# The Effect of Different EEG Signal Processing Techniques Applied to Hand Movement BCIs

Luiz Renato Rodrigues Carneiro Faculdade de Engenharia Mecânica Universidade Federal de Uberlândia Uberlândia, Brazil ORCID: 0000-0003-1218-1585 Victor Affonso Albuquerque Pacheco Faculdade de Engenharia Mecânica Universidade Federal de Uberlândia Uberlândia, Brazil ORCID: 0000-0001-6155-331X Alcimar Barbosa Soares Faculdade de Engenharia Elétrica Universidade Federal de Uberlândia Uberlândia, Brazil ORCID: 0000-0003-1100-3533

Abstract— This paper reports the research on the influence of different methods for feature extraction and classification of Electroencephalogram (EEG) data and how they can affect the accuracy of a Brain-Computer Interface (BCI) on interpreting the EEG signal due to Motor Imagery (MI). For this purpose, a BCI was built and programmed to run multiple Feature Extraction methods in time-domain for different ranges of frequencies and two Machine Learning classifiers. It was compared the Linear Discriminant Analysis (LDA) and the Support Vector Machine (SVM) machine learning techniques. It was concluded that EEG frequency bands have a great influence in the performance of the BCI, mainly for Alpha and Beta bands. Furthermore, by utilizing the SVM method, it was possible to identify differences on the performances of the feature extraction methods evidencing higher accuracies for the Willison Amplitude (WAMP) and a combination of WAMP and the Root Mean Square (RMS) techniques. Contrastingly, the LDA demonstrated a higher mean accuracy and lower standard deviation between the subjects, however compared to the SVM, it was not possible to detect meaningful differences between the time domain techniques. Therefore, it was verified that LDA is more robust with greater accuracy, in addition to having lower computational cost compared to SVM for the feature extraction techniques used in this article.

## Keywords — Brain-Computer Interface, Feature Extraction, Classification, Electroencephalogram, Motor Imagery

## I. INTRODUCTION

Individuals who suffer from a spinal cord trauma on the cervical region can lose partially or completely the movement of the superior members. In case the damage is located in between the C1 and the C3 vertebrae, the ability of speaking, breathing, and moving the head can also be compromised. This spine lesion results in biologic tissue loss, including myelinated nerve fibers that are responsible for transmitting motor and sensory impulses [1]. The people affected by this condition experience difficulty to accomplish essential tasks that require using the superior members, such as feeding.

It is estimated that in Brazil, there are 40 new cases of spinal cord trauma per one million people, resulting in 6 to 8 thousand new cases every year [2]. Due to these circumstances, the development and improvement of the control for rehabilitation devices, prosthetics and orthotics have been a major challenge to improve the life quality of people with this type of condition. The control for this kind of technology is usually done by reading and interpreting data captured Electromyography by (EMG), Electroencephalography (EEG) or Electroretinography (ERG) [3-4]. Therefore, the EEG enables the use of a Brain-Computer Interface (BCI) to create a new communication pathway from the brain to assistive devices that could help a person with physical disabilities [5].

The state of the art for a BCI would be to classify the intention of movement and motor control of fine movements such as for the fingers of the hand. However, to achieve this level of result it is necessary to better understand the classification of hand movements [6]. In this way, several different approaches can be applied for the signal processing of a BCI involving a group of preprocessing, feature extraction and classification techniques, thus it is relevant to evaluate among a set of techniques which one would bring the best results. Hence, the goal of this research is to investigate the hypothesis that different methods of feature extraction in time domain as the Mean Absolute Value (MAV). Line Length (LL), Nonlinear energy (NE), Root Mean Square (RMS), Willison Amplitude (WAMP) and a combination of WAMP and RMS (WAMP+RMS) calculated to different EEG bands, combined with different types of classification as Support Vector Machine (SVM) and Linear Discriminant Analysis (LDA) may affect the accuracy of the results in order to improve the performance of the Brain Computer Interfaces applied to hand movement. The software employed in this research were entirely designed for this specific purpose and the signals analyzed are the ones generated from Motor Imagery (MI) for the simple task of moving the right or the left hand. Additionally, the development of a hardware for EEG was also exemplified in the form of an electric circuit to better understand how a BCI works.

## II. MATERIAL AND METHODS

## A. The Physiology of the Electroencephalogram

The EEG captured the neurophysiologic signals by means of electrodes attached to the scalp according to the international 10-20 system. Therefore, three monopolar channels located at CZ, C3 and C4 were utilized and the data obtained from these channels were then compared to the fourth electrode A1, which is the reference electrode. There is also a fifth electrode for ground that can be placed at Fz position. The signal captured by these electrodes descend mostly from the activity of the primary motor area (M1) within the brain cortex.

It is known that the primary motor area is topographically organized in a form called Somatotopic Arrangement in which each part of the body is controlled by a specific part of the M1. Additionally, the body parts that require finer and more accurate movements such as hands and pectoral region are controlled by larger regions of the cortex [1-4].

# B. The Brain-Computer Interface

The BCI can be usually divided into many stages, such as signal acquisition, signal processing (processing, feature extraction and classification) and application interface [7,8].



Fig. 1. Brain-Computer Interface Diagram with differents techniques of signal processing.

The specific elements used to build the BCI for this research can be shown in Fig. 1.

The signal acquisition stage is composed of the electrodes that capture the brain signal and the hardware which is made of amplification and filtering circuits. Moreover, there is the preprocessing task, which aims to attenuate the artifacts through frequency domain filtering and attenuate the EOG artifacts. After that, it is necessary to extract the signal features, and in this work were utilized the MAV, LL, NE, RMS, WAMP and a combination of WAMP and RMS WAMP+RMS. Then, these features were classified by two different algorithms, SVM and LDA. Finally, the application interface was responsible to receive the prediction of the subject's brain intention to drive an assistive device or a computer. All these stages will be better explained in the next sections.

## C. The Hardware

The hardware to be exemplified is composed of a Printed Circuit Board (PCB) to make the signal acquisition. The first stage of the PCB is an instrumentation amplifier which will make the subtraction between the C3 or Cz or C4 and the A1 electrode (depending on the channel). This subtraction reduces the common mode artifacts mainly due to electromagnetic interference. Furthermore, the instrumentation amplifier makes a pre-implication in the signal.

The next stage is to reduce the Direct Current (DC) level of the signal. It was achieved by using a passive high pass filter with a low cutoff frequency. To amplify the signal is necessary an operational amplifier. After that, another passive high pass filter is necessary. The last stage of amplification is done by a variable gain amplifier controlled by a microcontroller.

In order to read only the frequency range of the EEG signal, two 3-Pole Butther Active Filters were made to build a band pass filter with a range of frequency above 0.5 Hz and below 100 Hz. To digitize the signal, a 16 bits A/D converter was used with a sample frequency of 250 Hz. This digitized signal is read by the microcontroller and transferred to the computer for the Signal Processing stages.

Moreover, the ground of the circuit is connected to the A1 electrode, and each channel C3, Cz and C4 has its own circuit which starts in the instrumentation amplifier and ends in the

3-Pole Butterworth Active High Pass filter. For more details of the circuit [9]. In addition, a block diagram of the circuit is shown in Fig. 2.

#### D. Preprocessing

Preprocessing is designed to lower the noise ratio that can be considerably high for EEG by consequence of low amplitudes that make the signal susceptible to artifacts generated by other sources such as eyes and head muscles movements [8]. The signal captured by the EEG appears to be stochastic and its amplitude range is in the order of microvolts for electrodes positioned at the scalp. In addition, the relevant frequency interval for Electroencephalography can be divided into bands as shown by Table 1 [10, 11].

Additionally, the dataset utilized was the 2b Motor Imagery Dataset from the BCI competition IV [12], that provides the signals for C3, C4 and CZ and also the signal for the Electrooculography (EOG) that can be useful to reduce artifacts generated by the blinking [13]. The data was collected in a section in which each individual was asked to execute tasks such as moving, closing and keeping their eyes open. In this way, the signal obtained can be used to estimate the linear regression coefficients that relate the EOG to the EEG patterns [14], enabling to estimate the EEG signal with lower noise due to eye artifacts.

Another preprocessing technique employed was a pass band digital filter implemented by Fast Fourier Transform (FFT) [8] that can divide the signal into a discrete frequency domain. This way is possible to separate only the frequency ranges that are useful to detect the intention to move the hand. According to the literature, the frequency range that is related the most with motor execution is 16 to 22 Hz in the Beta band while the one most related to the motor imagery is 10 to 14 Hz in the Alpha/Mu range [15]. About the region of activation, in both cases the primary motor area is the part of the brain that generates signals of activity [15].

TABLE I. AMPLITUDE AND FREQUENCY OF EEG BANDS

Bands	Frequencies (Hz)	Amplitudes ( $\mu V$ )
Delta (δ)	0.5-4	20-200
Theta $(\theta)$	4-8	<20
Alpha (α)	8-14	2-100
Beta (β)	14-30	5-10
Gamma (γ)	30-100	5-10



Fig. 2. Block diagram of EEG acquisition circuit.

Thus, the FFT was utilized to divide the signal into the five EEG bands as shown by Table 1. Furthermore, two sub bands were created inside the alpha band and four sub bands inside the beta band to increase the number of features extracted and improve the input for the classifier.

## E. Feature Extraction

Feature extraction is one of the most important parts in signal processing. This stage aims to extract the relevant information in the EEG signal in order to describe the mental states while rejecting the non-relevant information and noise [16], moreover it makes a dimension reduction of data for the classification stage [17]. There are many feature extraction techniques employed for different applications. They are grouped in time domain, frequency domain, time-frequency domain, non-linear parameters, spatial filter and many others [7,8].

In this paper the time domain techniques were utilized due to the simplicity and similarity between them. Most complex techniques such as Common Spatial Pattern (CSP) which is a spatial filter give better results but demand a higher computational power and is more recommended for multichannel EEG [18], which is not the case for this research.

Hence, the feature extractions compared in this paper are Root Mean Square (RMS) [5,17], Mean Absolute Value (MAV) [5, 18], Willison Amplitude (WAMP) [5, 17], Line Length (LL) [17, 18] and Nonlinear energy (NE) [18]. All these techniques were applied in a window of 2 seconds, the recording starts 0,5 s before the MI is requested to the individual and stops 1.5 seconds later. The reason why these characteristics were chosen is the simplicity of their calculations, as well as their similarity due to the fact that all of them are time domains, making their comparison fairer, in addition to having a close computational cost.

These five techniques of extraction features were calculated for each EEG band (Delta, Theta, Alpha, Beta, Gamma), for the whole/raw signal (between 0.5 Hz and 100 Hz) and for the six sub bands of Alpha and Beta. Where N is the size of the time window, x is the amplitude of the signal in respective time.

$$RMS = \sqrt{\frac{1}{N} \sum_{i=1}^{N} x_i^2}$$
(1)

$$WAMP = \sum_{i=1}^{N-1} f(|x_i - x_{i+1}|) \to f(x) = \begin{cases} 1 & x > \varepsilon \\ 0 & otherwise \end{cases}$$
(2)

$$NE = \sum_{i=1}^{N-2} (x^{2}[i] - x[i+1]x[i-1])$$
(3)

$$MAV = \frac{1}{N} \sum_{i=1}^{N} |x_i| \tag{4}$$

$$LL = \sum_{i=1}^{N-1} |x[i] - x[i-1]|$$
(5)

# F. Classification

Classification stage is intended to convert the features extracted from the signals into brain activity patterns [8] that, for this study, are the intention to move the right or the left hand. The classification methods employed were the Linear Discriminant Analysis (LDA) and the Support Vector Machine (SVM) [8]. About the differences of these approaches, the LDA can only classify patterns by applying linear separability, which requires lower computational power, the SVM can alternatively classify patterns that are not suitable for linear separability by means of the kernel function. These machine learning techniques were chosen due to their simplicity, thus reducing the computational cost and making the BCI easier to be embedded in a microcontroller, besides making it more suitable for real-time applications. Another technique that could be used in the comparison is neural networks due to their better ability to classify nonlinearly separable classes [8], however this technique was not utilized because their greater complexity also increases the computational cost.

The reason Scikit-learn library was used to classify the features instead of an original code implementation is the fact that it is a reliable open-source framework and widely employed in python language for machine learning techniques, as SVM and LDA. The specific kernel function employed for the SVM method was the Radial Basis Function (RBF) and the solver for the LDA was the Least Square Solution. Combining the Machine Learning techniques together with the different feature extractions, this research intends to compare the variation of the accuracy on the results. For training and testing of Artificial Intelligence (AI), the Dataset was destined 70 % for training data and 30 % for testing data.

#### III. RESULTS

To analyze the results of these distinct feature extraction techniques, each one of them were combined with different types of frequency bands, and then tested by the two classifiers. For this task, the EEG signal collected from a group of nine subjects was utilized to train the classifiers for each one of the combinations. An example of the result of one of these trainings is shown in Fig. 4. In order to have more accurate results, each model was calculated twenty times for each subject and the mean accuracy was used, as done in other studies [17]. Moreover, the test protocol performed in this research were similar to those performed by other authors to compare different processing techniques [5, 15, 17, 19].

The mean accuracy and standard deviation of accuracy of each model composed by 9 subjects were calculated, and their result is shown Fig. 3. In this graphic each color indicates a different feature extraction in the time domain. Furthermore, this graphic also indicates on the Y axis the group of frequency bands utilized in the classification.

The "Raw" indicates the data without any pass-band filter, " $\delta\theta\alpha\beta\gamma$ " indicates the signal divided into Delta, Theta, Alpha, Beta and Gamma bands, " $\alpha\beta$ " indicates the signal divided only into Alpha and Beta bands, and " $\alpha\beta$ \*" indicates the signal divided into two sub bands of Alpha frequencies and four sub bands of Beta frequencies, resulting in a total of six sub bands.

By analyzing Fig. 3 it is evident the lower accuracy of Raw data compared with the other groups of bands. Furthermore,



Fig. 3. (a) Accuracy of LDA. (b) Standard deviation of accuracy of LDA. (c) Accuracy of SVM. (d) Standard deviation of accuracy of SVM. Each graphic was made with all trained models, composed by different conditions of a group of bands and six different feature extraction.

the  $\alpha\beta^*$  sub bands have an average of greater accuracy than  $\alpha\beta$  which in turn is greater than  $\delta\theta\alpha\beta\gamma$ .

Moreover, the accuracy of LDA was greater than SVM for most cases, and the standard deviation was lower. Besides that, SVM had a bigger difference of accuracy between the feature extraction techniques, indicating the relevance of choosing the technique when using SVM.

In the Fig. 4 it is possible to visualize the accuracy of each subject with  $\alpha\beta^*$ , applying EMS feature extraction and LDA classifier. From this graphic it is possible to observe a greater accuracy and lower standard deviation for the subjects four and five, and lower accuracy and higher standard deviation for the two, three, seven and eight subjects.

#### IV. DISCUSSION

In this work the result shows that the range of frequencies, the feature extraction method and the algorithm of classification affect the performance of a BCI. It also explains how these variables impact the accuracy of the Interface.

The low accuracy of Raw signal is coherent because it has many other frequencies that are not relevant for the hand movement classification. Furthermore, the EEG bands are associated with different mental tasks [15, 20], hence dividing the signal into these bands to extract the features of each band



Fig. 4. Accuracy of the model composed by alpha and beta subbands, RMS feature extraction and LDA classifier.

separately makes the classification easier to the classifier. The use of only Alpha and Beta bands gave a slightly better result, because they are associated with motor imagery and motor execution respectively [15], and the dataset used is a set of IM signals. Therefore, using only these two groups of range of signals reduces the amount of irrelevant data, which can enhance the performance of the BCI [7, 22]. Moreover, the improvement of accuracy by using sub bands inside Alpha and Beta allows extracting more information inside the same range of frequency which may slightly increase the accuracy, but also increase the number of extracted features and the dimensionality, which makes the classification problem more complex [7, 22].

The use of the extracting feature is crucial to make dimensionality reduction of the data [7, 18, 19, 22], because in this case it used only one extracted feature for each selected band, except for WAMP+RMS which used two features. These time domain techniques show in Fig. 3 a great impact in the accuracy for SVM classifiers, WAMP and WAMP+RMS gave better results in most cases, other authors have reported better results for WAMP method when using SVM [17, 5]. However, for LDA the extracted feature did not seem to influence considerably in the accuracy, so it seems more robust, with mean accuracy greater and standard deviation was lower than SVM, indicating that LDA can be more suitable when the chosen features are used. Nevertheless, LDA is not recommended for more complex classifications where linear separation is not possible. In this case, the SVM can lead to better accuracy and better generalization properties [8, 17]. For more complex classifications, the Artificial Neural Networks (ANN) can achieve better results [8], however, the better accuracy of the LDA compared to SVM indicates that extraction feature results in a linearly separable problem, and because of that, ANN would probably not improve considerably the accuracy of the model as occurred in another test [23].

Moreover, the high difference of accuracy that can be seen between the subjects shown in Fig. 4 causes a high standard deviation of accuracy illustrated by Fig. 4(b, d). This variation of accuracy between the subjects can be explained by several variables. The employment of scalp electrodes generates signals 20 to 100 times worse in terms of quality in comparison to invasive electrodes [21]. Besides that, depending on how the subjects perform the IM, different signal patterns may be generated [15]. Another factor that might generate the high standard deviation on results is the different number of samples for each subject, as the dataset utilized provides distinct amounts of signals for the individuals. Additionally, the machine learning techniques could have a better performance with a larger dataset because the capability of learning and generalizing the problem of the LDA and SVM is qualitatively proportional to the amount of data employed [24].

#### V. CONCLUSION

In this paper, we expose the effects of using different signal processing techniques to classify motor imagery EEG data. The range of the utilized frequencies may improve the accuracy of the BCI, the employment of only Alpha and Beta bands can provide good results with less information, which can reduce computational complexity. Furthermore, the timedomain feature extraction selected may have different performances depending on the classifier. For the SVM classifier, WAMP, RMW and MAV gave better results, however for LDA all tested feature extractions had similar results. Moreover, for the condition tested in this research, LDA had on average better results than SVM. In addition, LDA needs less computational requirements and is simpler to use compared to SVM [8,16]. All this information can be useful to improve the performance of BCI, by having greater accuracy with a lower computational demand.

For future studies, the results of this article about timedomain feature extractions can be useful to compare the accuracy and computational cost in a future research about techniques in frequency domain and time-frequency domain.

## ACKNOWLEDGMENT

The authors are grateful to R. Leeb, C. Brunner, G. R. Muller-Putz, A. Schlogl and G. Pfurtscheller of Graz University of Technology for making the dataset used available in BCI Competition IV.

#### REFERENCES

- [1] Guyton, A.C.; Hall, J.E. Tratado de Fisiologia Médica, 12th ed. Rio de Janeiro: Elsevier, 2011, pp 693-717.
- [2] Ministério da Saúde, "Diretrizes de Atenção à Pessoa com Lesão Medular", Brasília - Distrito Federal, BR, 2013. Available: https://bvsms.saude.gov.br/bvs/publicacoes/diretrizes\_atencao\_pessoa \_lesao\_medular.pdf. Acesses on: 07/28/2021.
- [3] Webster, J. G.; Beebe, D. Encyclopedia of Medical Devices and Instrumentation., 2nd ed., vol. 6. John Wiley & Sons, 2006.
- [4] Lent, R. Cem Bilhões de Neurônios: Conceitos Fundamentais de Neurociência. Rio de Janeiro: Atheneu, 2004, pp 446-743.
- [5] Hamedi, M., Salleh, S. H., Noor, A. M., & Mohammad-Rezazadeh, I (2014) Neural network-based three-class motor imagery classification using time-domain features for BCI applications, IEEE TENSYMP, Dhaka, Bangladesh, 2014, IEEE Region 10 Symposium, 2014, pp. 204–207. https://doi.org/10.1109/tenconspring.2014.6863026.
- [6] Millán, J. D. R. et al. (2010) Combining brain-computer interfaces and assistive technologies: State-of-the-art and challenges, Frontiers in Neuroscience 4, Singapore, 2010. https://doi.org/10.3389/fnins.2010.00161.
- [7] Aboalayon, Khald A.I., Miad Faezipour, Wafaa S. Almuhammadi, and Saeid Moslehpour. 2016. "Sleep Stage Classification Using EEG Signal Analysis: A Comprehensive Survey and New

Investigation" *Entropy* 18, no. https://doi.org/10.3390/e18090272.

272.

9.

- [8] Aggarwal, S., & Chugh, N. "Signal processing techniques for motor imagery brain computer interface: A review". Array, vol. 1–2, January– April 2019. https://doi.org/10.1016/j.array.2019.100003.
- [9] Webster, J. G., Medical Instrumentation Application and Design, 4th ed., John Wiley & Sons 2009, pp.91–291.
- [10] Aboalayon, K.A.; Faezipour, M (2014) Multi-class SVM based on sleep stage identification using EEG signal, Proceedings of the IEEE Healthcare Innovation Conference (HIC), Seattle, WA, USA, 8–10 October 2014, pp. 181–184. https://doi.org/10.1109/HIC.2014.7038904.
- [11] Bahri, Z., Abdulaal, S., & Buallay, M. (2014) Sub-band-power-based efficient Brain Computer Interface for wheelchair control, 2014 World Symposium on Computer Applications & Research (WSCAR). https://doi.org/10.1109/wscar.2014.6916840.
- [12] Leeb, R., Brunner, C.,Muller-Putz, G., Schlogl, A.,Pfurtscheller, G., "BCI Competition 2008 – Graz data set B", Available: bbci.de/competition/iv/desc\_2b.pdf. Access on: 22/04/2021.
- [13] Abo-Zahhad, M., Ahmed, S. M., Abbas, S. N., "A New EEG Acquisition Protocol for Biometric Identification Using Eye Blinking Signals" International Journal of Intelligent Systems Technologies and Applications, May 2015, pp. 48-54. https://doi.org/10.5815/ijisa.2015.06.05.
- [14] Lotte, F. (2014) A Tutorial on EEG Signal-processing Techniques for Mental-state Recognition in Brain–Computer Interfaces. In: Miranda E., Castet J. (eds) Guide to Brain-Computer Music Interfacing. Springer, London. https://doi.org/10.1007/978-1-4471-6584-2\_7.
- [15] Selim S., Tantawi M., Shedeed H., Badr A. (2017) Reducing Execution Time for Real-Time Motor Imagery Based BCI Systems. In: Hassanien A., Shaalan K., Gaber T., Azar A., Tolba M. (eds) Proceedings of the International Conference on Advanced Intelligent Systems and Informatics 2016. AISI 2016. Advances in Intelligent Systems and Computing, vol 533. Springer, Cham. https://doi.org/10.1007/978-3-319-48308-5\_53.
- [16] Blankertz, B. et al. "Optimizing Spatial filters for Robust EEG Single-Trial Analysis." IEEE Signal Processing Magazine 25, pp. 41-56, 2008.
- [17] Khorshidtalab, A., Salami, M. J. E., & Hamedi, M. (2012) Evaluation of time-domain features for motor imagery movements using FCM and SVM, JCSSE 2012 - 9th International Joint Conference on Computer Science and Software Engineering, Bangkok, Thailand, 2012, pp. 17– 22. https://doi.org/10.1109/JCSSE.2012.6261918.
- [18] Boonyakitanont, P., Lek-uthai, A., Chomtho, K., & Songsiri, J. "A review of feature extraction and performance evaluation in epileptic seizure detection using EEG", Biomedical Signal Processing and Control. Elsevier Ltd, vol. 57, March 2020. https://doi.org/10.1016/j.bspc.2019.101702.
- [19] Islam, M., & Ahmad, M. (2016) Modeling of human emotion with effective frequency band during a test of sustained mental task, IEEE International WIE Conference on Electrical and Computer Engineering, WIECON-ECE, Dhaka, Bangladesh, 2015, pp. 403–406. https://doi.org/10.1109/WIECON-ECE.2015.7443951.
- [20] Bashashati, A., Fatourechi, M., Ward, R. K., & Birch, G. E, "A survey of signal processing algorithms in brain-computer interfaces based on electrical brain signals" Journal of Neural Engineering, vol. 4, 2007. https://doi.org/10.1088/1741-2560/4/2/R03.
- [21] Lotte, F. "A Tutorial on EEG Signal Processing Techniques for Mental State Recognition in Brain-Computer Interfaces", Eduardo Reck Miranda; Julien Castet, Guide to Brain-Computer Music Interfacing, Springer, 2014.
- [22] Ball, T., Kern, M., Mutschler, I., A. Aertsen, A. Schulze-Bonhage, "Signal quality of simultaneously recorded invasive and non-invasive EEG", NeuroImage, vol. 46, pp. 708–716, July 2009. https://doi.org/10.1016/j.neuroimage.2009.02.028.
- [23] Bozhkov, L., Georgieva, P., & Trifonov, R. (2014) Brain Neural Data Analysis Using Machine Learning Feature Selection and Classification Methods, In Communications in Computer and Information Science, vol. 459 CCIS, pp. 123–132. Springer Verlag. https://doi.org/10.1007/978-3-319-11071-4\_12.
- [24] Ajiboye, A. R., Abdullah-Arshah, R., Qin, H., Isah-Kebbe, H., "Evaluating The Effect of Dataset Size on Predictive Model Using Supervised Learning Technique", International Journal of Computer Systems & Software Engineering, vol. 1, pp. 75–84, February 2015.