

# Feature Weighting for Parkinson's Identification using Single Hidden Layer Neural Network

**SALWA KHALID ABDULATEEF<sup>1</sup>, AHMED NASER ISMAEL<sup>2</sup>, MOHANAD DAWOOD SALMAN<sup>1</sup>**

<sup>1</sup>Department of computer science, Computer Science and Mathematics College. Tikrit University. Iraq

<sup>2</sup>Department of management information systems, Administration and Economics College. Basrah University. Iraq

Corresponding author: Salwa Khalid Abdulateef (e-mail: khalid.salwa@tu.edu.iq)

**ABSTRACT** The diagnosis of Parkinson has become easier with the existence of machine learning. It includes using existing features from the biometric dataset generated by the person to identify whether he has Parkinson or not. The features differ in their discrimination capability and they suffer from redundancy. Hence, researchers have recommended using feature selection for Parkinson's identification. The feature selection aims at finding the most important and relevant features to produce an efficient and effective model. In this article, we present entropy-based Parkinson classification. The goal is to select only 50% of the most relevant features for Parkinson prediction. Two variants of neural networks are used for evaluation, the first one is a feed-forward Extreme Learning Machine ELM and the second one is Fast Learning Machine FLN. Also, the K-Nearest Neighbor KNN algorithm is used for evaluation. The results show the superiority of ELM and FLN when the model of feature selection is used with an accuracy of 80% compared with only 78% when the model is not used.

**KEYWORDS** Parkinson; feature selection; fast learning machine; feature weighting; extreme learning machine; identification.

## I. INTRODUCTION

PARKINSON is regarded as one of the major diseases that affects the population with a percentage of 2-3% for people over 65 and older [1] as it is provided by Parkinson Disease (PD) foundation, about 7-10 million people worldwide suffer from Parkinson's. The reason for Parkinson is the depletion of dopaminergic nigrostriatal neurons [2, 3]. It affects not only the articulators but also the voice and speech of the users. Some of the symptoms of Parkinson disease are tremors, gestures loss while talking or communicating, non-capability of doing fast movements, independent joint control, dragging in walking or small steps, and others. Also, it leads to effects on language, cognition, and mood. This disease is classified as neurodegenerative that is caused by a genetic mutation. Diagnosis of this disease is essential to avoid its major development. The voice signal is the most important element for doing early Parkinson diagnostic [4].

The diagnosis of Parkinson disease is becoming easier with the existence of Neural Network (NN) models that are trained on existing datasets of the disease [5-7]. The datasets can be built from various types of biometric signals recorded from the patients. Considering that voice signals play an important role in the discrimination of the diagnostic disease [8-10], it is

possible to build models for training on this data and using it for prediction. The features differ in their level of discrimination and they have an issue of redundancy. Hence, selecting the most powerful features is more effective for building an accurate model and assuring more efficiency in terms of computation. The research of feature selection has been applied in many fields for the goal of identification or classification such as Intrusion Detection System (IDS) [11], ear recognition [12, 13], face recognition [14-16], and gait classification [17,18]. In the area of Parkinson disease, applying feature selection for increasing accuracy is promising.

Single Hidden Layer Feed-Forward Neural Network (SLFN) is a neural network combined with an input layer, output layer, and single hidden layer. At the output of the hidden layer, various types of activation functions can be used. This NN has two forms, namely, a simple form without connections between the input and output layers and a parallel form with connections between the input and the output layers [19]. The former is trained using Extreme Learning Machine (ELM) [20-23] and the latter is trained using Fast Learning Machine (FLN) [24, 25]. The two training approaches are more effective than the classical Back-Propagation (BP) training algorithm that uses the error gradient.

The goal of this article is to explore two classification models with the assistance of feature weighting and selection for Parkinson disease identification. The approach is based on a new method for using entropy for feature selection and testing two models: the first model is based on using an extreme learning machine with whole/half features, the second model is based on using fast learning and KNN with whole/half features. Below is an organization of the remaining text. We outline the literature review in Section 2. The methodology is then presented in Section 3. The experimental strategy and outcomes are then presented in Section 4. Finally, Section 5 provides the conclusion and future work.

## II. RELATED WORK

In [26], a non-invasive sensing approach for Parkinson disease was proposed based on gait analysis. The approach is based on the wavelet transform of spatiotemporal gait variables. For features, the algorithm uses computational simplified features such as minimum and maximum values, mean, variance, and energy variables. Furthermore, the approach considers the evaluation of various gait parameters. For classification, the approach uses a support vector machine. However, the approach does not incorporate feature selection. In some approaches genetic optimization of the number of neurons and the parameters of Wavelet Kernel Extreme Learning Machine (WKELM) were considered [27]. Comparing this work with an earlier work that had considered the features weighting method called Subtractive Clustering Features Weighting (SCFW) showed an added performance [28]. Using voice patterns for the diagnosis of Parkinson disease was considered in [29]. The approach is based on eight different pattern ranking methods using a support vector machine. Bayesian optimization technique was used for optimizing the radial basis function of the support vector machine. Some researchers focused on developing a hybrid algorithm for Parkinson disease. For example, Multi-Layer Perceptron (MLP) was integrated with feature selection and their moduli for feature ranking [30].

In [31], the authors used the ensemble methods, categorical boosting method extreme gradient boosting, and random forest to select the most discriminating features for identifying PD. For the best results in PD prediction, the impact of these factors at various thresholds was investigated.

In the study of Yuvaraj et al. [32], Electroencephalography (EEG) signals were used and the usage of Higher-Order Spectra (HOS) for the diagnosis of Parkinson Disease was investigated. The obtained features were sorted and ranked using the  $t$  value, and highly ranked features were selected. The latter is a single value, which can distinguish the two classes. As well, the ranked features were fed one by one to the multiple classifiers, namely Support Vector Machine (SVM), Naive Bayes (NB), K-Nearest Neighbor (KNN), Probabilistic Neural Network, and Decision Tree (DT).

Overall, feature selection using meta-heuristic for Parkinson identification has not been tackled yet in the literature. This article proposes an entropy-based approach for selecting the most important features of Parkinson diagnostics.

## III. RESEARCH METHOD

This section presents the developed methodology for Parkinson identification using optimized features selection and classifier structure. We provide the framework in sub-section A. Next, the data description is presented in sub-section B. Afterwards, we provide extracted features in subsection C. The classifiers

ELM and FLN are explained in subsection D. Next, the feature weighting and selection is in subsection E. The evaluation metrics are presented in subsection F.

### A. FRAMEWORK

The framework of Parkinson identification is provided in Fig. 1. The received data are fed into a feature reduction block which reduces the feature and projects it from its original space to reduced space. Next, the reduced space for important features is provided to the classifier.

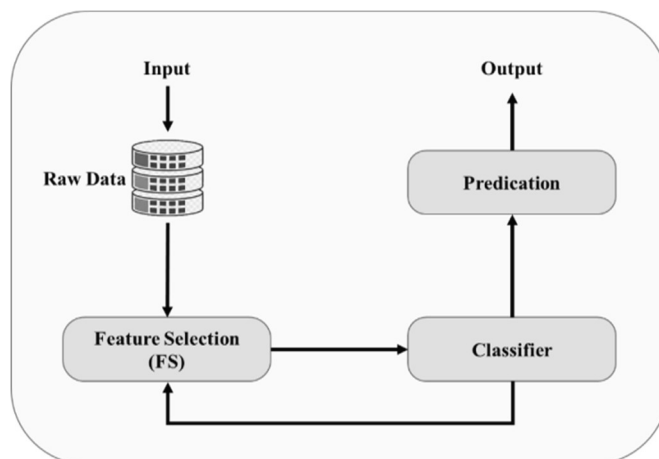


Figure 1. A framework for feature weighting and selection for Parkinson classification.

### B. DATA DESCRIPTION

A dataset for Parkinson disease created by the University of Oxford in collaboration with the national center for voice and speech [33] is used in the article. The data includes 31 people with 23 only having PD. Each column corresponds to one particular voice measure. In addition, the labeling of the data has one of two values: 0 for healthy and 1 for non-healthy.

### C. EXTRACTED FEATURES

Extracted several features from the datasets, including:

- 1- Fundamental frequency (F0) and its variability (standard deviation of F0);
- 2- Jitter and shimmer (measures of frequency and amplitude variability);
- 3- Harmonics-to-noise ratio (HNR) which measures the ratio of harmonic to non-harmonic components in the voice signal;
- 4- Spectral tilt, which measures the slope of the spectral envelope. The obtained features were sorted ranking using the  $t$  value, and highly ranked features were selected. The obtained features were sorted ranking using the  $t$  value, and highly ranked features were selected. The obtained features were sorted ranking using the  $t$  value, and highly ranked features were selected.
- 5- Formant frequencies, which represent the resonant frequencies of the vocal tract;
- 6- Pause and voice onset time (VOT), measures the duration of silence before speaking and the duration

between the onset of voicing and the release of a plosive sound.

All these features were calculated from the digitally normalized voice signals recorded from the subjects in an IAC sound-treated booth using a head-mounted microphone (AKG C420) positioned at 8 cm from the lips, and processed using CSL 4300B hardware (Kay Elemetrics).

#### D. EXTREME LEARNING MACHINE

Extreme Learning Machine (ELM) is a machine learning algorithm used for classification, regression, and feature learning tasks. ELM is a type of feedforward neural network with a single hidden layer, and its learning is based on randomly generated input weights and biases. ELM is a popular algorithm in the machine learning community due to its simplicity, fast learning speed, and high accuracy.

The ELM algorithm can be described in three main steps: input layer, hidden layer, and output layer. In the input layer, data is fed into the network. In the hidden layer, the weights and biases of the input data are randomly generated. The output of the hidden layer is then fed into the output layer, which produces the final output. The hidden layer of the ELM is responsible for the non-linear mapping of the input data into a higher-dimensional space, where the linear regression is performed.

One of the most significant advantages of ELM is its fast learning speed. Unlike traditional neural networks, where the weights and biases are adjusted through iterative training, ELM randomly generates input weights and biases and trains the output layer using a linear regression approach. This means that ELM requires fewer iterations to converge and can handle large datasets with high computational efficiency.

Another significant advantage of ELM is its ability to avoid overfitting. Overfitting occurs when a machine learning model is too complex and is tailored too closely to the training data, leading to poor performance on new data. ELM overcomes this problem by generating random input weights and biases, which increases the diversity of the model and reduces the likelihood of overfitting.

ELM has been successfully applied in various fields, including computer vision, speech recognition, and finance. ELM has also been used in medical diagnosis, where it has been shown to have high accuracy in the classification of diseases such as breast cancer, diabetes, and heart disease [34, 35].

Fast learning machine (FLN) is a new variant of extreme learning machine with weights that connect the input and output layer which enables connecting the input and output layer in a single hidden layer neural network that allows the network to learn complex nonlinear relationships between the input and output variables. This is because the hidden layer serves as a sort of "feature extractor", transforming the input data into a new representation that is better suited for predicting the output.

In particular, the connections between the input layer and the hidden layer allow the network to learn important features of the input data that are relevant to the prediction task, while the connections between the hidden layer and the output layer allow the network to use these features to make accurate predictions. The benefit of this architecture is that it can capture highly nonlinear relationships between the input and output variables, which may not be possible with simpler models like linear regression [36, 37]. The process of calculating the weights or training is as follows:

1. Calculate the weight in the input-hidden layer randomly.
2. Calculate the hidden output matrix using the activation functions at the hidden layer and the input data.
3. Calculate the parallel weights using the Moore-Penrose model.

#### E. FEATURE WEIGHTING and SELECTION

In feature weighting, each feature will be associated with weight. The weight is calculated based on the entropy of the feature. The entropy is calculated in a formula (1) and is used for weighting the features. The features that have more entropy are prioritized over the features that have less entropy. This leads to the concept of half feature classifier, we mean half feature classifier, that classifier will work on the most important 50% of features instead of working on the whole features. Hence, the most important 50% of features are selected. This implies avoidance of over-fitting behavior and more accuracy.

$$E = \sum_{i=1}^N -p_i \log(p_i), \quad (1)$$

where:  $i$  denotes an index of the sample,  
 $N$  denotes the number of samples,  
 $P$  denotes the probability of occurrence of a random sample.

Then the features are ordered in descending manner based on the entropy and the most important half of the features are selected.

#### F. EVALUATION

For evaluation, the following metrics are generated as in equations (2, 3, 4, 5, 6).

$$Precision (\%) = (TP / (TP + FP)) * 100\%, \quad (2)$$

$$Recall (\%) = (TP / (TP + FN)) * 100\%, \quad (3)$$

$$F\text{-measure}(\%) = 2 \left( \frac{Recall * Precision}{Precision + Recall} \right) * 100\%, \quad (4)$$

$$Accuracy(\%) = \left( \frac{TP + TN}{TP + TN + FP + FN} \right) * 100\%, \quad (5)$$

$$G\text{-mean} = \sqrt{(Precision * Recall)}, \quad (6)$$

where TP (True Positive) indicates the correct prediction for positive, TN (True Negative) indicates the number of correct predictions for negative. FN (False Negative) indicates some false predictions for negative, and FP (False Positive) indicates several false predictions for positive.

#### IV. RESULTS AND ANALYSIS

The evaluation has been conducted based on four activation functions, namely, hardlim, sigmoid, sin, and trib. The reason for selecting these activation functions is their common usage in ELM classification. Observing Fig. 2, we find that FLN-half has provided higher performance metrics in terms of accuracy, recall, and G-mean. However, we find that ELM-half has

produced higher precision which indicates the bias toward negative prediction which has been the reason for elevating its precision. For more confirmation, we have included KNN-half and KNN-full which corresponds to the k-nearest neighbor based on half features and full features set respectively. The results reveal that KNN-half has outperformed slightly KNN-full which confirms the effectiveness of our features weighting and selecting.

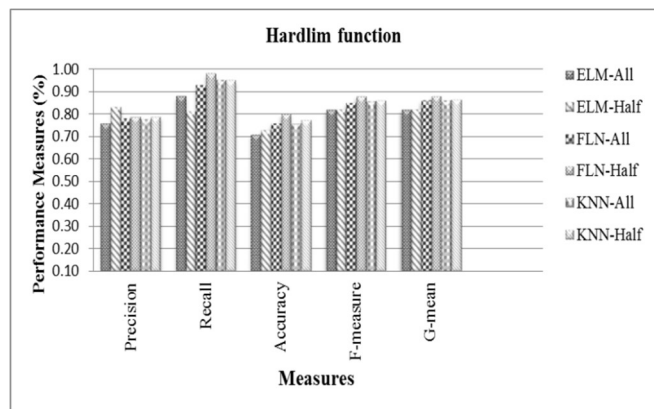


Figure 2. Classification measures for evaluation of Parkinson based on ELM, KNN, and FLN for both full and half features using Hardlim

The same behavior that is monitored in hardlim is also monitored in the sigmoid activation functions. However, we find in Fig. 3, that ELM-half has increased its performance compared with the hardlim activation function. This shows the role of deselecting the non-relevant features in increasing the performance of the classifier.

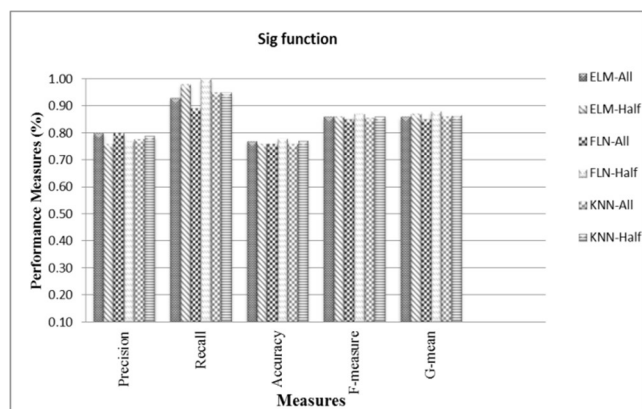


Figure 3. Classification measures for evaluation of Parkinson based on ELM, KNN, and FLN for both full and half features using sig

Similar to the performance that is monitored for hardlim and sigmoid is also monitored for sin where FLN-half has provided the highest accuracy, recall, F-measure, and G-mean in Fig. 4. We also observed that its precision was lower because of some false positive predictions.

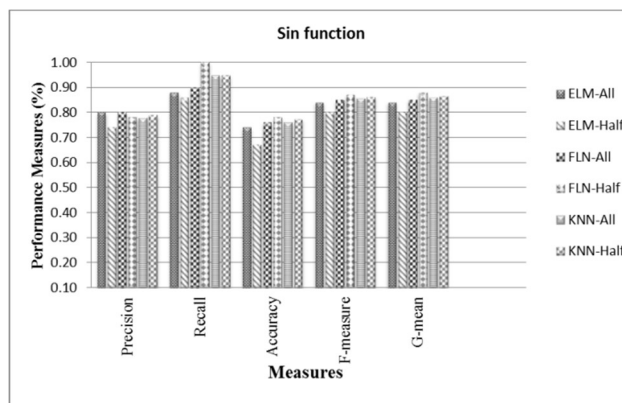


Figure 4. Classification measures for evaluation of Parkinson based on ELM, KNN, and FLN for both full and half features using sin

The only case where ELM-half has outperformed FLN half is the case of the tribis function depicted in Fig. 5. This is interpreted by the effect of the activation function on the learning capability of the classifier. Our finding is that sigmoid, hardlim, and sin are more suitable for FLN while tribis is more suitable for ELM when the best half of the features are used.

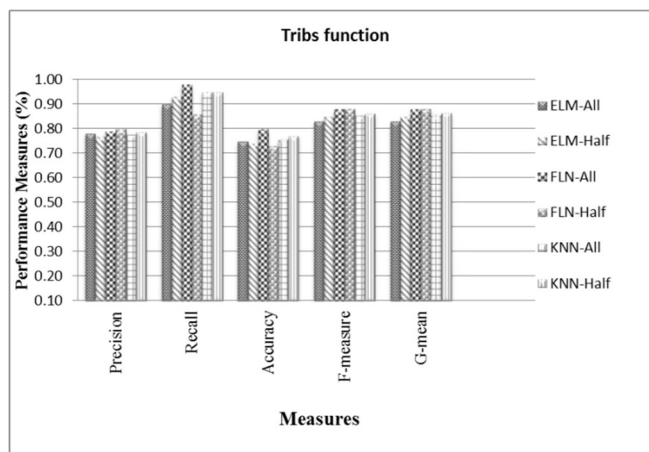


Figure 5. Classification measures for evaluation of Parkinson based on ELM, KNN, and FLN for both full and half features using Tribis

### V. CONCLUSIONS

This article has provided feature selection for finding the most important and relevant features to produce an efficient and effective model. The model is based on entropy-based Parkinson classification. It selects the top 50% of features. Two variants of neural networks are used for evaluation, the first one is ELM and the other one is FLN. In addition, the K-Nearest Neighbor KNN algorithm is applied for evaluation. The results show the superiority of ELM and FLN when the model of feature selection is used with an accuracy of 80% compared with only 78% when the model is not used. Future work will be devoted to exploring the effect of the percentage of selected features on the predicted results.



**ACKNOWLEDGMENT**

The authors would like to thank the reviewers for taking the time to read the manuscript.

**References**

[1] S. Przedborski, "The two-century journey of Parkinson disease research," *Nature Reviews Neuroscience*, vol. 18, issue 4, pp. 251-259, 2017, <https://doi.org/10.1038/nrn.2017.25>.

[2] D. S. Goldstein, P. Sullivan, C. Holmes, D. C. Mash, I. J. Kopin, & Y. Sharabi, "Determinants of denervation-independent depletion of putamen dopamine in Parkinson's disease and multiple system atrophy," *Parkinsonism & Related Disorders*, vol. 35, pp. 88-91, 2017, <https://doi.org/10.1016/j.parkreldis.2016.12.011>.

[3] J. J. Lee, J. S. Oh, J. H. Ham, D. H. Lee, I. Lee, Y. H. Sohn, & P. H. Lee, "Association of body mass index and the depletion of nigrostriatal dopamine in Parkinson's disease," *Neurobiology of Aging*, vol. 38, pp. 197-204, 2016, <https://doi.org/10.1016/j.neurobiolaging.2015.11.009>.

[4] S. L. Oh, Y. Hagiwara, U. Raghavendra, R. Yuvaraj, N. Arunkumar, M. Murugappan, & U. R. Acharya, "A deep learning approach for Parkinson's disease diagnosis from EEG signals," *Neural Computing and Applications*, vol. 32, pp. 10927-10933, 2020, <https://doi.org/10.1007/s00521-018-3689-5>.

[5] I. Kollia, A. G. Stafylopatis, & S. Kollias, "Predicting Parkinson's disease using latent information extracted from deep neural networks," *Proceedings of the 2019 International Joint Conference on Neural Networks (IJCNN)*, Budapest, July 14-19, 2019, pp. 1-8. <https://doi.org/10.1109/IJCNN.2019.8851995>.

[6] M. M. Khan, A. Mendes, & S. K. Chalup, "Evolutionary wavelet neural network ensembles for breast cancer and Parkinson's disease prediction," *PLoS One*, vol. 13, issue 2, pp. 1-15, 2018, <https://doi.org/10.1371/journal.pone.0192192>.

[7] O. Klempfir, & R. Krupicka, "Machine learning using speech utterances for Parkinson disease detection," *Lékař a Technika-Clinician and Technology*, vol. 48, issue 2, pp. 66-71, 2018. <https://doi.org/10.14311/CTJ.2019.2.05>.

[8] L. Ali, C. Zhu, M. Zhou, & Y. Liu, "Early diagnosis of Parkinson's disease from multiple voice recordings by simultaneous sample and feature selection," *Expert Systems with Applications*, vol. 137, pp. 22-28, 2019, <https://doi.org/10.1016/j.eswa.2019.06.052>.

[9] M. Nasrolahzadeh, Z. Mohammadpoory, & J. Haddadnia, "Higher-order spectral analysis of spontaneous speech signals in Alzheimer's disease," *Cognitive Neurodynamics*, vol. 12, issue 6, pp. 583-596, 2018, <https://doi.org/10.1007/s11571-018-9499-8>.

[10] C. Vitale, V. Marcelli, T. Abate, A. Pianese, R. Allocca, M. Moccia, & M. Cavaliere, "Speech discrimination is impaired in parkinsonian patients: Expanding the audiologic findings of Parkinson's disease," *Parkinsonism & Related Disorders*, vol. 22, pp. S138-S143, 2016, <https://doi.org/10.1016/j.parkreldis.2015.09.040>.

[11] O. Almomani, "A feature selection model for network intrusion detection system based on PSO, GWO, FFA and GA algorithms," *Symmetry*, vol. 12, issue 6, pp. 1-20, 2020. <https://doi.org/10.3390/sym12061046>.

[12] S. Sajadi, & A. Fathi, "Genetic algorithm based local and global spectral features extraction for ear recognition," *Expert Systems with Applications*, vol. 159, pp. 1-21, 2020, <https://doi.org/10.1016/j.eswa.2020.113639>.

[13] M. Hassaballah, H. A. Alshazly, & A. A. Ali, "Ear recognition using local binary patterns: A comparative experimental study," *Expert Systems with Applications*, vol. 118, pp. 182-200, 2019, <https://doi.org/10.1016/j.eswa.2018.10.007>.

[14] Q. Al-Tashi, H. M. Rais, S. J. Abdulkadir, S. Mirjalili, & H. Alhussian, "A review of grey wolf optimizer-based feature selection methods for classification," In *Evolutionary Machine Learning Techniques. Algorithms for Intelligent Systems*, E-Publishing Inc, Springer, Singapore, 2020, pp. 273-286, [https://doi.org/10.1007/978-981-32-9990-0\\_13](https://doi.org/10.1007/978-981-32-9990-0_13).

[15] M. Sajjad, M. Nasir, K. Muhammad, S. Khan, Z. Jan, A. K. Sangaiah, S. W. Baik, "Raspberry Pi assisted face recognition framework for enhanced law-enforcement services in smart cities," *Future Generation Computer Systems*, vol. 108, pp. 995-1007, 2020, <https://doi.org/10.1016/j.future.2017.11.013>.

[16] T. Ozseven, "A novel feature selection method for speech emotion recognition," *Applied Acoustics*, vol. 146, pp. 320-326, 2019. <https://doi.org/10.1016/j.apacoust.2018.11.028>.

[17] V. B. Semwal, J. Singha, P. K. Sharma, A. Chauhan, & B. Behera, "An optimized feature selection technique based on incremental feature analysis for bio-metric gait data classification," *Multimedia Tools and*

*Applications*, vol. 76, issue 22, pp. 24457-24475, 2017, <https://doi.org/10.1007/s11042-016-4110-y>.

[18] M. Muaz, & R. Mayrhofer, "An analysis of different approaches to gait recognition using cell phone based accelerometers," *Proceedings of the International Conference on Advances in Mobile Computing & Multimedia*, December, 2013, pp. 293-300. <https://doi.org/10.1145/2536853.2536895>.

[19] M. H. M. Zaman, M. M. Mustafa, M. A. Hamman, & A. Hussain, "Neural network based prediction of stable equivalent series resistance in voltage regulator characterization," *Bulletin of Electrical Engineering and Informatics*, vol. 7, issue 1, pp. 134-142, 2018, <https://doi.org/10.11591/eei.v7i1.857>.

[20] G. B. Huang, Q. Y. Zhu, & C. K. Siew, "Extreme learning machine: theory and applications," *Neurocomputing*, vol. 70 issue (1-3), pp. 489-501, 2006. <https://doi.org/10.1016/j.neucom.2005.12.126>.

[21] S. K. Abdulateef, T. A. N. Abdali, M. D. S. Alroomi, & M. A. A. Altaha, "An optimise ELM by league championship algorithm based on food images," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 20, issue 1, pp. 132-137, 2020. <https://doi.org/10.11591/ijeecs.v20.i1.pp132-137>.

[22] S. K. Abdulateef, M. Mahmuddin and N. H. Harun, "Aggregated features association classifier for multiple food items identification," *Journal of Engineering and Applied Sciences*, vol. 12, issue 8, pp. 2200-2206, 2017. <https://doi.org/10.14569/IJACSA.2017.080809>.

[23] I. Bahiuddin, et al., "Comparing the linear and logarithm normalized extreme learning machine in flow curve modeling of magnetorheological fluid," *Indonesian Journal of Electrical and Computer Engineering*, vol. 13, issue 3, pp. 1065-1072, 2019. <https://doi.org/10.11591/ijeecs.v13.i3.pp1065-1072>.

[24] M. H. Ali, B. A. D. Al Mohammed, A. Ismail, & M. F. Zolkipli, "A new intrusion detection system based on fast learning network and particle swarm optimization," *IEEE Access*, vol. 6, pp. 20255-20261, 2018, <https://doi.org/10.1109/ACCESS.2018.2820092>.

[25] G. Li, X. Qi, B. Chen, Y. Ma, P. Niu, & Z. Chen, "Fast learning network with parallel layer perceptrons," *Neural Processing Letters*, vol. 47, issue 2, pp. 549-564, 2018. <https://doi.org/10.1007/s11063-017-9667-6>.

[26] D. Joshi, A. Khajuria, & P. Joshi, "An automatic non-invasive method for Parkinson's disease classification," *Computer Methods and Programs in Biomedicine*, vol. 145, pp. 135-145, 2017, <https://doi.org/10.1016/j.cmpb.2017.04.007>.

[27] D. Avci, & A. Dogantekin, "An expert diagnosis system for Parkinson disease based on genetic algorithm-wavelet kernel-extreme learning machine," *Parkinson's Disease*, vol. 2016, pp. 1-9, 2016, <https://doi.org/10.1155/2016/5264743>.

[28] C. Ma, J. Ouyang, H. L. Chen, & X. H. Zhao, "An efficient diagnosis system for Parkinson's disease using kernel-based extreme learning machine with subtractive clustering features weighting approach," *Computational and Mathematical Methods in Medicine*, vol. 2014, pp. 1-14, 2014, <https://doi.org/10.1155/2014/985789>.

[29] S. Lahmiri, & A. Shmuel, "Detection of Parkinson's disease based on voice patterns ranking and optimized support vector machine," *Biomedical Signal Processing and Control*, vol. 49, pp. 427-433, 2019. <https://doi.org/10.1016/j.bspc.2018.08.029>.

[30] L. Parisi, N. RaviChandran, & M. L. Manaog, "Feature-driven machine learning to improve early diagnosis of Parkinson's disease," *Expert Systems with Applications*, vol. 110, pp. 182-190, 2018. <https://doi.org/10.1016/j.eswa.2018.06.003>.

[31] M. Al-Sarem, F. Saeed, W. Boulila, A. H. Emara, M. Al-Mohaimed, & M. Errais, "Feature selection and classification using CatBoost method for improving the performance of predicting Parkinson's disease," *Proceedings of the International Conference of Advanced Computing and Informatics*, Morocco, April 2020, pp. 189-199. [https://doi.org/10.1007/978-981-15-6048-4\\_17](https://doi.org/10.1007/978-981-15-6048-4_17).

[32] R. Yuvaraj, U. R. Acharya, & Y. Hagiwara, "A novel Parkinson's disease diagnosis index using higher-order spectra features in EEG signals," *Neural Computing and Applications*, vol. 30, issue 4, pp. 1225-1235, 2018. <https://doi.org/10.1007/s00521-016-2756-z>.

[33] M. Little, P. McSharry, E. Hunter, J. Spielman, & L. Ramig, "Suitability of dysphonia measurements for telemonitoring of Parkinson's disease," *IEEE Transactions on Biomedical Engineering*, vol. 56, pp. 1015-1022, 2009, <https://doi.org/10.1109/TBME.2008.2005954>.

[34] J. Cao, L. Lin, Y. Zhang, & Y. Lv, "Application of extreme learning machine algorithm in medical diagnosis," *Journal of Healthcare Engineering*, 2021.

[35] Q. Liu, H. Li, H. Tang, X. Zhang, & X. Li, "A survey on applications of extreme learning machine," *Neurocomputing*, 338, pp. 261-276, 2019.

- [36] I. Goodfellow, Y. Bengio, & A. Courville, Deep learning. single hidden layer neural networks and their benefits for nonlinear mapping, MIT Press. 2016.
- [37] L. R. Sultan, S. K. Abdulateef, & B. A. Shtayt, "Prediction of student satisfaction on mobile-learning by using fast learning network," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 27, issue 1, pp. 488-495, 2022. <https://doi.org/10.11591/ijeecs.v27.i1.pp488-495>.



**SALWA KHALID ABDULATEEF** received a B.Sc. degree in Computer Science from Basra University, Iraq, and an M.Sc. degree in Computer Science from Basra University, Iraq. Currently, she is working lecturer in the department of computer science at Tikrit University. Her research interests are artificial intelligence, computer vision, image processing, data science, and information hiding.



**AHMED NASER ISMAEL** received his B.Sc. degree in Computer Science from Tikrit University, Iraq. M.Sc degree from Utara Malaysia University, Malaysia. His research area is artificial intelligence, image processing, and pattern recognition.



**MOHANAD DAWOOD SALMAN AL-ROOMI** received his B.Sc. degree in Computer Science from Shatt Al-Arab University College, Basra, Iraq, in 2001. Also received a Master of Science (information technology) from Universiti Utara Malaysia (UUM) in 2017. Currently, works assistant lecturer in the department of computer science – College of computer science and mathematics – Tikrit University. His interest is in artificial intelligence of machine learning and compilers.

...