ELECTROENCEPHALOGRAPHY SIGNAL CLASSIFICATION USING NEURAL NETWORK, DECISION TREE AND ENSEMBLE BAGGED TREE FOR EPILEPSY DISEASE

NUR SYAHIRAH BINTI ABDUL AZIZ

A thesis submitted in fulfilment of the
requirements for the award of the degree of
Master of Philosophy

Faculty of Science
Universiti Teknologi Malaysia

AUGUST 2022

DEDICATION

This thesis is dedicated to my beloved mother and father, for their kindness, devotion and endless support.

ACKNOWLEDGEMENT

First and foremost, I would like to take this opportunity to express my appreciation to my supervisor, Assoc. Prof Dr. Norma Alias for encouragement, advice, guidance, and motivation.

I am also indebted to Universiti Teknologi Malaysia (UTM), all lecturers and staff who have guided and inspired me to develop this thesis. Librarians at UTM also deserve special thanks for their assistance in supplying the relevant literatures.

My sincere appreciation also extends to MybrainSc for funding my master's studies. I would also like to thank my seniors and friends Ms. Alia Rafiza, Ms. Kamila, Ms. Azura, Mr. Ainol, and Mr. Lim for their guidance in my thesis. It is not possible to list all of them in this limited space. I am grateful to all my family member for providing me with unfailing support and motivation throughout my years of study.

ABSTRACT

Epilepsy is a brain disease caused by abnormal brain activities. Machine learning algorithms are usually applied in the classification and identification of epilepsy at an early stage. This study's primary objective is to classify the Electroencephalography (EEG) signal dataset of epileptic seizures using a machine learning algorithm and to evaluate the performance using the Plot Confusion Matrix and area under the receiver operating characteristic (AUC-ROC) curve. The plot confusion matrix method will give an array that depicts the True Positives, False Positives, False Negatives, and True Negatives. Besides, the AUC-ROC curve is a performance measurement for classification problems at various threshold settings. These methods can be used to check or visualize the performance of the multi-class classification problem. This thesis involves a collection of datasets containing 200 healthy individuals and 300 epilepsy patients. Next, features were extracted from these datasets. Feature extractions help to reduce data dimensionality and eliminate noise, while its output is used as the input for classifier-based epileptic classification. This study selected the Discrete Wavelet Transform (DWT) and Statistical Features as feature extraction methods. In addition, multiple machine learning techniques are presented in this study. Feed Forward Neural Network (FFNN), Back Propagation Neural Network (BPNN), Decision Tree, and Ensemble Bagged Tree (EBT) were used as classifiers. Furthermore, Linear Discriminant Analysis (LDA) has been selected as the benchmark standard. Therefore, five classifiers were trained for classification purposes. Each classifier is combined with DWT and Statistical Features. The proposed feature extraction, DWT combined with BPNN, gives the highest accuracy of 91.2% and a shorter duration of training.

ABSTRAK

Epilepsi adalah penyakit otak yang disebabkan oleh aktiviti otak yang tidak normal. Algoritma pembelajaran mesin biasanya digunakan dalam klasifikasi dan pengenalpastian epilepsi pada peringkat awal. Objektif utama kajian ini adalah untuk mengklasifikasikan set data isyarat Electroencephalogram (EEG) bagi sawan epilepsi menggunakan algoritma pembelajaran mesin dan menilai prestasi menggunakan Matriks Kekeliruan Plot dan kawasan di bawah lengkung ciri pengendalian penerima (AUC-ROC). Kaedah matriks kekeliruan plot akan memberikan tatasusunan yang menggambarkan Positif Benar, Positif Palsu, Negatif Palsu dan Negatif Benar. Selain itu, keluk AUC-ROC ialah pengukuran prestasi untuk masalah pengelasan pada pelbagai tetapan ambang. Kaedah ini boleh digunakan untuk menyemak atau menggambarkan prestasi masalah pengelasan berbilang kelas. Tesis ini melibatkan sekumpulan set data yang mengandungi 200 individu yang sihat dan 300 orang pesakit epilepsi. Seterusnya, ciri telah diekstrak daripada set data ini. Pengekstrakan ciri membantu mengurangkan dimensi data dan meminimumkan bunyi, serta keluaran daripada pengekstrakan ciri ini digunakan sebagai input untuk pengelasan epilepsi dengan menggunakan pengelas. Kajian ini memilih Transformasi Gelombang Diskret (DWT) dan Ciri Statistik sebagai kaedah pengekstrakan ciri. Selain itu, pelbagai teknik pembelajaran mesin dibentangkan dalam kajian ini. Rangkaian Neural Suap Maju (FFNN), Rangkaian Neural Propagasi Belakang (BPNN), Pepohon Keputusan dan Pepohon Beg Ensembel (EBT) digunakan sebagai pengelas. Tambahan pula, Analisis Diskriminasi Linear (LDA) telah dipilih sebagai piawaian penanda aras. Oleh itu, lima pengelas telah dilatih bagi tujuan pengelasan. Setiap pengelas akan digabungkan dengan DWT dan Ciri Statistik. Pengekstrakan ciri yang dicadangkan, DWT digabungkan dengan BPNN memberikan ketepatan tertinggi iaitu 91.2% dan tempoh latihan yang lebih singkat.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ANN	-	Artificial Neural Network
ABC	-	Artificial Bee Colony
ACO	-	Ant Colony Optimization
AI	-	Artificial Intelligence
BCI	-	Brain Computer Interface
BPNN	-	Back Propagation Neural Network
CWT	-	Continuous Wavelet Transform
DT	-	Decision Tree
DWT	-	Discrete Wavelet Transform
EBT	-	Ensemble Bagged Tree
EEG	-	Electroencephalography
ELM	-	Ensemble Learning Method
FFNN	-	Feed Forward Neural Network
LDA	-	Linear Discriminant Analysis
LM	-	Lavernberg-Marquardt
ML	-	Machine Learning
RBF	-	Radial Basis Function
SF	-	Statistical Features

LIST OF SYMBOLS

p_i	-	Binary Input
w	-	Weight
a	-	Activation Function
b	-	Bias
y	-	Output Value
∇	-	Gradient
C	-	Cost Function
$g(n)$	-	High Pass Filter
$h(n)$	-	Low Pass Filter
L	-	Last Layer
$L-1$	-	Second Last Layer
n	-	Number of Weight and Biases
$\mu(x)$	-	Mean of Raw EEG Signal
S_i	-	EEG Signal Data Point

CHAPTER 1

INTRODUCTION

1.1 Research Background

Epilepsy is one of the brain disorders caused by abnormal brain activities. There are several symptoms to detect epilepsy such as unusual behaviour, confusion, and loss of awareness (Andrzejak et al., 2001). Moreover, epileptic symptoms in an individual may lead in many cases of injuries due to falls, involuntary muscular contraction, biting one's tongue and accompanied strong pain (Shorvon et al., 2012; Siuly, Li, and Zhang, 2016). Seizures can be analysed by using Electroencephalography (EEG), the signals which will provide useful information that can be used for monitoring individuals who suffer from this disorder (Bongiorni and Balbinot, 2020). Many studies have been conducted on the EEG of patients with diseases, such as Alzheimer's, Parkinson's, and depression, as well as healthy patients and patients with epilepsy (Almustafa, 2020).

There are several ways to predict the seizures such as from their clinical analysis to EMG (Vandecasteele et al., 2017) and the monitoring of diverse electrical variables (Jory et al., 2016). However, automatically inspecting and classifying the seizures accurately and effectively might be a time-consuming, costly and laborious process. Moreover, detecting a seizure beforehand is difficult for many researchers. It is also hard to determine the accuracy of the manual classification (Bongiorni and Balbinot, 2020). According to Chamorro-Atalaya *et al.* (2021), there are various technologies used to obtain predictive models, which use data from virtual platforms. Within these technologies is the branch of Artificial Intelligence that within its fields houses Machine Learning (Pedrero et al., 2021). Machine Learning is a set of algorithms capable of learning to perform certain tasks from the generalization of examples. AI is an intelligent computational technique, usually fed and trained with

EEG data with the ultimate goal of extracting important features from the data and learning how to generalize them to new inputs of the same type (Bongiorni and Balbinot, 2020).

According to Almustafa (2020), an epileptic seizure can be classified by using a special class of Artificial Neural Network (ANN) such as Recurrent Neural Network (RNN) that can hold an internal structure with a feedback loop. Other studies also decided to use different machine learning algorithms to classify the epileptic seizure dataset. Few studies have been conducted on the classification of epilepsy in real-time applications (Vidyaratne and Iftekharuddin, 2017). In addition, this thesis presents a bagging ensemble learning technique to improve epilepsy classification performance.

Initially, the methodology used will be detailed, and then the validation of the algorithm will be determined. The datasets will then undergo preprocessing in preparation for feature extraction. In this sense, the main objective of this thesis is to determine the predictive model using Machine Learning by comparing FFNN, BPNN, Decision Tree, and enhancement of decision trees which is the Ensemble Bagged Trees (EBT) algorithm. EBT is intended to enhance the performance of DT classifiers based on their classification accuracy. These four techniques were utilised for the classification of epileptic seizure datasets using predictive analysis.

Furthermore, a standard Linear Discriminant Analysis will be set up as a benchmark in this thesis. Then, the technique that achieves the highest classification accuracy and surpasses the benchmark will be chosen as the best prediction model for this thesis. The chosen methods for this thesis have been proven in two ways, using plot confusion to confirm the accuracy of the trained network and plot receiver operating characteristic curve (ROC) analysis to confirm the performance of the trained network.

The total performance metrics of the chosen algorithm, such as accuracy, precision, sensitivity, and specificity, can then be evaluated using the Plot Confusion Matrix. The research contribution focuses on applying a novel technique through machine learning by using the data and information collected, which compares machine learning methods to identify a predictive model. The predictive model with the highest classification accuracy and shortest training time will be chosen.

1.2 Problem Statement

Seizures caused by abnormal brain activity associated with an epileptic disease have a variety of symptoms, including unusual behaviour, confusion, and loss of awareness. EEG is a commonly used clinical approach for detecting epilepsy. Manual inspection of EEG brain signals can be done. However, this task will be a time-consuming, costly and laborious process. Furthermore, because the majority of seizures occur unexpectedly, it is difficult to detect a possible seizure before it occurs. It is also quite hard to determine the accuracy of manual classification. In this study, therefore, feature extraction and classification algorithms are used to predict whether a person will have a seizure or not. This study will aid in classifying the epileptic seizure dataset more precisely and efficiently. Machine Learning has been chosen to solve these classification issues.

In recent years, Machine Learning techniques have been successfully used in clinical approaches and a wide range of human endeavours to design algorithms, methods, and models (Weng, 2020). As classifiers for the epileptic seizures datasets, DT and ANN were selected. In this context, classifiers may be over-fitted when they perform very well with training data but poorly with test data. In addition, overfitting occurs when the training data are not cleaned and contain noise. This may also occur in other machine learning approaches, however, to address this problem, we trained the model with sufficient data, applied a data pre-processing method, extracted features, and adopted ensembling techniques known as Ensemble Bagged Tree (EBT). EBT is one of the ensemble approaches utilised to enhance the performance evaluation of DT based on percentages of classification accuracy. All of these classifiers will then be evaluated and analyzed using plots of the confusion matrix and receiver operating characteristic curve (ROC) analysis.

1.3 Research Goal

The goal of our research is to determine the most appropriate and superior classification algorithm to classify the epileptic seizure dataset whether a person will have a seizure or not by employing various classification techniques, namely Artificial Neural Network (ANN) and Decision Trees (DT). The Ensemble Bagged Tree was then presented in this thesis to improve the decision tree's performance in terms of accuracy and training time. The performance evaluations of these classifiers will then be analysed and differentiated.

1.3.1 Research Objectives

The objectives of the research are:

(a) To extract the epileptic seizure dataset of Electroencephalography (EEG) using feature extraction techniques.

(b) To classify the epileptic seizure dataset using ANN and Decision Trees (DT).

(c) To improve the performance of the Decision Tree by employing the Ensemble Bagged Tree (EBT) approach and compare the performance of ANN, DT, and EBT based on their respective performance evaluations.

(d) To perform performance evaluation using Area under the Receiver Operator Characteristic (AUC-ROC) curve and Plot Confusion technique.

1.4 Scope of research

This thesis mainly focuses on classifying the EEG datasets of epileptic seizures. Firstly, the datasets will be preprocessed to remove noises and reduce the dimensionality of the data. Then, the datasets will be extracted by using two types of feature extractions, which are discrete wavelet transform (DWT) and statistical features. In this thesis, Machine Learning Techniques are used to classify the dataset of epileptic seizures. FFNN, BPNN, and DT were initially used to classify the data. The DT will then be upgraded to Ensemble Bagged Tree. In addition, a fundamental Linear Discriminant Analysis (LDA) will be included as one of the classifiers and will stand as the benchmark for this thesis. The performance of these classifiers will be tested using a plot confusion matrix and plot AUC-ROC curve. The performance evaluation of these methods will be analysed and compared based on their accuracy and duration of training. Plot Confusion Matrix will also be used to analyse the algorithm's entire performance metrics, including specificity, precision, sensitivity, recall, and F1-score. Last but not least, Matlab software was used for implementation and coding in this study.

1.5 Significance of research

Data analysis and classification of two or more different classes is an important skill for the clinical sector during this age of technological advancement. The need to possess appropriate skills and training, especially in artificial intelligence (AI) is a must. The classification of epileptic seizures can be achieved by using Machine Learning techniques. Artificial Neural Network (ANN), Decision Trees, and Ensemble Bagged Trees are the major focus for the classification of epileptic seizures in this thesis. These classifiers are subtopics of AI and fundamental approaches in the data learning network. Furthermore, these classifiers are simple to implement while outperforming any other classical statistical approach. One of the machine learning techniques chosen as the benchmark in this thesis is the Linear Discriminant Analysis. The benchmark test functions have been used for evaluating the performance of machine learning algorithms. The algorithm that performs well on a set of numerical optimization problems is considered one of the effective methods for solving classification problems.

To prove the effectiveness of these methods, a performance evaluation must be conducted. By using the Plot AUC-ROC curve and Plot confusion technique, the performance of classifiers that classified the epileptic seizure can be rated. The main purpose of the performance evaluation is to evaluate the accuracy, sensitivity, specificity, precision, f1 score and training time of the data. Moreover, the changes in brain activity can be determined by using the EEG device and it might be useful in diagnosing brain disorders, especially epilepsy and other seizure disorder. An EEG might also help treat or diagnose brain damage from a head injury and brain tumour.

In recent years, ANN, decision trees and Ensemble Bagged Tree have become a trend in various fields of study as evidenced by the increase in annual publications of these fields in the Scopus platform. This proves that the research of epilepsy classification by using ANN, decision trees and Ensemble Bagged Tree will be a beneficial move. Besides, there are several advantages when using MATLAB software. MATLAB also known as the automated driving toolbox, provides tools and algorithms for designing, simulating, testing, and autonomous driving systems. MATLAB is also easier and simple to use to remember syntax and the time required from idea to implementation is shorter. This thesis not only provides a deeper understanding of machine learning but also assists algorithm providers in analysing a variety of EEG signal data using machine learning techniques.

For this research, ANN and DT are the primary focus of the classification process. Both methods are becoming a trend in recent years in various fields of study and can be analyzed because both related fields of studies published in the Scopus platform increased yearly. Figure 1 shows the increasing graph of documents published in Scopus from 2010 to 2021 for ANN and DT studies increasing dramatically.

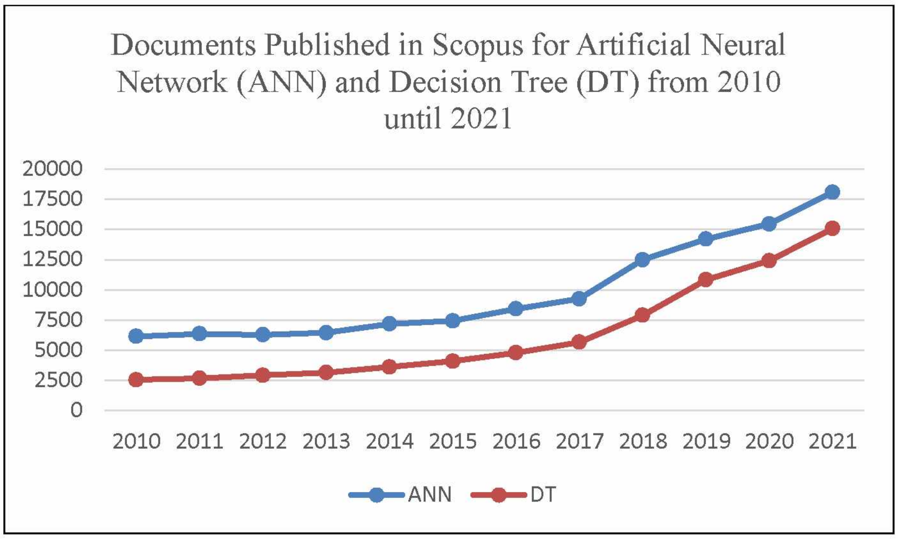

Figure 1.1 Documents Published in Scopus Platform for ANN and DT Studies (Scopus,2021)

Besides both methods stated above, this research also introduced another enhanced technique to produce a more significant result which is the Ensemble Bagged Tree (EBT). By using EBT, the results obtained for the classification process will become more reliable and accurate. EBT is also a new branch emerging from the bagging technique that has yet to be discovered further for the performance optimization of the result. The documents published in Scopus regarding this potential technique came on the rise from 2017 onwards, as shown in Figure 1.2.

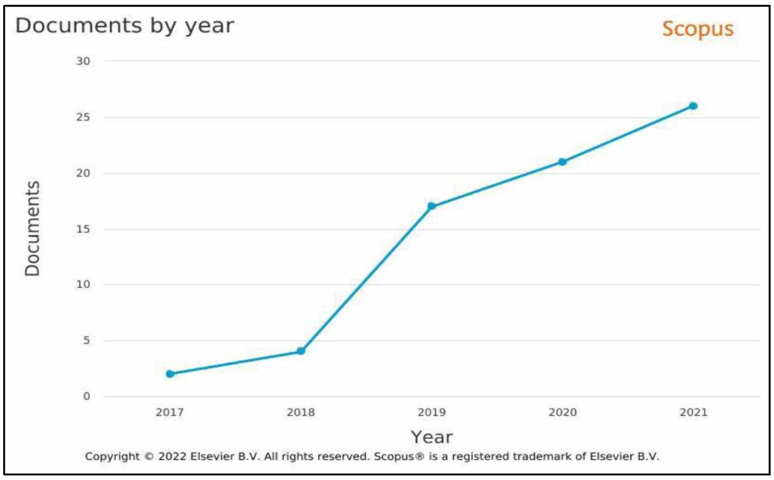

Figure 1.2 Documents Published in Scopus Platform for Decision Trees Studies (Scopus, 2020).

In this era of technological advancement, data analysis and classifications play a significant role in all related industries, including medical analysis. In a sense, this study will assist in resolving common issues faced by medical practitioners when conducting medical assessments on patients, especially early epilepsy detection.

1.6 Research Outline

In the following chapters, the thesis is organized as follows:

Chapter 2: Provides an overview of past studies and sequences for classifying data. The literature included the preprocessing technique, the relevant feature extraction method as well as the option for classifying the dataset by using various Machine Learning techniques.

Chapter 3: Provides the details of the chosen method based on the previously studied variety of options from chapter 2. The details are arranged in a way to represent the process of data preprocessing, extraction features, and choosing the suitable classifiers to classify the epileptic seizure dataset and proposed method. Lastly, the methodology for designing the programming will be elaborated.

Chapter 4: This chapter thoroughly explains the implementation of methods and Matlab coding used in this thesis.

Chapter 5: Presents the outcome of the feature extraction data analysis. The performance results of classifiers will be compared in terms of training time and classification accuracy and presents the discussion based on outcomes in Chapter 4.

Chapter 6: The conclusion from the performance of the classifiers. Further improvements are also discussed in this chapter.

REFERENCES

Al-Fahoum, A. S., & Al-Fraihat, A. A. (2014). Methods of EEG signal features extraction using linear analysis in frequency and time-frequency domains. *International Scholarly Research Notices, 2014*.

Almustafa, K. M. (2020). Classification of epileptic seizure dataset using different machine learning algorithms. *Informatics in Medicine Unlocked, 21*, 100444.

AlSharabi, K., Ibrahim, S., Djemal, R., & Alsuwailem, A. (2016). *A DWT-entropy-ANN based architecture for epilepsy diagnosis using EEG signals.* Paper presented at the 2016 2nd international conference on advanced technologies for signal and image processing (ATSIP).

Alsharif, A., Salleh, N., Baharun, R., & Safaei, M. (2020). Neuromarketing approach: An overview and future research directions. *Journal of Theoretical and Applied Information Technology, 98*(7), 991-1001.

Andrzejak, R. G., Lehnertz, K., Mormann, F., Rieke, C., David, P., & Elger, C. E. (2001). Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state. *Physical Review E, 64*(6), 061907.

Ankushkuwar. (2019) *ML Voting Classifier using Sklearn*. [Online]. Available at: https://www.geeksforgeeks.org/ml-voting-classifier-using-sklearn/ (Accessed: 25 May 2021).

Anunaya, S. (2021) *Data Preprocessing in Data Mining*. [Online]. Available at: https://www.analyticsvidhya.com/blog/2021/08/data-preprocessing-in-data-mining-a-hands-on-guide/ (Accessed: 12 June 2021)

Behrang, M., Assareh, E., Ghanbarzadeh, A., & Noghrehabadi, A. (2010). The potential of different artificial neural network (ANN) techniques in daily global solar radiation modeling based on meteorological data. *Solar Energy, 84*(8), 1468-1480.

Bheemalingaiah, M., Swamy, G. R., Vishvapathi, P., Babu, P. V., Rao, E. N., & Rao, P. N. (2021). Detection Of Heart Disease By Using Reliable Boolean Machine Learning Algorithm. *Journal of Theoretical and Applied Information Technology, 99*(15).

Bongiorni, L., & Balbinot, A. (2020). Evaluation of recurrent neural networks as epileptic seizure predictor. *Array, 8*, 100038.

Browniee, J. (2016) Time Series Forecasting as Supervised Learning. [Online]. Available at: https://machinelearningmastery.com/time-series-forecasting-supervised-learning/ (Accessed: December 2020)

Chamorro-Atalaya, O., Chávez-Herrera, C., Anton-De los Santos, M., Anton-De los Santos, J., Torres-Quiroz, A., Leva-Apaza, A., . . . Peralta-Eugenio, G. (2021). Machine Learning Model through Ensemble Bagged Trees in Predictive Analysis of University Teaching Performance. *Machine Learning, 12*(12).

Chen, F., Li, X., & Liu, L. (2013). *Improved C4. 5 decision tree algorithm based on sample selection.* Paper presented at the 2013 IEEE 4th International Conference on Software Engineering and Service Science.

Cheong, L. C., Sudirman, R., & Hussin, S. S. (2015). Feature extraction of EEG signal using wavelet transform for autism classification. *ARPN Journal of Engineering and Applied Sciences, 10*(19), 8533-8540.

CR, H., & MP, P. (2014). Classification of eye movements using electrooculography and neural networks. *International Journal of Human Computer Interaction (IJHCI), 5*(4), 51.

Cvetkovic, D., Übeyli, E. D., & Cosic, I. (2008). Wavelet transform feature extraction from human PPG, ECG, and EEG signal responses to ELF PEMF exposures: A pilot study. *Digital signal processing, 18*(5), 861-874.

Djemal, R., AlSharabi, K., Ibrahim, S., & Alsuwailem, A. (2017). EEG-based computer aided diagnosis of autism spectrum disorder using wavelet, entropy, and ANN. *BioMed research international, 2017*.

Elnemr, H. A., Zayed, N. M., & Fakhreldein, M. A. (2016). Feature extraction techniques: fundamental concepts and survey. In *Handbook of research on emerging perspectives in intelligent pattern recognition, analysis, and image processing* (pp. 264-294): IGI Global.

Faust, O., Acharya, U. R., Adeli, H., & Adeli, A. (2015). Wavelet-based EEG processing for computer-aided seizure detection and epilepsy diagnosis. *Seizure, 26*, 56-64.

Gokgoz, E., & Subasi, A. (2014). Effect of multiscale PCA de-noising on EMG signal classification for diagnosis of neuromuscular disorders. *Journal of medical systems, 38*(4), 1-10.

Graimann, B., Allison, B. Z., & Pfurtscheller, G. (2010). *Brain-computer interfaces: Revolutionizing human-computer interaction*: Springer Science & Business Media.

Graupe, D. (2013). *Principles of artificial neural networks* (Vol. 7): World Scientific.

Gupta, S. S., Soman, S., Raj, P. G., Prakash, R., Sailaja, S., & Borgohain, R. (2012). *Detecting eye movements in EEG for controlling devices.* Paper presented at the 2012 IEEE International Conference on Computational Intelligence and Cybernetics (CyberneticsCom).

Habib, A., Alalyani, M., Hussain, I., & Almutheibi, M. (2015). Brief review on sensitivity, specificity and predictivities. *IOSR J Dent Med Sci, 14*(4), 64-68.

Harris, J. M., Ciorciari, J., & Gountas, J. (2018). Consumer neuroscience for marketing researchers. *Journal of consumer behaviour, 17*(3), 239-252.

Islam, R., & Shahjalal, M. A. (2019). *Soft Voting-Based Ensemble Approach to Predict Early Stage DRC Violations.* Paper presented at the 2019 IEEE 62nd International Midwest Symposium on Circuits and Systems (MWSCAS).

Jaffrin, L. C. (2021). Applying Ensemble Learning Algorithm in Early Prognosis of Heart Illness. *Turkish Journal of Computer and Mathematics Education (TURCOMAT), 12*(2), 1769-1774.

Jory, C., Shankar, R., Coker, D., McLean, B., Hanna, J., & Newman, C. (2016). Safe and sound? A systematic literature review of seizure detection methods for personal use. *Seizure, 36*, 4-15.

Kumar, S. (2015). Neuromarketing: the new science of advertising. *Universal Journal of Management, 3*(12), 524-531.

Kumari, S., Kumar, D., & Mittal, M. (2021). An ensemble approach for classification and prediction of diabetes mellitus using soft voting classifier. *International Journal of Cognitive Computing in Engineering, 2*, 40-46.

Lakshmi, B., Indumathi, T., & Ravi, N. (2016). A Study on C. 5 decision tree classification algorithm for risk predictions during pregnancy. *Procedia Technology, 24*, 1542-1549.

Luthra, S. (2018) Machine Learning: An Automated Learning Approach. *International Journal of Computer Engineering and Applications,* 12: 156-161.

Machado, J., & Balbinot, A. (2014). Executed movement using EEG signals through a Naive Bayes classifier. *Micromachines, 5*(4), 1082-1105.

Mavrovouniotis, M., & Yang, S. (2015). Training neural networks with ant colony optimization algorithms for pattern classification. *Soft Computing, 19*(6), 1511-1522.

Mehta, A. *Machine Learning*. Digital Vidya, [online]. Available at: https://www.digitalvidya.com/blog/supervised-learning/. [Accessed: 7 April 2021].

Mienye, I. D., Sun, Y., & Wang, Z. (2019). Prediction performance of improved decision tree-based algorithms: a review. *Procedia Manufacturing, 35*, 698-703.

Moubayed, A., Injadat, M., Nassif, A. B., Lutfiyya, H., & Shami, A. (2018). E-learning: Challenges and research opportunities using machine learning & data analytics. *IEEE Access, 6*, 39117-39138.

Muslim, M., Rukmana, S., Sugiharti, E., Prasetiyo, B., & Alimah, S. (2018). *Optimization of C4. 5 algorithm-based particle swarm optimization for breast cancer diagnosis.* Paper presented at the Journal of Physics: Conference Series.

Mutlag, W. K., Ali, S. K., Aydam, Z. M., & Taher, B. H. (2020). *Feature Extraction Methods: A Review*. Paper presented at the Journal of Physics: Conference Series.

Neuroinform, F. (2015) The PREP pipeline: standardized preprocessing for large-scale EEG analysis. [Online]. Available at: https://www.frontiersin.org/articles/10.3389/fninf.2015.00016 /full. (Accessed: 12 June 2021)

Nicolaou, N., & Georgiou, J. (2012). Detection of epileptic electroencephalogram based on permutation entropy and support vector machines. *Expert Systems with Applications, 39*(1), 202-209.

Nikookar, E., & Naderi, E. (2018). Hybrid ensemble framework for heart disease detection and prediction. *IJACSA) International Journal of Advanced Computer Science and Applications, 9*(5).

Nunes, T. M., Coelho, A. L., Lima, C. A., Papa, J. P., & De Albuquerque, V. H. C. (2014). EEG signal classification for epilepsy diagnosis via optimum path forest–A systematic assessment. *Neurocomputing, 136*, 103-123.

Panchal, F. S., & Panchal, M. (2014). Review on methods of selecting number of hidden nodes in artificial neural network. *International Journal of Computer Science and Mobile Computing, 3*(11), 455-464.

Pedrero, V., Reynaldos-Grandón, K., Ureta-Achurra, J., & Cortez-Pinto, E. (2021). Overview of machine learning and its application in the management of emergency services. *Revista Medica de Chile, 149*(2), 248-254.

Pirotti, F., Sunar, F., & Piragnolo, M. (2016). Benchmark Of Machine Learning Methods For Classification Of A Sentinel-2 Image. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences, 41*.

Priya, S. S., & Iqbal, M. H. (2015). Solar radiation prediction using artificial neural network. *International Journal of Computer Applications, 116*(16).

Qi, L. J., & Alias, N. (2018). *Comparison of ANN and SVM for classification of eye movements in EOG signals.* Paper presented at the Journal of Physics: Conference Series.

Rajendran, N. A., & Vincent, D. R. (2021). Heart Disease Prediction System using Ensemble of Machine Learning Algorithms. *Recent Patents on Engineering, 15*(2), 130-139.

Ravish, D., Devi, S. S., Krishnamoo, S., & Karthikeya, M. (2013). Detection of epileptic seizure in EEG recordings by spectral method and statistical analysis. *Journal of Applied Sciences, 13*(2), 207-219.

Rogerson, J. (2015) *Designs and Applied Principles of Artificial Neural Networks. Clanrye International.*

Senthilmurugan, M., Latha, M., & Malmurugan, N. (2011). Classification in EEG-based brain computer interfaces using inverse model. *International Journal of Computer Theory and Engineering, 3*(2), 274.

Shorvon, S., Guerrini, R., Cook, M., & Lhatoo, S. (2012). *Oxford textbook of epilepsy and epileptic seizures*: OUP Oxford.

Shubham. (2019) *Decision Tree.*[Online]. Available at: https://www.hackerearth.com/practice/machine-learning/machine-learning-algorithms/ml-decision-tree/tutorial/ (Accessed: 10 January 2021)

Singh, A. (2018). A Comprehensive Guide to Ensemble Learning. [Online]. Available at: https://www.analyticsvidhya.com/blog/2018/06/comprehensive-guide-for-ensemble-models/ . (Accessed: 7 May 2021).

Siuly, S., Li, Y., & Zhang, Y. (2016). EEG signal analysis and classification. *IEEE Trans Neural Syst Rehabilit Eng, 11*, 141-144.

Sivasankari, K., & Thanushkodi, K. (2014). An improved EEG signal classification using neural network with the consequence of ICA and STFT. *Journal of Electrical Engineering and Technology, 9*(3), 1060-1071.

Song, Y., Crowcroft, J., & Zhang, J. (2012). Automatic epileptic seizure detection in EEGs based on optimized sample entropy and extreme learning machine. *Journal of neuroscience methods, 210*(2), 132-146.

TechVidvan. *Use of Decision Trees and Random Forest in Machine Learning An Insight into Supervised Learning for Classification Problems*. [Online]. Available at: https://techvidvan.com/tutorials/supervised-learning/. [Accessed: 7 April 2021].

Temko, A., Thomas, E., Marnane, W., Lightbody, G., & Boylan, G. (2011). EEG-based neonatal seizure detection with support vector machines. *Clinical Neurophysiology, 122*(3), 464-473.

Tigran Gevorgyan. *Adoption and inclusion of Artificial Intelligence in digitalization strategies of organizations*. Master. Aalborg University; 2019.

Vandecasteele, K., De Cooman, T., Gu, Y., Cleeren, E., Claes, K., Paesschen, W. V., . . . Hunyadi, B. (2017). Automated epileptic seizure detection based on wearable ECG and PPG in a hospital environment. *Sensors, 17*(10), 2338.

Vidyaratne, L. S., & Iftekharuddin, K. M. (2017). Real-time epileptic seizure detection using EEG. *IEEE Transactions on Neural Systems and Rehabilitation Engineering, 25*(11), 2146-2156.

Wen, T., & Zhang, Z. (2017). Effective and extensible feature extraction method using genetic algorithm-based frequency-domain feature search for epileptic EEG multiclassification. *Medicine, 96*(19).

Weng, W.-H. (2020). Machine learning for clinical predictive analytics. In *Leveraging data science for global health* (pp. 199-217): Springer, Cham.

Zhang, Y., Phillips, P., Wang, S., Ji, G., Yang, J., & Wu, J. (2016). Fruit classification by biogeography-based optimization and feedforward neural network. *Expert Systems, 33*(3), 239-253.

Zhou, M., Tian, C., Cao, R., Wang, B., Niu, Y., Hu, T., . . . Xiang, J. (2018). Epileptic seizure detection based on EEG signals and CNN. *Frontiers in neuroinformatics*, 95

LIST OF PUBLICATION

Non-Indexed Conference Proceedings

1. **Aziz, N.S.A.,** & Alias, N. (2020). Artificial Neural Network versus Support Vector Machine for Classifying the Epileptic Seizure by Using Electroencephalography Signals. *Proceedings of Science and Mathematics, Vol. 1, pp.55-65*.

Paper Presented in Conferences

1. Nur Syahirah Binti Abdul Aziz and Norma Alias. Artificial Neural Network versus Support Vector Machine for Classifying the Epileptic Seizure by Using Electroencephalography Signals. 7th International Conference and Workshop on Basic and Applied Sciences (ICOWOBAS 2019). July 16-17, 2019, Johor Bahru.