - Absolute Percentage Error Normalization formula 2, ARCAPE 1 - Arctan APE formula 1, ARCAPE 2 - Arctan APE formula 2, GN - Gaussian Normalization, RSSV - Relative Sum Squared Value, and RLSSV - Relative Logarithmic Sum Squared Value.

Table 4.10 The statistical analysis result of p-value, F -value and the best Normalization Method (NM) for heart rate dataset.

No.	**Heart Rate**											
	PreHR				**ExpHR**				**PostHR**			
	NM	***p*-value**	***F*-value**	**10 Best NM**	**NM**	***p*-value**	***F*-value**	**10 Best NM**	**NM**	***p*-value**	***F*-value**	**10 Best NM**
1	RSSV	0	0	LS	RSSV	0	0	ARCAPE 1	RSSV	0	0	MMS
2	RLSSV	0.001			RLSSV	0.001			RLSSV	0.001		
3	BNN	0.016	0.822	MMS	ARCAPE 2	0.002	0.0149	APE 1	ARCAPE 2	0.002	0.015	BPN
4	ARCAPE 2	0.018			RMN	0.017			RMN	0.017		
5	RMN	0.018	0.999	RMN	ARCAPE 1	0.017	1.064	BNN	BNN	0.017	1.050	BNN
6	APE2	0.018			APE1	0.018			ARCAPE 1	0.018		
7	ARCAPE 1	0.019	0.929	APE 2	BNN	0.019	0.640	MMS	APE 1	0.018	0.721	ARCAPE 1
8	APE1	0.019			MMS	0.023			LS	0.022		
9	LS	0.020	0.999	ARCAPE 1	LS	0.023	0.309	LS	MMS	0.022	3.240	APE 1
10	MMS	0.020			BPN	0.042			BPN	0.039		
11	BPN	0.037	NM not accepted	APE1	RIN	0.096	NM not accepted	BPN	RIN	0.099	NM not accepted	LS
12	RIN	0.121			RSDN	0.127			ZS	0.127		
13	ZS	0.127		BNN	ZS	0.127		ARCAPE 2	RSDN	0.127		ARCAPE 2
14	RSDN	0.127			APE 2	0.163			APE 2	0.170		
15	VN	0.212		ARCAPE 2	VN	0.163		RMN	VN	0.170		RMN
16	TS	0.985			TS	0.985			TS	0.985		
17	RN	17.683		RSSV	RN	15.800		RSSV	RN	15.869		RSSV
18	GS	29.966			GS	20.499			GS	20.990		
19	DMN	228.207		RLSSV	DMN	175.094		RLSSV	DMN	177.157		RLSSV
20	DILSN	2.416E4			DILSN	498.258			DILSN	477.643		

PreP – Pulse recorded before 5G exposure, ExpP - Pulse recorded during 5G exposure, PostP - Pulse recorded after 5G exposure, NM – Normalization Method, ZS - Z-score Normalization, LS - Linear Scaling, BNN - Binary Normalization, BPN - Bipolar Normalization, MMS - Min-Max Scaling Normalization, TS - t-score Normalization, DILSN - Decimal Inverse Logarithmic Scaled Normalization, RMN - Relative Mean Normalization, RSDN - Relative Standard Deviation Normalization, VN - Variation Normalization, RN - Robust Normalization, RIN - Relative Interquartile Normalization, DMN - Differential Moment Normalization, APE 1 - Absolute Percentage Error Normalization formula 1, APE 2 - Absolute Percentage Error Normalization formula 2, ARCAPE 1 - Arctan APE formula 1, ARCAPE 2 - Arctan APE formula 2, GN - Gaussian Normalization, RSSV - Relative Sum Squared Value, and RLSSV - Relative Logarithmic Sum Squared Value.

Table 4.11 The statistical analysis result of p-value, F -value and the best Normalization Method (NM) for DSPAN dataset.

No.	DSPAN			
	NM	***p*-value**	***F*-value**	**10 Best NM**
1	DILSN	0	51.061	DILSN
2	RSSV	0		
3	GS	0.003	0.777	RSSV
4	RLSSV	0.003		
5	ARCAPE 1	0.026	1.064	ARCAPE 2
6	APE1	0.027		
7	ARCAPE 2	0.027	1.097	APE2
8	APE2	0.029		
9	RMN	0.029	1.093	RMN
10	BNN	0.030		
11	LS	0.038	0.999	BNN
12	MMS	0.038		
13	VN	0.042	NM not accepted	ARCAPE 1
14	BPN	0.068		
15	RIN	0.094		APE 1
16	RSDN	0.127		
17	ZS	0.127		LS
18	RN	0.375		
19	TS	0.985		MMS
20	DMN	2.735		

RT-A – Controlled for Accuracy, RT-AI - Congruent minus incongruent conditions, NM – Normalization Method, ZS - Z-score Normalization, LS - Linear Scaling, BNN - Binary Normalization, BPN - Bipolar Normalization, MMS - Min-Max Scaling Normalization, TS - t-score Normalization, DILSN - Decimal Inverse Logarithmic Scaled Normalization, RMN - Relative Mean Normalization, RSDN - Relative Standard

Deviation Normalization, VN - Variation Normalization, RN - Robust Normalization, RIN - Relative Interquartile Normalization, DMN - Differential Moment Normalization, APE 1 - Absolute Percentage Error Normalization formula 1, APE 2 - Absolute Percentage Error Normalization formula 2, ARCAPE 1 - Arctan APE formula 1, ARCAPE 2 - Arctan APE formula 2, GN - Gaussian Normalization, RSSV - Relative Sum Squared Value, and RLSSV - Relative Logarithmic Sum Squared Value.

Table 4.12 Statistical analysis result of p-value, F -value and the best Normalization Method (NM) for Flanker Task dataset.

No.	**FLANKER TASK**							
	RT-A				**RT-AI**			
	NM	***p*-value**	***F*-value**	**10 Best NM**	**NM**	***p*-value**	***F*-value**	**10 Best NM**
1	RSSV	0	1273.281	RSSV	GS	0	0.843	BPN
2	GS	0			RSSV	0		
3	RLSSV	0	0.003	GS	RLSSV	0.007	0.169	ARCAPE 2
4	ARCAPE 1	0.011			BNN	0.016		
5	APE 1	0.011	0.981	APE 2	LS	0.021	1	ARCAPE 1
6	ARCAPE 2	0.011			MMS	0.021		
7	APE2	0.011	1	RMN	BPN	0.037	1.605	APE 2
8	RMN	0.011			ARCAPE 2	0.047		
9	BNN	0.013	0.640	APE1	ARCAPE 1	0.050	1.214	LS
10	MMS	0.016			APE2	0.054		
11	LS	0.016	0.309	ARCAPE 2	RMN	0.054	0.769	MMS
12	BPN	0.029			APE 1	0.062		
13	RIN	0.099		BNN	RIN	0.102	NM not accepted	GS
14	RSDN	0.127			ZS	0.131		
15	ZS	0.127		MMS	RSDN	0.131		RSSV
16	VN	0.424255877			TS	1.020236646		
17	TS	0.984969703	NM not accepted	LS	VN	1.045552835		RMN
18	RN	240.5576704			RN	62.65370185		
19	DMN	4535.347921		BPN	DMN	271.6460544		APE 1
20	DILSN	3.60862E+58			DILSN	283775.1257		

RT-A – Controlled for Accuracy, RT-AI - Congruent minus incongruent conditions, NM – Normalization Method, ZS - Z-score Normalization, LS - Linear Scaling, BNN - Binary Normalization, BPN - Bipolar Normalization, MMS - Min-Max Scaling Normalization, TS - t-score Normalization, DILSN - Decimal Inverse Logarithmic Scaled Normalization, RMN - Relative Mean Normalization, RSDN - Relative Standard Deviation Normalization, VN - Variation Normalization, RN - Robust Normalization, RIN - Relative Interquartile Normalization, DMN - Differential Moment Normalization, APE 1 - Absolute Percentage Error Normalization formula 1, APE 2 - Absolute Percentage Error Normalization formula 2, ARCAPE 1 - Arctan APE formula 1, ARCAPE 2 - Arctan APE formula 2, GN - Gaussian Normalization, RSSV - Relative Sum Squared Value, and RLSSV - Relative Logarithmic Sum Squared Value.

Table 4.13 The statistical analysis result of p-value, F -value and the best Normalization Method (NM) for Berg's Card Sorting Task with three measured outcomes of Correct Percentage (C %), Percentage of Perseverative Error (PE) and Percentage of Non-Perseverative Error (NPE) Dataset.

No.	BERG'S CARD SORTING GAME											
	C %				PE				NPE			
	NM	***p*-value**	***F*-value**	**10 Best NM**	**NM**	***p*-value**	***F*-value**	**10 Best NM**	**NM**	***p*-value**	***F*-value**	**10 Best NM**
1	RSSV	0	2108.615	RSSV	RSSV	0	41.231	DILSN	RSSV	0	6.791	RSSV
2	GS	0			GS	0			GS	0		
3	RLSSV	0.001	0.007	GS	DILSN	0	403.920	BNN	DILSN	0.004	0.055	GS
4	ARCAPE 2	0.017			BNN	0.018			BNN	0.018		
5	BNN	0.018	0.973	BNN	LS	0.023	1	RSSV	LS	0.023	1	LS
6	RMN	0.018			MMS	0.023			MMS	0.023		
7	APE 2	0.018	0.865	RMN	BPN	0.041	NM not accepted	GS	BPN	0.041	NM not accepted	MMS
8	ARCAPE 1	0.019			ARCAPE 1	0.051			ARCAPE 1	0.072		
9	APE 1	0.020	0.848	APE2	ARCAPE 2	0.054		LS	ARCAPE 2	0.072		DILSN
10	MMS	0.022			APE1	0.065			APE1	0.095		
11	LS	0.022	0.309	ARCAPE 1	RMN	0.0720		MMS	RMN	0.116		BNN
12	BPN	0.039			APE2	0.0720			APE2	0.116		
13	RIN	0.0999	NM not accepted	APE 1	ZS	0.127		BPN	ZS	0.127		
14	ZS	0.127			RSDN	0.127			RSDN	0.127		
15	RSDN	0.127		MMS	RIN	0.182			RIN	0.131		
16	VN	0.196			VN	0.484			VN	0.933		
17	TS	0.985		LS	TS	0.985			TS	0.985		
18	RN	19.747			RN	4.011			RN	8.010		
19	DMN	201.238		BPN	DMN	31.048			DMN	37.502		
20	DILSN	454.498			RLSSV	None			RLSSV	None		

C% – Correct Percentage, PE – Percentage of Perseverative Error, NPE - Percentage of Non-perseverative error, NM – Normalization Method, ZS - Z-score Normalization, LS - Linear Scaling, BNN - Binary Normalization, BPN - Bipolar Normalization, MMS - Min-Max Scaling Normalization, TS - t-score Normalization, DILSN - Decimal Inverse Logarithmic Scaled Normalization, RMN - Relative Mean Normalization, RSDN - Relative Standard Deviation Normalization, VN - Variation Normalization, RN - Robust Normalization, RIN - Relative Interquartile Normalization, DMN - Differential Moment Normalization, APE 1 - Absolute Percentage Error Normalization formula 1, APE 2 - Absolute Percentage Error Normalization formula 2, ARCAPE 1 - Arctan APE formula 1, ARCAPE 2 - Arctan APE formula 2, GN - Gaussian Normalization, RSSV - Relative Sum Squared Value, and RLSSV - Relative Logarithmic Sum Squared Value.

Table 4.14 The statistical analysis result of p-value, F -value and the best Normalization Method (NM) for the Tower of London task with two measured outcomes of Percentage of Success (S %) and the time needed until first move for each problem (FM) dataset.

No.	TOWER OF LONDON TASK							
	S%				FM			
	NM	***p*-value**	***F*-value**	**10 Best NM**	**NM**	***p*-value**	***F*-value**	**10 Best NM**
1	RSSV	0	300.595	RLSSV	GS	0	0	RLSSV
2	GS	0			RSSV	0		
3	RLSSV	0.002	3. 766E12	ARCAPE 2	RLSSV	0.002	25.715	BNN
4	ARCAPE 2	0.025			BNN	0.011		
5	APE 2	0.026	1	RSSV	MMS	0.014	0.999	MMS
6	RMN	0.026			LS	0.014		
7	ARCAPE 1	0.026	0.955	GS	BPN	0.025	0.329	LS
8	BNN	0.027			ARCAPE 1	0.044		
9	APE 1	0.027	0.661	APE2	ARCAPE 2	0.046	NM not accepted	BPN
10	MMS	0.033			APE 1	0.050		
11	LS	0.033	NM not accepted	RMN	APE 2	0.070		ARCAPE 1
12	BPN	0.060			RMN	0.070		
13	RIN	0.104		ARCAPE 1	ZS	0.127		GS
14	RSDN	0.127			RSDN	0.127		
15	ZS	0.127		BNN	RIN	0.157		RSSV
16	VN	0.395			TS	0.985		
17	TS	0.985		APE1	VN	346.802		
18	RN	36.538			DILSN	626.730		
19	DMN	289.975		MMS	RN	2.509E6		
20	DILSN	3764.602			DMN	24.403E6		

S% – Percentage of Success, FM – Time needed until first move for each problem, NM – Normalization Method, ZS - Z-score Normalization, LS - Linear Scaling, BNN - Binary Normalization, BPN - Bipolar Normalization, MMS - Min-Max Scaling Normalization, TS - t-score Normalization, DILSN - Decimal Inverse Logarithmic Scaled Normalization, RMN - Relative Mean Normalization, RSDN - Relative Standard Deviation Normalization, VN - Variation Normalization, RN - Robust Normalization, RIN - Relative Interquartile Normalization, DMN - Differential Moment Normalization, APE 1 - Absolute Percentage Error Normalization formula 1, APE 2 - Absolute Percentage Error Normalization formula 2, ARCAPE 1 - Arctan APE formula 1, ARCAPE 2 - Arctan APE formula 2, GN - Gaussian Normalization, RSSV - Relative Sum Squared Value, and RLSSV - Relative Logarithmic Sum Squared Value.

The presented statistical analysis table compiles the lowest p-value and highest F-value across all tables, summarizing the top three normalization methods suitable for each dataset. Table 4.7 reveals that, for the physiological dataset on body temperature, the top three normalization methods for PreBT are MMS, LS, and ARCAPE 2; for ExpBT, GS, RSSV, and BPN; and for PostBT, ARCAPE 2, APE 2, and MMS. In Table 4.8, focusing on the systolic blood pressure dataset, LS, MMS, and RMN emerge as the top three normalization methods for pre-exposure data. ARCAPE 1, APE 1, and APE 2 are favored for exposure data, while APE 1, MMS, and ARCAPE 1 top the list for PostSYS. Moving to the diastolic blood pressure dataset in Table 4.9, the top three normalization methods are ARCAPE 2, APE 2, and RMN for the overall dataset, while ARCAPE 1, APE 2, and MMS are preferred for ExpDIA. For PostDIA, ARCAPE 1, BNN, and APE 1 stand out as the top three normalization methods. In Table 4.10, heart rate data analysis indicates that LS, MMS, and RMN are the top three normalization methods for preHR, whereas ARCAPE 1, APE 1, and BNN are preferred for ExpHR. PostHR, on the other hand, favors MMS, BPN, and BNN.

Turning to cognitive datasets in Table 4.11, DILSN, RSSV, and ARCAPE 2 are the top three normalization methods for the DSPAN dataset. For the Flanker task, RSSV, GS, and APE 2 are recommended for RT-A, while BPN, ARCAPE 2, and ARCAPE 1 are suitable for RT-AI. Table 4.13 reveals that for the BCST dataset, RSSV, GS, and BNN are the top three normalization methods for C %, and for PE, DILSN, BNN, and RSSV are preferred. NPE dataset normalization suggests RSSV, GS, and LS as the top three methods. Finally, Table 4.14 outlines the normalization methods for the TOL cognitive dataset, indicating RLSSV, ARCAPE 2, and RSSV as the top three for S%, and RLSSV, BNN, and MMS for FM datasets.

4.4.1 Classification of Subject and Exposure Result for Physiological Parameter

The primary metrics which result from the confusion matrix are used to evaluate the classifier's performance include accuracy, precision, F1-score, sensitivity, and specificity from PNN classifier. The comparison of analysis approach without MSFS and with MSFS are shown in Table 4.15 and Table 4.16 with average result from the proposed classifer. These two tables tabulated the classification results for to differentiate between subject involved of EHS and Non-EHS in the datasets as well as the exposure classification which consist of Sham, 5G 700 MHz, 5G 3.5 GHz and 28 GHz. In Table 4.15, it was observed that for classification exposure, the normalization methods MMS and LS exhibited significantly increased specificity as well as accuracy. The normalization method named BNN for the category of subject classification accounted for most enhanced results in terms of specificity, accuracy, sensitivity, and precision. Next, it was shown that exposure classification with the normalization methods of MMS and LS featured higher data metrics of sensitivity and precision. For BCST (S), BCST (PE), and TOL (FM) data parameters, ZS normalization boasted the utmost level of specificity. Turning to Table 4.16, the physiological dataset results demonstrated that LS and BNN normalization methods achieved specificity, precision, and accuracy that were remarkably elevated. In the case of subject classification, the LS normalization method consistently achieved the most improved precision and specificity values. The results show that good predictive accuracy can be achieved when using feature selection methods. This study further confirmed that supervised ML is a viable strategy for discovering features best characterizing the RF-EMF exposure scenarios. Technologies are changing with time and, therefore, utilizing and recognizing the time of the study as a feature is significant (Halgamuge 2020).

Table 4.15 Classification of subject and exposure result for physiological data parameter

No	Data Parameter	Normalization Method	Classification	Presence of MSFS	Accuracy	Precision	F1-score	Sensitivity	Specificity
1	Body Temp (Pre)	ZS, BNN, DILSN, RMN, RSDN	Subject	Yes	0.952	0.923	0.960	0.923	0.923
			Exposure		0.952	0.800	0.889	0.857	0.857
			Subject	No	0.482	0.484	0.632	0.484	0.452
			Exposure		0.226	0.197	0.053	0.455	0.445
2	Body Temp (Exp)	LS, BPN, MMS, TS, RSDN	Subject	Yes	0.952	0.923	0.960	0.923	0.889
			Exposure		0.952	0.889	0.941	0.889	0.800
			Subject	No	0.485	0.487	0.623	0.487	0.471
			Exposure		0.233	0.233	0.051	0.467	0.470
3	Body Temp (Post)	ZS, LS, BNN, MMS, DILSN	Subject	Yes	0.952	0.900	0.947	0.900	0.923
			Exposure		0.905	0.875	0.875	0.875	0.778
			Subject	No	0.489	0.485	0.223	0.485	0.489
			Exposure		0.224	0.233	0.107	0.479	0.538
4	Dias (Pre)	ZS, LS, BNN, MMS, RSDN	Subject	Yes	0.905	0.846	0.917	0.846	0.923
			Exposure		0.905	0.667	0.800	0.800	0.833
			Subject	No	0.475	0.479	0.485	0.479	0.471
			Exposure		0.223	0.274	0.120	0.597	0.455
5	Dias (Exp)	LS, BNN, BPN, MMS, TS	Subject	Yes	0.952	0.900	0.947	0.900	0.917
			Exposure		0.905	0.952	0.857	0.833	0.889
			Subject	No	0.479	0.475	0.507	0.475	0.485
			Exposure		0.228	0.225	0.231	0.481	0.480
6	Dias (Post)	BNN, BPN, MMS, RMN, RSDN	Subject	Yes	0.952	0.875	0.889	0.875	0.846
			Exposure		0.952	0.857	0.923	0.833	0.750
			Subject	No	0.473	0.471	0.616	0.471	0.497
			Exposure		0.231	0.211	0.124	0.462	0.474
7	Sys (Pre)		Subject	Yes	0.952	0.917	0.952	0.867	0.909

Table 4.15 Classification of subject and exposure result for physiological data parameter

No	Data Parameter	Normalization Method	Classification	Presence of MSFS	Accuracy	Precision	F1-score	Sensitivity	Specificity
		LS, BNN, BPN, MMS, RSDN	Exposure		0.952	0.875	0.933	0.875	0.857
			Subject	No	0.490	0.490	0.489	0.490	0.490
			Exposure		0.231	0.243	0.361	0.500	0.543
8	Sys (Exp)	ZS, BNN, BPN, MMS, RSDN	Subject	Yes	0.952	0.900	0.947	0.900	0.889
			Exposure		0.905	0.909	0.952	0.909	0.905
			Subject	No	0.476	0.475	0.514	0.475	0.478
			Exposure		0.223	0.201	0.196	0.431	0.359
9	Sys (Post)	BNN, BPN, MMS, TS, RMN	Subject	Yes	0.952	0.857	0.952	0.889	0.923
			Exposure		0.905	0.714	0.833	0.800	0.800
			Subject	No	0.488	0.481	0.533	0.481	0.497
			Exposure		0.228	0.232	0.357	0.478	0.500
10	Pulse (Pre)	LS, BPN, MMS, RMN, RSDN	Subject	Yes	0.952	0.929	0.941	0.905	0.923
			Exposure		0.952	0.857	0.923	0.857	0.800
			Subject	No	0.491	0.491	0.510	0.491	0.490
			Exposure		0.231	0.238	0.261	0.483	0.513
11	Pulse (Exp)	LS, BNN, MMS, TS, RSDN	Subject	Yes	0.905	0.917	0.917	0.917	0.889
			Exposure		0.905	0.800	0.947	0.800	0.889
			Subject	No	0.481	0.484	0.569	0.484	0.475
			Exposure		0.221	0.224	0.318	0.464	0.452
12	Pulse (Post)	LS, BNN, MMS, TS, RSDN	Subject	Yes	0.952	0.889	0.960	0.917	0.889
			Exposure		0.905	0.818	0.900	0.900	0.800
			Subject	No	0.479	0.480	0.486	0.480	0.478
			Exposure		0.236	0.222	0.107	0.459	0.434

Table 4.16 Classification of subject and exposure result for cognitive data parameter

No	Data Parameter	Normalization Method	Classification	Presence of MSFS	Accuracy	Precision	F1-score	Sensitivity	Specificity
1	DSPAN	ZS, LS, RIN, RMN, RSDN	Subject	Yes	0.905	0.889	0.889	0.889	0.917
			Exposure		0.941	0.800	0.933	0.800	0.833
			Subject	No	0.477	0.479	0.530	0.479	0.474
			Exposure		0.222	0.219	0.246	0.452	0.463
2	FLANKER (RT-A)	BNN, MMS, DILSN, RIN, RMN	Subject	Yes	0.905	0.917	0.917	0.917	0.889
			Exposure		0.952	0.889	0.941	0.889	0.857
			Subject	No	0.488	0.489	0.626	0.489	0.483
			Exposure		0.230	0.235	0.299	0.475	0.476
3	BCST (S)	ZS, BNN, BPN, DILSN, RSDN	Subject	Yes	0.952	0.923	0.960	0.923	0.909
			Exposure		0.952	0.833	0.923	0.905	0.750
			Subject	No	0.485	0.484	0.556	0.484	0.488
			Exposure		0.222	0.194	0.097	0.456	0.594
4	BCST (PE)	BNN, MMS, TS, RMN, RSDN	Subject	Yes	0.905	0.818	0.900	0.818	0.923
			Exposure		0.905	0.875	0.875	0.900	0.750
			Subject	No	0.498	0.498	0.558	0.498	0.499
			Exposure		0.228	0.230	0.337	0.465	0.456
5	BCST (NPE)	ZS, BPN, DILSN, RMN, RSDN	Subject	Yes	0.905	0.846	0.917	0.846	0.818
			Exposure		0.905	0.667	0.800	0.750	0.875
			Subject	No	0.478	0.473	0.275	0.473	0.480
			Exposure		0.227	0.227	0.236	0.476	0.463
6	TOL (S)	BNN, MMS, DILSN, RMN, RSDN	Subject	Yes	0.952	0.923	0.960	0.923	0.909

Table 4.16 Classification of subject and exposure result for cognitive data parameter

No	Data Parameter	Normalization Method	Classification	Presence of MSFS	Accuracy	Precision	F1-score	Sensitivity	Specificity
			Exposure		0.952	0.833	0.909	0.800	0.952
			Subject	No	0.492	0.491	0.640	0.491	0.500
			Exposure		0.226	0.264	0.068	0.487	0.508
7	TOL (FM)	ZS, LS, MMS, RMN, RSDN	Subject	Yes	0.952	0.917	0.957	0.917	0.923
			Exposure		0.952	0.857	0.923	0.778	0.857
			Subject	No	0.483	0.491	0.594	0.491	0.457
			Exposure		0.228	0.236	0.315	0.490	0.409

4.4.2 Classifier using Multi-Stage Feature Selection Based on Supervised Machine Learning

The dataset in its raw form is limited in size, posing a challenge when applying it to machine learning models due to insufficient data for training and testing. To address this issue, various normalization techniques are employed on the data. These normalization methods aim to standardize the data and enhance its suitability for training machine learning models. Following normalization, the datasets undergo thorough statistical inspection. This scrutiny is conducted to identify sets of data that are both non-redundant and unaffected by distortions resulting from the normalization process. The objective is to carefully select datasets that maintain their integrity and informational content post-normalization. These meticulously chosen datasets are then utilized in the training phase of the machine learning model. This approach ensures that despite the initial limitations in the size of the raw dataset, the normalization techniques, and subsequent data selection contribute to the robustness and effectiveness of the machine learning model (Elkhouly et al., 2023). In terms of exposure classification and accuracy percentages for physiological data parameters, the SVM classifier demonstrates exceptional performance with the highest accuracy rates: 99.89% for Body Temperature and 99.583% for Diastolic Blood Pressure, Systolic Blood Pressure, and Heart Rate. The subject classification results for physiological datasets reveal that the Ensemble classifier excels, achieving an accuracy of 99.583% for Body Temperature, Heart Rate, and Systolic Blood Pressure. Additionally, the SVM classifier achieves the highest accuracy of 99.89% for Diastolic Blood Pressure.

Moving on to cognitive datasets, the exposure classification analysis indicates that the SVM classifier is optimal for the BCST dataset with an accuracy of 99.583% (C), TOL task with 99.583% (FM and S), and DSPAN task with an impressive accuracy value of 99.917%. Meanwhile, the Ensemble classifier performs exceptionally well for NPE and PE datasets. In the Flanker task, the RTA dataset achieves the highest accuracy of 99.583%. However, for subject classification, the KNN classifier proves to be the most effective for the BCST task, boasting accuracies of 99.987% (C), 99.167% (NPE), and 99.583% (PE). The Ensemble model excels in classifying subjects for the DSPAN dataset, achieving an accuracy of 99.983%, and the TOL cognitive task with an accuracy of 99.583%. Lastly, SVM stands out for subject classification in the Flanker Task with an accuracy of 99.583% and the TOL task with an accuracy of 99.97%. In summary, the choice of the best classifier depends on the specific task and dataset, with SVM, Ensemble, and KNN demonstrating superior performance in different contexts. In comparison to the findings of Halgamuge in 2020, where the Random Forest from the Ensemble method demonstrated an accuracy of 83.56%, our research takes a machine learning approach that incorporates MSFS. This novel approach surpasses the accuracy achieved by Halgamuge 2020, particularly in the classification of subjects into Electromagnetic Hypersensitivity (EHS) and Non-EHS categories, as well as distinguishing between different exposure scenarios (Sham, 5G 700 MHz, 5G 3.5 GHz, and 5G 28 GHz). The machine learning model employed in this study utilizes SVM classifier, and the results are presented in detail in Table 4.17 and Table 4.18. Notably, the average highest accuracy achieved through our approach with MSFS surpasses the accuracy obtained by Halgamuge 2020's Random Forest model. This suggests that the integration of MSFS enhances the classification performance, especially in the context of discerning electromagnetic sensitivity and exposure scenarios. These results contribute to the advancement of accurate and effective classification methods within the field of electromagnetic health studies.

Table 4.17 Classification subject and exposure accuracy percentage result for physiological data parameter.

	Exposure Classification (%)				Subject Classification (%)			
	Body Temp	Dias BP	HR	Sys BP	Body Temp	Dias BP	HR	Sys BP
Quadratic SVM	99.167	99.583	99.167	99.583	98.750	98.750	98.750	99.583
Cubic SVM	99.899	98.750	99.583	99.583	99.583	99.583	99.583	99.583
Fine Gaussian SVM	77.500	75.000	81.667	84.583	93.333	90.000	92.500	93.333
Medium Gaussian SVM	95.833	92.500	96.250	95.000	99.167	99.167	97.083	96.667
Coarse Gaussian SVM	97.917	97.917	98.750	98.750	98.750	99.890	99.167	99.167
Fine KNN	62.917	73.333	77.083	84.583	98.750	97.083	97.917	97.083
Coarse KNN	69.583	69.167	72.917	63.333	98.333	97.917	98.750	99.167
Weighted KNN	72.917	75.417	77.500	82.917	98.750	99.167	97.500	97.917
Ensemble (Bagged Trees)	99.583	90.000	99.583	99.167	99.167	98.333	97.917	99.167
SVM Kernel	90.000	99.167	97.917	99.583	97.917	97.917	99.167	99.583
Cosine KNN	64.583	67.083	66.667	67.500	98.333	98.333	97.917	97.917
Medium KNN	69.583	71.250	69.167	72.917	98.750	98.750	98.333	97.917
Cubic KNN	64.167	63.750	64.583	70.417	98.750	98.750	97.917	97.083
Kernel Naive Bayes	55.833	54.583	52.917	54.167	70.833	75.417	67.083	67.917
Ensemble (Subspace KNN)	77.083	90.000	83.750	86.250	80.833	90.833	92.917	89.167
Ensemble (Subspace Discriminant)	30.417	28.750	24.167	30.417	49.167	55.000	55.000	50.417
Ensemble (Boosted Trees)	25.000	25.000	25.000	25.000	50.000	50.000	50.000	50.000

Ensemble (RUSBoosted Trees)	25.000	25.000	25.000	25.000	50.000	50.000	50.000	50.000

Table 4.18 Classification subject and exposure accuracy percentage result for cognitive data parameter.

	Exposure Classification (%)							Subject Classification (%)						
	C	NPE	PE	DSPAN	RTA	FM	S	C	NPE	PE	DSPAN	RTA	FM	S
Quadratic SVM	98.333	98.750	98.750	98.333	99.583	95.000	99.583	99.583	90.000	99.583	95.833	97.917	93.333	99.583
Cubic SVM	99.583	99.167	97.500	90.833	99.167	98.750	99.167	93.333	99.167	93.333	95.000	99.167	96.250	99.920
Fine Gaussian SVM	80.833	85.417	83.750	99.583	77.083	99.583	95.000	96.667	95.833	93.750	98.750	92.500	94.583	99.970
Medium Gaussian SVM	95.833	94.583	92.917	99.917	96.250	99.167	98.333	99.00	96.667	97.500	99.167	99.583	99.167	99.583
Coarse Gaussian SVM	99.167	98.333	96.667	98.333	98.750	99.583	98.750	97.500	98.750	98.00	99.583	99.583	97.500	99.583
Fine KNN	68.333	82.917	81.250	90.833	77.083	97.083	99.167	99.967	97.500	96.250	99.167	97.500	97.800	99.583
Coarse KNN	75.000	65.417	78.750	99.167	63.333	97.083	97.500	99.583	99.167	99.583	99.300	99.167	99.583	99.583
Weighted KNN	73.750	83.750	85.000	99.167	75.833	97.917	99.167	99.987	97.917	98.750	99.583	97.500	99.583	96.250
Ensemble (Bagged Trees)	99.167	99.583	99.583	99.167	99.583	99.167	92.500	99.167	93.333	95.000	99.983	99.167	99.583	97.500
SVM Kernel	97.083	98.333	93.333	78.750	99.167	78.750	69.167	98.333	98.750	98.750	98.333	97.500	98.333	95.833
Cosine KNN	57.083	68.750	75.833	67.083	95.000	82.083	82.500	98.333	96.667	95.000	96.250	97.083	97.500	95.000
Medium KNN	62.917	67.500	81.667	65.417	63.750	65.000	49.167	98.333	97.500	96.250	94.583	97.500	97.083	96.667
Cubic KNN	56.667	67.500	80.833	64.167	63.750	69.583	55.833	94.167	97.083	95.833	92.917	97.500	95.833	96.250
Kernel Naive Bayes	59.583	57.083	55.417	62.500	56.250	97.083	60.833	83.750	71.250	65.833	75.000	97.500	93.750	50.000
Ensemble (Subspace KNN)	79.583	79.167	65.000	53.333	63.750	57.083	55.833	71.667	75.833	66.250	74.167	72.917	70.000	65.417
Ensemble (Subspace Discriminant)	24.583	26.667	23.333	25.417	26.250	25.000	25.000	55.000	50.000	50.417	50.000	46.250	52.917	50.000
Ensemble (Boosted Trees)	25.000	25.000	25.000	25.000	25.000	25.000	25.000	50.000	50.000	50.000	50.000	50.000	50.000	50.000

Ensemble (RUSBoosted Trees)	25.000	25.000	25.000	25.000	25.000	21.667	20.000	50.000	50.000	50.000	39.583	50.000	50.000	49.583

4.5 Summary

The assessment of adult health focuses on cognitive performance, well-being symptoms, and physiological parameters. This study aims to examine various measures related to adults' physiological, cognitive, and well-being aspects, collected before, during, and after exposure to each 5G signal, including a Sham signal. The analysis of the results, with a p-value > 0.05, suggests that there are no statistically significant effects observed in terms of cognitive function and physiological parameters due to short-term 5G radiation exposure in adults. The results of this study hypothetically have important implications for public health and safety policies related to the deployment of 5G technology with the application of machine learning algorithms, particularly supervised learning in the scope of prediction models, with the goal of developing high accuracy classifiers for predicting the potential impact of RF-EMF exposure on humans in epidemiological studies. The high value of accuracy, precision, recall and f1-score were obtained by hybrid dataset from the outcome of MSFS in which included the several normalization methods in pre-processing phase.

Selecting the most effective classifier is contingent upon the unique characteristics of the task at hand and the nature of the dataset being analysed. In the context of exposure classification for physiological data, the SVM classifier emerges as a top performer, showcasing exceptional accuracy rates of 99.89% for Body Temperature and 99.583% for a combined dataset comprising Diastolic Blood Pressure, Systolic Blood Pressure, and Heart Rate. The robust performance of SVM in accurately categorizing physiological parameters makes it a favourable choice for tasks involving these specific data types. When it comes to subject classification within physiological datasets, the Ensemble classifier takes the lead, achieving a noteworthy accuracy of 99.583% for Body Temperature, Heart Rate, and Systolic Blood Pressure. As for the cognitive datasets, the SVM classifier stands out as the preferred choice for exposure classification. It demonstrates superior accuracy across various cognitive tasks, including the BCST dataset (99.583% accuracy), the TOL task (99.583% accuracy for FM and S), and the DSPAN task (99.917% accuracy). The adaptability and high accuracy of SVM in handling diverse cognitive datasets highlight its versatility in capturing complex patterns inherent in cognitive data. However, in subject classification for the BCST task, the KNN classifier proves to be the most effective, boasting impressive accuracies of 99.987% for C, 99.167% for NPE, and 99.583% for PE. This emphasizes the need for considering the idiosyncrasies of each cognitive task when selecting a classifier, as different algorithms may excel in capturing distinct patterns within the data. Furthermore, the Ensemble classifier demonstrates its efficacy in subject classification, particularly for the DSPAN dataset (99.983% accuracy) and the TOL cognitive task (99.583% accuracy). The ability of Ensemble methods to combine multiple base classifiers enhances their performance in capturing complex relationships and variations present in cognitive datasets.

CHAPTER 5 : CONCLUSION

5.1 Summary

In this work, RF-EMF in humans were studied using signals from base station operating in the Fifth Generation (5G) low band at 700 MHz, sub-6 band at 3.5 GHz and millimeter Wave (mmWave) at 28 GHz. These mimicked emissions from 5G BSs were used to verify whether a relation exists between Electromagnetic Field (EMF)s and these parameters on a total of 60 self-reported EHS and Non-EHS subjects. This study applied a counter-balanced, randomized, and double-blinded experimental approach with four sessions (Sham (No exposure), 5G 700 MHz, 5G 3.5 GHz and 5G 28 GHz). Prospective respondents were subjected to an analysis that included 57 health symptoms, EHS symptoms questionnaires scale to identify the EHS or Non-EHS subject category. The effects on adult health are evaluated in terms of cognitive performance, well-being symptoms and physiological parameters. The purpose of this study is determined as adult physiological and cognitive measures gathered before, during, and after exposure to each 5G signal, including Sham signal. Based on the p-value ($p>0.05$) result analysis, the findings indicated that there are no statistically significant effects from short-term 5G radiation exposure from adults in terms of cognitive function and physiological parameter. The technique for analyzing the impact of short-term 5G base station exposure on the cognitive performance and physiological parameters of adults from the level of human exposure to 5G RF-EMF exposure assessment from base station sources operating at low band 5G at 700 MHz, Sub-6 Band 5G at 3.5 GHz, and Millimeter Wave (mmWave) 5G at 28 GHz. The research findings conclude that there are no significant effects of short-term radiation exposure emitted from the 5G base station antenna signals and that the short-term radiation exposure emitted from the 5G base station signals does not cause adverse health effects to Malaysian adults in the areas of cognitive performance and physiological parameters (body temperature, blood pressure, and heart rate).

The proposed Hybridized MSFS classification approach is used to select the most relevant features for a given classification problem. The approach is said to be "hybridized" as it combines multiple feature selection methods in a series of cascading stages. The goal of this approach is to improve the performance and accuracy of the classification model by selecting the most informative features, while also reducing the potential for overfitting or generalization errors. Overall, this scientific study holds significant value as it utilizes machine learning and statistical methodologies to analyze the effects of short-term 5G exposure on human health and cognitive function, thereby predicting the potential impact of RF-EMF exposure on humans in epidemiological studies and application. The results of this study hypothetically have important implications for public health and safety policies related to the deployment of 5G technology. The high value of accuracy, precision, recall and f1-score were obtained by hybrid using PNN and PCA machine learning approach, in which included the several normalization methods in pre-processing phase with the presence of MSFS.

5.2 Recommendations for Future Work

1. It is recommended to explore additional datasets, including electroencephalogram (EEG) data, well-being parameters, and EMF perception, to enhance the breadth of information gathered during the assessment of short-term 5G base station

exposure. This expanded dataset could contribute to a more comprehensive understanding of the potential effects and provide a more nuanced analysis of the impact of 5G radiation on various physiological and perceptual aspects.

2. A deep learning approach integrated to the MSFS methodology is recommended. This combination could offer a more sophisticated and nuanced exploration of the field, potentially uncovering deeper insights and patterns in the context of the impact of 5G radiation. Integrating deep learning techniques could enhance the model's capacity to learn complex relationships within the data, while MSFS ensures a strategic and refined selection of features, contributing to improved model performance and interpretability. The incorporation of a deep learning approach along with MSFS holds the potential to streamline the research process, significantly saving time for researchers. This combination is likely to empower researchers to make more informed and precise decisions, as it can efficiently handle complex patterns within the data and strategically select relevant features.

REFERENCES

(FDA), T. U. S. F. and D. A. (2018). Current research result on cell phones.

[ITU-T, K. 61. (2014). Guidance on Measurement and Numerical Prediction of Electromagnetic Fields for Compliance with Human Exposure Limits for Telecommunication Installations. Recommendation ITU-T K.61 (2018), 52.

A-Infomw-NF Broadband Horn Antenna LB-660 Datasheet. (n.d.). Inc, A-Info.

Aerts, S., Verloock, L., Van Den Bossche, M., Colombi, D., Martens, L., Tornevik, C., & Joseph, W. (2019). In-situ measurement methodology for the assessment of 5G NR massive MIMO base station exposure at sub-6 GHz frequencies. IEEE Access, 7, 184658–184667. https://doi.org/10.1109/ACCESS.2019.2961225

Alwohaibi, M., Alzaqebah, M., Alotaibi, N. M., Alzahrani, A. M., & Zouch, M. (2021a). A hybrid multi-stage learning technique based on brain storming optimization algorithm for breast cancer recurrence prediction. Journal of King Saud University - Computer and Information Sciences, xxxx. https://doi.org/10.1016/j.jksuci.2021.05.004

Andrew, Allan M, Zakaria, A., Mad Saad, S., & Md Shakaff, A. Y. (2016). Multi-Stage Feature Selection Based Intelligent Classifier for Classification of Incipient Stage Fire in Building. In Sensors (Vol. 16, Issue 1). https://doi.org/10.3390/s16010031

Andrianome, S., Gobert, J., Hugueville, L., Stéphan-Blanchard, E., Telliez, F., & Selmaoui, B. (2017b). An assessment of the autonomic nervous system in the electrohypersensitive population: a heart rate variability and skin conductance study. Journal of Applied Physiology (Bethesda, Md. : 1985), 123(5), 1055–1062. https://doi.org/10.1152/japplphysiol.00229.2017

ARPANSA. (2017). Guide for Radiation Protection in Existing Exposure Situations, Radiation Protection Series G-2. Radiation Protection Series, September.

Awad, M., & Khanna, R. (2015). Efficient learning machines: Theories, concepts, and applications for engineers and system designers. In Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers (Issue April). https://doi.org/10.1007/978-1-4302-5990-9

Bogers, R. P., van Gils, A., Clahsen, S. C. S., Vercruijsse, W., van Kamp, I., Baliatsas, C., Rosmalen, J. G. M., & Bolte, J. F. B. (2018). Individual variation in temporal relationships between exposure to radiofrequency electromagnetic fields and non-specific physical symptoms: A new approach in studying ‘electrosensitivity.’ Environment International, 121(May), 297–307. https://doi.org/10.1016/j.envint.2018.08.064

Bosch-Capblanch, X., Esu, E., Dongus, S., Oringanje, C. M., Jalilian, H., Eyers, J., Oftedal, G., Meremikwu, M., & Röösli, M. (2022). The effects of radiofrequency electromagnetic fields exposure on human self-reported symptoms: A protocol for a systematic review of human experimental studies. Environment International, 158, 106953. https://doi.org/10.1016/j.envint.2021.106953

Cameron. (2020). Bits to Beams : RF Technology Evolution for 5G Millimeter Wave Radios. Analog Dialogue, pp. 1-6.

Chekroud, A. M., Bondar, J., Delgadillo, J., Doherty, G., Wasil, A., Fokkema, M., Cohen, Z., Belgrave, D., DeRubeis, R., Iniesta, R., Dwyer, D., & Choi, K. (2021). The promise of machine learning in predicting treatment outcomes in psychiatry. World Psychiatry, 20(2), 154–170. https://doi.org/https://doi.org/10.1002/wps.20882

Chen, C., Twycross, J., & Garibaldi, J. M. (2017). A new accuracy measure based on bounded relative error for time series forecasting. PLoS ONE, 12(3), 1–23. https://doi.org/10.1371/journal.pone.0174202

Chiaraviglio, L., Elzanaty, A., & Alouini, M. S. (2021). Health risks associated with 5G exposure: A view from the communications engineering perspective. IEEE Open Journal of the Communications Society, 2, 2131–2179. https://doi.org/10.1109/OJCOMS.2021.3106052

Choi, S. B., Kwon, M. K., Chung, J. W., Park, J. S., Chung, K., & Kim, D. W. (2014a). Effects of short-term radiation emitted by WCDMA mobile phones on teenagers and adults. BMC Public Health, 14(1), 1–9. https://doi.org/10.1186/1471-2458-14-438

Ciaburro, G. (2017). MATLAB for Machine Learning: Practical examples of regression, clustering and neural networks.

Cinel, C., Boldini, A., Fox, E., & Russo, R. (2008). Does the Use of Mobile Phones Affect Human Short-Term Memory or Attention? Applied Cognitive Psychology, 22, 1113–1125. https://doi.org/10.1002/acp.1425

Cisco. (2020). Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2018–2023. 2020. https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internetreport%0A/white-paper-c11-741490.html

Curcio, G., Ferrara, M., De Gennaro, L., Cristiani, R., D’Inzeo, G., & Bertini, M. (2004). Time-course of electromagnetic field effects on human performance and tympanic temperature. NeuroReport, 15(1), 161–164. https://doi.org/10.1097/00001756-200401190-00031

Curcio, G., Ferrara, M., Moroni, F., D’Inzeo, G., Bertini, M., & De Gennaro, L. (2005). Is the brain influenced by a phone call? An EEG study of resting wakefulness. Neuroscience Research, 53(3), 265–270. https://doi.org/10.1016/j.neures.2005.07.003

Curcio, G., Valentini, E., Moroni, F., Ferrara, M., De Gennaro, L., & Bertini, M. (2008). Psychomotor performance is not influenced by brief repeated exposures to mobile phones. Bioelectromagnetics, 29(3), 237–241. https://doi.org/10.1002/bem.20393

de Bruijne, M. (2016). Machine learning approaches in medical image analysis: From detection to diagnosis. Medical Image Analysis, 33, 94–97. https://doi.org/https://doi.org/10.1016/j.media.2016.06.032

Elkhouly, A., Andrew, A. M., Rahim, H. A., Abdulaziz, N., Malek, M. F. A., & Siddique, S. (2023). Data-driven audiogram classifier using data normalization and multi-stage feature selection. Scientific Reports, 13(1), 1–14. https://doi.org/10.1038/s41598-022-25411-y

Eltiti, S., Wallace, D., Ridgewell, A., Zougkou, K., Russo, R., Sepulveda, F., & Fox, E. (2009).

Short-Term Exposure to Mobile Phone Base Station Signals Does Not Affect Cognitive Functioning or Physiological Measures in Individuals Who Report Sensitivity to Electromagnetic Fields and Controls. Bioelectromagnetics, 30, 556–563. https://doi.org/10.1002/bem.20504

Eltiti, S., Wallace, D., Ridgewell, A., Zougkou, K., Russo, R., Sepulveda, F., Mirshekar-Syahkal, D., Rasor, P., Deeble, R., & Fox, E. (2007). Does short-term exposure to mobile phone base station signals increase symptoms in individuals who report sensitivity to electromagnetic fields? A double-blind randomized provocation study. Environmental Health Perspectives, 115(11), 1603–1608. https://doi.org/10.1289/ehp.10286

ETS-Lindgren's Model 3117 Double-Ridged Waveguide Datasheet. (n.d.). ETS-LINDGREN. Retrieved February 1, 2021, from https://avalontest.com/images/uploaded/3117 Double-Ridged Guide Antenna.pdf

FCC. (1997). Code of Federal Regulations CFR Title 47, Part 1.1310, Radiofrequency Radiation Exposure Limits, Federal Commun. Commission, Washington, DC, USA.

Fernandes, L. C. (2017). The Design and Development of an App to Compute Exposure to Electromagnetic Fields [EM Programmer's Notebook]. IEEE Antennas and Propagation Magazine, 59(2), 149–160. https://doi.org/10.1109/MAP.2017.2655527

Gläscher, J., Adolphs, R., Damasio, H., Bechara, A., Rudrauf, D., Calamia, M., Paul, L. K., & Tranel, D. (2012). Lesion mapping of cognitive control and value-based decision making in the prefrontal cortex. Proceedings of the National Academy of Sciences, 109(36), 14681 LP – 14686. https://doi.org/10.1073/pnas.1206608109

Halgamuge, M., & Davis, D. (2019). Lessons learned from the application of machine learning to studies on plant response to radio-frequency. Environmental Research, 178, 108634. https://doi.org/10.1016/j.envres.2019.108634

Halgamuge, M. N. (2020). Supervised machine learning algorithms for bioelectromagnetics: Prediction models and feature selection techniques using data from weak radiofrequency radiation effect on human and animals cells. International Journal of Environmental Research and Public Health, 17(12), 1–27. https://doi.org/10.3390/ijerph17124595

Halim, A. A. A., Andrew, A. M., Mustafa, W. A., Mohd Yasin, M. N., Jusoh, M., Veeraperumal, V., Abd Rahman, M. A., Zamin, N., Mary, M. R., & Khatun, S. (2022a). Optimized Intelligent Classifier for Early Breast Cancer Detection Using Ultra-Wide Band Transceiver. Diagnostics, 12(11). https://doi.org/10.3390/diagnostics12112870

Haris, H. (2021). Kecoh 5G Penyebab Kanser & Covid-19, Ini Perkara Sebenar Anda Perlu Tahu! Gempak. https://gempak.com/berita-terkini/sembanggajet-kecoh-5g-penyebab-kanser-covid19-ini-perkara-sebenar-anda-perlu-tahu-49482

Hernán, M. A., Hsu, J., & Healy, B. C. (2019). A Second Chance to Get Causal Inference Right: A Classification of Data Science Tasks. CHANCE, 32, 42–49.

Hinrikus, H., Koppel, T., Lass, J., Orru, H., Roosipuu, P., & Bachmann, M. (2022). Possible health effects on the human brain by various generations of mobile telecommunication: a review based estimation of 5G impact. International Journal of Radiation Biology, 98(7), 1210–1221. https://doi.org/10.1080/09553002.2022.2026516

Huang, P.-C., Chiang, J., Cheng, Y.-Y., Cheng, T.-J., Huang, C.-Y., Chuang, Y.-T., Hsu, T., & Guo, H.-R. (2022). Physiological changes and symptoms associated with short-term exposure to electromagnetic fields: a randomized crossover provocation study. Environmental Health, 21(1), 31. https://doi.org/10.1186/s12940-022-00843-1

Huber, R., Treyer, V., Borbély, A. A., Schuderer, J., Gottselig, J. M., Landolt, H.-P., Werth, E., Berthold, T., Kuster, N., Buck, A., & Achermann, P. (2002). Electromagnetic fields, such as those from mobile phones, alter regional cerebral blood flow and sleep and waking EEG. Journal of Sleep Research, 11(4), 289–295. https://doi.org/https://doi.org/10.1046/j.1365-2869.2002.00314.x

IBM. (2019). Artificial Intelligence Analyst 2019. [Course training].

ICNIRP. (2020). Guidelines for Limiting Exposure to Electromagnetic Fields (100 kHz to 300 GHz). Health Physics, 118(5), 483–524. https://doi.org/10.1097/HP.0000000000001210

IEEE Standard for Safety Levels with Respect to Human Exposure to Electric, Magnetic, and Electromagnetic Fields, 0 Hz to 300 GHz. (2019). IEEE Standard C95.1-2019 (Revision of IEEE Std C95.1-2005/Incorporates IEEE Std C95.1-2019/Cor 1-2019).

Isaksson, L. J., Raimondi, S., Botta, F., Pepa, M., Gugliandolo, S. G., De Angelis, S. P., Marvaso, G., Petralia, G., De Cobelli, O., Gandini, S., Cremonesi, M., Cattani, F., Summers, P., & Jereczek-Fossa, B. A. (2020). Effects of MRI image normalization techniques in prostate cancer radiomics. Physica Medica, 71(January), 7–13. https://doi.org/10.1016/j.ejmp.2020.02.007

Ismail, A., Din, N. M., Jamaluddin, M. Z., & Balasubramaniam, N. (2009). Electromagnetic assessment for mobile phone base stations at major cities in Malaysia. 2009 IEEE 9th Malaysia International Conference on Communications (MICC), 150–153. https://doi.org/10.1109/MICC.2009.5431485

ITU. (2022). 5G - Fifth generation of mobile technologies. https://www.itu.int/en/mediacentre/backgrounders/Pages/5G-fifth-generation-of-mobile-technologies.aspx#:~:text=The number of connected devices,in smart networked communication environments.

Jain, S., Shukla, S., & Wadhvani, R. (2018). Dynamic selection of normalization techniques using data complexity measures. Expert Systems with Applications, 106, 252–262. https://doi.org/10.1016/j.eswa.2018.04.008

Jiang, T., Gradus, J. L., & Rosellini, A. J. (2020). Supervised Machine Learning: A Brief Primer. Behavior Therapy, 51(5), 675–687. https://doi.org/https://doi.org/10.1016/j.beth.2020.05.002

Karipidis, K., Mate, R., Urban, D., Tinker, R., & Wood, A. (2021). 5G mobile networks and health—a state-of-the-science review of the research into low-level RF fields above 6 GHz. Journal of Exposure Science & Environmental Epidemiology, 31(4), 585–605. https://doi.org/10.1038/s41370-021-00297-6

Kim, Seungmo, & Nasim, I. (2020). Human electromagnetic field exposure in 5G at 28 GHz. IEEE Consumer Electronics Magazine, 9(6), 41–48. https://doi.org/10.1109/MCE.2019.2956223

Kim, Sungil, & Kim, H. (2016). A new metric of absolute percentage error for intermittent demand forecasts. International Journal of Forecasting, 32(3), 669–679. https://doi.org/10.1016/j.ijforecast.2015.12.003

Kleinlogel, H., Dierks, T., Koenig, T., Lehmann, H., Minder, A., & Berz, R. (2008). Effects of weak mobile phone - Electromagnetic fields (GSM, UMTS) on well-being and resting EEG. Bioelectromagnetics, 29(6), 479–487. https://doi.org/10.1002/bem.20419

Koivisto, M, Krause, C. M., Revonsuo, A., Laine, M., & Hämäläinen, H. (2000). The effects of electromagnetic field emitted by GSM phones on working memory. Neuroreport, 11(8), 1641–1643. https://doi.org/10.1097/00001756-200006050-00009

Koivisto, Mika, Revonsuo, A., Krause, C., Haarala, C., Sillanmäki, L., Laine, M., & Hämäläinen, H. (2000). Effects of 902 MHz electromagnetic field emitted by cellular telephones on response times in humans. NeuroReport, 11(2), 413–415. https://doi.org/10.1097/00001756-200002070-00038

Książek, W., Abdar, M., Acharya, U. R., & Pławiak, P. (2019). A novel machine learning approach for early detection of hepatocellular carcinoma patients. Cognitive Systems Research, 54, 116–127. https://doi.org/https://doi.org/10.1016/j.cogsys.2018.12.001

KumarSingh, B., Verma, K., & S. Thoke, A. (2015). Investigations on Impact of Feature Normalization Techniques on Classifier's Performance in Breast Tumor Classification. International Journal of Computer Applications, 116(19), 11–15. https://doi.org/10.5120/20443-2793

Kwon, M. K., Choi, J. Y., Kim, S. K., Yoo, T. K., & Kim, D. W. (2012a). Effects of radiation emitted by WCDMA mobile phones on electromagnetic hypersensitive subjects. Environmental Health: A Global Access Science Source, 11(1), 1–8. https://doi.org/10.1186/1476-069X-11-69

Lee, D.-H. (2013). Pseudo-label: The simple and efficient semi-supervised learning method for deep neural networks. ICML 2013 Workshop: Challenges in Representation Learning, 1–6. https://www.kaggle.com/blobs/download/forum-message-attachment-files/746/pseudo_label_final.pdf

Liang, S., Ma, A., Yang, S., Wang, Y., & Ma, Q. (2018). A Review of Matched-pairs Feature Selection Methods for Gene Expression Data Analysis. Computational and Structural Biotechnology Journal, 16, 88–97. https://doi.org/10.1016/j.csbj.2018.02.005

Maccartney, G. R., Rappaport, T. S., Sun, S., & Deng, S. (2015). Indoor Office Wideband Millimeter-Wave Propagation Measurements and Channel Models at 28 and 73 GHz for Ultra-Dense 5G Wireless Networks. IEEE Access, 3, 2388–2424. https://doi.org/10.1109/ACCESS.2015.2486778

Malek, F., Rani, K. A., Rahim, H. A., & Omar, M. H. (2015). Effect of Short-Term Mobile Phone Base Station Exposure on Cognitive Performance, Body Temperature, Heart Rate and Blood Pressure of Malaysians. Scientific Reports, 5(1), 13206. https://doi.org/10.1038/srep13206

Markov, M. (2019). Mobile communications and public health.

Masrakin, K., Rahim, H. A., Soh, P. J., Abdulmalek, M., Adam, I., Warip, M. N. B. M., Abbasi,

Q. H., & Yang, X. (2019a). Assessment of worn textile antennas' exposure on the physiological parameters and well-being of adults. IEEE Access, 7, 98946–98958. https://doi.org/10.1109/ACCESS.2019.2928343

MCMC. (2019). NATIONAL 5G TASK FORCE FINDINGS AND RECOMMENDATIONS TO THE GOVERNMENT. August. https://www.mcmc.gov.my/skmmgovmy/media/General/pdf/National-5G-Task-Force-Public-Consultation-Document.pdf

Mueller, S. T., & Piper, B. J. (2014). The Psychology Experiment Building Language (PEBL) and PEBL Test Battery. Journal of Neuroscience Methods, 222, 250–259. https://doi.org/10.1016/j.jneumeth.2013.10.024

Nasim, I., & Kim, S. (2019a). Adverse Impacts of 5G Downlinks on Human Body. Conference Proceedings - IEEE SOUTHEASTCON, 2019-April(ii), 1–6. https://doi.org/10.1109/SoutheastCon42311.2019.9020454

Nasim, I., & Kim, S. (2019b). Mitigation of human EMF exposure in downlink of 5G. Annales Des Telecommunications/Annals of Telecommunications, 74(1–2), 45–52. https://doi.org/10.1007/s12243-018-0696-6

NRPB. (2004). ICNIRP: Advice on Limiting Exposure to Electromagnetic Fields (0–300 GHz). 15(2).

Nyberg, R., & Hardell, L. (2017). EU 5G Appeal – Scientists warn of potential serious health effects of 5G. Jrs Eco Wireless, 1–11.

Oftedal, G., Straume, A., Johnsson, A., & Stovner, L. J. (2007a). Mobile phone headache: A double blind, sham-controlled provocation study. Cephalalgia, 27(5), 447–455. https://doi.org/10.1111/j.1468-2982.2007.01336.x

Olsen, C. R., Mentz, R. J., Anstrom, K. J., Page, D., & Patel, P. A. (2020). Clinical applications of machine learning in the diagnosis, classification, and prediction of heart failure. American Heart Journal, 229, 1–17. https://doi.org/10.1016/j.ahj.2020.07.009

Pawlak, R., Krawiec, P., & Żurek, J. (2019). On Measuring Electromagnetic Fields in 5G Technology. IEEE Access, 7, 29826–29835. https://doi.org/10.1109/ACCESS.2019.2902481

Pickering, T. G., Hall, J. E., Appel, L. J., Falkner, B. E., Graves, J., Hill, M. N., Jones, D. W., Kurtz, T., Sheps, S. G., & Roccella, E. J. (2005). Recommendations for blood pressure measurement in humans and experimental animals: part 1: blood pressure measurement in humans: a statement for professionals from the Subcommittee of Professional and Public Education of the American Heart Association Co. Circulation, 111(5), 697–716. https://doi.org/10.1161/01.CIR.0000154900.76284.F6

Pophof, B., Burns, J., Danker-Hopfe, H., Dorn, H., Egblomassé-Roidl, C., Eggert, T., Fuks, K., Henschenmacher, B., Kuhne, J., Sauter, C., & Schmid, G. (2021). The effect of exposure to radiofrequency electromagnetic fields on cognitive performance in human experimental studies: A protocol for a systematic review. Environment International, 157, 106783. https://doi.org/10.1016/j.envint.2021.106783

Preece, A. W., Iwi, G., Davies-Smith, A., Wesnes, K., Butler, S., Lim, E., & Varey, A. (1999). Effect of a 915-MHz simulated mobile phone signal on cognitive function in man. International Journal of Radiation Biology, 75(4), 447–456. https://doi.org/10.1080/095530099140375

Radák, Z. (2018). Chapter 6 - Speed as a Complex Conditional Ability. In Z. Radák (Ed.), The Physiology of Physical Training (pp. 111–118). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-815137-2.00006-1

Raju, V. N. G., Lakshmi, K. P., Jain, V. M., Kalidindi, A., & Padma, V. (2020). Study the Influence of Normalization/Transformation process on the Accuracy of Supervised Classification. Proceedings of the 3rd International Conference on Smart Systems and Inventive Technology, ICSSIT 2020, Icssit, 729–735. https://doi.org/10.1109/ICSSIT48917.2020.9214160

Regel, S. J., Negovetic, S., Röösli, M., Berdiñas, V., Schuderer, J., Huss, A., Lott, U., Kuster, N., & Achermann, P. (2006). UMTS base station-like exposure, well-being, and cognitive performance. Environmental Health Perspectives, 114(8), 1270–1275. https://doi.org/10.1289/ehp.8934

Regrain, C., Caudeville, J., de Seze, R., Guedda, M., Chobineh, A., de Doncker, P., Petrillo, L., Chiaramello, E., Parazzini, M., Joseph, W., Aerts, S., Huss, A., & Wiart, J. (2020). Design of an Integrated Platform for Mapping Residential Exposure to Rf-Emf Sources. International Journal of Environmental Research and Public Health, 17(15). https://doi.org/10.3390/ijerph17155339

Robert F., C. J., Sylvar, D. M., & Ulcek, J. L. (1997). Evaluating compliance with FCC guidelines for human exposure to radiofrequency electromagnetic fields. FCC OET Bulletin 65, Edition 97-01, August, 1–79.

Röösli, M., Dongus, S., Jalilian, H., Feychting, M., Eyers, J., Esu, E., Oringanje, C. M., Meremikwu, M., & Bosch-Capblanch, X. (2021). The effects of radiofrequency electromagnetic fields exposure on tinnitus, migraine and non-specific symptoms in the general and working population: A protocol for a systematic review on human observational studies. Environment International, 157, 106852. https://doi.org/10.1016/j.envint.2021.106852

RT-RF HA-1840GA1-NF Double Ridged Broadband Waveguide Horn Antenna. (n.d.). Fei Teng Wireless Technology CO.LTD. Retrieved February 1, 2021, from https://horn.ft-rf.com.tw/18ghz-to-40ghz-double-ridged-broadband-waveguide-horn-antenna

Saad, S. M., Shakaff, A. Y. M., Saad, A. R. M., Yusof, A. M., Andrew, A. M., Zakaria, A., & Adom, A. H. (2017). Analysis of feature selection with Probabilistic Neural Network (PNN) to classify sources influencing indoor air quality. AIP Conference Proceedings, 1808(March). https://doi.org/10.1063/1.4975275

Sani, N. M., & Tay, K. X. (2018). Resident's Perception of the health impact in living proximity of telecommunication base station: A case study of Malaysia. Pacific Rim Property Research Journal, 24(3), 213–223. https://doi.org/10.1080/14445921.2018.1551714

Sarker, I. H. (2021). Machine Learning: Algorithms, Real-World Applications and Research Directions. SN Computer Science, 2(3), 160. https://doi.org/10.1007/s42979-021-00592-x

Sauter, C., Eggert, T., Dorn, H., Schmid, G., Bolz, T., Marasanov, A., Hansen, M.-L., Peter, A., & Danker-Hopfe, H. (2015). Do signals of a hand-held TETRA transmitter affect cognitive performance, well-being, mood or somatic complaints in healthy young men? Results of a randomized double-blind cross-over provocation study. Environmental Research, 140, 85–94. https://doi.org/10.1016/j.envres.2015.03.021

Shahriyari, L. (2019). Effect of normalization methods on the performance of supervised learning algorithms applied to HTSeq-FPKM-UQ data sets: 7SK RNA expression as a predictor of survival in patients with colon adenocarcinoma. Briefings in Bioinformatics, 20(3), 985–994. https://doi.org/10.1093/bib/bbx153

Simplilearn. (2020). Data Preprocessing - Machine Learning: Simplilearn. Simplilearn.Com. https://www.simplilearn.com/data-preprocessing-tutorial.

Singh, Anushikha, Singh, N., Jindal, T., Rosado, A., & Dutta, M. (2020). A novel pilot study of automatic identification of EMF radiation effect on brain using computer vision and machine learning. Biomedical Signal Processing and Control, 57, 101821. https://doi.org/10.1016/j.bspc.2019.101821

Singh, Asmita, Halgamuge, M., & Lakshmiganthan, R. (2017). Impact of Different Data Types on Classifier Performance of Random Forest, Naïve Bayes, and K-Nearest Neighbors Algorithms. International Journal of Advanced Computer Science and Applications, 8. https://doi.org/10.14569/IJACSA.2017.081201

Stanislawska, K., Krawiec, K., & Kundzewicz, Z. W. (2012). Modeling global temperature changes with genetic programming. Computers and Mathematics with Applications, 64(12), 3717–3728. https://doi.org/10.1016/j.camwa.2012.02.049

Stöckel, T., Wunsch, K., & Hughes, C. M. L. (2017). Age-Related Decline in Anticipatory Motor Planning and Its Relation to Cognitive and Motor Skill Proficiency. Frontiers in Aging Neuroscience, 9, 283. https://doi.org/10.3389/fnagi.2017.00283

Suthaharan, S. (2015). Machine Learning Models and Algorithms for Big Data Classification: Thinking with Examples for Effective Learning (Vol. 36). https://doi.org/10.1007/978-1-4899-7641-3

Tahvanainen, K., Niño, J., Halonen, P., Kuusela, T., Laitinen, T., Länsimies, E., Hartikainen, J., Hietanen, M., & Lindholm, H. (2004). Cellular Phone Use Does Not Acutely Affect Blood Pressure or Heart Rate of Humans. Bioelectromagnetics, 25, 73–83. https://doi.org/10.1002/bem.10165

Tasneem Sofri, Hasliza A Rahim, Allan Melvin Andrew, Ping Jack Soh, Latifah Munirah Kamarudin, & Nishizaki Hiromitsu. (2023). Data Normalization Methods of Hybridized Multi-Stage Feature Selection Classification for 5G Base Station Antenna Health Effect Detection. Journal of Advanced Research in Applied Sciences and Engineering Technology, 30(2 SE-Articles), 133–140. https://doi.org/10.37934/araset.30.2.133140

Thors, B., Colombi, D., Ying, Z., Bolin, T., & Tornevik, C. (2016). Exposure to RF EMF from Array Antennas in 5G Mobile Communication Equipment. IEEE Access, 4(c), 7469–7478. https://doi.org/10.1109/ACCESS.2016.2601145

Tran, B., Xue, B., & Zhang, M. (2018). A New Representation in PSO for Discretization-Based

Feature Selection. IEEE Transactions on Cybernetics, 48(6), 1733–1746. https://doi.org/10.1109/TCYB.2017.2714145

Trunk, A., Stefanics, G., Zentai, N., Bacskay, I., Felinger, A., Thuróczy, G., & Hernádi, I. (2015). Effects of concurrent caffeine and mobile phone exposure on local target probability processing in the human brain. Scientific Reports, 5, 14434. https://doi.org/10.1038/srep14434

UK Publishes 5G Conspiracy Guide to Quash Misinformation. (2020). E&T Magazine.

United States Government Accountability Office. (2012). Telecommunications: exposure and testing requirements for mobile phones should be reassessed. GAO-12-771(Aug).

van Moorselaar, I., Slottje, P., Heller, P., van Strien, R., Kromhout, H., Murbach, M., Kuster, N., Vermeulen, R., & Huss, A. (2017). Effects of personalised exposure on self-rated electromagnetic hypersensitivity and sensibility – A double-blind randomised controlled trial. Environment International, 99, 255–262. https://doi.org/10.1016/j.envint.2016.11.031

Vecsei, Z., Knakker, B., Juhász, P., Thuróczy, G., Trunk, A., & Hernádi, I. (2018a). Short-term radiofrequency exposure from new generation mobile phones reduces EEG alpha power with no effects on cognitive performance. Scientific Reports, 8(1), 18010. https://doi.org/10.1038/s41598-018-36353-9

Vecsei, Z., Knakker, B., Juhász, P., Thuróczy, G., Trunk, A., & Hernádi, I. (2018b). Short-term radiofrequency exposure from new generation mobile phones reduces EEG alpha power with no effects on cognitive performance. Scientific Reports, 8(1), 1–12. https://doi.org/10.1038/s41598-018-36353-9

Verbeek, J., Oftedal, G., Feychting, M., van Rongen, E., Rosaria Scarfì, M., Mann, S., Wong, R., & van Deventer, E. (2021). Prioritizing health outcomes when assessing the effects of exposure to radiofrequency electromagnetic fields: A survey among experts. Environment International, 146, 106300. https://doi.org/https://doi.org/10.1016/j.envint.2020.106300

Vijayasarveswari, V., Andrew, A. M., Jusoh, M., Sabapathy, T., Raof, R. A. A., Yasin, M. N. M., Ahmad, R. B., Khatun, S., & Rahim, H. A. (2020). Multi-stage feature selection (MSFS) algorithm for UWB-based early breast cancer size prediction. PLoS ONE, 15(8 August), 1–21. https://doi.org/10.1371/journal.pone.0229367

Wali, S. Q., Sali, A., Allami, J. K., & Osman, A. F. (2022). RF-EMF Exposure Measurement for 5G Over Mm-Wave Base Station With MIMO Antenna. IEEE Access, 10, 9048–9058. https://doi.org/10.1109/ACCESS.2022.3143805

Wallace, D., Eltiti, S., Ridgewell, A., Garner, K., Russo, R., Sepulveda, F., Walker, S., Quinlan, T., Dudley, S., Maung, S., Deeble, R., & Fox, E. (2010). Do TETRA (Airwave) base station signals have a short-term impact on health and well-being? a randomized double-blind provocation study. Environmental Health Perspectives, 118(6), 735–741. https://doi.org/10.1289/ehp.0901416

Wallace, J., Andrianome, S., Ghosn, R., Blanchard, E. S., Telliez, F., & Selmaoui, B. (2020). Heart rate variability in healthy young adults exposed to global system for mobile communication (GSM) 900-MHz radiofrequency signal from mobile phones. Environmental Research, 191, 110097.

https://doi.org/https://doi.org/10.1016/j.envres.2020.110097

Walpole, R. E., Myers, R. H., Myers, S. L., & Ye, K. (2016). Probability & statistics for engineers & scientists. Boston: Prentice Hall.

Wang, C. X., Wu, S., Bai, L., You, X., Wang, J., & I, C. L. (2016). Recent advances and future challenges for massive MIMO channel measurements and models. Science China Information Sciences, 59(2), 1–16. https://doi.org/10.1007/s11432-015-5517-1

Wang, Y., Chu, Y. M., Thaljaoui, A., Khan, Y. A., Chammam, W., & Abbas, S. Z. (2021). A multi-feature hybrid classification data mining technique for human-emotion. BioData Mining, 14(1), 1–20. https://doi.org/10.1186/s13040-021-00254-x

Wei, Y., Wang, H., Tsang, K. F., Liu, Y., Wu, C. K., Zhu, H., … Hung, F. H. (2019). Proximity Environmental Feature Based Tree Health Assessment Scheme Using Internet of Things and Machine Learning Algorithm. Sensors, 19(14), 31.

Woods, D. L., Wyma, J. M., Yund, E. W., Herron, T. J., & Reed, B. (2015). Factors influencing the latency of simple reaction time. Frontiers in Human Neuroscience, 9. https://doi.org/10.3389/fnhum.2015.00131

World Health Organization (WHO). (n.d.). Definition of key terms. Retrieved March 15, 2022, from http://www.who.int/hiv/pub/guidelines/arv2013/intro/keyterms/en/

Xue, B., Zhang, M., Member, S., & Browne, W. N. (2016). A Survey on Evolutionary Computation Approaches to Feature Selection. IEEE Transactions on Evolutionary Computation, 20(4), 606–626. https://doi.org/10.1109/TEVC.2015.2504420

Zhang, R., Nie, F., Li, X., & Wei, X. (2018). Feature Selection with Multi-view Data: A Survey. Information Fusion, 50. https://doi.org/10.1016/j.inffus.2018.11.019

Zhou, X., Zhang, R., Yang, K., Yang, C., & Huang, T. (2020). Using hybrid normalization technique and state transition algorithm to VIKOR method for influence maximization problem. Neurocomputing, 410, 41–50. https://doi.org/10.1016/j.neucom.2020.05.084

Ziegelberger, G., Croft, R., Feychting, M., Green, A. C., Hirata, A., d'Inzeo, G., Jokela, K., Loughran, S., Marino, C., Miller, S., Oftedal, G., Okuno, T., van Rongen, E., Röösli, M., Sienkiewicz, Z., Tattersall, J., & Watanabe, S. (2020). Guidelines for limiting exposure to electromagnetic fields (100 kHz to 300 GHz). In Health Physics (Vol. 118, Issue 5). https://doi.org/10.1097/HP.0000000000001210

Zwamborn, A. P. M., Vossen, I. H. J. A. B. J. A. M., Leersum, I. M. A. Van, Ouwens, W. N., Mäkel, & Ongerubriceerd, O. (2003). Effects of Global Communication system radio-frequency fields on Well Being and Cognitive Functions of human subjects with and without subjective complaints.

APPENDIX A: ETHICAL APPROVAL FROM UNIMAP

UNIVERSITI MALAYSIA PERLIS
UniMAP

JABATAN CANSELORI
PEJABAT TIMBALAN NAIB CANSELOR (PENYELIDIKAN & INOVASI)

Alamat Surat Menyurat :

Pejabat Timbalan Naib Canselor
(Penyelidikan & Inovasi)
Universiti Malaysia Perlis
Aras 3, Bangunan Canselori,
Kampus UniMAP Pauh Putra,
02600 Arau Perlis,
MALAYSIA.

Pusat Pengurusan Penyelidikan
(RMC)
Universiti Malaysia Perlis,
Aras 3, Bangunan Canselori,
Kampus UniMAP Pauh Putra,
02600 Arau Perlis,
MALAYSIA.

Pusat Pengurusan Keselamatan,
Kesihatan Dan Persekitaran
Pekerjaan (COSHE)
Universiti Malaysia Perlis,
No.2, Blok D, Taman Pertiwi
Indah, Jalan Kangar-Alor Setar,
Seriab 01000 Kangar Perlis,
MALAYSIA.
Tel : 04-9797756
Faks : 04-9798351

Pusat Inovasi & Pengkomersilan
(CIC)
Universiti Malaysia Perlis
Tingkat 4, Bangunan KWSP,
01000 Kangar Perlis,
MALAYSIA.

Institut Agroteknologi Lestari
(IN...)
Universiti Malaysia Perlis
Kampus Sg. Ch...
...adang...
MAL...
: 04-9...
: 04-9...

...(COSCEM)
...ia Perlis,
...Pelajar,
...engaji... ...teraan
...tronik

Ruj. Kami : UniMAP/PTNC(P&)I/100-1 ( )
Tarikh : 22 Mac 2021

Prof. Madya Ts. Dr. Hasliza A Rahim@Samsuddin
Fakulti Teknologi Kejuruteraan Elektronik
Universiti Malaysia Perlis (UniMAP)

YBrs. Prof. Madya Ts. Dr.,

KEPUTUSAN JAWATANKUASA ETIKA PENYELIDIKAN MANUSIA BAGI PENYELIDIKAN BERTAJUK *EFFECTS OF SHORT-TERM 5G BASE STATION SIGNAL EXPOSURE ON COGNITIVE PERFORMANCE, SUBJECTIVE WELL-BEING, PHYSIOLOGICAL PARAMETERS AND EEG IN MALAYSIAN ADULTS: A DOUBLE BLIND, RANDOMIZED, COUNTERBALANCED, CROSS-OVER STUDY*

Dengan segala hormatnya saya merujuk perkara di atas.

2. Seperti pihak YBrs. Prof. Madya Ts. Dr. sedia maklum, Pejabat Timbalan Naib Canselor (Penyelidikan dan Inovasi) telah mengadakan Mesyuarat Jawatankuasa Etika Penyelidikan Manusia Bil. 1/2021 pada 12 Mac 2021 (Jumaat) secara atas talian (Google Meet).

3. Ahli Jawatankuasa yang hadir adalah seperti berikut :

BIL	NAMA DAN ORGANISASI	
1.	Prof. Ts. Dr. Mohd Mustafa Al Bakri Abdullah Menjalankan Fungsi, Timbalan Naib Canselor (Penyelidikan dan Inovasi) Universiti Malaysia Perlis	Pengerusi
2.	Prof. Dr. Mohd Foad Sakdan Menjalankan Fungsi, Pengarah Pusat Kesihatan Universiti Universiti Malaysia Perlis	Ahli
3.	Dr. Mohd Shahfi Fauzee Ismail Pegawai Perubatan, Pusat Kesihatan Universiti Universiti Malaysia Perlis	Ahli
4.	Dr. Ruzita Jamaluddin Ketua Jabatan Jabatan Psikiatri dan Kesihatan Mental Hospital Tuanku Fauziah, Kangar, Perlis	Ahli
5.	Dr. Karniza Khalid Pegawai Perubatan, Pusat Penyelidikan Klinikal Hospital Tuanku Fauziah, Kangar, Perlis	Ahli

STARS™ RATING SYSTEM

ILMU KEIKHLASAN KECEMERLANGAN

6.	Prof. Madya Ir. Ts. Dr. Shahriman Abu Bakar Ketua, Pusat Kecemerlangan - *Center Of Automotive & Motorsport* Universiti Malaysia Perlis	Ahli
7.	Prof. Madya Dr. Ahmad Faizal Salleh Ketua, Pusat Kecemerlangan - *Sport Engineering Research Centre* Universiti Malaysia Perlis	Ahli
8.	Prof. Madya Dr. Huzili Hussin Ketua, Pusat Kecemerlangan - Social Innovation and Sustainability Universiti Malaysia Perlis	Ahli
9.	Prof. Madya Dr. Mohd Najib Mohd Yasin Ketua, Pusat Kecemerlangan - *Advanced Communication Engineering Centre* Universiti Malaysia Perlis	Ahli
10.	Prof. Madya Dr. Mohd Sobri Idris Ketua, Pusat Kecemerlangan – *Frontier Materials Research* Universiti Malaysia Perlis	Ahli
11.	Prof. Madya Dr. Ku Syahidah Ku Ismail Ketua, Pusat Kecemerlangan – *Biomass Utilization* Universiti Malaysia Perlis	Ahli
12.	Prof. Madya Dr. Muzamir Isa Ketua, Pusat Kecemerlangan – *Renewable Energy* Universiti Malaysia Perlis	Ahli
13.	Prof. Madya Dr. Ruzelita Ngadiran Ketua, Pusat Kecemerlangan - *Advanced Computing* Universiti Malaysia Perlis	Ahli
14.	Dr. Wan Mohd Khairy Adly Wan Zaimi Timbalan Dekan (Penyelidikan dan Pascasiswazah) Fakulti Sains Gunaan & Kemanusiaan Universiti Malaysia Perlis	Ahli Ganti

4. Justeru itu, Jawatankuasa setelah berbincang bersetuju seperti ketetapan di bawah:

Tajuk Penyelidikan	*Effects of Short-Term 5G Base Station Signal Exposure on Cognitive Performance, Subjective Well-Being, Physiological Parameters and EEG in Malaysian Adults: A Double Blind, Randomized, Counterbalanced, Cross-Over Study*
Ketua Penyelidik	Prof. Madya Ts. Dr. Hasliza A Rahim@Samsuddin

Fakulti / Jabatan	Fakulti Teknologi Kejuruteraan Elektronik
Cadangan Penambahbaikan	i. Perkemaskan proposal kajian dengan menambah elemen *contingency plan* sekiranya peserta kajian mendapat kesan sampingan semasa intervensi (exposure Risk Factor). Cadangan *contingency plan* termasuk adanya pasukan perubatan *stand-by* semasa dan selepas intervensi dan pelan untuk *follow-up* subjek yang melaporkan kesan sampingan. ii. Berkenaan *randomization* pula, tidak ada elemen *randomization* dalam kajian *design* beliau. *Randomise* dari segi *sequency study arm* tidak dikira sebagai *randomization* dalam *clinical trial*. Dicadangkan boleh dikeluarkan *randomization* dalam *study design*.
Keputusan Jawatan Etika UniMAP	**LULUS**

5. Ketua Penyelidik disarankan untuk pertimbangan berkenaan dengan cadangan penambahbaikan seperti dalam jadual di atas.

Sekian , terima kasih.

"PRIHATIN RAKYAT: DARURAT MEMERANGI COVID-19"

"BERKHIDMAT UNTUK NEGARA"
"Ilmu, Keikhlasan, Kecemerlangan"

Saya yang menjalankan amanah

(PROF. TS. DR. MOHD MUSTAFA AL BAKRI ABDULLAH)
Menjalankan Fungsi,
Timbalan Naib Canselor (Penyelidikan dan Inovasi)
Universiti Malaysia Perlis (UniMAP)

APPENDIX B: NORMALIZATION TECHNIQUES AND THEIR EXPRESSIONS

No.	Normalization	Equation	Reference
1	Z- Score Normalization	$x' = \frac{x - \mu}{\sigma}$	(Raju et al., 2020)
2	Linear scaling	$x' = \frac{(x - min)}{max - min}$	(Raju et al., 2020)
3	Binary Normalization	$x' = \frac{0.8(x - min)}{max - min} + 0.1$	(Stanislawska et al., 2012)
4	Bipolar Normalization	$x' = \frac{1.8(x - min)}{max - min} - 0.9$	(Stanislawska, Krawiec, & Kundzewicz, 2012)
5	MM scaling	$x' = x/(max - \min)$	(KumarSingh et al., 2015)
6	t- Score Normalization	$x' = \frac{x - \mu}{\sigma/\sqrt{n}}$	(Walpole, R. E., Myers, R. H., Myers, S. L., & Ye, 2016)
7	Differential Moment Normalization	$M_i = \frac{1}{N^2}(\sum_{i=1}^{N} x_i)^2 - x_i{}^2$	(Saad et al., 2017)
8	Variation Normalization	$C_{V,i} = \frac{\sigma}{\mu} x_i$	(Walpole, R. E., Myers, R. H., Myers, S. L., & Ye, 2016)
9	Decimal Inverse Logarithmic Scaled Normalization	$x' = 10^{-12}\, 10^{0.1x} * 10^7$	(Zhou et al., 2020)

10	Absolute Percentage Error Normalization (APE) formula 1	$x' = (\frac{\bar{x} - x_i}{(\bar{x} + x_i)/2})$	(Chen et al., 2017)
11	Absolute Percentage Error Normalization (APE) formula 2	$x' = (\frac{\bar{x} - x_i}{\bar{x}})$	(Sungil Kim & Kim, 2016)
12	Arctan APE formula 1	$x' = arctan(\frac{\bar{x} - x_i}{(\bar{x} + x_i)/2})$	(Sungil Kim & Kim, 2016)
13	Arctan APE formula 2	$x' = arctan(\frac{\bar{x} - x_i}{\bar{x}})$	(Sungil Kim & Kim, 2016)
14	Gaussian Normalization	$x' = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\,(-\frac{(x_i - \mu)^2}{2\sigma^2})$	(Walpole, R. E., Myers, R. H., Myers, S. L., & Ye, 2016)
15	Relative Sum Squared Value (RSSV)	$x' = \frac{x}{\sum x^2}$	(Andrew et al., 2016)
16	Relative Logarithmic Sum Squared Value (RLSSV)	$x' = \frac{\log\,(x)}{\log\,(\sum x^2)}$	(Andrew et al., 2016)
17	Relative Mean normalization	$x' = x/mean$	(Saad et al., 2017)
18	Relative Standard deviation normalization	$x' = x/std$	(Saad et al., 2017)
18	Relative Interquartile normalization	$x' = x/IQR$	(Saad et al., 2017)
20	Robust normalization	$x' = (x - median)/IQR$	(Raju et al., 2020)

APPENDIX C: THE RF SHIELDED ROOM