IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 13, No. 1, March 2024, pp. 1160∼1171
ISSN: 2252-8938, DOI: 10.11591/ijai.v13.i1.pp1160-1171

A novel ensemble model for detecting fake news

Nissrine Bensouda, Sanaa El Fkihi, Rdouan Faizi
ENSIAS, Mohammed V University in Rabat, Rabat, Morocco

Article Info

Article history:

Received Jan 20, 2023
Revised Oct 20, 2023
Accepted Nov 7, 2023

Keywords:

Bagging
Blending
Boosting
Deep learning
Ensemble learning
Fake news
Machine learning

ABSTRACT

Due the growing proliferation of fake news over the past couple of years, our objective in this paper is to propose an ensemble model for the automatic classification of article news as being either real or fake. For this purpose, we opt for a blending technique that combines three models, namely bidirectional long short-term memory (Bi-LSTM), stochastic gradient descent classifier and ridge classifier. The implementation of the proposed model (i.e. BI-LSR) on real world datasets, has shown outstanding results. In fact, it achieved an accuracy score of 99.16%. Accordingly, this ensemble learning has proven to do perform better than individual conventional machine learning and deep learning models as well as many ensemble learning approaches cited in the literature.

Corresponding Author:

Nissrine Bensouda
ENSIAS, Mohammed V University in Rabat
Rabat, Morocco
Email: nissrine_bensouda@um5.ac.ma

1. INTRODUCTION

Given the increasing proliferation of social media platforms over the past decade, the quantity of fake news has witnessed an exponential growth. This represents a real threat to humanity since this malicious information targets diverse fields [1] and has the power to harm not only individuals but institutions as well. As a consequence, the battle against this phenomenon has taken the form of a research area led mainly by researchers who investigate the use of various machine learning algorithms and natural language processing techniques to detect and, thus, stop such types of news.

As is widely known, a wide range of machine learning algorithms is commonly used. However, some are more used than others depending on the nature of the data studied, that is whether labelled, unlabelled or even semi-labelled. In fact, the use of a specific algorithm is also influenced by the type of the issue it is trying to solve, that is whether it is a classification or regression problem. Such algorithms are used for multiple purposes, one of which is in the detection of bogus reviews. This includes supervised methods as in [2]-[5] and unsupervised methods as in [6]-[8].

Rather than being confined to conventional learning algorithms, other researchers have, however, chosen to detect fake news through the use of deep learning algorithms, which is a well-known brand of machine learning and the most widely used in recent years. Such algorithms rely on an artificial network of neurons inspired by the human brain. This network is made up of hundreds of neuron layers, each receiving and interpreting information from the previous layer. Several works which have opted for deep learning methods in order to classify fake news from real ones can be cited in this context [9]-[11].

Our objective in this article is to propose a new method that combines algorithms from both deep learning and traditional machine learning. For this purpose, we will firstly introduce the concept of ensemble learning whose fundamental idea [12] consists of grouping a set of models together in order to obtain better performance.

The remainder of this paper is organized as follows. Section 2 provides a general overview of the studies that have been interested in the detection of fake reviews using ensemble learning methods. In section 3, we present our proposed approach. Section 4 is concerned with the experiments that are carried to justify the choice of the ensemble technique and its performance. Section 5 discusses the major findings of the present study. Finally, section 6 provides a brief conclusion of the paper.

2. RELATED WORK

Over the past decade, many research works have proposed various techniques and approaches that aim at classifying news as being real or fake. In the present section, we provide an overview of the major studies that have been concerned with fake news classification. Nevertheless, we will be restricted to works that are based on ensemble learning, which is the combination of several learning algorithms. Such approaches have proven to obtain better predictions and performance than those achieved by using individual algorithms. In this respect, three different techniques of ensemble learning [13] have been widely used. These are bagging, blending, and boosting. Each of these ensemble methods is detailed below.

Bagging [14] consists of creating several copies of the same model and then training each copy on a random part of the dataset using a special technique called ”bootstrapping” to choose the training data. This assembly method is widely used to resolve classification problems such as sentiment analysis [15], cancer detection [16], gender recognition [17]. As far as fake news classification is concerned, Patel *et al.* [18] used a comprehensive approach for the automatic categorization of news articles. In this context, the authors worked on textual features to distinguish between fake content and honest one. Their proposed approach was tested on a real world dataset and results show that Bagging got 89% accuracy in fake news classification. Opting for the same method of ensemble learning, Ser *et al.* [19] trained a converted text data using the bagging ensemble of bidirectional Echo State Networks. Experiments were conducted on two different fake news datasets and the findings revealed that an accuracy of 63% was achieved with LIAR dataset and 90% with Fake or Real News dataset.

Unlike Bagging, Blending [20] allows training machine learning models to recognize what is wrong and what is right in a holdout set of models. The specificity of blending is to implement a validation set that varies between 10% and 20% to train the subsequent layer in order to improve the overall performance even more. In their fake news detection approach, Hansrajh *et al.* [21] opted for blending ensemble learning models established from support vector machine, logistic regression, stochastic gradient descent, ridge regression, and linear discriminant analysis. They implemented it on available LIAR and ISOT public datasets to affirm their proposal, then they evaluated the proposed model with other standard machine learning models through different performance measures, and obtained 60.8% and 98.4% as accuracy values with LIAR and ISOT datasets, respectively.

In the same context, Pham *et al.* [22] also proposed a Blending ensemble learning approach in the detection of fake news. They examined the integration of bidirectional encoder representations from transformers (BERT) as a significant method that helps to improve performance in automatic language processing. They combined it with two machine learning algorithms, namely gradient-boosting trees and neural network classifiers. They also extracted text characteristics and formed credibility scores from news title information. The suggested model was implemented using a Chinese dataset translated into English and achieved an accuracy of 88.29%. Based on the same grouping method in spotting fake reviews, Agarwal *et al.* [23] presented a deep learning model, blending both convolutional and recurrent neural networks. They worked on a Kaggle fake news dataset but confined to analysing only reviews and titles. After testing, their proposed model got an accuracy of 97.21%.

As opposed to the two previous methods, Boosting [24] allows to build various models in a sequence one after the other, while asking each model to correct the errors of its predecessor. It is worth noting here that Adaboost and Gradient Boosting are among the most widely used algorithms in this assembly technique. In this vein, Ahmad *et al.* [25] proposed using the learning package approach for the automated classification of press articles. Their study explored different textual properties that can be used to distinguish fake content from real one. They trained the data using a combination of different machine-learning algorithms tested by various methods. Their approach was evaluated using four real-world datasets such as ISOT. After testing, the XGBoost overall learner was found out to perform better than other learning models with an accuracy of 98%

on all the implemented performance measures.

Using the same technique of boosting, most researchers have treated the problem of fake news as a binary classification problem, but there are many other classes of prediction. In Kaliyar *et al.* [26], experiments were conducted using an ensemble machine learning tree-based algorithm, namely Gradient Boosting with optimized parameters, which combines content and context features for fake news detection. The study was carried out using a multi-class dataset and various machine-learning models. The experimental results demonstrated the effectiveness of the Gradient Boosting algorithm with an accuracy of 86%. In another study, Gravanis *et al.* [27] also opted for Boosting techniques where they evaluated Adaboost, Bagging, and Machine learning algorithms. The experiments were done on linguistic text feature using five datasets: Kaggle-EXT, McIntire, BuzzFeed, Politifact, and UNBiased and the best score was achieved by Adaboost over all datasets. Similarly, Espinoza *et al.* [28] used Adaboost with support vector machine (SVM) and MLP to detect fake reviews in a restaurant dataset. Results of their study showed that their ensemble model achieved an accuracy of 77.30%. For their parts, Sharma *et al.* [29] opted for the same grouping method to classify fake tweets using a combination of ML with XGBoost classifier. This method obtained an accuracy of 81% on FakeNewsNet dataset that contains tweets with user characteristics.

However, this does not mean that boosting techniques are used only with traditional machine learning algorithms. In fact, deep learning algorithms are also made use of. In this respect, Aslam *et al.* [30] employed an ensemble-based model to separate bogus news from authentic ones. This study was carried out based on two main attributes: textual ”statement” used with Bi-LSTM-gated recurrent unit (GRU)-dense model, and remaining attributes with the dense deep learning model. The approach was tested on the LIAR dataset and showed a precision of 91.3% and an accuracy of 89.8%. The results also demonstrated the efficiency of the model and showed its superiority over other works that used the same dataset for the same problem.

From the overview above, it is quite clear that a lot of research studies have attempted to classify bogus reviews from truthful ones using ensemble learning techniques. The section that follows describes our proposed method as well as the experiments carried out to illustrate the different types of grouping methods.

3. OUR APPROACH

In this article, our objective is to present a novel ensemble learning method that can allow us to efficiently detect false reviews and subsequently avoid their negatives effects. In this we propose Blending-BI-LSR as a new approach qualified within the Blending group method. For this purpose, we used a list of classifiers mainly composed of Bi-LSTM as the stronger learner, Stochastic Gradient Descent Classifier (SGDC) and Ridge Classifier (RG) as the weak learners in our study. Hence, we obtained six predictions in totality, two outputs for each model: training data predictions and testing predictions. Finally, we concatenated the results to end up with a matrix named the ”meta model” which consists of the three test predictions and the three training predictions obtained before.

This meta model was used to create the final predictions using ”the base model”; In our case we used Random Forest Classifier since it gave us better results compared to the mostly used model which is logistic regression. Accordingly, we generated a function to fit our meta model based on the predictions of both training and testing sets. Then we generated final predictions. Figure 1 illustrates our model architecture.

Before moving on, it is worth noting that the Bi-LSTM integrated in our study takes a word embeddings in the input layer that consists of 100 neurons and 20000 tokens. However, since not all the sentences have the same length, as some are longer or shorter than others, we need to have the entries of the same size; thus, each token in a circulated array is reformed to 350 dimensions. The embedding dimension which is defined as the length of the unique vector used is equal to 50, the batch size which represents the number of units in a cycle is set to 32 and the number of total epochs is 10.

4. EXPERIMENTAL SET UP

To justify the choice of the selected machine learning algorithms as well as that of the grouping method, we used three different ensemble learning techniques, namely Bagging, Blending, and Boosting and we implemented them using different traditional algorithms together with LSTM and Bi-LSTM models.

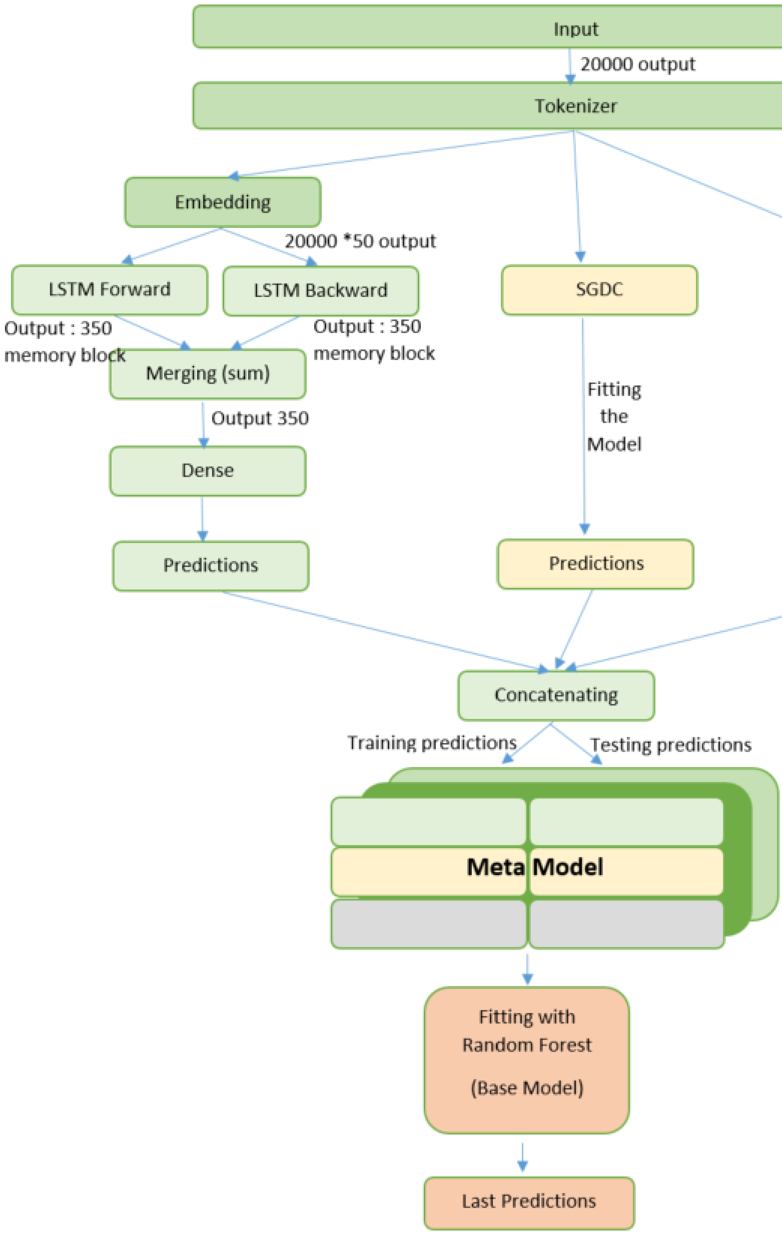


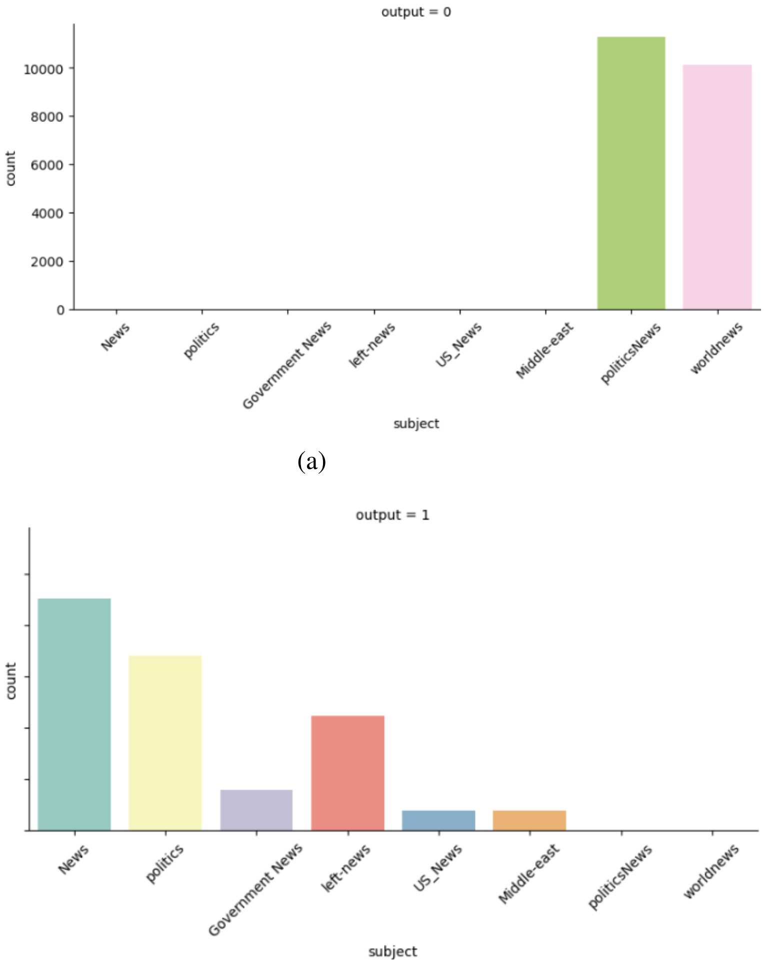
Figure 1. Blending BI-LSR model architecture

4.1. Dataset description

For the particular needs of the present study, we used the ISOT fake news dataset [31] and [32] which is made up of two different types of articles, namely real and fake. The first set contains 21417 real news articles crawled from the news website "Reuters.com" while the second contains 23481 fake articles collected from different unreliable websites that were flagged by Wikipedia and "Politifact", a fact-checking organisation. Both datasets are mainly composed of articles on diverse topics most of which focus on World news and politics. Each of these articles includes: title, text, subject, and publication date. Figure 2 gives the considered dataset distribution according to the different subjects divided into two subfigures, (a) represents real news while (b) represents fake news.

4.2. Data pre-processing

In order to obtain effective results, cleaning our dataset is a major prerequisite. For this purpose, various pre-processing tasks have been carried put. This includes the removal of stops words, which are a set of terms that are commonly used in the text but they do not contribute in natural language processing tasks and as such they are not important in our analysis. These words include prepositions, pronouns, articles, etc. In addition to removing stop words, nulled values as well as words which are less than two characters were also deleted. The cleaned data was then subject to two morphological processes, namely stemming and lemmatization. The first chops all affixes to get the stem or the root of the word. The second is quite similar to stemming but brings the word to its dictionary form.



(b)

Figure 2. Dataset distribution according to (a) the different subjects of truthful news and (b) the different subjects of untruthful news

After pre-processing our dataset, we carried out feature selection to detect the most relevant characteristics based on the chi-squared test. The latter was implemented to choose the features which got the highest chi-squared values. This test measures the dependence between stochastic variables, and shows the most likely independent features of the class and consequently inappropriate for our classification. Therefore, we calculated the dependency between 'text', 'content', 'subject' and 'date'. As shown in Figure 3, the results indicate that the feature subject has no dependence on the other variables.

We combined the columns 'text', 'title' and 'date' to obtain a complete news article with its headline, and removed the 'subject' column since it is irrelevant in our analysis. But before feeding the data into the machine learning algorithms, we converted the articles into integers using tokenizers to get them prepared for the embedding matrix. Moreover, we divided the dataset into sets: 85% for training and 15% as a test data to be implemented in the predictions evaluation.

4.3. Data visualisation

Before conducting the experiments, we will proceed to data visualization which allows us to visualize the particular intricacies of our dataset by using graphs. This task was carried out before deleting the 'subject' column. Probably one of the most striking features of our dataset is that it is almost balanced as it contains approximately the same number of real and fake news articles. Real news articles exceed fake news only by 2.4%; this is illustrated in the Figure 4.

Subsequently, we extracted the polarity, which is the sentiment degree expressed the text be it negative, positive, or neutral depending on the assigned value which varies between -1.0 and 1.0. The subjectivity of the text returns a number between 0.0 and 1.0, where the value 0.0 is assigned to the subjective while the value 1.0 is allocated to the objective. And finally we calculated the number of words of each news item. This was executed on the news of each class independently. This is illustrated in both Figures 5 and 6 where the subfigures (a), (b), and (c) expose sentiment, subjectivity, and word count, respectively.

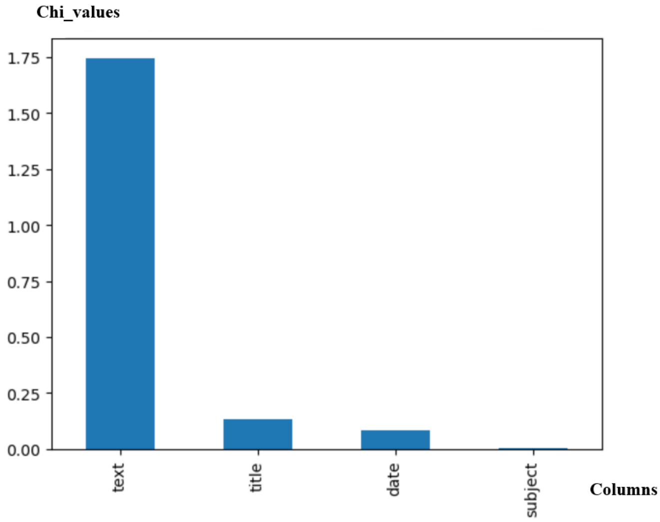


Figure 3. Chi-squared test

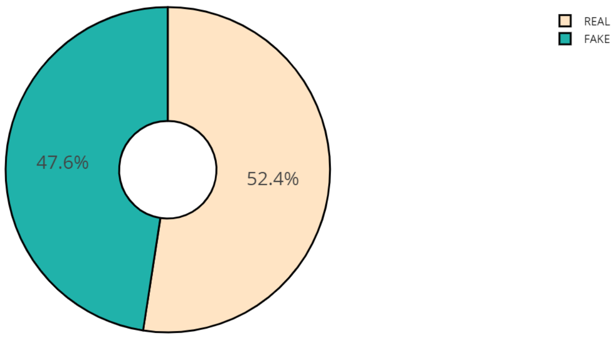
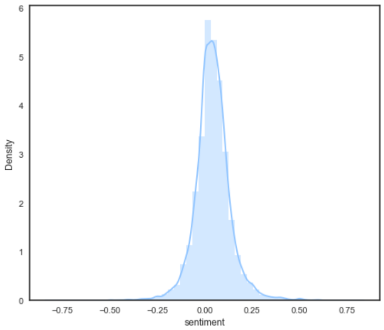


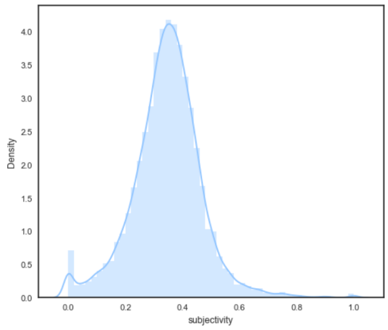
Figure 4. Dataset balanced

As is noticeable, the polarity of truthful news articles ranges between -0.25 and 0.25 while the polarity of fake news varies between -0.5 and 0.5 which indicates the use of words carrying more emotion even if the majority of both classes sentences carry a neutral polarity.

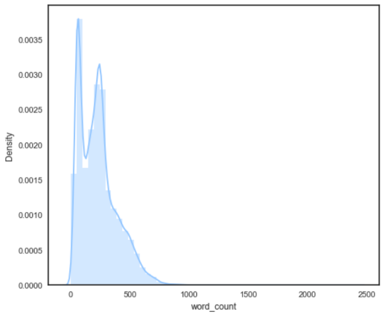
As far as subjectivity is concerned, both news classes seem to be different. For the truthful news articles, the values of subjectivity vary between 0.0 and 0.8, where the majority of the articles are between 0.2 and 0.4, whereas in fake news, subjectivity values diverge between 0.0 and 1.0, where the largest mass is placed between 0.4 and 0.6. Thus, it can be concluded that real news articles are likely to be more objective while bogus news tends to be more subjective.



(c)

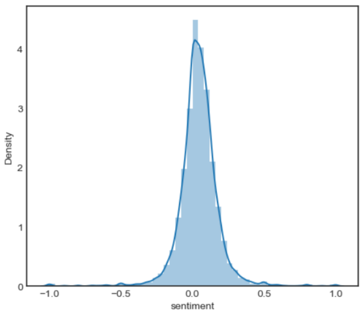


(d)

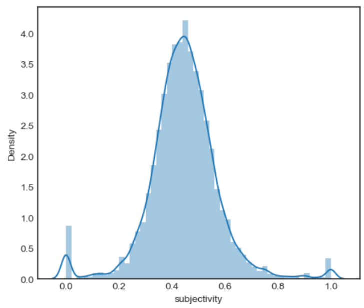


(e)

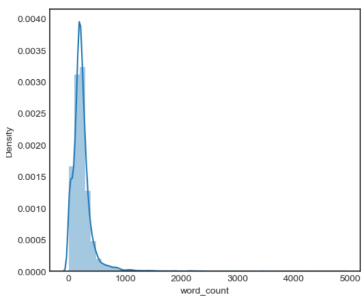
Figure 5. Dataset distribution according to the different subjects of truthful news (a) Sentiment, (b) subjectivity, and (c) word count



(f)



(g)



(h)

Figure 6. Dataset distribution according to (a) the different subjects of untruthful news (a) Sentiment, (b) subjectivity, and (c) word count

As far as word count is concerned, the figures above clearly show that the number of words in each real news item varies between 0 and 750, and that most paragraphs contain less than 500 words. On the other hand, the total number of words containing each fake news article alternates between 0 and 1000, where the majority are around 500 words. These results allow us to conclude that fake news are likely to involve more words in order to better express the idea and give the maximum information that could mislead the reader.

4.4. Experiments

For the purpose of validating our Blending-BI-LSR model, we conducted a set of experiments based on combinations of algorithms that consist of traditional machine learning algorithms and deep learning ones. Therefore, we integrated Bi-LSTM as the basic learner of all the experimentations since it is known for its performance in classification problems. We also used some traditional machine learning algorithms which are listed in the subsequent Table 1.

Table 1. Findings of the experiments

Method	Model (s)	Accuracy With 'date'	Accuracy without 'date'
Single	K-Nearest Neighbors (KNN)	59.93%	57.92%
	DecisionTree (DT)	83.07%	82.94%
	Bidirectional Long Short-Term Memory (BI-LSTM)	99.04%	95.60%
	Long Short-Term Memory (LSTM)	98.64%	59.38%
	RandomForest (RF)	86.14%	82%
	ExtraTrees (ET)	80.13%	72.93%
	Support Vector Classifier (SVC)	71.66%	68.54%
	Stochastic Gradient Descent Classifier (SGDC)	58.42%	55.48%
	Logistic regression (LG)	67%	59.84%
	Ridge Classifier (RG)	64.69%	60%
Blending	LSTM + LG	98.72%	98.67%
	LSTM +DT + ET	92.59%	84.63%
	BI-LSTM +DT + ET	92.82%	84.63%
	LSTM + RG	98.73%	97.20%
	BI-LSTM + SGDC	99.04%	98.33%
	LSTM + SVC	98.76%	98.52%
	Blending-BI-LSR	99.16%	98.24%
	LSTM + SVC + ET	87.05%	78.38%
	LSTM + ET + RG	82.19%	74.29%
	BI-LSTM + ET + SGDC	97%	73.89%
	LSTM + SGDC + RG + KN	98.74%	98.63%
	BI-LSTM + ET + RG + SGDC	96.31%	73.27%
	ET + SGDC +RG + LSTM	85.48%	77.78%
	BI-LSTM + ET + RG + SGDC + SVC	90.41%	74.83%
	SVC+ RG + SGDC + LG + LSTM + ET	88.42%	79.22%
Boosting	Ada boost (Ada)	79.48%	75.89%
	Gradient boosting (Grad)	73.76%	72.92%
	XGB Classifier	81.13%	78.52%
	LSTM + DT	89%	88%
	BI-LSTM + KNN	86%	84%
	LSTM + DT + KNN	86%	84%
	BI-LSTM + ET + SVC	85%	82%
	BI-LSTM + RG + SGDC	79%	79%
	LSTM + RG + SGDC	79%	79%
Bagging	LG	67.07%	60.39%
	DT	90.13%	89.62%
	LSTM + RF	89%	88%
	LSTM + SGDC	89%	88%
	BILSTM + SGDC	89%	89%
	BILSTM + RG	89%	88%
	BILSTM + LG	89%	89%
	LSTM + KNN	89%	88%
	LSTM + SGDC + SVC	90%	90%
	BI-LSTM + SGDC + SVC	90%	89%
	LSTM + ET + SVC	90%	90%
	LG + RF + KNN + RG + BI-LSTM	90%	90%
	ET + RF + KNN + SVC + RG + LSTM	90%	90%

After training each algorithm separately, we created several collections using three different methods, each representing a classification model for ensemble learning. Thus, the first ensemble method is bagging We constructed independent estimators on diverse samples of our dataset, then we voted on all the individual predictions for obtaining the final value. We repeated these operations on different collections of algorithms containing varied numbers and content.

The second method is boosting in which algorithms are executed subsequently and the predictions are made based on 'XGB Classifier' since we obtained the greatest value through its implementation compared to 'Ada Boost' and 'Gradient Boosting'. Accordingly, multiple simples are created from our original dataset. Before producing the first model we gave all the observations identical weights, then the predictions were made over the whole dataset. In the same way, the subsequent models are produced by correcting the preceding model's errors to finally obtain the strong learner model by calculating the weighted mean of the others. Likewise, multiple collections of different algorithms are generated.

The third applied method is blending, where the meta-model is trained on predictions done on a hold-out dataset using the base models. Therefore, we created a list of models that we integrated in our experiments. For this purpose, we used logistic regression as the final learner for all the proposed sets of models. Then, we created a function that fits all the models and allows us to acquire predictions via the validation dataset. These predictions were restructured and joined to a list that was stacked for the meta-model features. The same steps were reproduced for each nominated assortment. All the experiments were evaluated to calculate their efficiency. However, each experiment was conducted twice. First, we implemented only two characteristics, namely 'text' and 'title', while in the second, three features, namely 'text', 'title' and 'date' were used in order to illustrate the impact of the 'date' column in the results of our classification. The Table 1 shows more details about the findings associated with each model.

5. DISCUSSION

Taking into consideration the findings of the experiments that were carried out in our work, classified by type of grouping, it is quite clear that the implementation of an ensemble learning method allows us to obtain better results compared to the single application of a separate model. Specially, our Blending BI-LSR model achieved the best score among all the implemented models. In fact, it has been confirmed that the best results in our study were obtained using the blending method through its various collections. In contrast, low results were obtained using the boosting grouping method despite using the same dataset as well as the same learning algorithms. On the other hand, the performance of bagging experiments lies between the two previous methods and demonstrates its weakness compared to the blending method and its capability against boosting method. We also noted the impact of 'date' column that is shown in the final results in all the experiments, in that it leads to an improvement of our analysis performance.

Moreover, we noted that the number of algorithms constituting a group of model does not influence the final efficiency, even though the algorithms used influence effectively the final performance either by improving the results or by worsening them. Let us take the example of two Boosting experiments. We grouped long short term memory (LSTM) with decision tree (DT) and an accuracy of 89% was obtained. In another experiment, where we kept the two previous algorithms and added the K-nearest neighbors (KNN), an accuracy value of 86% was achieved. Accordingly, the addition of a third algorithm only reduced the model performance. While running LSTM, DT and KNN models separately, we got the accuracy of 98.64%, 83.07% and 59.93%, respectively. Hence, we can say that the combination of these models allowed the improvement of the results compared to KNN and DT but decreased the performance of LSTM results.

To confirm the effectiveness and the outstanding results of our proposal, the Table 2 illustrates the performance of various approaches cited in the literature in comparison to ours. This table clearly shows that our approach is ranked first compared to other cited studies. In fact, with an accuracy score of 99.16%, our ensemble model has outperformed other previous approaches despite opting for similar assembling techniques in some cases and using the same dataset in other cases but never implementing the same features.

Table 2. Comparison of our approach and some other methods

Study	Model	Dataset	Year	Accuracy	Features
[18]	SVM Bagging	Real word database	2021	89%	Headlines
[19]	Bagging ensemble of bidirectional Echo State Networks	LIAR/ Fake o Real News dataset	2022	63% / 90%	Text
[21]	Blending (logistic regression, support vector machine, linear discriminant analysis, stochastic Gradient Descent, and Ridge Regression)	LIAR/ ISOT	2021	60.8% / 98.4%	Text
[22]	Blending (Bert + gradient-boosting trees + neural network classifiers)	Chinese dataset translated into English	2019	88.29%	Title
[23]	Blending (CNN and RNNs)	Kaggle fake news dataset	2020	97.21%	Title and text
[25]	Boosting (XGBoost overall learner)	ISOT and Kaggle datasets	2020	98%	Text
[26]	Boosting (Bi-LSTM-GRU-dense model)	LIAR	2019	86%	Text
[27]	AdaBoost	Kaggle-EXT, McIntire, BuzzFeed, Politifact, UNBiased	2019	95%	Linguistic text features
[28]	Adaboost with SVM and MLP	Restaurant Dataset	2020	77.3%	Text
[29]	XGBoost with ML algorithms	FakeNewsNet	2022	81%	tweet content and user characteristics
Our approach	Blending (BI-LSTM +SGDC + RG) using RF as meta-model	ISOT	2023	99.16%	Title, text and date

6. CONCLUSION

Our objective in this work was to propose a novel approach that can automatically classify news articles as being real or fake. In this respect, we put forward a new ensemble learning approach that groups traditional machine learning and deep learning methods using blending techniques. The overall findings of the experiments that were carried out demonstrated that the proposed Blending BI-LSR method achieved an accuracy score of 99.16%. This has proven its outstanding performance over other models used in the literature.

REFERENCES

[1] N. Bensouda, S. El Fkihi, and R. Faizi, "Opinion spam detection: A review of the literature", in Proc. International Conference on Learning and Optimization Algorithms: Theory and Applications (LOPAL), Mai 2018, pp. 1–6, doi: 10.1145/3230905.3230922.

[2] A. M. Elmogy, U. Tariq, A. Mohammed and A. Ibrahim, "Fake reviews detection using supervised machine learning" *International Journal of Advanced Computer Science and Applications(IJACSA)*, 12(1), 2021, doi :10.14569/IJACSA.2021.0120169.

[3] R. Mohawesh, S. Xu, S. N. Tran, T. Ollington, M. Springer, Y. Jararwehs, and S. Maqsood, "Fake reviews detection: A survey", *IEEE Access*, vol. 9, pp. 65771-65802, 2021, doi: 10.1109/ACCESS.2021.3075573.

[4] S. N. Alsubari, S. N. Deshmukh, A. A. Alqarni, N. Alsharif, T. H. H. Aldhyani, F. W. Alsaade, and O. I. Khalaf, "Data analytics for the identification of fake reviews using supervised learning", *CMC-Computers, Materials & Continua*, vol. 70, no 2, pp. 3189-3204, 2022, doi: 10.32604/cmc.2022.019625.

[5] N. Kumar, D. Venugopal, L. Qiu, and S. Kumar, "Detecting review manipulation on online platforms with hierarchical supervised learning ", *Journal of Management Information Systems*, 35:1, pp. 350-380, doi: 10.1080/07421222.2018.1440758.

[6] J. SONI, and N. PRABAKAR, "Effective machine learning approach to detect groups of fake reviewers", *in Proc. 14th international conference on data science (ICDATA'18)*, pp. 3-9, 2018.

[7] S. Yu, J. Ren, S. Li, M. Naseriparsa , and F. Xia, "Graph learning for fake review detection", *Frontiers in Artificial Intelligence*, vol. 5, 2022, doi: 10.3389/frai.2022.922589.

[8] A. NEISARI, L. RUEDA, and S. SAAD, "Spam review detection using self-organizing maps and convolutional neural networks", *Computers & Security*, vol. 106, 2021, doi: 10.1016/j.cose.2021.102274.

[9] G. Bathla and A. Kumar, "Opinion spam detection using deep learning", in Proc. 8th International Conference on Signal Processing and Integrated Networks (SPIN), 2021, pp. 1160-1164, doi: 10.1109/SPIN52536.2021.9566048.

[10] D. Baishya,J. J. Deka, G. Dey, and P. K. Singh, "SAFER: Sentiment analysis-based fake review detection in E-Commerce using deep learning", *SN Computer Science*. 2, 2021, doi: 10.1007/s42979-021-00918-9.

[11] J. Tao, X. Fang, and L. Zhou, "Unsupervised deep learning for fake content detection in social media", In Proc. the Annual Hawaii International Conference on System Sciences, 2020, IEEE Computer Society, doi: 10.24251/hicss.2021.032.

[12] M. A. Ganaie, M. Hu, A. K. Malik, M. Tanveer, and P. N. Souganthan, "Ensemble deep learning: A review", Engineering Applications of Artificial Intelligence, 2021, doi:10.1016/j.engappai.2022.105151

[13] I. K. Nti, A. F. Adekoya, and B. A. Weyori, ”A comprehensive evaluation of ensemble learning for stock-market prediction”, J Big Data 7, 2020, doi: 10.1186/s40537-020-00299-5.
[14] G. Tüysüzoğlu, and D. Birant, ”Enhanced Bagging (eBagging): A novel approach for ensemble learning”, The International Arab Journal of Information Technology, 17, 2020, pp. 515-528, doi: 10.34028/iajit/17/4/10.
[15] S. Sharma, S. Srivastava, A. Kumar, and A. Dangi, ”Multi-class sentiment analysis comparison using support vector machine (SVM) and BAGGING technique-an ensemble method”, in Proc. International Conference on Smart Computing and Electronic Enterprise (ICSCEE), 2018, pp. 1-6, doi: 10.1109/ICSCEE.2018.8538397.
[16] G. Valentini, M. Muselli, and F. Ruffino, ”Cancer recognition with bagged ensembles of support vector machines”, *Neurocomputing*, Volume 56, 2004, Pages 461-466, doi: 10.1016/j.neucom.2003.09.001.
[17] T. Danisman, ”Bagging ensemble for deep learning based gender recognition using test-timeaugmentation on large-scale datasets”, *Turkish Journal of Electrical Engineering and Computer Sciences*, Vol. 29, No. 4, 2021, Article 14, doi: 10.3906/elk-2008-166.
[18] P. Patel, S. Patel, and Y. Patel, ”Ensemble based approach for fake news detection”, *International Research Journal of Engineering and Technology (IRJET)*, Volume: 08, Issue: 05, May 2021, pp. 2733-2737.
[19] J. D. Ser, M. N. Bilbao, I. LAÑA, K. Muhammad, and D. Camacho, ”Efficient fake news detection using bagging ensembles of bidirectional echo state networks”, in Proc. International Joint Conference on Neural Networks (IJCNN), IEEE, 2022, pp. 1-7, doi: 10.1109/IJCNN55064.2022.9892331
[20] T. Wu, W. Zhang, X. Jiao, W. Guo, and Y. A. Hamoud, ”Evaluation of stacking and blending ensemble learning methods for estimating daily reference evapotranspiration”, Computers and Electronics in Agriculture, Volume 184, 2021, 106039, doi: 10.1016/j.compag.2021.106039.
[21] A. Hansrajh, T. T. Adeliyi, and J. Wing, ”Detection of online fake news using blending ensemble learning”, Scientific Programming, vol. 2021, Article ID 3434458, 10 pages, 2021, doi: 10.1155/2021/3434458.
[22] L. Pham, ”Transferring, transforming, ensembling: the novel formula of identifying fake news”, in Proc. 12th ACM International Conference on Web Search and Data Mining, Melbourne, 2019, pp. 11-15,
[23] A. Agarwal, M. Mittal, A. Pathak, and L. M. Goyal, ”Fake news detection using a blend of neural networks: an application of deep learning”, *SN Computer Science*, 1, 143, 2020, doi: 10.1007/s42979-020-00165-4.
[24] N. J. Ezeora, S. E. Abhadiomhen, R. Nzeh, and B. C. Uzo, G. E. Anichebe, C. Asogwa, I. Uzo, I., and N. E. Ogbene, ”Combination of bagging and boosting ensemble learning methods: A solution to model underfitting and overfitting”, in Proc. the LASUSTECH 30th iSTEAMS Multidisciplinary Innovations Conference, 2022, pp. 149-156, doi:10.22624/AIMS/iSTEAMS/LASUSTECH2022V30P13.
[25] I. Ahmad, M. Yousaf, S. Yousaf, and M. O. Ahmad, ”Fake news detection using machine learning ensemble methods”, Complexity, vol. 2020, Article ID 8885861, 11 pages, 2020, doi: 10.1155/2020/8885861.
[26] R. K. Kaliyar, A. Goswami and P. Narang, ”Multiclass fake news detection using ensemble machine learning”, in Proc. 9th International Conference on Advanced Computing (IACC), IEEE, 2019, pp. 103-107, doi: 10.1109/IACC48062.2019.8971579.
[27] G. Gravanis, A. Vakali, K. Diamantaras, and B. Karadais, ”Behind the cues: A benchmarking study for fake news detection”, Expert Systems with Applications, 2019, vol. 128, p. 201-213, doi: 10.1016/j.eswa.2019.03.036.
[28] L. G. Espinoza, F. Abri, A. S. Namin, K. S. Jones, D. R. W. Sears, ”Fake reviews detection through ensemble learning”, 2020, doi: 10.48550/ARXIV.2006.07912.
[29] S. SHARMA, M. SARASWAT, and A. DUBEY, ”Fake news detection on Twitter”. *International Journal of Web Information Systems*, 2022, doi: 10.48550/arXiv.2006.07912.
[30] N. Aslam, I. U. Khan, F. S. Alotaibi, L. A. Aldaej, and A. K. Aldubaikil, ”Fake detect: A deep learning ensemble model for fake news detection”, Complexity, vol. 2021, Article ID 5557784, 8 pages, 2021, doi: 10.1155/2021/5557784.
[31] H. Ahmed, I. Traore, and S. Saad, ”Detecting opinion spams and fake news using text classification”, *Journal of Security and Privacy*, Volume 1, Issue 1, dec. 2017, doi: 10.1002/spy2.9.
[32] H. Ahmed, I. Traore, and S. Saad, ”Detection of online fake news using N-Gram analysis and machine learning techniques”, in Proc. International conference on intelligent, secure, and dependable systems in distributed and cloud environments. Springer, Cham, volume 10618, 2017, pp. 127-138, doi: 10.1007/978-3-319-69155-8_9.

BIOGRAPHIES OF AUTHORS

Nissrine Bensouda is a PHD student at the Higher National School of Computer Science and Systems Analysis (ENSIAS), Mohammed V University of Rabat, Morocco. She is working on the issue of the automatic detection of fake news and opinions. She can be contacted at email: nissrine_bensouda@um5.ac.ma.

Sanaa El Fkihi received her Ph.D. degree in computer science from Mohammed V Rabat University and the University of Science and Technology of Lille in 2008. She is currently a Full Professor of computer engineering at the Higher National School of Computer Science and Systems Analysis (ENSIAS), Mohammed V University, Morocco. Her research interests include multimedia and information retrieval and processing. She can be contacted at email: sanaa.elfkihi@ensias.um5.ac.ma.

Rdouan Faizi received his Ph.D. from Mohammed V University in Rabat. He is currently a Full Professor at the Higher National School of Computer Science and Systems Analysis (ENSIAS), Mohammed V University, Rabat, Morocco. His research interests include Natural Language Processing and E-learning. He can be contacted at email: rdouan.faizi@ensias.um5.ac.ma.