Comparative Performance of Ensemble Machine Learning for Arabic Cyberbullying and Offensive Language Detection

Marwa Khairy
Minia University
Tarek M. Mahmoud
Minia University Faculty of Science
Ahmed Omar
Minia University Faculty of Science
Tarek Abd El-Hafeez (tarek@mu.edu.eg)
Minia University Faculty of Science https://orcid.org/0000-0003-1785-1058

Research Article

Keywords: OSN, Arabic Offensive Language, Cyberbullying, NLP, Ensemble Machine learning

Posted Date: June 22nd, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1730412/v1

Abstract

In recent years, research on abusive language and cyberbullying detection have gained a great deal of interest since it affects both individual victims and societies. Hateful communications, bullying, sexism, racism, aggressive content, harassment, toxic comments, and other forms of abuse have all increased dramatically due to the ease of access to social media platforms such as Facebook, Instagram, Twitter, and others. As a result, there is a significant need to identify, manage, and restrict the spread of offensive content on social networking sites, prompting us to perform this study to automate the detection of offensive language or cyberbullying. Having a balanced data set for a model would generate higher accuracy models, thus we build a new Arabic balanced data set to be used in the process of offensive detection. Lately, Ensemble Machine Learning has been used to enhance the performance of single classifiers. The aim of this study is to compare the performance of different single and ensemble machine learning algorithms in detecting Arabic text containing cyberbullying and offensive language. For this purpose, we have chosen three machine learning classifiers and three ensemble models and apply them to three Arabic datasets two of them are offensive datasets that are publicly available, and the third one which we constructed. The results showed that the ensemble machine learning methodology outperforms the single learner machine learning approach. Voting performs is the best among the trained ensemble machine learning classifiers, with accuracy scores of (71.1%, 76.7%, and 98.5%) for the three used datasets respectively, exceeding the score obtained by the best single learner classifier (65.1%, 76.2%, and 98%) for the same datasets. Finally we use hyperparameters tunning on the Arabic cyberbylluing data set to optize the performance of the voting technique.

1. Introduction

For many users, online social networks (OSNs) are becoming the most prevalent and interactive media. The majority of individuals use social media without considering the impact these networks have on our lives, whether beneficial or harmful. However, along with valuable and interesting content, these networks can also broadcast inappropriate or harmful content, such as cyberbullying, hate speech, and insults [22] [23]. The detection of such language is essential because it may cause emotional distress and affect the mental health of social media users [1].

The Arabic language is the fifth most spoken in the world, with more than 420 million speakers [2]. The use of the Arabic language in social media is widespread and continually increasing. As of 2017, the Arabic Social Media Report estimates that Facebook users from the Arab region constitute 8.4% of all Facebook users which is more than 150 million Arab users [3]. Because the linguistic format might be sophisticated or slang, classifying Arabic social media texts is a difficult task. The Arabic language has multiple dialects with different lexicons and structures, making high-performance classification difficult.

Ensemble Machine Learning is a machine learning methodology that integrates multiple distinct prediction models into a single model to improve performance. It has to be considered whenever good predictive accuracy is demanded [4]. In addition to Ensemble classifiers have been shown to be more

effective than data resampling techniques to enhance the classification performance of imbalanced data [25]. Having a balanced data set for a model would generate higher accuracy models, higher balanced accuracy, and a balanced detection rate. The results obtained by Qiong W. et. al [26] demonstrates that using balanced training data (50% neutral and 50% deleterious) results in the highest balanced accuracy (the average of True Positive Rate and True Negative Rate). Hence, it is important to have a balanced data set for a classification model. Thus, we build a new Arabic balanced data set to be used in the process of offensive detection. Also, we will apply Ensemble Machine Learning models on detecting Arabic offensive posts for both balanced and imbalanced datasets. We have used several Machine Learning Models and Feature Extraction techniques to compare the performance of using a Single ML Classifier and ensemble ML in Arabic abusive language and cyberbullying detection.

The remainder of the paper is structured as follows: We discuss related work in section 2. After that a background covering the Ensemble Machine Learning models in section 3. Section 4 presents the proposed method to Arabic Abusive language and cyberbullying Detection. Finally, results and discussion will be presented in section 5.

2. Related Work

Research on Arabic abusive language detection has recently drawn much attention; Hamdy et al. have many efforts in the Arabic language field, especially in Offensive Language detection. In [6], they show how to use popular trends in offensive and rude communications to build and extend a list of offensive words and hashtags using an automated tool. Also, Twitter users were ranked based on whether or not they use any of these offensive terms in their tweets. Using this classification, they expand the list of bad words and present the results on a newly created dataset of labeled Arabic tweets (obscene, offensive, and clean). Also, they publish a large corpus of classified user comments publicly, which were removed from a famous Arabic news site due to rule violations and guidelines of this site. In [14], they present a rapidly creating training dataset for identifying offensive tweets using a seed list of offensive words. They trained a deep learning classifier based on charterer n-grams that can efficiently classify tweets with a 90% F1 score. They recently make a new dialectal Arabic news comment dataset public, which was collected from a variety of social media platforms, including Twitter, Facebook, and YouTube. In abusive comments, they investigate the unique lexical content in connection with the use of Emojis. The results show that the data set model of multi-platform news commentary can capture diversity in various dialects and domains. In addition to evaluating the models' generalization power, they also presented an in-depth analysis of emojis usage in offensive comments. Findings suggest emojis in the animal category are exploited in offensive comments, like lexical observation [15].

Abozinadah E. et al. [7] evaluates various machine learning algorithms for detecting abusive accounts with Arabic tweets, the data set for this analysis was collected based on the top five Arabic swearing words, from the total result we ended up having 255 unique users. Naïve Bayes (N.B.) classifier with 10 tweets and 100 features has the best performance with a 90% accuracy rate. An Arabic word correction method was also suggested in [8] to tackle internet censorship systems and content-filtering

vulnerabilities. This method achieved an accuracy of 96.5%. A statistical learning approach was used in [9] to detect adult accounts with the Arabic language in social media. The uses of obscenity, vulgarity, slang, and swear words in Arabic content on Twitter were examined in order to identify abusive accounts. With this statistical method, a predictive precision of 96% was achieved, and the imitations of the bag-of-word (BOW) approach were overcome.

Alakrot et al. [10] build an Arabic dataset of YouTube comments to detect offensive language in a machine learning context. The data were collected by the principles of availability, variety, representativeness, and balance, thus ensuring that predictive analytical models for identifying the abusive language in online communication in Arabic can be implemented for training. Lately, Haddad B. et al. [11] address the issue of abusive language and hate speech detection. They suggest a method for data pre-processing and rebalancing, and then they used the bidirectional Gated Recurrent Unit (GRU) and Convolutional Neural Network (CNN) models. The bidirectional GRU model augmented with attention layer generated the best results among proposed models on a labeled dataset of Arabic tweets were achieved 85.9%F1 scores for offensive language detection and 75% F1 scores for the purpose of detecting hate speech.

Ensemble machine learning is a powerful machine learning algorithm, the result obtained from an ensemble, a combination of machine learning models can be more accurate than any single member of the group [5].

Regarding using Ensemble machine learning methods in Arabic offensive language and cyberbullying detection; Haidar et al. [12] present Ensemble-Machine-Learning as another solution for Arabic cyberbullying detection to enhance their previous work. They accomplish an improvement on Precision, Recall, and F1-Score. Also, Fatemah, H. [13] examined the impact of applying a single learner machine learning approach (SVM, logistic regression, and decision tree) and ensemble machine learning approach (Bagging, AdaBoost, and random forest) on Arabic offensive language detection. The study shows that applying ensemble machine learning techniques rather than single learner machine learning approaches has a significant effect. With an F1 of 88%, which exceeds the best single learner classifier score by 6 percent, Bagging performs the best among the qualified ensemble machine learning classifiers in offensive language detection.

3. Ensemble Machine Learning

Ensemble Machine Learning, a strong machine learning technology that is used by data science experts in industries as it is considered the state-of-the-art solution for many machine learning problems [16]. The result obtained from an ensemble, a combination of machine learning models, can be more accurate than any single member of the group [4] [5]. See Fig. 1.

The three most popular methods for combining the predictions from different models are [16] [17]:

1. Bagging: Stands for Bootstrap Aggregation which takes multiple samples from the training dataset and several weak models, aggregating the predictions to select the best prediction. It has models such as Bagged Decision Trees, Random Forest.
2. Boosting: is a general ensemble technique that combines a number of weak classifiers to produce a strong classifier. This is accomplished by developing a model from the training data, then attempting to fix the faults in the first model with a second model. AdaBoost and Stochastic Gradient Boosting are the two most popular boosting ensemble machine learning techniques.
3. Voting: is a model that combines the predictions from multiple other and simple statistics are used to combine predictions models to achieve better performance.

4. Proposed Method

In this section, the proposed method is discussed in detail with a diagram. The datasets, the preprocessing, the classification methods used, and the performance measures are described below. Figure 2 illustrates the method that we follow in this study.

4.1 Dataset Description

In this paper three Arabic datasets are used, two are offensive datasets that are publicly available at [18] and the third one is a balanced dataset we decide to collect. We used a set of offensive keywords that represent different types of offensive and cyberbullying meaning, we used these keywords to search for tweets on Twitter and posts on Facebook, then we developed a web crawler to collect the Facebook search results automatically. In Twitter, we use Twitter API to collect the tweets generated by the search. Final all the collected posts and tweets are stored in a text file without repetition.

Non-Arabic letters, URLs, and emoji are examples of non-useful text that have a negative impact on categorization performance. As a result, the gathered data must be cleaned and filtered. The filtering phase removes all non-useful text. For the newly collected dataset, we manually labeled the filtered posts to 1 for cyberbullying and 0 for non-cyberbullying posts. At the end of the preprocessing, the final dataset contained 6,000 non- cyberbullying instances and 6,000 cyberbullying ones. See Fig. 3. After finishing the annotation process the data had to be tokenized and stemmed.
Table 1 gives the original distribution of the used datasets in terms of the source, size of the dataset, number of their majority and minority instances, and their imbalance ratio (IR).

Table 1
The used Arabic Datasets

Dataset	Source	Size	Class 0	Class 1	Imbalanced Proportion
Arabic1 [19]	Twitter	1,100	453	647	0.7: 1
Arabic2 [20]	YouTube	8,577	5332	3245	1.64: 1
The proposed Dataset	Facebook & Twitter	12,000	6,000	6,000	1: 1

4.2 Classification Method

After splitting the dataset (80% training and 20% testing), we generate the 'TF-IDF' and passing the features to our Classification method which is based on two approaches. The first approach is Single Machine Learning Model (SML). The second one is Ensemble Machine Learning Model (EML). To evaluate the model's performance, we use various metrics (Accuracy, F1, Recall, and Precision). All models are implemented using the Python library Scikit-learn.

4.2.1 Single Learner Approach (SML)

In this approach, we employ three single machine learning classification classifiers (Linear Support Vector (Linear SVC), logistic regression (LR), and KNeighbors (KNN)).

4.2.2 Ensemble machine learning approach (EML)

For this approach, we select the three different ensembles machine learning models (bagging, voting, and boosting). For the bagging we select the random forest model, for the boosting, we select the AdaBoost method. For the voting, we used the three single learners used in the first approach. We select three different models each use a different ensemble machine learning method. The random forest was trained using 100 maximum numbers of trees. Hard voting was used to create the ensemble because it can validate the result with more confidence than individually applied algorithms [19].

- # 4.3 Performance Measures

The metrics measured used to analyze the performance of each classifier are Accuracy, Precision, Recall, and F1_score. The definitions of those metrics are as follows:

A. Accuracy: The accuracy is the percentage of instances that were correctly classified into their respective classes. It is also called sample accuracy.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

1

B. Precision: Precision is used to measure the exactness of the classifier. Moreover, it refers to the fraction of predicted positive which are actually positive. The formula for precision is the number of positive predictions divided by the total number of positive class values predicted.

$$\text{Precision} = \frac{\text{TP}}{(\text{TP} + \text{FP})}$$

2

C. Recall: refers to the fraction of those that are actually positive that were predicted as positive. The formula for the recall is the number of positive predictions divided by the number of positive class values in the test data.

$$\text{Recall} = \frac{\text{TP}}{(\text{TP} + \text{FN})}$$

3

D. The F-measure (or F-score): is used to measure the accuracy of the test by considering both precision and recall in computing the score. It conveys a balance between precision and recall wherein it reaches its best value at 1 and its worst value at 0.

- $$\text{F1\_Score} = \frac{(2 \times \text{Precision} \times \text{Recall})}{(\text{Precision} + \text{Recall})}$$

  4

5. Experimental Results And Discussion

This section presents the results of the experiments, which were carried out to compare the performance of different single and ensemble machine learning algorithms in detecting Arabic messages containing cyberbullying and abusive language.

Table 2 have shown the performance metrics for single and ensemble ML models for the first Arabic dataset. From Table 2 we can notice that Voting outperforms all used single and ensemble classifiers in all performance metrics. Logistic Regression (LR) outperforms the other single classifiers with an Accuracy of 65.1%. After that Linear SVC and KNeighbors achieved an Accuracy of 64.2%.

The Voting EML classifier outperforms the other EML classifiers with an Accuracy of 71.1%. After that, the Bagging EML achieved an accuracy of 60.5% and AdaBoost EML of 56.4%. These findings prove the efficiency of using ensemble machine learning methods; Accuracy rises from 65.1% using the best SML classifier to 71.1% using the best EML model. Our results are inconsistent with [13].

Table 2
Comparative Analysis of Single and Ensemble ML Classifiers Performance for Arabic1 Dataset

Arabic1 Dataset						
Performance Metrics	**Single Learner**			**Ensemble Learner**		
	Logistic Regression	KNeighbors	Linear SVC	Bagging (Random Forest)	Voting	Boosting (Adaboost)
Accuracy	**65.1**	64.2	64.2	60.5	**71.1**	56.4
F1_score	**62.7**	64.1	63.4	58.6	**70.9**	56
Recall	**65.1**	64.2	64.2	60.5	**71.1**	56.4
Precision	**64.3**	64	63.4	58.4	**70.9**	64.6

Table 3 have shown the performance metrics for single and ensemble ML models for the second Arabic dataset. From Table 3 we can notice that Voting also outperforms all used single and ensemble classifiers for all performance measures. From Table 1 we can notice that this dataset is partially imbalanced, so we will use the F1 _score to evaluate the performance instead of using accuracy as it can be misleading.

The SML classifier (Linear SVC) outperforms the other SML classifiers with F1_score of 75.2% then the KNeighbors (KNN) achieved F1_score of 74.5% and the Logistic Regression (LR) with F1_score of 74.4%. The voting classifier outperforms the other EML with F1_score of 75.8%. After that the Bagging EML achieved F1_score of 72.8% then the Boosting EML achieved F1_score of 70.2%. These findings prove the efficiency of using ensemble machine learning methods; F1_score rises from 75.2% using the best SML model to 75.8% using the best EML model. Our results are inconsistent with [13].

Table 3
Comparative Analysis of Single and Ensemble ML Classifiers Performance for Arabic2 Dataset

Arabic2 Dataset						
Performance Metrics	**Single Learner**			**Ensemble Learner**		
	Logistic Regression	KNeighbors	Linear SVC	Bagging (Random Forest)	Voting	Boosting (Adaboost)
Accuracy	76.2	75.4	76.1	75.3	76.7	73
F1_score	74.4	74.5	**75.2**	72.8	**75.8**	70.2
Recall	76.2	75.4	76.1	75.3	76.7	73
Precision	76.5	75.3	75.8	77.2	78	74.4

Table 4 shows the performance metrics for single and ensemble ML models for the new Arabic dataset. Our new dataset was balanced so we will consider the Accuracy for comparing the results. From Table 4 we can notice that Voting outperforms all used single and ensemble classifiers in all performance metrics.

Linear SVC classifier outperforms the other SML Classifiers with an Accuracy of 98. After that the Logistic Regression (LR) achieved an Accuracy of 97.4% then KNeighbors (KNN) achieved Accuracy of 96.3%. The voting classifier outperforms the other EML Classifiers with an accuracy of 98.5%. After that, the Bagging achieved an accuracy of 96.1%. % then the Boosting achieved an accuracy of 94.8%. These findings prove the efficiency of using ensemble machine learning methods; Accuracy rises from 98% using the best SML model to 98.5% using the best EML model. Our results are inconsistent with [13].

Table 4
Comparative Analysis of Single and Ensemble ML Classifiers Performance for New Arabic Dataset

New Dataset						
Performance Metrics	**Single Learner**			**Ensemble Learner**		
	Logistic Regression	KNeighbors	Linear SVC	Bagging (Random Forest)	Voting	Boosting (Adaboost)
Accuracy	97.4	96.3	**98**	96.1	**98.5**	94.8
F1_score	97.2	96.2	**97.8**	96	**98.3**	94.5
Recall	97.4	96.3	**98**	96.1	**98.5**	94.8
Precision	97.5	96.4	**98.2**	96.2	**98.7**	95.1

Our findings partly support the hypothesis that ensemble models naturally do better in comparison to single classifiers, but not in all cases. In some cases, Logistic Regression, which is a single classifier can achieve better results than ensemble classifier (Bagging and Boosting). It is dependent on the characteristics of the data set being examined.

5.1 Hyperparameter tuning

To optimize the previous results, which obtained by using the default parameters for each classifier, we apply hyperparameter tuning in the best obtained results (Voting). The best parameters are shown in Table 5:

Table 5
Best parameters obtained by Hyperparameter tuning

Classifier	parameter	Value
KNeighbors	n_neighbors	**8**
Logistic Regression	C	**20**
Linear SVC	C	**20**
And the optimized performance is shown in Table 6.		

Table 6
Comparison between Best performance obtained by Hyperparameter tuning and default values

Classifier	Tuning	Default
Accuracy	**98.6**	**98.5**
F1_score	**98.4**	**98.3**
Recall	**98.6**	**98.5**
Precision	**98.8**	**98.7**

6. Conclusion

Ensemble machine learning is a meta-learning machine learning method that aims to improve single learner classifier's performance by combining predictions from multiple single learner classifiers. In this study, we investigate the effect of applying three single learner machine learning approach (Linear SVC, logistic regression, and KNeighbors) and three ensemble machine learning approach (bagging- random forest, Voting, and Boosting-Adaboost) on offensive language and cyberbullying detection for the Arabic language.

The ensemble machine learning methodology outperforms the single learner machine learning approach in terms of impact. Among the trained ensemble machine learning classifiers, Voting performs the best in offensive language and cyberbullying detection with an Accuracy score of (71.1%, 76.7%, and 98.5%) for the three used datasets respectively which exceeds the score obtained by the best single learner classifier (65.1%, 76.2%, and 98%) for the three used datasets respectively. Finally, we used hyperparameters tuning to optimize the performance of the voting approach and the accuracy become 98.6%.

Declarations

Funding: Not applicable.

Availability of data and material and code : https://github.com/omammar167/Arabic-Abusive-Datasets

Disclosure of potential Conflict of Interest: The authors declare that they have no conflict of interest.

Ethical Statement: “All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.”

Consent Statement: “Informed consent was obtained from all individual participants included in the study.”

References

[1] M. Mengü and S. Mengü, "Violence and Social Media," Athens Journal of Mass Media and Communications, vol. 1, pp. 211-228, 2015.

[2] http://istizada.com/complete-list-of-arabic-speaking-countries-2014/, Accessed 16-2-2021.

[3] F. Salem, "The Arab Social Media Report 2017: Social Media and the Internet of Things: Towards Data-Driven Policymaking in the Arab World," MBR School of Government, Dubai, 2017.

[4] Džeroski S., Panov P., Ženko B. (2009) Machine Learning, Ensemble Methods in. In: Meyers R. (eds) Encyclopedia of Complexity and Systems Science. Springer, New York, NY. https://doi.org/10.1007/978-0-387-30440-3_315

[5] Dietterich T.G. (2000) Ensemble Methods in Machine Learning. In: Multiple Classifier Systems. MCS 2000. Lecture Notes in Computer Science, vol 1857. Springer, Berlin, Heidelberg. https://doi.org/10.1007/3-540-45014-9_1

[6] Hamdy Mubarak, Kareem Darwish, and Walid Magdy. 2017. Abusive language detection on Arabic social media. In Proceedings of the First Workshop on Abusive Language Online. Vancouver, Canada, pages 52–56.

[7] Abozinadah, E. A., Mbaziira A. V, Jones, J. H. ., Jr, (2015) "Detection of abusive accounts with Arabic tweets," Int. J. Knowl. Eng., 1(2): 113–119.

[8] Abozinadah, E. A., Jones J. H., Jr, (2016) "Improved Micro-Blog Classification for Detecting Abusive Arabic Twitter Accounts". Int. J. Data Min. Knowl. Manag. Process, 6(6):17–28

[9] Abozinadah E. A., Jones, J. H., (2017) "A Statistical Learning Approach to Detect Abusive Twitter Accounts," Proc. Int. Conf. Comput. Data Anal. - ICCDA '17, pp. 6–13.

[10] Alakrot A., Murray L., Nikolov N. S., (2018) "Dataset Construction for the Detection of Anti-Social Behavior in Online Communication in Arabic," Procedia Comput. Sci., vol. 142, pp. 174–181

[11] Bushr, H., Zoher, O., Anas, A., Nada, G., (2020) "Arabic Offensive Language Detection with Attention-based Deep Neural Networks*. In Proceedings of the 4th Workshop on Open-Source Arabic Corpora and Processing Tools,* pp.76–81.

[12] Haidar B., Chamoun, M., Serhrouchni, A., (2019) "Arabic Cyberbullying Detection Enhancing Performance by Using Ensemble Machine Learning". International conference of Internet of Things. pp. 323-327

[13] Husain F. Arabic Offensive Language Detection Using Machine Learning and Ensemble Machine Learning Approaches. ArXiv Prepr ArXiv200508946 2020.

[14] Mubarak, H. and Darwish, K. (2019) "Arabic offensive language classification on twitter" .In: International Conference on Social Informatics, pp. 269–276. Springer.

[15] Shammur, A., Hamdy, M., Ahmed, A., Soongyo, J., Beard, J., Joni S., (2020) "A Multi-Platform Arabic News Comment Dataset for Offensive Language Detection" In Proceedings of the 12th Conference on Language Resources and Evaluation (LREC 2020), pp. 6203–6212 Marseille, 11–16.

[16] https://courses.analyticsvidhya.com/courses/ensemble-learning-and-ensemble-learning-techniques , last access 2-6-2021.

[17] Jason Brownlee,"Machine Learning Mastery With Python" book 2021

[18] https://github.com/omammar167/Arabic-Abusive-Datasets

[19] Alam, K. S., Bhowmik, S., & Prosun, P. R. K. (2021). Cyberbullying Detection: An Ensemble Based Machine Learning Approach. 2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV), 710–715.

[20] Nadali, S., Murad, M., Sharef, N., Mustapha, A. and Shojaee, S. (2013). A review of cyberbullying detection: An overview. 13th International Conference on Intellient Systems Design and Applications, Bangi, pp. 325-330.

[21] Kaati, L., Omer, E., Prucha, N., and Shrestha, A. (2015). Detecting Multipliers of Jihadism on Twitter. 2015 IEEE International Conference on Data Mining Workshop (ICDMW), 954–960.

https://doi.org/10.1109/ICDMW.2015.9

[22] Khairy M., Mahmoud T.M., Abd-El-Hafeez T., Mahfouz A. (2021) User Awareness of Privacy, Reporting System and Cyberbullying on Facebook. In: Hassanien AE., Chang KC., Mincong T. (eds) Advanced Machine Learning Technologies and Applications. AMLTA 2021. Advances in Intelligent Systems and Computing, vol 1339. Springer, Cham. https://doi.org/10.1007/978-3-030-69717-4_58

[23] Marcin Michał Mironczuk and Jarosław Protasiewicz. 2018. A recent overview of the state-of-the-art elements of text classification. Exp. Syst. Applic. 106 (2018), 36–54. DOI:https://doi.org/10.1016/j.eswa.2018.03.058.

[25] Wei F. ID , Wenjiang H. and Jinchang R. ,Class Imbalance Ensemble Learning Based on theMargin Theory, Appl. Sci. **2018**, 8, 815; doi:10.3390/app8050815.

[26] Wei Q, Dunbrack RL Jr (2013) The Role of Balanced Training and Testing Data Sets for Binary Classifiers in Bioinformatics. PLoS ONE 8(7): e67863. https://doi.org/10.1371/journal.pone.0067863

Figures

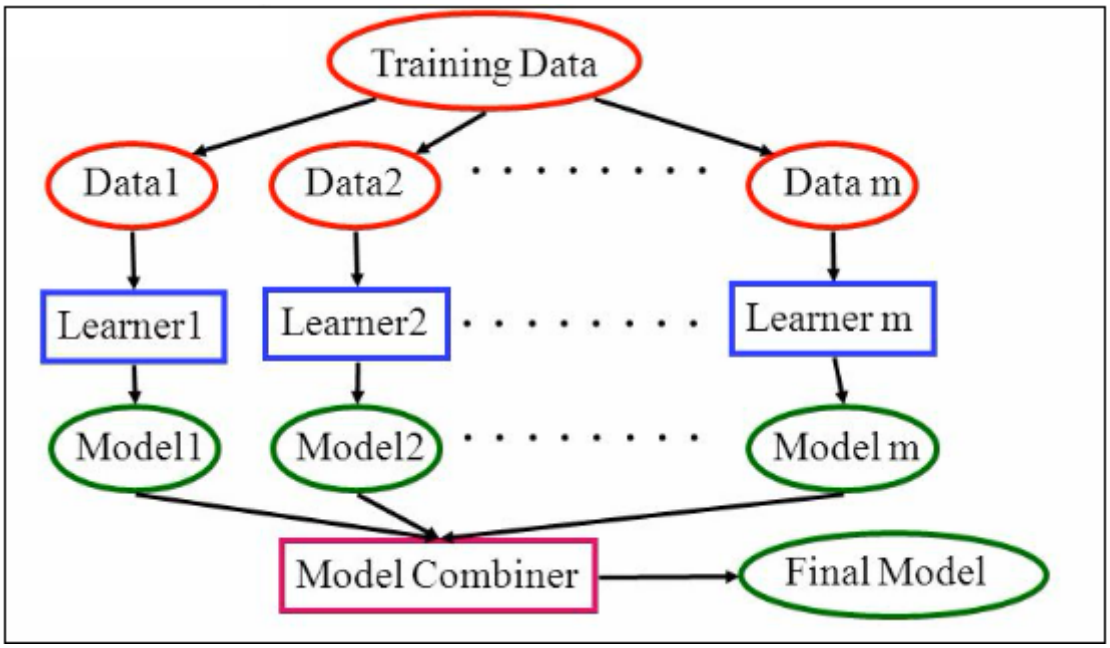

Figure 1

Ensemble Machine learning

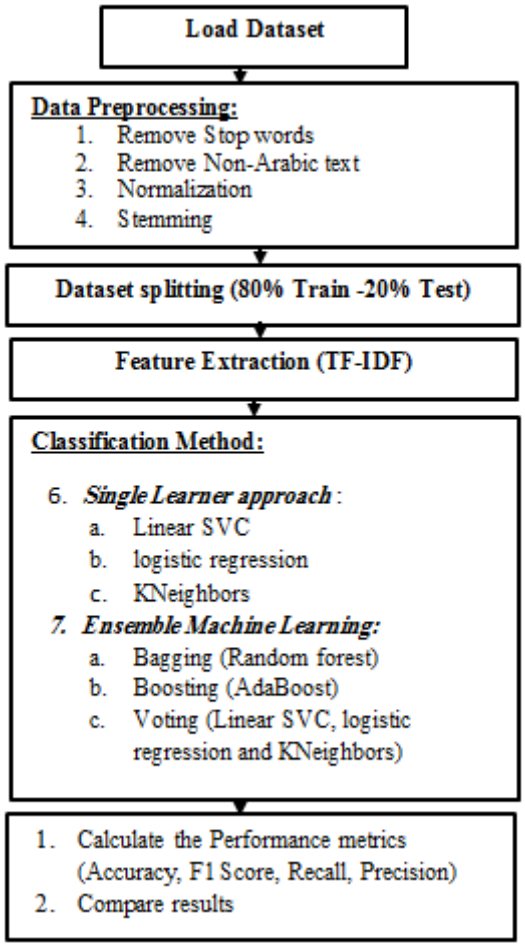

Figure 2

The Proposed Method