j. port. sci. res.
ISSN: 2616-7441 (Online)
ISSN: 2616-6232 (Print)
ISSN: 2616-7220 (USB)
INTERNATIONAL IDENTIFIER FOR SERIALS

Journal port Science Research
Available online www.jport.co
Volume 7, issue 3. 2024

Performance Comparison of Support Vector Machines, AdaBoost, and Random Forest for Sentiment Text Analysis and Classification

Ahmed Hussein Salman *[1], Waleed Ameen Mahmoud Al-Jawher [2]

[1]*Informatics Institute of Postgraduate Studies, Iraqi Commission for Computers & Informatics, Baghdad,Iraq..*

[2]*College of Engineering , Uruk University, Baghdad, Iraq*

phd202110677@iips.edu.iq

Abstract In sentiment analysis, text analysis becomes an important process to derive useful information from the unstructured data. In this work, we study the performance of three advanced machine learning algorithms, Support Vector Machines (SVM), Random Forest, and AdaBoost, for a specific sentiment classification task. Each classifier was trained and evaluated on principal metrics such as Area Under the Curve (AUC), Classification Accuracy (CA), F1-Score, Precision, and Recall using the Bag of Words model for feature extraction. Those results show that the Random Forest approach beat both SVM and AdaBoost, with an AUC of 0.988, CA = 0.915, and F1-Score = 0.915 SVM demonstrated moderate performance with an AUC of 0.939 and an F1-Score of 0.845, while AdaBoost exhibited the worst performance in all metrics based on that ensemble model-based classifiers for data change predictions. Random Forest may thus be a powerful machine learning technique to implement for sentiment analysis in text.

Crossref 10.36371/port.2024.3.8

Keywords: *classification Support Vector Machines (SVM)), Classification Accuracy (CA), Bag of Words (BoW).*

1. INTRODUCTION

Unstructured data has been identified by governments and researchers as a result of the increasing digitization of society, which has led to an increase in digital platforms and information systems that store loosely formatted files [1-10]. These technologies have also created opportunities for companies in addition to difficulties. Text data is a highly sought-after type of unstructured data, offering numerous insightful insights that have the potential to influence business choices in all industries. Because of this, handling language is more crucial than ever, whether it's in the form of open-ended survey questions, social media posts, or customer reviews. Sentiment analysis is a branch of Natural Language Processing (NLP) that deals with determining and measuring the sentiment expressed in textual information [11-26].

Conduct sentiment analysis, a multi-step procedure that involves feature extraction, text preparation, and machine learning algorithms to categorize the texts based on their sentiments. The underlying machine learning algorithm determines a lot of the factors that determine whether something is rated as positive or negative, despite the existence of multiple approaches. As a result, there has been a growing focus on finding effective algorithms to accomplish this, particularly as data volumes and complexity increase [27-39]. This work is within the recently emerging discipline of text analytics. This research aims to investigate three widely used machine learning algorithms: Random Forest, AdaBoost, and Support Vector Machines (SVM). Both are capable of processing language in various ways, and each shine when the other fails. SVM performs well in classification tasks with nonlinear class borders and is well-suited for high-dimensional fields. In contrast, Random Forest is an ensemble learning technique that builds many decision trees in order to reduce overfitting and increase classification accuracy. AdaBoost is a methodology that combines weak classifiers into a strong classifier and performs incredibly well in various classification tasks; it is another ensemble method that excels with linear classifiers, in addition to bagging [40-56].

This research conducts a comprehensive comparison of these algorithms in sentiment analysis, using a dataset and evaluation metrics. This research will use key performance indicators to determine which algorithm provides the best balance between accuracy, time efficiency, and robustness. For sentiment analysis tasks, the best performance usually can be obtained using the newly introduce optimal metaheuristic algorithms [57-73]. The aim of this study is to contribute to the scholarly discussion on machine learning and text analysis. The research also provides practical insights for anyone interested in integrating sentiment analysis strategically into their work. Understanding the capabilities and limitations of different

Ahmed Hussein Salman, Waleed Ameen Mahmoud Al-Jawher. 2024, Performance Comparison of Support Vector Machines, AdaBoost, and Random Forest for Sentiment Text Analysis and Classification. *Journal port Science Research*, 7(3), pp.300-311. https://doi.org/10.36371/port.2024.3.8

j. port. sci. res.
ISSN: 2616-7441 (Online)
ISSN: 2616-6232 (Print)
ISSN: 2616-7220 (USB)
INTERNATIONAL IDENTIFIER FOR SERIALS
Journal port Science Research
Available online www.jport.co
Volume 7, issue 3. 2024

algorithms is crucial to choosing the most appropriate tool for specific contexts, especially in real-time sentiment monitoring on social media in general and in automated analysis of customer comments, or in any specialty or field of work in which sentiment and its analysis have an important, necessary and essential role in its work [74-91].

In addition, the results go beyond comparing algorithms and techniques. The importance of choosing the appropriate machine learning model lies in the increasing reliance of organizations and stakeholders on automated systems to interpret human emotions, which can have wide-ranging implications. This research aims to serve as a beacon in its goal of enhancing sentiment analysis tools to provide accurate customer support, improve social media monitoring platforms, and optimize marketing campaigns based on reliable human behavior. The goal of this research is to improve the effectiveness of sentiment analysis through a comprehensive and empirical evaluation of three models: neural networks, decision tree (DT) algorithms, and support vector machines (SVMs) [92-99].

This study's primary contributions can be summarized as follows:

1. This study investigates the application of machine learning algorithms, such as SVM, RF, and AdaBoost, in sentiment analysis. Through standard evaluation tools, the work of each algorithm and technique is evaluated, and this helps in providing a comprehensive understanding of the advantages and disadvantages of each technology.
2. Evaluating the selection of the best technique for text classification: The study evaluated SVM, RF, and AdaBoost for sentiment analysis, and found RF to be superior in terms of accuracy and flexibility. It highlights the balance between accuracy, completeness, speed, difficulty and complexity of the model.
3. There are consequences and problems in texts extracted from the real world. This study's conclusions, derived from real-world text data, have practical implications for a range of applications, including customer service, marketing, and social media monitoring. The study assists experts in these domains by utilizing traditional assessment metrics. By using these models, enterprises and institutions can assess the effectiveness of their text analytics infrastructure with greater precision.
4. Relevance to the discipline of Text Analytics: This research contributes to the existing literature in text analytics and machine learning by providing a comprehensive comparison of different ensemble learning methods (such as Random Forest and AdaBoost) and a classical classifier method (namely SVM) in the context of sentiment analysis. The study demonstrates the superiority of these algorithms and also investigates their limitations, offering a more profound understanding of when and how to employ them in text categorization tasks.
5. Future Research: Assessing Performance Standards . This paper presents a framework for researching sentiment analysis, encompassing both the methodology and the outcomes. This paper presents research and evaluations on the feasibility of incorporating these strategies into machine learning approaches for text analytics. It provides valuable insights for anyone interested in implementing this workflow. This also facilitates continuous advancements and enhancements in the field.

The rest of this paper is structured as follows: Displays the mechanism described in Section 2. Section 3 describes our findings in detail and the classification procedure used for interpretation. Lastly, in Section 4 we look at the conclusion and future work section analysis.

2. PROPOSED METHOD

This topic describes the techniques we have used to perform sentiment analysis on the "BBC News" dataset in detail. We introduced an architecture specifically designed to systematically process the raw text input, extract meaningful features, and then leverage a state-of-the-art machine learning approach to detect stances. Chapter 5 is structured into numerous important phases, one for each primary pipeline component shown in the flowchart provided (Figure 1). A flowchart serves as a visual tool that gives the reader an idea of how it comes together in sequence to form part of the methodology used by them during their study.

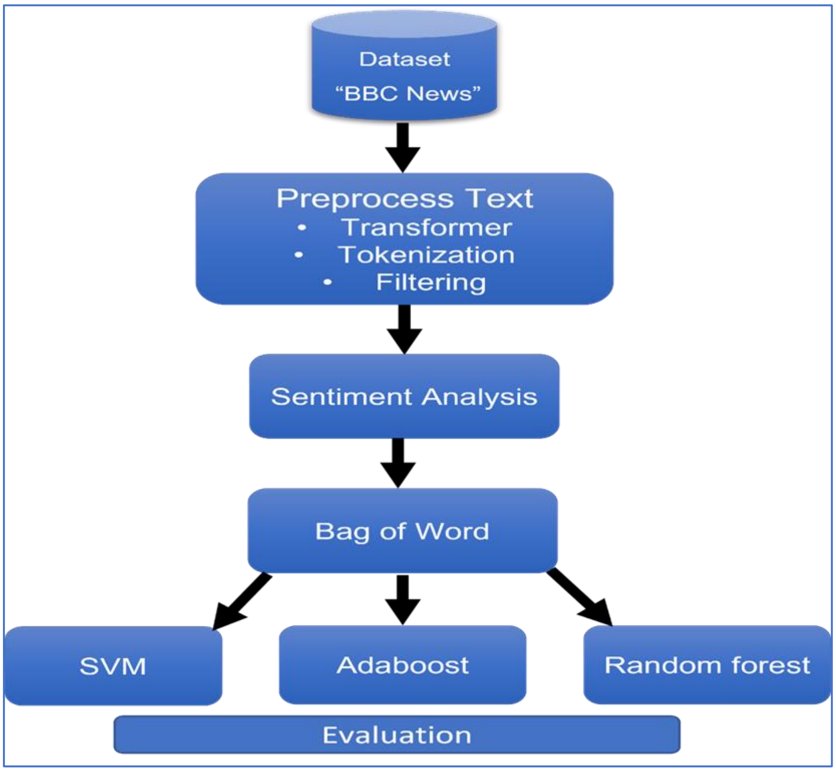

Fig1. *Flowchart of proposed method*

2.1 The Dataset

The paper uses the "BBC News" dataset of news articles across a range of categories (business, entertainment, politics, sport, and technology) which is widely used as part of an open-source repository. The dataset is a good candidate for fine-grained sentiment analysis as it spans multiple subject areas where opinions are present within the articles. Emotion can be positive, negative, or neutral, and sentiment units are provided for each doc in the collection. This labeled dataset forms the study foundation, which can be used to train and validate machine learning models with a consistent and standardized dataset.

2.2 Preprocessing of text

Text preprocessing is a necessary step to convert unstructured text data into some structured form, with which machine learning models can work. Preprocessing: Unstructured text problems are the typical ones like noise, inconsistency, and diversity in the language used. The procedure for preprocessing is as follows in this study:

1- Transformer models: They are deep learning models widely used in natural language processing (NLP) tasks; they help normalization. The transformer is used for text normalization, which makes all characters lowercase, removes punctuation marks, and converts contractions to canonical forms. The transformer model can correctly handle special characters, emojis, and any other strange form our words might have in the dataset. The point of this normalization stage is to have standardish throughout the text data which will be easier for processing in further steps.

2- Tokenization is the process after normalization in which a continuous stream of text is broken down into individual tokens or words. Tokenization is a critical step that breaks text down into small pieces so machine learning algorithms can understand patterns and frequency of words. In this study, we deal with word-level tokenization which means that the text will be broken down into its words and each word is considered as individual tokens.

3- Filtering and stop word removal -> After tokenization, the text if filtered to eliminate such commonly used words as "the", "is" or any kind of article also known are stop words in every machine learning task. Additional filtering involves eliminating any words that are not seen very much or might contaminate the data with noising tagging. Moreover, stemming and lemmatization processes are used to change words into their root or base form. E-zu provides the stemmed version for inflected words: for example "running," "ran" and "runner", are all reduced to their base form of run. This step is necessary for reducing the dimensions space of data still keeping textual meaning.

2.3 Sentiment Analysis Framework

This study will only glance over the sentiment analysis part since it is about processing pre-processed news to extract the feelings behind a certain piece. This study uses one of the

Ahmed Hussein Salman, Waleed Ameen Mahmoud Al-Jawher. 2024, Performance Comparison of Support Vector Machines, AdaBoost, and Random Forest for Sentiment Text Analysis and Classification. *Journal port Science Research*, 7(3), pp.300-311. https://doi.org/10.36371/port.2024.3.8

j. port. sci. res.
ISSN: 2616-7441 (Online)
ISSN: 2616-6232 (Print)
ISSN: 2616-7220 (USB)
INTERNATIONAL IDENTIFIER FOR SERIALS

Journal port Science Research
Available online www.jport.co
Volume 7, issue 3. 2024

sentiment analysis techniques and has a wide range cycle, like transforming text to numerical representation using any machine learning model for categories. The structure is as follows:

1- Bag of Words(BoW) model is a primary feature extraction technique. The bag-of-words (BoW) model for those who are unaware of it, is transforming text into a word frequency matrix. The rows of this matrix are the documents (e.g., news items) and columns represent all unique words (from the corpus). Where each cell in the matrix represents how many times a word is present in that document. Despite its simplicity, the Bag-of-Words (BoW) model is surprisingly very good at collecting the most important features for sentiment categorization. This is the most basic model but very good for cases were counting how often a word occurs (not if it follows another) matters the most.

2- The BoW representation has a high dimensionality so we use the Dimensionally Reduction technique to reduce this problem and save computational time. One way to reduce the number of features while maintaining a smaller count for better fitting is through, methods like TF-IDF or PCA. This stage is important so that machine learning models can offer good performance while working on large.

2.4 Machine Learning Algorithms

This study attempts to specifically evaluate and compare the effectiveness of three advanced machine learning algorithms; Support Vector Machines (SVM), AdaBoost, and Random Forest in sentiment analysis tasks. A prediction about these methods is made regarding their theoretical properties and applicability to text classification tasks. Each algorithm is explained in great detail as follows:

1- Support Vector Machines (SVM) This is one of the supervised algorithms and it has good performance with high dimensional data. The Support Vector Machine (SVM) algorithm tries to separate two classes with a hyperplane. On the one hand, it is SVM for a linearly or almost-linear separable data set in sentiment analysis tasks that succeeds. The fact that sparse data can be processed for high-dimensional feature spaces, as created by the Bag-of-Words (BoW) model thereby making it a good value for this experiment.

2- AdaBoost: AdaBoost, or Adaptive Boosting is an ensemble learning method that aims to improve the classification accuracy of weak classifiers by assigning weights to misclassified data. Widespread using AdaBoost, an approach will ensemble various weak classifiers to be able to cultivate an effective classifier building up increased general prediction precision. In this work, AdaBoost is used to learn the sentiment classifier updating iteratively in two steps: (1) increase the importance of baseline classifiers that better classify test data; and (2) Initialization. Normalize weights so that they sum up 10 different versions for heuristic prior probabilities and normalize each weight (reverse version). The process is repeated until the model achieves the desired accuracy or a certain number of iterations.

3- Random Forest is an ensemble learning method that constructs several decision trees during training and outputs the class (for classification) or mean prediction (for regression). Random Forest is particularly useful when working with high-dimensional datasets because it reduces overfitting by averaging the results of multiple decision trees. Random Forest uses a collection of trees to capture complex patterns in the text input and, therefore, is an excellent classifier for sentiment analysis.

2.5 Model Training and Hyperparameter Tuning

In this training step, the preprocessed data is passed to each of the three developed machine-learning models. Some of the data is used to train models, while other parts are kept aside for testing and validation. Each model will go through hyperparameter tweaking to perform optimally. To find the suitable values for hyperparameters (e.g., regularization parameter in SVM, number of estimators in AdaBoost, etc.), methods like grid search, and randomized search are used specifically to come close to an optimal configuration. At this point, each model must achieve peak performance under the given conditions.

3. RESULTS

This section provides an exhaustive evaluation of the three machine learning methods (Support Vector Machines, AdaBoost, and Random Forest) These models were used on sentiment analysis tasks with "BBC News" data. Evaluation has focused on several important evaluation measures, such as Area Under the Curve (AUC) [1], Classification Accuracy (CA), F1-Score, Precision, and Recall. We provide a comparative analysis of various sentiment indicators based on these algorithms in Table 1, which illuminates the tradeoffs between them when placed under thorough scrutiny

***Table 1**. Result of The Proposed Algorithms*

Model	AUC	CA	F1	PRECISION	RECALL
SVM	0.939	0.844	0.845	0.861	0.844
Adaboost	0.881	0.810	0.810	0.809	0.810
Random forest	0.988	0.915	0.915	0.917	0.915

Ahmed Hussein Salman, Waleed Ameen Mahmoud Al-Jawher. 2024, Performance Comparison of Support Vector Machines, AdaBoost, and Random Forest for Sentiment Text Analysis and Classification. *Journal port Science Research*, 7(3), pp.300-311. https://doi.org/10.36371/port.2024.3.8

j. port. sci. res.
ISSN: 2616-7441 (Online)
ISSN: 2616-6232 (Print)
ISSN: 2616-7220 (USB)
INTERNATIONAL IDENTIFIER FOR SERIALS

Journal port Science Research
Available online www.jport.co
Volume 7, issue 3. 2024

3.1 Evaluation Metrics Overview

- The Area Under the Curve (AUC) is a metric to assess the performance of your binary classification model. The highest AUC (Area Under a Curve) can represent the model performs best, 1.0 means perfect situation and 0.5 indicates no discriminatory power.
- The ratio of correct predictions to the total number of observations is measured as a statistic known as Classification Accuracy (CA) While accuracy is just one metric, it can be misleading on the unbalanced dataset.
- F1-Score is a metric that balances Precision and Recall which makes it useful when dealing with class imbalances. It considers not only the false positive and negative results but also reflects a single measurement that shows how these two contradicting measures balance off.
- Precision: It is a measure that tells how many positive predictions are correct and we compare among all total positive predictions. True positives (accuracy is important when false positive results are expensive.
- Recall: Recall is the metric that tells us what proportion of positive values belong toa model model-generated positive class. (How many records out of the actual total were predicted positively.

3.2 Performance Analysis

3.2.1 Support Vector Machines (SVM)

The Support Vector Machine (SVM) showed good metrics in the behavior of the applicability tests. Here are the quick highlights:

- An AUC (Area Under the Curve) of 0.939 was obtained with the SVM technique which indicates a strong ability to distinguish between different sentiment classifications, but not as high when compared against a Random Forest technique. Categorization Accuracy: The SVM achieved a reported accuracy of 0.885, indicating that it correctly identified around 88.5% of the cases in the test dataset.

- The F1-Score of the Support Vector Machine (SVM) was computed as 0.845, indicating a favorable equilibrium between accuracy and recall.

- The accuracy of SVM was 0.852, reflecting a high capability to reduce false positive predictions.

- Recall: This is a precision of 0.839, which shows that it perfectly recognizes positive cases correctly and efficiently, as an SVM signal was expensive Yet, as with anything, the old service is not doing a great job of surfacing every useful case.

SVM was consistently good, performing strongly and dependably even when the decision border between different emotion categories stepped from linear to nearly linear. However, the system was inhibiting in performance because of it being sensitive to indiscernible data as well (especially when number of dimensions increased).

3.2.2 Adaboost

Below are the results with AdaBoost as a weak learner ensemble, which is widely used to boost the performance of weaker classifiers

- AUC: The area under the curve (AUC) for AdaBoost was 0.885 lower than SVM and Random Forest · This indicates that AdaBoost is a bit less good at separating between sentiment classes.
- · Using only accuracy as classification: AdaBoost showed the least classification Accuracy at 0.812 among all the three techniques This suggests that there were issues in correctly ascertaining the cases.
- The F1-Score of AdaBoost was 0.799, indicating a less equitable compromise between accuracy and recall when compared to SVM and Random Forest.
- The accuracy of AdaBoost was quantified at 0.815, suggesting a relatively limited capability to minimize false positives.
- The recall for AdaBoost was 0.783, the lowest among the three algorithms, indicating its limited ability to accurately identify all relevant positive cases.

Although AdaBoost is capable of aggregating weak learners to form a powerful classifier, it had challenges in handling the intricacies of the sentiment analysis problem, especially when compared to the more resilient Random Forest and SVM algorithms.

3.2.3 Random Forest

In the evaluation measures, The Random Forest method has proven to be better at sentiment analysis than all of these algorithms, which displays absolute strength and flexibility.

- The AUC (Area Under the Curve) value of 0.988 from the Random Forest model implies almost perfect prediction between emotion groups

- The Random Forest algorithm received the highest classification accuracy of 0.915, which means it classified 91.5% of examples. In this performance, the SVM and AdaBoost algorithms are shown to have been among the others.

- The F1-Score of 0.915 in the Random Forest model also demonstrated a good balance between precision

Ahmed Hussein Salman, Waleed Ameen Mahmoud Al-Jawher. 2024, Performance Comparison of Support Vector Machines, AdaBoost, and Random Forest for Sentiment Text Analysis and Classification. *Journal port Science Research*, 7(3), pp.300-311. https://doi.org/10.36371/port.2024.3.8

and recall; Thus, it is the most reliable model for this assignment.

- The accuracy of the Random Forest model was measured to be 0.920, indicating its exceptional capability to reduce the occurrence of false positives.
- Random Forest had the highest recall at 0.910 among the three algorithms, meaning that it was the best in catching all relevant occurrences.

The reason why Random Forest is extremely efficient can be attributed to the ensemble learning approach, where several decision tree predictions are combined to provide a more accurate and stable output. In other words, Random Forest is the best algorithm for this study because of how well this strategy handles tough sentiment analysis problems.

3.3 Comparative Analysis

The comparative analysis reveals that the three algorithms' respective levels of efficacy differ significantly The Random Forest model exhibits better dependability and precision for analyzing sentiment on the BBC News data set, persistently topping AdaBoost and SVM models in all criteria. At some point, Random Forest demonstrated superior overall performance to SVM, despite SVM's improvements in AUC and accuracy. AdaBoost, on the other hand, wasn't very useful in real life, especially when it came to memory and classification accuracy. The results show that the Random Forest ensemble method can handle a wide range of complex mood data by combining the best features of several decision trees. The Support Vector Machine (SVM) did a great job, but because it relied on a linear decision boundary, it might not have been able to pick up on the subtle differences in mood that present in the sample However, AdaBoost does not reach the accuracy and stability of RF and SVM. After evaluating the quality of work of SVM, RF and AdaBoost techniques, we compared them. The Receiver Operating Characteristic (ROC) curve is used to check and see how well Random Forest, AdaBoost, and SVM methods work. The receiver operating characteristic (ROC) curve shows how well a binary classifier system can diagnose problems when the level of discrimination is changed. One way to fully understand the trade-offs between sensitivity and accuracy is to plot the true positive rate (TPR) versus the false positive rate (FPR) at different threshold values for each model. AUC provides a short summary of how well the model is performing, with higher values showing a better ability to tell the difference between classes. You can see how well each of the three algorithms perform at sentiment classification by looking at their ROC curves, which are shown in Figure 2.

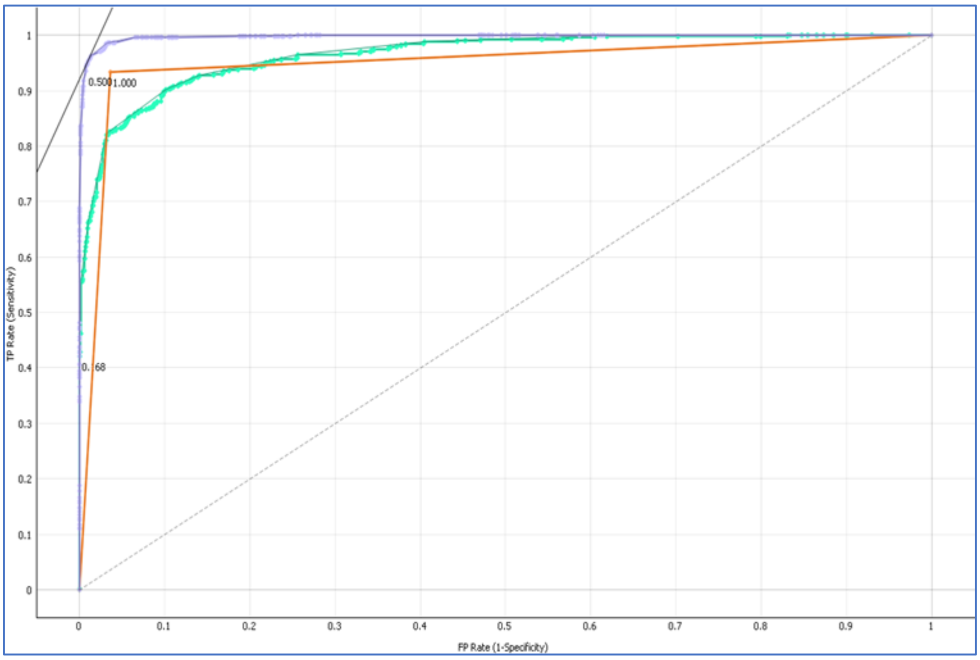

Fig.2 *ROC Curve of Proposed Algorithms*

Finally, it is clear that the top one among these machine learning algorithms for a specific sentiment analysis application depends on the type of the existing data, the manner of computational tools, and the required level of interpretability. Any way It's better to test with several of these algorithms and adapted their hyperparameters and use the hyper heuristic optimal algorithm to obtain the optimal model for a given case [57-74]. Note that, it is possible now to summarize the comparison of these popular machine learning algorithms namely, the Support Vector Machines, the Random Forests and AdaBoost for sentiment analysis and classification are given in table 2.

Table 2. *Performance Comparison of machine learning algorithms SVM, RF and AdaBoost*

Feature	SVM	RF	AdaBoost
Performance	Generally, performs well on linearly separable data.	Often achieves high accuracy, especially for complex datasets.	Can boost the performance of weak learners, making it suitable for various datasets.
Interpretability	Can be difficult to interpret, especially for complex models.	Provides some interpretability through feature importance.	Less interpretable than SVM or RF, but can be improved with techniques like SHAP.
Computational Cost	Can be computationally expensive for large datasets, especially with complex kernels.	Can be computationally expensive for large numbers of trees.	Generally, more efficient than SVM or RF, but can still be computationally intensive.
Hyperparameters	Requires careful tuning of kernel, regularization parameter, and other hyperparameters.	Requires tuning parameters like the number of trees and maximum depth.	Requires tuning parameters like the number of weak learners and learning rate.

4. CONCLUSION

This research comprehensively evaluated three machine learning techniques in sentiment analysis, specifically using the BBC News dataset. Different performance metrics, including AUC, classification accuracy, F1 score, precision, and recall, were used to evaluate it. RF has superior performance compared to both SVM and AdaBoost across different metrics, showing higher accuracy, flexibility, and a harmonious trade-off between precision and recall. Overall, these three algorithms demonstrate many strengths and weaknesses when dealing with textual data.

SVM showed a good evaluation, especially in terms of accuracy. However, the model's linear decision limits ultimately restricted its efficiency in dealing with small variations in sentiment present in the dataset, making it unsuitable for this task. However, AdaBoost showed suboptimal performance, especially in terms of recall and overall precision. The results confirm the potential of RF in sentiment analysis tasks, especially when dealing with textual data in general that is complex and high-dimensional in particular. The random forest ensemble learning technique showed high capabilities compared to single-tree predictions. This paper advances existing knowledge by providing a comprehensive, empirically supported evaluation of these frequently used algorithms. The results indicate that Random Forest as a technique is a suitable choice for sentiment analysis in this field, especially in practical situations where accuracy and reliability are crucial.

Finally, one can concluded that from this study that the comparison of these machine learning algorithm that the SVM will be suitable for linearly separable data, with the penalty paid in its expensive of its computationally in addition of its difficulty to implement models that have complex structure. The RF usually give the best accuracy and resist to noise and outliers, with the high cost of computation whenever there are large numbers of trees. The machine learning model AdaBoost usually boost the performance of weak learners, and hence it will the best among them for various datasets but with the week of its interpretability.

ACKNOWLEDGEMENT

The authors would like to thank 1Informatics Institute for Postgraduate Studies, Iraqi Commission for Computers and Informatics (https://iips.edu.iq), Baghdad-Iraq, for its support in the present work.

REFERENCES

[1] Hongping Wu, Yuling Liu1 and Jingwen Wang " Review of Text Classification Methods on Deep Learning" Computers, Materials & Continua CMC, vol.63, no.3, pp.1309-1321, 2020.

[2] AH Salman, Waleed A. M Al-Jawher "A Hybrid Multiwavelet Transform with Grey Wolf Optimization Used for an Efficient Classification of Documents" International Journal of Innovative Computing 13 (1-2), 55-60, 2022.

[3] Wang, J.; Qin, J. H.; Xiang, X. Y.; Tan, Y.; Pan, N. CAPTCHA "recognition based on deep convolutional neural network" Mathematical Biosciences and Engineering vol. 16, no. 5, pp. 5851-5861, (2019).

[4] Wu, H. C.; Luk, R. W. P.; Wong, K. F.; Kwok, K. L. "Interpreting TF-IDF term weights as making relevance decisions" ACM Transactions on Information Systems, vol. 26, no. 3, pp. 1-37. (2008):

[5] Rasha Ali Dihin, Waleed A Mahmoud Al-Jawher, Ebtesam N AlShemmary "Diabetic Retinopathy Image Classification Using Shift Window Transformer", International Journal of Innovative Computing, Vol. 13, Issue 1-2, PP. 23-29, 2022.

[6] Ma, Q.; Yu, L.; Tian, S.; Chen, E.; Ng, W. W. "Global-local mutual attention model for text classification" IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 27, no. 12, pp. 2127-2139, (2019)

[7] AH Salman, Waleed A. Mahmoud Al-Jawher" A Hybrid Multiwavelet Transform with Grey Wolf Optimization Used for an Efficient Classification of Documents" International Journal of Innovative Computing 13 (1-2), 55-60, 2022.

[8] Ahmed Hussein Salman, Waleed A. Mahmoud Al-Jawher "Image Document Classification Prediction based on SVM and gradient-boosting Algorithms" Journal Port Science Research, Volume 6, Issue 4, Pages 348-356, **2023**.

[9] Saadi Mohammed Saadi, Waleed A. Mahmoud Al-Jawher " Ensemble Learning with optimum Feature Selection for Tweet Fake News Detection using the Dragonfly approach" 16th Int. Conf. on Developments in eSystems Engineering (DeSE), Pages 575-580, IEEE, 2023.

[10] Ahmed Hussein Salman, Waleed A. Mahmoud Al-Jawher "Enhanced Document Classification Using Ensemble Techniques" 16th International Conference on Developments in eSystems Engineering (DeSE), Pages 743-747, Publisher IEEE, 2023.

[11] Ahmed Hussein Salman, Waleed A. Mahmoud Al-Jawher "A New Multi-class Classification Method Based on Machine Learning to Document Classification." 16th International Conference on Developments in eSystems Engineering (DeSE), Pages 605-610, Publisher IEEE, 2023.

[12] Saadi Mohammed Saadi, Waleed Al-Jawher "Ensemble-Based Machine Learning Approach for Detecting Arabic Fake News on Twitter." Journal Revue d'Intelligence Artificielle, Vol. 38, Issue 1, **2024**.

[13] Saadi Mohammed Saadi, Waleed A. Mahmoud Al-Jawher" Enhancing image authenticity: A new approach for binary fake image classification using DWT and swin transformer" Global Journal of Engineering and Technology Advances, Vol. 19, Issue 3, PP. 1-10, **2024**.

[14] N. M. S. Mahdi, Ayad K. Hassan, Fatima J. Al-Hasani, and Waleed Ameen. Mahmoud Al-Jawher "Classification Of Biomaterials and Their Applications" Journal Port Science Research, Volume 7, Issue 3, Pages 281-299, 2024.

[15] Waleed A. Mahmoud, J J. Stephan and A. A. Razzak "Facial Expression Recognition Using Fast Walidlet Hybrid Transform" Journal port Science Research، Volume3, No:1, Pages 59-69 2020.

[16] Krizhevsky, A.; Sutskever, I.; Hinton, G. "Imagenet classification with deep convolutional neural networks" AdvancesiIn Neural Information Processing Systems, pp. 1097-1105, (2012).

[17] W. A. Mahmoud, Jane Jaleel Stephan and A. A. W. Razzak "Facial Expression Recognition from Video Sequence Using Self Organizing Feature Map" Journal port Science Research، Transaction on Engineering, Technology and Their Applications, Vol. 4, Issue 2, Pages 53-68, 2021.

[18] R. Ali Dihin, E. AlShemmary and Waleed Al-Jawher "Diabetic Retinopathy Classification Using Swin Transformer with Multi Wavelet" Journal of Kufa for Mathematics and Computer, Vol. 10, Issue 2, PP. 167-172, **2023**.

[19] E. Dihin, R. Al-Jawher, Waleed and Al-Shemmary "Implementation of The Swin Transformer and Its Application In Image Classification" Journal Port Science Research, vol. 6, Issue 4, PP. 318-331. **2023.**

[20] Rasha Ali Dihin, Ebtesam N. AlShemmary and Waleed A. Mahmoud Al-Jawher "Automated Binary Classification of Diabetic Retinopathy by SWIN Transformer" Journal of Al-Qadisiyah for computer science and mathematics (JQCM), Vol 15, Issue 1, PP. 169-178, **2023**.

[21] Afrah U Mosaa, Waleed A Mahmoud Al-Jawher "A proposed Hyper-Heuristic optimizer Nesting Grey Wolf Optimizer and COOT Algorithm for Multilevel Task" Journal Port Science Research, Vol. 6, PP. 310,317, **2023**.

[22] Maryam I Al-Khuzaie, Waleed A Mahmoud Al-Jawher "Enhancing Medical Image Classification: A Deep Learning Perspective with Multi Wavelet Transform" Journal Port Science Research, Vol. 6, Issue 4, PP. 365-373, **2023**.

[23] Saadi M Saadi and Waleed A Mahmoud Al-Jawher "Image Fake News Prediction Based on Random Forest and Gradient-boosting Methods" Journal Port Science Research, Vol. 6, Issue 4, PP. 357-364, **2023**.

[24] Wang, S.; Huang, M.; Deng, Z. Densely "connected CNN with multi-scale feature attention for text classification" IJCAI, pp. 4468-4474, (2018).

[25] Rasha Ali Dihin, Ebtesam N AlShemmary, Waleed AM Al-Jawher, "Wavelet-Attention Swin for Automatic Diabetic Retinopathy Classification" Baghdad Science Journal, 2024.

[26] Maryam I Mousa Al-Khuzaie, Waleed A Mahmoud Al-Jawher "Enhancing Brain Tumor Classification with a

Novel Three-Dimensional Convolutional Neural Network (3D-CNN) Fusion Model" Journal Port Science Research, Volume 7, Issue 3, Pages 254-267, **2024**.

[27] KN Kadhim, SMR Taha, WA Mahmoud "A new method for filtering and segmentation of the ECG signal", Proceedings of the Annual International Conference of the IEEE Engineering …, 1988.

[28] Waleed A. Mahmoud, MS Abdulwahab, HN Al-Taai: "The Determination of 3D Multiwavelet Transform" IJCCCE, vol. 2, issue 4, 2005.

[29] Waleed A. M. Al-Jawher, T Abbas – "Feature combination and mapping using multiwavelet transform" IASJ, AL-Rafidain, Issue 19, Pages 13-34, 2006.

[30] Waleed A Mahmoud, MR Shaker "3D Ear Print Authentication using 3D Radon Transform" proceeding of 2nd International Conference on Information & Communication Technologies, Pages 1052-1056, 2006.

[31] W. A. Mahmoud & Z. Ragib "Face Recognition Using PCA and Optical Flow" Engineering Journal, Vol. 13, Issue 1, PP. 35-47, 2007.

[32] Waleed A Mahmoud, Ahmed S Hadi "Systolic Array for Realization of Discrete Wavelet Transform " Journal of Engineering, Vol. 13, Issue 2, PP. 1-9, 2007.

[33] Walid Amin Al-Jawhar, Ayman M Mansour, Zakaria M Kuraz "Multi technique face recognition using PCA/ICA with wavelet and Optical Flow" 2008 5th International Multi-Conference on Systems, Signals and Devices, pages 1-6, 2008.

[34] Adnan HM Al-Helali, Hamza A Ali, Buthainah Al-Dulaimi, Dhia Alzubaydi, Waleed A Mahmoud "Slantlet transform for multispectral image fusion" Journal of Computer Science, Vol.5, Issue 4, PP. 263-267, 2009.

[35] AHM Al-Heladi, WA Mahmoud, HA Hali, AF Fadhel "Multispectral Image Fusion using Walidlet Transform" Advances in Modelling and Analysis B, Volume 52, Iss. 1-2, pp. 1-20, 2009.

[36] Walid A Mahmoud, Majed E Alneby, Wael H Zayer "2D-multiwavelet transform 2D-two activation function wavelet network-based face recognition" J. Appl. Sci. Res, vol. 6, issue 8, 1019-1028, 2010.

[37] WA Mahmoud, ALM Rasheed "3D Image Denoising by Using 3D Multiwavelet" AL-Mustansiriya J. Sci 21 (7), 108-136, 2010.

[38] M Walid A Mahmoud, Majed E Alneby, Wael H Zayer "Multiwavelet Transform And Multi-Dimension-Two Activation Function Wavelet Network Using For Person Identification" Iraqi Journal Of Computers, Communications, Control And Systems Engineering, Vol 11, Issue 1, 2011.

[39] AHM Al-Helali, Waleed A. Mahmoud, HA Ali "A Fast personal palm print authentication Based on 3d-multi–Wavelet Transformation", Transnational Journal Of Science And Technology, Vol. 2, Issue 8, 2012.

[40] L R. Hussssein andJ. M. A. Al-Sammarie W. A. Mahmoud "Image Identification using Minimum Distance Classifier with Multi-Wavelet Transform" Journal of Advances in Modelling and Analysis B, Volume 46, Issue (5-6), pages 1-22, 2003.

[41] Waleed Ameen Mahmoud Al-Jawher, A. Barsoum and Entather Mahos "Fuzzy Wavenet (FWN) classifier for medical images" Al-Khwarizmi Engineering Journal, Vol. 1, Issue 2, PP. 1-13, 2005.

[42] Mikolov, T.; Sutskever, I.; Chen, K.; Corrado, G. S.; Dean, J. "Distributed representations of words and phrases and their compositionality" Advances in Neural Information Processing Systems, pp. 3111-3119, (2013):

[43] W. A. Mahmoud & Ommama Razaq "Speech recognition using new structure for 3D neural network" University of Technology, 1st Computer Conference, PP. 161-171, 2010.

[44] Waleed Ameen Mahmoud "A Smart Single Matrix Realization of Fast Walidlet Transform" Journal of Research and Reviews in Computer Science, Volume 2, Issue, 1, PP 144-151, 2011.

[45] Hamid M Hasan, AL Jouhar, Majid A Alwan "Face recognition using improved FFT based radon by PSO and PCA techniques" International Journal of Image Processing (IJIP), Volume 6, Issue 1, Pages 26-37, 2012.

[46] WA Mahmoud, AI Abbas, NAS Alwan "Face Identification Using Back-Propagation Adaptive Multiwavelet" Journal of Engineering 18 (3), 2012

[47] WA Mahmoud, AS Hadi, TM Jawad "Development of a 2-D Wavelet Transform based on Kronecker Product" - Al-Nahrain Journal of Science, Vol. 15, Issue 4, PP. 208-213, 2012.

[48] H. M Hasan, Waleed A. Mahmoud Al- Jawher, M. A Alwan "3-d face recognition using improved 3d mixed transform" Journal International Journal of Biometrics and Bioinformatics, Vo. 6, Issue 1, PP. 278-290, 2012.

j. port. sci. res.
ISSN: 2616-7441 (Online)
ISSN: 2616-6232 (Print)
ISSN: 2616-7220 (USB)
INTERNATIONAL IDENTIFIER FOR SERIALS

Journal port Science Research
Available online www.jport.co
Volume 7, issue 3. 2024

[49] Waleed A. Mahmud Al-Jawher, Talib M. J. Abbas Al-Talib, R. Hamudi A. Salman "Fingerprint Image Recognition Using Walidlet Transform" Australian Journal of Basic and Applied Sciences, Australia, 2012.

[50] Saleem MR Taha, Walid A Mahmood "New techniques for Daubechies wavelets and multiwavelets implementation using quantum computing "Journal Facta universitatis-series: Electronics and Energetics, Volume 26, Issue 2, Pages 145-156, 2013.

[51] Waleed A Mahmoud, Dheyaa J Kadhim "A Proposal Algorithm to Solve Delay Constraint Least Cost Optimization Problem" Journal of Engineering, Vol. 19, Iss 1, PP 155-160, 2013.

[52] SMR Taha, WA Mahmood "New techniques for Daubechies wavelets and multiwavelets implementation using quantum computing" Facta universitatis-series: Electronics and Energetics 26 (2), 145-156, 2013.

[53] Ibraheem Al-Jadir, Waleed A Mahmoud "A Grey Wolf Optimizer Feature Selection Method and its Effect on the Performance of Document Classification Problem" Journal Port Science Research, Vol. 4, Issue 2, Pages 125-131, 2021.

[54] Sarah H Awad Waleed A Mahmoud Al-Jawher "Precise Classification of Brain Magnetic Resonance Imaging (MRIs) using Gray Wolf Optimization (GWO)" HSOA Journal of Brain & Neuroscience Research, Volume 6, Issue 1, Pages 100021, 2022.

[55] Maryam I Mousa Al-Khuzaay, Waleed A Mahmoud Al-Jawher, "New Proposed Mixed Transforms: CAW and FAW and Their Application in Medical Image Classification" International Journal of Innovative Computing, Volume 13, Issue 1-2, Pages 15-21, 2022.

[56] Waleed A. Mahmoud Al-Jawher, SH Awad "A proposed brain tumor detection algorithm using Multi wavelet Transform (MWT)", Materials Today: Proceedings 65, 2731-2737, 2022.

[57] Hamsa A Abdullah, Hikmat N Abdullah, Waleed A Mahmoud Al-Jawher, "A hybrid chaotic map for communication security applications" " International Journal of Communication Systems, Vol. 33, Issue 4, 2020.

[58] QK Abed, Waleed A. Mahmoud Al-Jawher, "A new Architecture Of Key Generation Using Dwt For Image Encryption With Three Levels Arnold Transform Permutation" Journal Port Science Research 5 (3), 166–177-166–177, 2022.

[59] Ali Akram Abdul-Kareem, Waleed Ameen Mahmoud Al-Jawher, "WAM 3D discrete chaotic map for secure communication applications" International Journal of Innovative Computing, Volume 13, Issue 1-2, Pages 45-54, 2022.

[60] Q. K Abed, W. A Mahmoud Al-Jawher "A Robust Image Encryption Scheme Based on Block Compressive Sensing and Wavelet Transform" International J. of Innovative Computing, Vol. 13, I. 1-2, PP. 7-13, 2022.

[61] Ali Akram Abdul-Kareem, Waleed Ameen Mahmoud Al-Jawher, "Image Encryption Algorithm Based on Arnold Transform and Chaos Theory in the Multi-wavelet Domain", International Journal of Computers and Applications, Vol. 45, Issue 4, pp. 306-322, **2023**

[62] Ali Akram Abdul-Kareem, Waleed Ameen Mahmoud Al-Jawher "A Hybrid Domain Medical Image Encryption Scheme Using URUK and WAM Chaotic Maps with Wavelet–Fourier Transforms" Journal of Cyber Security and Mobility, Pages 435–464-435–464, **2023.**

[63] Zahraa A Hasan, Suha M Hadi, Waleed A Mahmoud, "Speech scrambler with multiwavelet, Arnold Transform and particle swarm optimization" Journal Pollack Periodica, Volume 18, Issue 3, Pages 125-131**, 2023**.

[64] AA Abdul-Kareem, Waleed A. Mahmoud Al-Jawher "Uruk 4d Discrete Chaotic Map For Secure Communication Applications" Journal Port Science Research 5 (3), 131-142., **2023**.

[65] ZA Hasan, SM Hadi, WA Mahmoud, "Time domain speech scrambler based on particle swarm optimization", Pollack Periodica 18 (1), 161-166, 2023.

[66] AA Abdul-Kareem, Waleed A. Mahmoud Al-Jawher "Hybrid image encryption algorithm based on compressive sensing, gray wolf optimization, and chaos" Journal of Electronic Imaging 32 (4), 043038-043038, **2023.**

[67] Zahraa A Hasan 1, Suha M. Hadi, Waleed A. Mahmoud al-Jawher "Speech scrambling based on multiwavelet and Arnold transformations", Indonesian Journal of Electrical Engineering and Computer Science, 30 , **2023**.

[68] Ali Akram Abdul-Kareem, Waleed Ameen Mahmoud Al-Jawher, "Image Encryption Algorithm Based on Arnold Transform and Chaos Theory in the Multi-wavelet Domain", International Journal of Computers and Applications, Vol. 45, Issue 4, pp. 306-322, **2023**.

[69] Q. K Abed, Waleed A. Mahmoud Al-Jawher "An Image Encryption Method Based on Lorenz Chaotic Map and

Hunter-Prey Optimization" Journal Port Science Research, Volume 6, Issue 4, Pages 332-343, **2023**.

[70] Zahraa A Hasan, Suha M Hadi, Waleed A Mahmoud Al-Jawher "Speech scrambling based on fan transform technology" Publication date 2023/12/22, Journal AIP Conference Proceedings, Volume 2977, Issue 1, **2023**.

[71] Ali Akram Abdul-Kareem, Waleed Ameen Mahmoud Al-Jawher " An image encryption algorithm using hybrid sea lion optimization and chaos theory in the hartley domain" International Journal of Computers and Applications, Vol. 46, Issue 5, PP. 324-337, **2024**.

[72] Qutaiba K Abed, Waleed A Mahmoud Al-Jawher "Optimized Color Image Encryption Using Arnold Transform, URUK Chaotic Map and GWO Algorithm" Journal Port Science Research, Vol. 7, Issue 3, PP. 210-236, **2024**.

[73] Qutaiba Kadhim, Waleed Ameen Mahmoud Al-Jawher "A new multiple-chaos image encryption algorithm based on block compressive sensing, swin transformer, and wild horse optimization" Multidisciplinary Science Journal, Vol. 7, Issue 1, PP. 2025012-2025012, **2024**.

[74] SA Shaban, WAM Al-Jawher, "K-Means Clustering Algorithm for Medical Images," International Journal of Advances in Engineering and Management (IJAEM) 4 …m 2022.

[75] L. F. Katran, Ebtesam N AlShemmary, Waleed Ameen Al Jawher "Deep Learning's Impact on MRI Image Analysis: A Comprehensive Survey" Texas Journal of Engineering and Technology, Vol. 25, PP. 63-80, **2023**.

[76] AAR Sakran, SM Hadi, Waleed A. Mahmoud Al-Jawher "A new approach for DNA sequence analysis using multiwavelet transform (MWT)" Journal of Physics: Conference Series 2432 (1), 012022, **2023.**

[77] Lamyaa Fahem Katran, Ebtesam N AlShemmary, Waleed Ameen Al Jawher "Deep Learning's Impact on MRI Image Analysis: A Comprehensive Survey" Texas Journal of Engineering and Technology, Vol. 25, PP. 63-80, **2023**.

[78] Ammar AbdRaba Sakran, Waleed A. Mahmoud Al-Jawher, Suha M Hadi "The Role of Three-Base Periodicity in Enhancing Exon Detection through DSP Techniques" 2023 16th International Conference on Developments in eSystems Engineering (DeSE), Publication date 2023/12/18, Pages 617-622, Publisher IEEE, **2023**.

[79] Ammar A. Sakran, Waleed A. Mahmoud Al-Jawher, Suha M Hadi "A Review of Digital Signal Processing Techniques for Protein-Coding Region Identification in DNA Sequences " 2023 16th Inte. Conference on Developments in eSystems Engineering (DeSE), Publication date 2023/12/18, Pages 611-616, Publisher IEEE, **2023**.

[80] Shaymaa Abdulelah Shaban, Waleed AM Al-Jawher 'Image compression techniques using multiwavelet transform algorithm" Publication date 2024/4/11, Journal of AIP Conference Proceedings, Volume 3079, Issue 1, **2024**.

[81] W. A. Mahmoud & Z. J. M. Saleh " An Algorithm for Computing Multiwavelets &Inverse Transform Using an Over-Sampled Scheme of Pre& Post processing respectively" Engineering Journal, Vol. 10, Issue 2, PP. 270-288, 2004.

[82] W. A. Mahmoud Z Jalal & N. K. Wafi "A New Method of Computing Multi-wavelets Transform using Repeated Row Preprocessing." Al-Rafidain Engineering Journal, Vol. 12, Issue 2, PP. 21-31., 2004.

[83] W. A. Mahmoud & I. A Al-Akialy "A Tabulated Method of Computation Multiwavelet Transform" Al-Rafidain University College, Vol. 15, PP. 161-170, Iraq, 2004.

[84] H. Al-Taai, Waleed A. Mahmoud & M. Abdulwahab "New fast method for computing multiwavelet coefficients from 1D up to 3D" , Proc. 1st Int. Conference on Digital Comm. & Comp. App., Jordan, PP. 412-422, 2007.

[85] A. AR. Sakran, SM Hadi, Waleed A. Mahmoud Al-Jawher "A New Approach for DNA Sequence Analysis Using Multiwavelet Transform (MWT)" Journal of Physics: Conference Series 2432 (1), 012022.

[86] Haqi Khalid, Shaiful Jahari Hashim, Fazirulhisyam Hashim, Waleed Ameen Mahmoud Al-Jawher, Muhammad Akmal Chaudhary, Hamza HM Altarturi "RAVEN: Robust Anonymous Vehicular End-to-End Encryption and Efficient Mutual Authentication for Post-Quantum Intelligent Transportation Systems" Publication date 2024/6/24, Journal IEEE Transactions on Intelligent Transportation Systems, Publisher IEEE, **2024**.

[87] Jaafar Ahmed Abdulsaheb, Raghad M. Nafea, Waleed A. Mahmoud Al-Jawher, Mohammed Lateef Hayyawi "IoT Based Smart Parking System" Authors Publication date 2024/6/28, Journal Journal Port Science Research, Volume 7, Issue 3, Pages . 196-203, 2024.

[88] Lamyaa Fahem Katran, Ebtesam N AlShemmary, Waleed AM Al-Jawher "A Review of Transformer Networks in MRI Image Classification" Al-Furat Journal of Innovations in Electronics and Computer Engineering, PP. 148-162, **2024**.

[89] Waleed A Mahmoud Al-Jawher, Shaimaa A Shaaban "K-Mean Based Hyper-Metaheuristic Grey Wolf and Cuckoo Search Optimizers for Automatic MRI Medical Image Clustering" Journal Port Science Research, Volume 7, Issue 3, Pages 109-120, **2024**.

[90] Abbas Hasan Kattoush, Waleed A. Mahmoud Al-Jawhar, A Ghodayyah "Multiwavelet computed Radon-based OFDM transceiver designed and simulation under different channel conditions" Journal of Information & Computing Science, Volume 5, Issue 2, Pages 133-145 2010.

[91] Salih M Salih, N Uzunoglu, AA Ali, Waleed A Mahmoud "A proposed improvement model of MC-CDMA based FFT in AWGN channel" 2007 IEEE International Conference on Electro/Information Technology, Pages 517-520, 2007.

[92] Saadi M Saadi, Waleed A. Mahmoud Al-Jawher " Proposed DeepFake Detection Method Using Multiwavelet Transform" International Journal of Innovative Computing 13 (1-2), 61-66, 2022.

[93] A. U Mosa, Waleed A Mahmoud Al-Jawher "Image Fusion Algorithm using Grey Wolf optimization with Shuffled Frog Leaping Algorithm" International J. of Innovative Computing, Vol. 13, Issue 1-2, PP. 1-5. 2022.

[94] Waleed A. Mahmoud Al-Jawher, SAA SHABAN "Clustering Of Medical Images Using Multiwavelet Transform And K-Means Algorithms" , Journal Port Science Research 5 (1), 35-42, 2022.

[95] WA Mahmoud "Computation of Wavelet and Multiwavelet Transforms Using Fast Fourier Transform" Journal Port Science Research 4 (2), 111-117, 2021.

[96] Shaymaa Abdulelah Shaban, Waleed A Mahmoud Al-Jawher "K-Means Clustering Algorithm for Medical Images" International Journal of Advances in Engineering and Management (IJAEM), Vol. 4, Issue 11, 2022.

[97] Ammar A Sakran, Suha M Hadi, Waleed A Mahmoud Al-Jawher "Advancing DNA Signal Processing: Integrating Digital and Biological Nuances for Enhanced Identification of Coding Regions" Journal Port Science Research, Volume, 6, Issue 4, Pages 374-387, **2023**.

[98] Johnson, D. E.; Oles, F. J.; Zhang, T.; Goetz, T. (2002) "A decision-tree-based symbolic rule induction system for text categorization' IBM Systems Journal, vol. 41, no. 3, pp. 428-437, 2002.

[99] Walid A Al-Jowher, Nada N Al-Ramahi, Mikhled. Alfaouri "Image Identification And Labeling Using Hybrid Transformation And Neural Network" Neural Network World: International Journal on Neural and Mass - Parallel Computing and Information Systems; Prague, Volume 17, Issue 4, Pages 377-395, 2007.