Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Fake news detection within online social media using supervised artificial intelligence algorithms

Feyza Altunbey Ozbay, Bilal Alatas *

Department of Software Engineering, Faculty of Engineering, Firat University, 23100, Elazig, Turkey

ARTICLE INFO

Article history:
Received 16 May 2019
Received in revised form 28 August 2019
Available online xxxx

Keywords:
Fake news detection
Online social media
Supervised artificial intelligence algorithm

ABSTRACT

Along with the development of the Internet, the emergence and widespread adoption of the social media concept have changed the way news is formed and published. News has become faster, less costly and easily accessible with social media. This change has come along with some disadvantages as well. In particular, beguiling content, such as fake news made by social media users, is becoming increasingly dangerous. The fake news problem, despite being introduced for the first time very recently, has become an important research topic due to the high content of social media. Writing fake comments and news on social media is easy for users. The main challenge is to determine the difference between real and fake news. In this paper, a two-step method for identifying fake news on social media has been proposed, focusing on fake news. In the first step of the method, a number of pre-processing is applied to the data set to convert un-structured data sets into the structured data set. The texts in the data set containing the news are represented by vectors using the obtained TF weighting method and Document-Term Matrix. In the second step, twenty-three supervised artificial intelligence algorithms have been implemented in the data set transformed into the structured format with the text mining methods. In this work, an experimental evaluation of the twenty-three intelligent classification methods has been performed within existing public data sets and these classification models have been compared depending on four evaluation metrics.

1. Introduction

The last technological developments and the spread of the Internet have caused an enormous impact on social interactions. Social media has become an increasingly popular way of obtaining information for people. Additionally, people share their personal activities, interests, and opinions on different social media platforms. Social media provides many advantages such as easy access to information, low cost, and rapid spread of information. Owing to these advantages, many people tend to search for news from social media rather than classical news sources such as television or newspaper. Consequently, social media news is quickly replacing classical news sources. Although social media has many advantages, news on online social media is not qualified when compared the classical news sources. Nevertheless, sometimes the contents of social media may be changed to achieve different goals. These websites may be seen in Macedonia, Romania, Russia, United States, United Kingdom, and many other countries [1]. For this reason, fake news and rumors spread very quickly and broadly [2]. This situation leads to the production and propagation of the news articles which are inaccurate.

* Corresponding author.
E-mail addresses: faltunbey@firat.edu.tr (F.A. Ozbay), balatas@firat.edu.tr (B. Alatas).

https://doi.org/10.1016/j.physa.2019.123174

In addition, without careful checking, fake news and misinformation are spread by well-meaning users. In social media, there are many websites that aim to produce only fake news.

Reviews, opinions, and news on social media have been playing an important role in the decisions of users. The spread of low-quality news, namely fake news, have a negative effect on opinions of society and individual. Fake news is not only harmful to individuals and society, but also to businesses and governments. For instance, fake news about the organization, which are emitted by spam or malicious users, can cause considerable damage. Therefore, fake news detection has become a significant research area.

In this study, a detection model containing two different steps has been proposed to detect fake news in social media. The proposed model is an approach that combines methods of text analysis and supervised artificial intelligence algorithms. In the first step of this work, text mining methods have been applied to the online news data set. The aim of text analysis methods and techniques is to obtain structured data from an unstructured news article. In the second step, supervised artificial intelligence algorithms (BayesNet, JRip, OneR, Decision Stump, ZeroR, Stochastic Gradient Descent (SGD), CV Parameter Selection (CVPS), Randomizable Filtered Classifier (RFC), Logistic Model Tree (LMT), Locally Weighted Learning (LWL), Classification Via Clustering (CvC), Weighted Instances Handler Wrapper (WIHW), Ridor, Multi-Layer Perceptron (MLP), Ordinal Learning Model (OLM), Simple Cart, Attribute Selected Classifier (ASC), J48, Sequential Minimal Optimization (SMO), Bagging, Decision Tree, IBk, and Kernel Logistic Regression (KLR)) have been applied to the structured news data sets. There are individual studies about using only one or two of the supervised algorithms in the literature. Furthermore, they are tested within limited data sets. Unlike other individual studies, the problem of fake and false news detection has been handled and modeled as a classification problem and twenty-three supervised artificial intelligence algorithms have been adapted for the first time to fake and false news detection problem in three real data sets in this study.

The outline of the paper is organized as follows. Previous works performed in the field of fake news detection has been briefly described in Section 2. Details of the proposed model, supervised artificial intelligence algorithms, text mining steps, and performance evaluation metrics have been described in Section 3. Section 4 describes data sets and experimental results obtained from twenty-three supervised artificial intelligence algorithms for three different data sets. Section 5 presents a discussion of the results obtained from the current work. Conclusions and future research directions have been discussed in Section 6.

2. Related works

Although the fake news detection problem has been introduced for the first time very recently, it has attracted considerable attention. Various approaches have proposed to detect fake news in various types of data. A review of existing and related works in the literature about fake news detection have been presented in this section.

The problem of fake news detection has been separated into three groups by Rubin et al. [3]. These groups are serious fabrication, large-scale hoaxes, and humorous fake news. Conroy et al. have used a hybrid method to propose a fake news detector. Their hybrid method combines linguistic cue approaches and network analysis approaches [4]. The vector space model has been used to verify news [5]. The authors aim to detect deceptive information in online news sources. Dadgar et al. [6] have aimed to classify news into various groups using TF-IDF and SVM.

Satirical Cues have been used to detect misleading or fake news in [7]. The authors have proposed an SVM based model and tested the model on 360 news articles. Jin et al. have identified conflicting viewpoints in social media to verify the news. They have used real-world data sets to test their model [8].

Tacchini et al. have presented two classification models for the problem of fake news detection. One of these models is based on logistic regression and the other is a boolean crowdsourcing algorithm [9]. In another work, the data mining algorithm for the problem of fake news detection, evaluation metrics, and data sets have been extensively presented [10]. For detecting opinion spam and fake news, Ahmed et al. have used machine learning classification techniques and n-gram analysis [11]. In their work, the authors have applied their methods to public data sets. Gilda has used classification algorithms, namely Random Forests, SVM, Bounded Decision Trees, Stochastic Gradient Descent, Gradient Boosting. The author has obtained the best performance with the Stochastic Gradient Descent method [12]. Ruchansky et al. [13] have adapted a hybrid algorithm called CSI for fake news detection. With their method, three characteristics have been combined for a more accurate prediction. These characteristics are Capture, Score, and Intagrate. Shu et al. [14] have proposed the tri-relationship fake news detection model, taking into account the correlation of publisher bias, news stance, and related user interactions. The authors have tested their model on real-world fake news detection data sets.

Long et al. [15] have utilized a novel hybrid algorithm considering attention-based long-short memory network for fake news detection problem. The performance of the method is tested on benchmark fake news detection data sets. Figueira and Oliveira [16] have reviewed the current state of fake news. They have proposed a solution and described two opposite approach for fake news. Janze and Risius [17] have introduced a detection model to automatically identify fake news. The authors have tested their models on the news posted on Facebook during the U.S. presidential election of 2016. Perez-Rosas et al. have adapted a new automated algorithm [18]. In this work, the authors have indicated the classification model based on the combination of lexical, syntactic, and semantic information. Buntain and Golbeck have proposed an automated system to detect fake news in popular twitter threads [19]. They have applied this method to three publicly available data sets. Bessi has performed a study on the statistical properties of fake news, hoaxes, and unproven claims in online social media [20].

Please cite this article as: F.A. Ozbay and B. Alatas, Fake news detection within online social media using supervised artificial intelligence algorithms, Physica A (2019) 123174, https://doi.org/10.1016/j.physa.2019.123174.

Table 1
Fake news detection in the literature.

2015	• Deceptive detection for news: three types of fakes [3] • Automatic deception detection [4] • News verification [5]
2016	• News classification with TF-IDF and Support Vector Machine [6] • Satirical cues for fake news [7] • Conflicting social viewpoints for news verification [8]
2017	• Automated fake news detection [9] • Data mining perspective for fake news detection [10] • Text classification for detecting opinion spam and fake news [11] • Machine learning methods for fake news detection [12] • A hybrid model for fake news detection [13] • Tri-relationship model for fake news detection [14] • Multi-perspective speaker profiles for fake news detection [15] • A survey: current state of fake news [16] • Automatic detection of fake news on social media platforms [17] • Automatic detection of fake news [18] • Identifying fake news in popular twitter threads [19]
2018	• Statistical properties of viral misinformation in online social media [20] • Information dissemination model for social media [21] • User profiles approach for fake news detection [22] • Fake news detection with crowd signals [23] • Tensor embeddings and label propagation for fake news [24]
2019	• Fake news via network analysis [25] • Geometric deep learning for fake news detection [26] • Fake news detection with task-generic features [27] • Emotions based fake news detection [28]

Zu et al. have proposed a competitive model, which focus on the relationship between original false information and updated information to reduce the impact of false information [21]. Shu et al. [22] have adapted a new algorithm for the problem of fake news detection, considering the trust of the users. Tschiatschek et al. [23] have used a crowd signal for the problem. For this reason, the authors have proposed a new Detective algorithm which performs Bayesian inference and jointly learns flagging accuracy of users over time. Content-based fake news detection model has been proposed by Guacho et al. The authors have formulated fake news detection as a semi-supervised method [24]. Shu et al. [25] have examined the kinds of social networks and proposed usage of these networks for detecting and mitigating fake news on social media.

Monti et al. have adapted the geometric deep learning-based model to detect fake news. The authors have used news stories verified Twitter to test their model [26]. In another work, the task-generic features have applied to tackle the detection of fake news [27]. Guo et al. have proposed emotion-based fake news detection method, which combines the publisher emotion and social emotion [28]. All of the works performed about fake news detection are demonstrated in Table 1.

3. Fake news detection model

This section provides details of the proposed model for fake news detection. It begins by pre-processing the data set by filtering the redundant terms or characters such as numbers, stop-words, etc. Feature extraction has been applied to the fake news data set for reducing the dimension of feature space. The terms in each document are weighted and the Document-Term Matrix is constructed. The last process of the model is to apply supervised artificial intelligence algorithms on the fake news data set. Twenty-three different supervised artificial intelligence algorithms, namely; BayesNet, JRip, OneR, Decision Stump, ZeroR, Stochastic Gradient Descent (SGD), CV Parameter Selection (CVPS), Randomizable Filtered Classifier (RFC), Logistic Model Tree (LMT), Locally Weighted Learning (LWL), Classification Via Clustering (CvC), Weighted Instances Handler Wrapper (WIHW), Ridor, Multi-Layer Perceptron (MLP), Ordinal Learning Model (OLM), Simple Cart, Attribute Selected Classifier (ASC), J48, Sequential Minimal Optimization (SMO), Bagging, Decision Tree, IBk, and Kernel Logistic Regression (KLR) have been used in this study. Fig. 1 outlines the process of the model.

3.1. Text mining

In the learning system, the representation of data greatly affects the accuracy of the results. Specifically, the text analysis problems need to be transformed into a representation which is suitable for the method to be applied. Text-based data shared by users on social media are generally in unstructured forms. For this reason, unstructured data extracted from social media should be transformed into a structural form with text mining methods. The problem of text mining can be defined as the extraction of meaningful, useful, and previously unknown information from textual data [29]. The text mining method begins with data pre-processing which contains three steps.

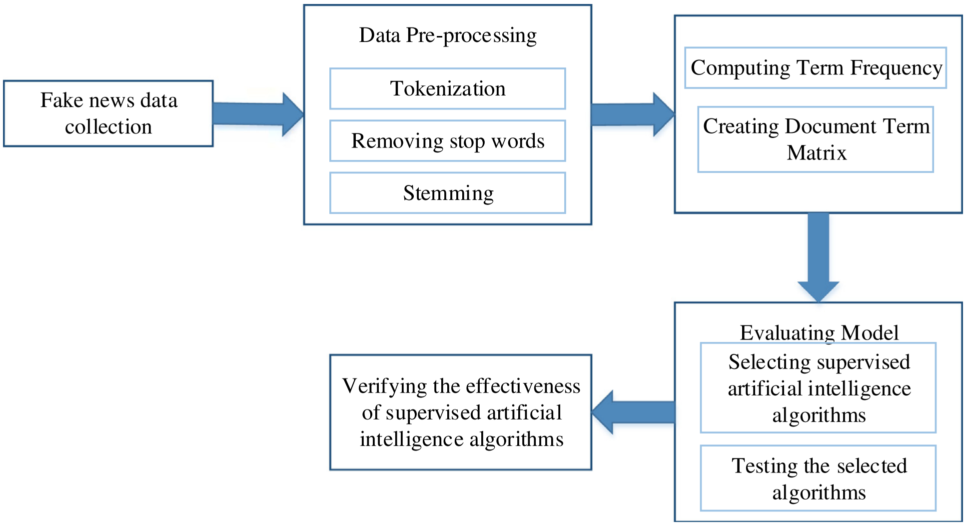


Fig. 1. The process of the proposed model.

Table 2
Data pre-processing.

Input: Given textual data **Output:** Pre-processed data
1. Remove numbers from textual data. 2. Delete punctuation characters from textual data. 3. Filter characters which contain character $< N$. 4. Apply case converter to textual data. 5. Remove stop words. 6. Stem textual data.

3.1.1. *Data pre-processing*

Tokenization

Tokenization process divides the given text into smaller parts which are called tokens and it removes all the punctuations from the textual data [30]. The number filter has been applied to remove terms which contain numbers. The case converter has been used to transform text data to lowercase or uppercase. In this paper, all the terms have been converted into lowercase. Finally, in this step, the N-chars filter has been used to delete the words, which consist of less than N characters.

Stop-words removal

Stop-words are not important words, although they are frequently used to complete the sentence structure and connect expressions. These words are language-specific words, which do not carry information. Conjunctions, pronouns, and prepositions are stop-words. There are about 400–500 stop-words in English [31]. Some of the stop words are a, an, about, by, but, that, does, on, above, once, after, until, too, again, when, where, what, all, am, and, any, against, and so on.

Stemming

In this process, different grammatical forms of a word like its adjective, adverb, noun, verb, etc. have been transformed into its root form. The aim of stemming is to obtain the basic forms of the words whose meanings are the same, but the word forms are different from each other. For example, the words, connection, connections, connective, connected, and connecting can be stemmed to the word 'connect'. Data pre-processing steps have been given in Table 2.

3.1.2. *Extraction and selection of features*

The biggest problem in text mining is high dimensional data. Therefore, it is necessary to remove unrelated and redundant features to improve model accuracy. In data pre-processing steps, features are extracted from high dimensional unstructured data. A feature selection method that selects the stem terms in the data sets frequency of which are bigger than the threshold is used in this study. After this process, the terms in the data set for each document have been weighted and each document has been converted into a vector of term weights. The basic representation is called the Vector Space Model (VSM). In VSM, each word is represented by a value that indicates the weight of the word in the document. Several different methods such as Term Frequency-Inverse Document Frequency (TF-IDF), Inverse Document Frequency (IDF), Term Frequency (TF), or binary representation have been developed to compute these weights. TF and TF-IDF are the best known of these methods.

$$\begin{array}{c} D_1 \\ D_2 \\ \vdots \\ D_m \end{array} \begin{pmatrix} W_{11} & W_{12} & W_{13} & \cdots & W_{1n} \\ W_{21} & W_{22} & W_{23} & \cdots & W_{2n} \\ \vdots & & & & \vdots \\ W_{m1} & W_{m2} & W_{m3} & \cdots & W_{mn} \end{pmatrix}$$

Fig. 2. Document-Term Matrix.

In this study, TF has been used to compute the weight for each word in each document. TF shows the number of times a word comes into view in a document [32]. TF is calculated using Eq. (1).

$$Term\ Frequency : TF = \frac{n_{ij}}{|d_i|} \tag{1}$$

d_i is the sum of all terms in the ith document. n_{ij} is the number of jth word in the ith document.

Document-Term Matrix (DTM) is created according to the weights of the words after TF value is calculated for each word in the document. The DTM is defined as in Fig. 2 and is an m x n matrix. In the matrix, each row represents the documents, columns represent the terms and cells are real numbers that indicate the weight of the terms in the document.

3.2. Supervised artificial intelligence algorithms

Supervised artificial intelligence algorithms are form of learning based on the training set. They assume that the labels of instances in the data sets are already known. They require a training set of labeled data and return a function that instances to the supervised data. The task of the function is to estimate the output having the minimum error rate from the input. In the following section, information about supervised artificial intelligence algorithms used in this study is given.

3.2.1. BayesNet

BayesNet (Bayesian Network) is a classifier that uses different search methods and various quality criteria. Bayes theorem based BayesNet [33] is a probabilistic graphical model used to represent conditional dependencies between multiple sets of variables. The generated Bayesian network consists of a directed graph and probability tables.

3.2.2. JRip

Jrip performs a propositional rule learner that is developed by [34]. In this algorithm, the samples of the data set are processed within increasing order and a set of rules are generated for the data set. Then the next class is processed and this process continues until all classes are covered.

3.2.3. OneR

OneR (One Rule) is a simple and fast algorithm proposed by Holt [35]. In this method, simple rules based on a single attribute are generated. The algorithm chooses one rule which has the minimum error rate for prediction [36]. If two or more rules have the same error rate, the rule is selected randomly.

3.2.4. Decision Stump

Decision Stump is a sophisticated classifier that has only one-level of the decision tree. It has one root that is connected to its leaves. Decision Stump is applied to ensemble algorithms like Adaboost [37]. The algorithm uses the values of a single input feature for prediction.

3.2.5. ZeroR

ZeroR is the simplest classifier that depends on the target class and ignores all predictors. ZeroR selects the highest-frequency label all the time [38].

3.2.6. Stochastic gradient descent

Stochastic gradient descent is a modern classifier that uses an iterative method to optimize an objective function. This classifier uses randomly selected samples to evaluate the gradients, for this reason, it is called stochastic [39].

Please cite this article as: F.A. Ozbay and B. Alatas, Fake news detection within online social media using supervised artificial intelligence algorithms, Physica A (2019) 123174, https://doi.org/10.1016/j.physa.2019.123174.

3.2.7. CV Parameter Selection

CV Parameter Selection is a method used to select parameters with cross-validation for any classifier [40].

3.2.8. Randomizable filtered classifier

The basic idea of this classifier is that the decision of a committee is better than that of an individual [41]. In this method, the training data are filtered through an optional filter and then a random classifier is applied to the data. The filter only works on training data. If a randomizable filter is used, each base classifier is generated with a random number seed. The final prediction is calculated by averaging the estimates generated by the individual base classifiers.

3.2.9. Logistic model tree

The classifier combines decision tree learning and logistic regression to obtain a single tree. LogitBoost algorithm is used to obtain a model of Logistic Regression at every node in the tree [42]. Logitboost is used to build a logistic model that is obtained using the whole data. A threshold value is used to split the data at the root. The split process continues until the stop condition is satisfied [43].

3.2.10. Locally weighted learning

In this method, an instance-based algorithm is used to assign the weight of instances. Then this algorithm is used for classification or regression. The algorithm creates a local model based on neighboring data for the entire function space stores the training data in the memory [44].

3.2.11. Classification via clustering

The method uses a cluster for classification. Clustering is the treat of grouping similar samples. Clustering can be regarded as a preprocessing step before classifying the data. Classification is based on the clustering process [45].

3.2.12. Weighted instances handler wrapper

In this method, the training data are passed through to the base classifier if it can handle instance weights. However, it is possible to force the use of resampling with weights as well.

3.2.13. Ridor

The algorithm creates a default rule and then the exception of the rule that has the least error rate. The most excellent exception is generated for each exception and in this way, a tree-like expansion of the exceptions is obtained. It continues to produce until the best exceptions for each exception is achieved and ends when the pure is obtained [46].

3.2.14. Multi-Layer Perceptron

The Multi-Layer Perceptron (MLP) algorithm was proposed by Rosenblatt in 1950. An MLP has at least three layers of nodes: an input layer, a hidden layer, and an output layer. This algorithm uses a supervised learning technique called backpropagation for training. MLP is discriminated from a linear perceptron by its multiple layers. In MLP, learning depends on the amount of error in the output. The weights of connection are changed after each piece of data is processed [47].

3.2.15. Ordinal Learning Model

The Ordinal Learning Model (OLM) was proposed by Ben-David et al. for ordinal classification with monotonicity constraints [48]. In the learning stage, each example is checked to each rule in a rule-base. If an example is not consistent to a rule in the rule base, one of them is randomly selected when discarding, but if the example is selected, it must be checked for consistency against all other monotonicity rules. If example passes the consistency test, it is assumed as a rule.

3.2.16. Simple Cart

The method was proposed by Leo Breiman [49] to be an alternative to conventional methods for data exploration and prediction. Simple Cart creates the binary decision tree for classification. In the pruning phase, cross-validation or a large test data is used for selecting the best tree. In this method, the data is separated into two subgroups that have a different outcome. This process is terminated when the subgroup size is minimum [50].

3.2.17. Attribute Selected Classifier

Attribute Selected Classifier algorithm decreases the dimensionality of training and test data before starting the classification process [51].

3.2.18. J48

The J48 algorithm is often the preferred algorithm for classification applications. In this respect, J48 is a statistical decision tree algorithm based on ID3 and C4.5 algorithms. J48 works on nodes with the logic of using one node from each tree and the child rows from the top row. In this regard, it is generally one of the fastest and highest accuracy algorithms among classification algorithms [52].

Please cite this article as: F.A. Ozbay and B. Alatas, Fake news detection within online social media using supervised artificial intelligence algorithms, Physica A (2019) 123174, https://doi.org/10.1016/j.physa.2019.123174.

3.2.19. Sequential Minimal Optimization (SMO)

The algorithm was first introduced in 1998 by Platt. SMO algorithm is mainly used to strengthen the training of Support Vector Machines (SVM). It is an algorithm has used to solve the Quadratic Programming (QP) problem that arises for training SVM. This algorithm is applied by the popular LIBSVM tool, especially while is using for training SVM. SMO's working logic divides a problem into a set of the smallest possible sub-problems, which are then solved analytically. Subsequently, due to the constraint of linear equality contained in the Lagrange multipliers, the smallest possible problem is subdivided into two factors. Then, the restrictions applied to any two factors are reduced. This reduction problem is solved analytically to find a minimum one-dimensional quadratic function. In each iteration, the terms in the constant constraint of equality are calculated as the negative of the sum in the rest [53].

3.2.20. Bagging

Bagging algorithm was proposed by L. Breiman in 1994. It is a method of retraining the basic learner by deriving new sets of training from an existing training set. During training, substituting sampling is performed. The training set is produced randomly by replacing the training set consisting of n samples with a sample set with n samples in Bagging. Each selected sample is returned to the training set. Some examples are not included in the new training set, while others may take place more than once. With the education clusters chosen randomly, successful basic learners are trained to ensure nonconformity. Thus, a collective success is achieved. Basic learners do not have to be decision trees. In this respect, any machine learning algorithm can be a basic learner. However, even in the change of the minimum set of training for Bagging, the selection of basic learners that will affect the outcome to the maximum will increase the success [54].

3.2.21. Decision tree

Tree-based learning algorithms are one of the most used supervised learning algorithms. Methods such as Decision Trees, Random Forests, and Gradient Boosting are widely used in all kinds of data science problems. Decision Trees have a predefined target variable. In terms of their structure, they offer a top-down strategy. A decision tree is a structure used to divide a data set containing a large number of records into smaller clusters by applying a set of decision rules. In other words, it is a structure used by dividing large amounts of records into very small groups of records by applying simple decision-making steps [55].

3.2.22. IBk

The IBk algorithm is inspired by the nearest neighbor algorithm (kNN). The IBk algorithm does not generate a model while classification but generates a prediction for a test sample just in time. The IBk algorithm uses a distance measure to find *k* "close" instances in the training data for each test sample and uses the selected instances to make an estimate. In incremental learning, it is necessary to define an instance-based framework for learning algorithms. The aim is to elaborate on the simplest IBL algorithm (IB1) and provide an analysis of which classes of concepts can be learned. In this step, the similarity function calculates the similarity in the form of numerical values between a training example and the examples in the concept description. The classification function takes the results and classification of the similarity function and it performs the recording of the instances in the concept description [56].

3.2.23. Kernel Logistic Regression

Kernel Logistic Regression (KLR) is a kernel version of Logistic Regression (LR) and it is used as a powerful classification technique in many classification problems. However, since this classification algorithm is costly in terms of calculation, it is not preferred especially in large scale data classification problems. In KLR, in contrast to SVM, the input vector is mapped to a high dimensional field. In addition, KLR can be expanded to address multi-class classification and requires only unconstrained optimization problems [57].

3.3. Performance evaluation metrics

Different evaluation metrics have been utilized to compare the performances of the supervised artificial intelligence algorithms for fake news detection. Evaluation metrics are frequently used in supervised artificial intelligence algorithms and allow us to test the effectiveness of the algorithm. A confusion matrix has been used to evaluate the performance of fake news detection as shown in Table 3. In this matrix, samples are classified as fake or real.

TP, FP, TN, and FN concepts are explained as follows:

- True Positive (TP): If the predicted fake news is actually fake news, the prediction is TP.
- False Positive (FP): If the predicted fake news is actually real news, the prediction is FP.
- True Negative (TN): If the predicted real news is actually real news, the prediction is TN.
- False Negative (FN): If the predicted real news is actually fake news, the prediction is FN.

Using the confusion matrix in Table 3, the following performance evaluation criteria are used [58]:

$$Accuracy = \frac{|TN| + |TP|}{|FN| + |FP| + |TN| + |TP|} \tag{2}$$

Please cite this article as: F.A. Ozbay and B. Alatas, Fake news detection within online social media using supervised artificial intelligence algorithms, Physica A (2019) 123174, https://doi.org/10.1016/j.physa.2019.123174.

Table 3
Confusion matrix for fake news.

Confusion matrix		Actual classes	
		Positive class	Negative class
Predicted classes	Positive class	True Positive (TP)	False Positive (FP)
	Negative class	False Negative (FN)	True Negative (TN)

Table 4
The selected features from the BuzzFeed Political News data set.

Data set	Selected terms
BuzzFeed Fake News	Political, presidential, people, America, nation, Donald, email, Hillary, election, vote, Clinton, Government, country, time, trump, candidate, American, Bill, report, white, tell, call, democrat, believe, women, told, Muslim, unit, support, look, sander, republican, former, campaign, donate, party, conservation, president

$$Precision = \frac{|TP|}{|FP| + |TP|} \tag{3}$$

$$Recall = \frac{|TP|}{|FN| + |TP|} \tag{4}$$

$$F - measure = 2 \times \frac{Recall \times Precision}{Recall + Precision} \tag{5}$$

In the fake news detection problem, the accuracy is the rate of correctly predicted news to all of the samples. Recall value shows the ratio of the fake news, which is correctly predicted over the total number of fake news. Precision metric measures the fake news, which is correctly predicted from the total predicted news in fake class. F-measure is the harmonic average value of the recall value and precision value obtained for fake news detection.

4. Experimental evaluations

In this paper, three different data sets have been used to evaluate intelligent classification algorithms. As declared in previous sections, twenty-three different supervised artificial intelligence algorithms; BayesNet, JRip, OneR, Decision Stump, ZeroR, Stochastic Gradient Descent (SGD), CV Parameter Selection (CVPS), Randomizable Filtered Classifier (RFC), Logistic Model Tree (LMT), Locally Weighted Learning (LWL), Classification Via Clustering (CvC), Weighted Instances Handler Wrapper (WIHW), Ridor, Multi-Layer Perceptron (MLP), Ordinal Learning Model (OLM), Simple Cart, Attribute Selected Classifier (ASC), J48, Sequential Minimal Optimization (SMO), Bagging, Decision Tree, IBk, and Kernel Logistic Regression (KLR) have been adapted to detect fake news.

70% of the data set has been used for training and 30% of the total data set has been used for testing of the algorithms within all data sets. Details of three data sets have been introduced and experimental results have been presented in the following sections.

4.1. Data set 1 – BuzzFeed Political News data set

This data set was collected by Horne and Adali from the fake election news article [59] in BuzzFeed 2016 [60]. Before the 2016 US Presidential Election, BuzzFeed has reviewed real and fake stories during the nine months to obtain these data. The data set contains 1627 news articles related to the 2016 U.S. election gathered from Facebook [61]. A section from the data set is shown in Fig. 3. The first column represents the number of news, the second column is the news in text, and the third column represents the label of the news as "Fake" or "Real". The used features by selecting the stem terms in this data set frequency of which are bigger than the threshold (30 for this data set) is given in Table 4.

Keywords related to the election have identified to label fake news in the BuzzFeed Political News data set. The news confirmed by reputed news platforms has been identified as real news.

The previously mentioned supervised artificial intelligence algorithms have been applied to the BuzzFeed Political News data set to predict whether the news is real or fake. TF is used to extract the feature from the data set. Table 5 shows the performance comparison for the different supervised artificial intelligence algorithms on the BuzzFeed Political News data set. Graphical representation of algorithm performances with respect to the accuracy, precision, recall, and F-measure metrics has been demonstrated in Fig. 4.

1	Hillary in hot water over her email server, again.	Fake
2	The list of Republicans supporting Hillary Clinton is still growing.	Real
3	Democratic nominee Hillary Clinton is in hot water again after nearly 5 milli...	Fake
4	For months, we have been keeping track of all of the GOP politicians, adm...	Real
5	The election commission has an emergency meeting scheduled for tomorro...	Fake
6	The most recent high-profile announced was from former secretary of sta...	Real
7	Ms. Clinton has already come under fire during this election cycle over usi...	Fake
8	General Powell said at a meeting of the Long Island Association that he w...	Real
9	Ms. Clinton?s public relations official released a brief statement saying, ?H...	Fake
10	Powell's support isn't much of a surprise, given his support for President O...	Real
11	These ballots could change the fortune of Bernie Sanders in his bid for the...	Fake
12	In his leaked emails from last month, Powell also appeared to be talking ov...	Real
13	South African Billionaire, Femi Adenugame, has released a statement offe...	Fake
14	But at least it's official now. And Powell is now the third Bush Cabinet offic...	Real
15	Concerns about Donald Trump becoming president has prompted a South ...	Fake
16	Kathleen Kennedy Townsend said in a Facebook post that Bush told her t...	Real
17	Americans have trouble trusting the Clintons for many reasons, but near t...	Fake
18	The president's office isn't confirming it, but she stands by it.	Real
19	Now, it appears that yet another name has been added to this list.	Fake
20	That?s what he said, she told Politico	Real
21	Former U.N. President John Ashe was reportedly found dead last Wednes...	Fake
22	a moderate Republican who is retiring this year, told Syracuse.com that h...	Real
23	Ashe?s death became even more suspicious when police learned that he ...	Fake
24	He cited Trump's criticism of Khizr Khan.	Real
25	It was later revealed that Seng also illegally funneled several hundred tho...	Fake
26	I think Trump is a national embarrassment, Hanna said. Is he really the gu...	Real
27	?It would have been very embarrassing,? the source added. ?His death w...	Fake
28	Loose lips sink ships. Got that, Trump? Loose lips sink ships. Warner adde...	Real

Fig. 3. A section from the BuzzFeed Political News data set.

Table 5
The performance of different supervised artificial intelligence algorithms for the BuzzFeed Political News data set.

	Accuracy	Precision	Recall	F-measure
BayesNet	0,620	0,640	0,582	0,610
JRip	0,589	0,592	0,626	0,609
OneR	0,507	0,514	0,639	0,569
Decision Stump	0,532	**0,747**	0,534	0,534
ZeroR	0,510	0,509	**1,000**	**0,675**
SGD	0,605	0,619	0,590	0,604
CVPS	0,509	0,509	**1,000**	**0,675**
RFC	0,604	0,621	0,574	0,604
LMT	0,619	0,627	0,621	0,627
LWL	0,558	0,642	0,558	0,490
CvC	0,501	0,507	0,777	0,613
WIHW	0,509	0,509	**1,000**	**0,675**
Ridor	0,562	0,567	0,592	0,579
MLP	0,638	0,640	0,639	0,639
OLM	0,538	0,573	0,573	0,489
SimpleCart	0,646	0,654	0,649	0,652
ASC	0,563	0,616	0,563	0,516
J48	**0,655**	0,655	0,681	0,668
SMO	0,619	0,629	0,616	0,622
Bagging	0,653	0,666	0,642	0,653
Decision Tree	0,634	0,626	0,707	0,664
IBk	0,513	0,441	0,480	0,460
KLR	0,521	0,481	0,583	0,527

From Table 5; it can be seen that in terms of accuracy, the J48 has performed the best for the data set with an accuracy of 0,655. In this data set, the worst accuracy of 0,501 has been achieved using CvC. In terms of precision, the Decision Stump algorithm has the highest precision among the twenty-three algorithms while again CvC has the lowest precision. On recall metric, ZeroR, CVPS, and WIHW have the highest performance with a precision of 1000. In terms of F-measure, the highest value has been achieved by ZeroR, CVPS, and WIHW (1000) while the lowest value has been achieved by IBk (0,460).

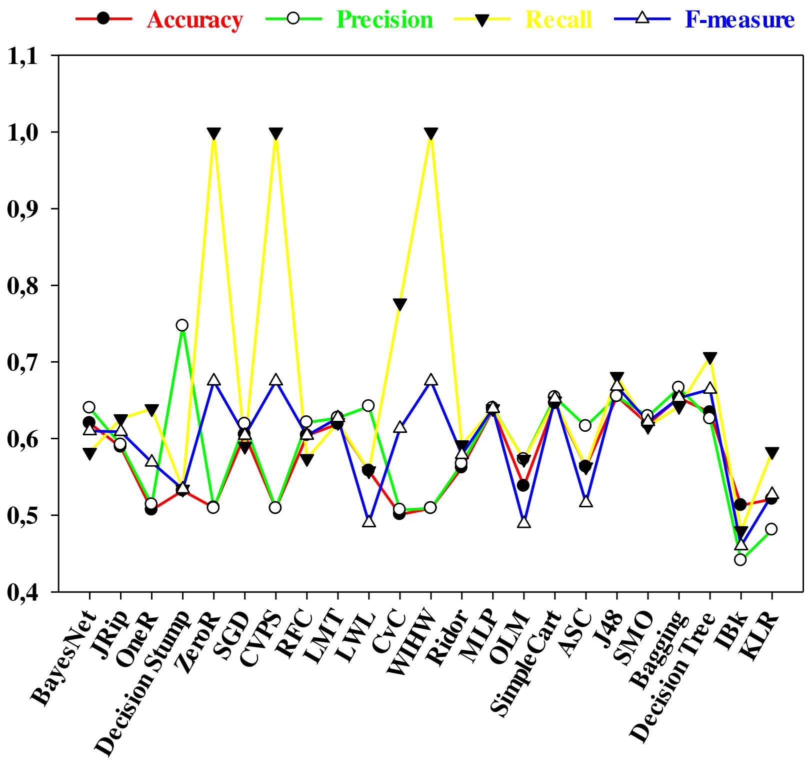
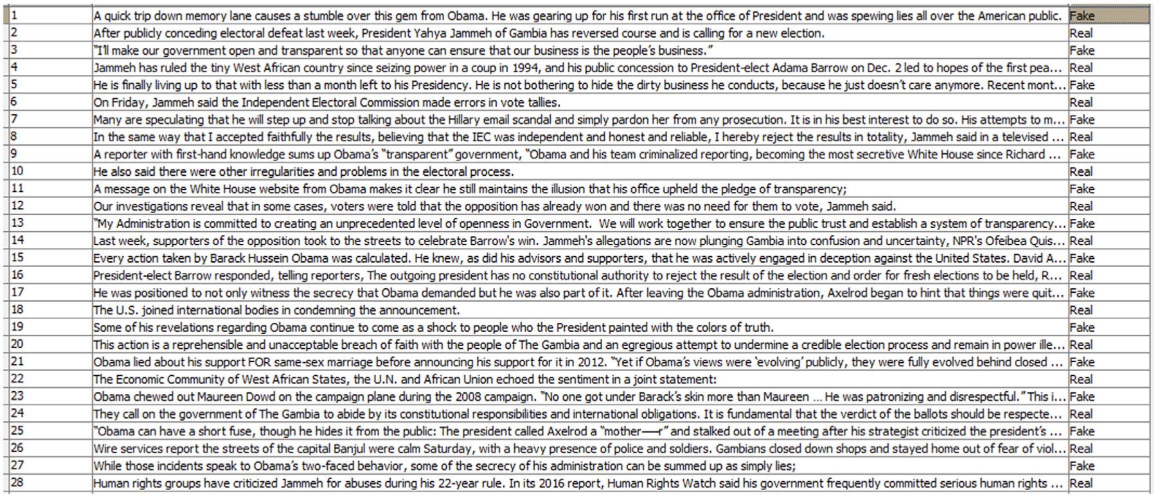


Fig. 4. Performances of artificial intelligence algorithms in BuzzFeed Political News data set.



1	A quick trip down memory lane causes a stumble over this gem from Obama. He was gearing up for his first run at the office of President and was spewing lies all over the American public.	Fake
2	After publicly conceding electoral defeat last week, President Yahya Jammeh of Gambia has reversed course and is calling for a new election.	Real
3	"I'll make our government open and transparent so that anyone can ensure that our business is the people's business."	Fake
4	Jammeh has ruled the tiny West African country since seizing power in a coup in 1994, and his public concession to President-elect Adama Barrow on Dec. 2 led to hopes of the first pea…	Real
5	He is finally living up to that with less than a month left to his Presidency. He is not bothering to hide the dirty business he conducts, because he just doesn't care anymore. Recent mont…	Fake
6	On Friday, Jammeh said the Independent Electoral Commission made errors in vote tallies.	Real
7	Many are speculating that he will step up and stop talking about the Hillary email scandal and simply pardon her from any prosecution. It is in his best interest to do so. His attempts to m…	Fake
8	In the same way that I accepted faithfully the results, believing that the IEC was independent and honest and reliable, I hereby reject the results in totality, Jammeh said in a televised …	Real
9	A reporter with first-hand knowledge sums up Obama's "transparent" government, "Obama and his team criminalized reporting, becoming the most secretive White House since Richard …	Fake
10	He also said there were other irregularities and problems in the electoral process.	Real
11	A message on the White House website from Obama makes it clear he still maintains the illusion that his office upheld the pledge of transparency;	Fake
12	Our investigations reveal that in some cases, voters were told that the opposition has already won and there was no need for them to vote, Jammeh said.	Real
13	"My Administration is committed to creating an unprecedented level of openness in Government. We will work together to ensure the public trust and establish a system of transparency…	Fake
14	Last week, supporters of the opposition took to the streets to celebrate Barrow's win. Jammeh's allegations are now plunging Gambia into confusion and uncertainty, NPR's Ofeibea Quis…	Real
15	Every action taken by Barack Hussein Obama was calculated. He knew, as did his advisors and supporters, that he was actively engaged in deception against the United States. David A…	Fake
16	President-elect Barrow responded, telling reporters, The outgoing president has no constitutional authority to reject the result of the election and order for fresh elections to be held, R…	Real
17	He was positioned to not only witness the secrecy that Obama demanded but he was also part of it. After leaving the Obama administration, Axelrod began to hint that things were quit…	Fake
18	The U.S. joined international bodies in condemning the announcement.	Real
19	Some of his revelations regarding Obama continue to come as a shock to people who the President painted with the colors of truth.	Fake
20	This action is a reprehensible and unacceptable breach of faith with the people of The Gambia and an egregious attempt to undermine a credible election process and remain in power ille…	Real
21	Obama lied about his support FOR same-sex marriage before announcing his support for it in 2012. "Yet if Obama's views were 'evolving' publicly, they were fully evolved behind closed …	Fake
22	The Economic Community of West African States, the U.N. and African Union echoed the sentiment in a joint statement:	Real
23	Obama chewed out Maureen Dowd on the campaign plane during the 2008 campaign. "No one got under Barack's skin more than Maureen … He was patronizing and disrespectful." This i…	Fake
24	They call on the government of The Gambia to abide by its constitutional responsibilities and international obligations. It is fundamental that the verdict of the ballots should be respecte…	Real
25	"Obama can have a short fuse, though he hides it from the public: The president called Axelrod a "mother——r" and stalked out of a meeting after his strategist criticized the president's …	Fake
26	Wire services report the streets of the capital Banjul were calm Saturday, with a heavy presence of police and soldiers. Gambians closed down shops and stayed home out of fear of viol…	Real
27	While those incidents speak to Obama's two-faced behavior, some of the secrecy of his administration can be summed up as simply lies;	Fake
28	Human rights groups have criticized Jammeh for abuses during his 22-year rule. In its 2016 report, Human Rights Watch said his government frequently committed serious human rights …	Real

Fig. 5. A section from Random Political News data set.

4.2. Data set 2 – Random Political News data set

There is only political news in data set 1. Therefore, Horne and Adali have created their political news data set. They have collected fake news from Zimdars' list of fake and misleading news websites [62] and real news from Business Insider's "Most Trusted" list [63]. This data set consists of 75 news articles collected from different sources. A section from this data set is demonstrated in Fig. 5. The used features by selecting the stem terms in this data sets frequency of which are bigger than the threshold (55 for this data set) is given in Table 6.

Table 6
The selected features from the Random Political News data set.

Data set	Selected terms
Random Fake News	Public, include, senate, call, white, time, people, news, email, government, told, day, campaign, elect, republican, report, nation, week, vote, office, Hillary, Clinton, support, presidential, Obama, administration, democrat, political, country, American, house, intelligent, Trump, Donald, elector, office, change, president-election, congress

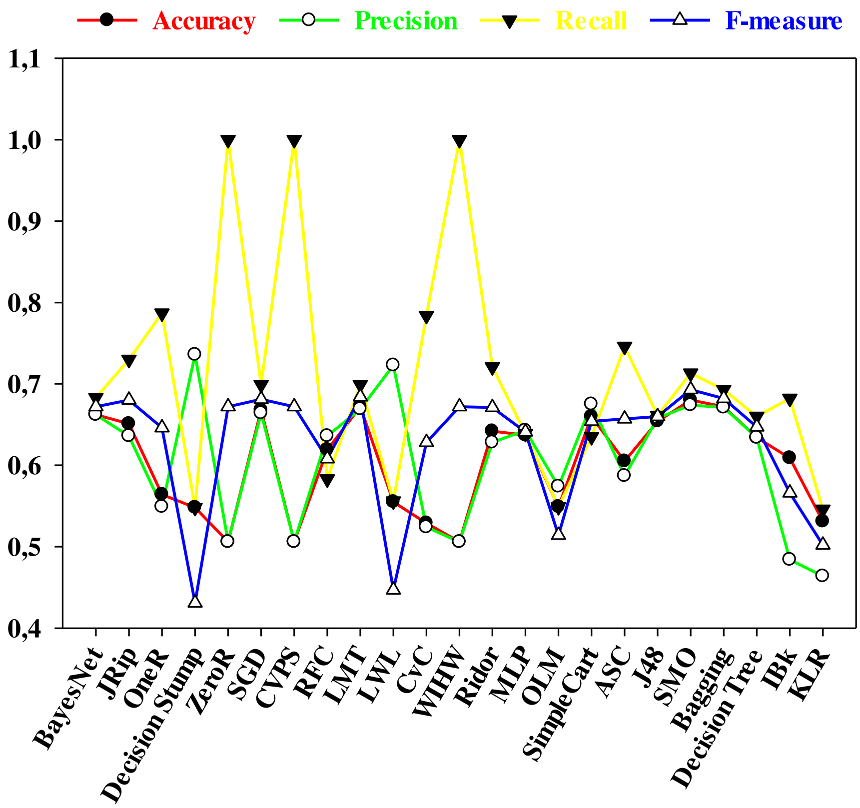


Fig. 6. Performances of artificial intelligence algorithms in Random Political News data set.

Table 7 gives the obtained results for the different supervised artificial intelligence algorithms on the random political news data set. Fig. 6 demonstrates the graphical representation of a comparison of the results obtained by twenty-three supervised artificial intelligence algorithms.

In terms of accuracy and F-measure, the SMO algorithm has the highest value among the twenty-three algorithms. Decision Stump seems the best algorithm in term of the precision metric. The lowest precision achieved is 0,464 when using KLR. ZeroR, CVPS, and WIHW have the highest performance with a recall of 1000 also in this data set. The worst recall has been achieved by KLR and the worst F-measure has been achieved by Decision Stump.

4.3. Data set 3 – ISOT Fake News data set

This data set contains two types of news, fake and real news. Fake and real news are obtained from real-world sources. The real news articles were collected from Reuters.com, while the fake news articles were collected from unreliable various websites such as Politifact and Wikipedia [64]. The data set consists of a total of 44 898 data, 21 417 real-labeled data and 23 481 false-labeled data. A section from this data set is shown in Fig. 7. A section from this data set is demonstrated in Fig. 5. The used features by selecting the stem terms in this data sets frequency of which are bigger than the threshold (1500 for this data set) is given in Table 8.

The performance comparison for the different supervised artificial intelligence algorithms on the ISOT Fake News data set is given in Table 9. The performances of these algorithms for this data set have been shown in Fig. 8. According to the obtained results this table and figure, the best values in terms of accuracy, precision, and F-measure have been obtained from the Decision Tree algorithm. The best recall value as 1000 has achieved by ZeroR, CVPS, and WIHW algorithms.

Table 7
The performance of different supervised artificial intelligence algorithms for the Random Political News data set.

	Accuracy	Precision	Recall	F-measure
BayesNet	0,662	0,662	0,683	0,672
JRip	0,651	0,636	0,730	0,680
OneR	0,564	0,549	0,787	0,646
Decision Stump	0,548	**0,736**	0,548	0,431
ZeroR	0,506	0,506	**1,000**	0,672
SGD	0,668	0,664	0,699	0,681
CVPS	0,506	0,506	**1,000**	0,672
RFC	0,619	0,636	0,583	0,608
LMT	0,672	0,669	0,699	0,684
LWL	0,555	0,723	0,556	0,447
CvC	0,529	0,524	0,784	0,628
WIHW	0,506	0,506	**1,000**	0,672
Ridor	0,642	0,628	0,721	0,671
MLP	0,637	0,643	0,639	0,641
OLM	0,549	0,574	0,549	0,514
SimpleCart	0,660	0,675	0,635	0,654
ASC	0,605	0,587	0,746	0,657
J48	0,654	0,658	0,661	0,660
SMO	**0,680**	0,674	0,713	**0,693**
Bagging	0,672	0,671	0,693	0,682
Decision Tree	0,634	0,634	0,660	0,647
IBk	0,609	0,484	0,682	0,566
KLR	0,531	0,464	0,546	0,502

Table 8
The selected features from the ISOT Fake News data set.

Data set	Selected terms
ISOT Fake News	Presidential, American, office, include, nation, accord, people, time, feature, imagine, day, donald, Trump, told, report, unit, election, month, republican, white, house, former, look, country, call, campaign, administration, tri, news, statement, week, issue, official, congress, support, senate, government, leader, democrat, help

1	This is Disturbing,Donald Trump just couldn t wish all Americans a Happy New Year and leave it at that. Instead, he had to give a shout out to his enemies, haters and the very dish...	Fake
2	House Intelligence Committee Chairman Devin Nunes is going to have a bad day. He s been under the assumption, like many of us, that the Christopher Steele-dossier was what pro...	Fake
3	On Friday, it was revealed that former Milwaukee Sheriff David Clarke, who was being considered for Homeland Security Secretary in Donald Trump s administration, has an email sca...	Fake
4	On Christmas day, Donald Trump announced that he would be back to work the following day, but he is golfing for the fourth day in a row. The former reality show star blasted for...	Fake
5	Pope Francis used his annual Christmas Day message to rebuke Donald Trump without even mentioning his name. The Pope delivered his message just days after members of the Uni...	Fake
6	The number of cases of cops brutalizing and killing people of color seems to see no end. Now, we have another case that needs to be shared far and wide. An Alabama woman by th...	Fake
7	Donald Trump spent a good portion of his day at his golf club, marking the 84th day he s done so since taking the oath of office. It must have been a bad game because just after th...	Fake
8	In the wake of yet another court decision that derailed Donald Trump s plan to bar Muslims from entering the United States, the New York Times published a report on Saturday morni...	Fake
9	Many people have raised the alarm regarding the fact that Donald Trump is dangerously close to becoming an autocrat. The thing is, democracies become autocracies right under the...	Fake
10	Just when you might have thought we d get a break from watching people kiss Donald Trump s ass and stroke his ego ad nauseam, a pro-Trump group creates an ad that s nothing b...	Fake
11	A centerpiece of Donald Trump s campaign, and now his presidency, has been his white supremacist ways. That is why so many of the public feuds he gets into involve people of colo...	Fake
12	Republicans are working overtime trying to sell their scam of a tax bill to the public as something that directly targets middle-class and working-class families with financial relief. Nothi...	Fake
13	Republicans have had seven years to come up with a viable replacement for Obamacare but they failed miserably. After taking a victory lap for gifting the wealthy with a tax break o...	Fake
14	The media has been talking all day about Trump and the Republican Party s scam of a tax bill	Fake
15	Abigail Disney is an heiress with brass ovaries who will profit from the GOP tax scam bill but isn t into f-cking poor people over. Ms. Disney penned an op-ed for USA Today in which sh...	Fake
16	Donald Trump just signed the GOP tax scam into law. Of course, that meant that he invited all of his craven, cruel GOP sycophants down from their perches on Capitol Hill to celebrat...	Fake
17	A new animatronic figure in the Hall of Presidents at Walt Disney World was added, where every former leader of the republic is depicted in an audio-animatronics show. The figure ...	Fake
18	Trump supporters and the so-called president s favorite network are lashing out at special counsel Robert Mueller and the FBI. The White House is in panic-mode after Mueller obtain...	Fake
19	Right now, the whole world is looking at the shocking fact that Democrat Doug Jones beat Republican Roy Moore in the special election to replace Attorney General Jeff Sessions in t...	Fake
20	Senate Majority Whip John Cornyn (R-TX) thought it would be a good idea to attack Special Counsel Robert Mueller over the Russia probe. As Mueller s noose tightens, Republicans a...	Fake
21	Trump Invites NRA To Xmas Party On Sandy Hook Anniversary	Fake
22	In this #METOO moment, many powerful men are being toppled. It spans many industries, from entertainment, to journalism, to politics and beyond. Any man that ever dared to abu...	Fake
23	As a Democrat won a Senate seat in deep-red Alabama, social media offered up everyone s opinion because that s what social media does. Democrat Doug Jones narrowly defeated ...	Fake
24	Alabama is a notoriously deep red state. It s a place where Democrats always think that we have zero chances of winning especially in statewide federal elections. However, that is ...	Fake
25	A backlash ensued after Donald Trump launched a sexist rant against Kirsten Gillibrand Thursday morning, saying that the Democratic Senator would do anything for a campaign con...	Fake
26	Donald Trump is afraid of strong, powerful women. He is a horrific misogynist, and has shown himself to be so over and over again. That is nothing new. He has mocked the weight of...	Fake
27	Ronald Reagan is largely seen as the Messiah of the Republican Party. Despite how long it has been since the man was president, he has always remained the high standard of GOP ...	Fake
28	Judge Jeanine Pirro has continued her screamy ragey meltdown over special counsel Robert Mueller s investigation into any possible collusion between the Trump campaign and Russ...	Fake
29	Donald Trump held a rally for Alabama Senate candidate and alleged pedophile Roy Moore in Pensacola, Florida on Friday night which he later claimed was packed to the rafters but ...	Fake
30	When Sen. Al Franken (D-MN) announced his plans to resign Thursday, he specifically called out Donald Trump over the Access Hollywood video and Roy Moore, an alleged pedophile...	Fake
31	In America, we have been having a conversation about police brutality against black Americans. Despite the countless black people murdered unjustly by police, there is usually no ju...	Fake
32	New questions are being asked about President Donald Trump s ability to lead after he slurred his words during a speech about his Jerusalem decision. Possible reasons for this includ...	Fake
33	On Wednesday, Donald Trump took a step no previous president had ever dared to consider. He has declared that the United States now officially recognizes the city of Jerusalem a...	Fake
34	Why He Is Recognizing Jerusalem Today,President Donald Trump announced yesterday that he plans to formally recognize Jerusalem as the capital of Israel. While this is left most p...	Fake
35	While on the campaign trail, Donald Trump promised to revive the coal industry and after he took power, he signed an executive order rolling back a temporary ban on mining coal an...	Fake

Fig. 7. A section from ISOT Fake News data set.

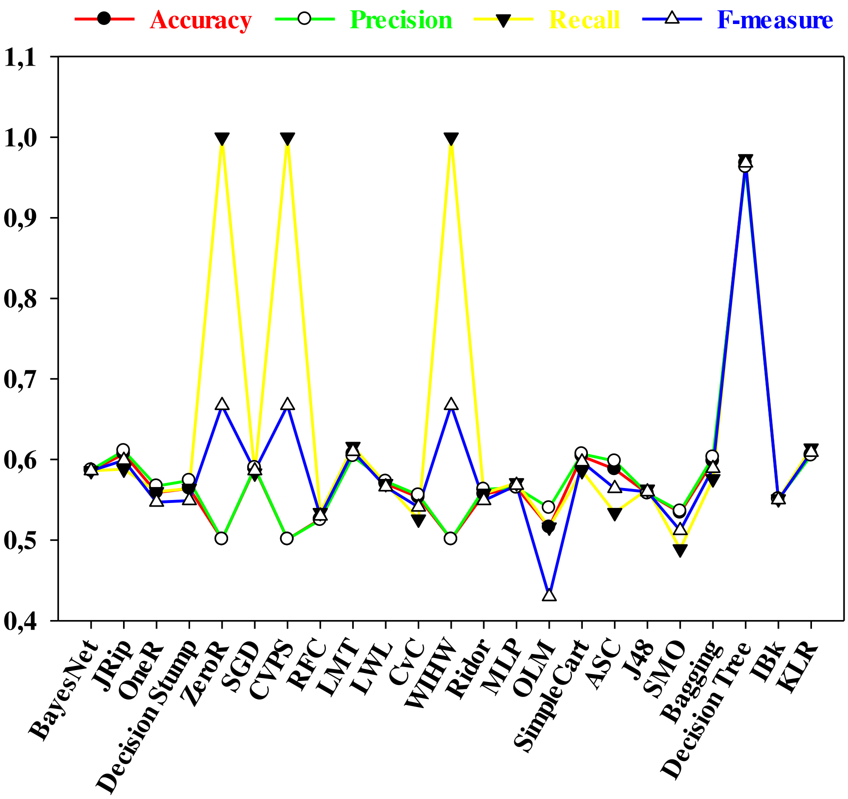


Fig. 8. Performances of artificial intelligence algorithms in ISOT Fake News data set.

Table 9
The performance of different supervised artificial intelligence algorithms for the ISOT Fake News data set.

	Accuracy	Precision	Recall	F-measure
BayesNet	0,586	0,587	0,586	0,586
JRip	0,607	0,611	0,588	0,599
OneR	0,559	0,567	0,560	0,547
Decision Stump	0,564	0,574	0,564	0,549
ZeroR	0,501	0,501	**1,000**	0,667
SGD	0,589	0,590	0,583	0,586
CVPS	0,501	0,501	**1,000**	0,667
RFC	0,526	0,525	0,534	0,530
LMT	0,607	0,604	0,616	0,610
LWL	0,570	0,573	0,570	0,566
CvC	0,553	0,556	0,526	0,541
WIHW	0,501	0,501	**1,000**	0,667
Ridor	0,557	0,563	0,558	0,549
MLP	0,565	0,565	0,571	0,568
OLM	0,516	0,540	0,516	0,430
SimpleCart	0,604	0,607	0,586	0,597
ASC	0,588	0,598	0,534	0,564
J48	0,558	0,558	0,563	0,560
SMO	0,534	0,536	0,489	0,512
Bagging	0,598	0,603	0,576	0,589
Decision Tree	**0,968**	**0,963**	0,973	**0,968**
IBk	0,551	0,551	0,551	0,550
KLR	0,606	0,605	0,614	0,609

5. Discussion

BayesNet, JRip, OneR, Decision Stump, ZeroR, Stochastic Gradient Descent (SGD), CV Parameter Selection (CVPS), Randomizable Filtered Classifier (RFC), Logistic Model Tree (LMT), Locally Weighted Learning (LWL), Classification Via

Table 10
Performance comparison of the algorithms for BuzzFeed Political News data set.

Accuracy	J48 > Bagging > SimpleCart > MLP > Decision Tree > BayesNet > LMT > SMO > SGD > RFC > JRip > ASC > Ridor > LWL > OLM > Decision Stump > KLR > IBk > ZeroR > CVPS > WIHW > OneR > CvC
Precision	Decision Stump > Bagging > J48 > SimpleCart > LWL > BayesNet > MLP > SMO > LMT > Decision Tree > RFC > SGD > ASC > JRip > OLM > Ridor > OneR > ZeroR > CVPS > WIHW > CvC > KLR > IBk
Recall	ZeroR = CVPS = WIHW > CvC > Decision Tree > J48 > SimpleCart > Bagging > OneR > MLP > JRip > LMT > SMO > Ridor > SGD > KLR > BayesNet > RFC > OLM > ASC > LWL > Decision Stump > IBk
F-measure	BayesNet > JRip > OneR > Decision Stump > ZeroR > SGD > CVPS > RFC > LMT > LWL > CvC > WIHW > Ridor > MLP > OLM > SimpleCart > ASC > J48 > SMO > Bagging > Decision Tree > IBk > KLR

Table 11
Performance comparison of the algorithms for Random Political News data set.

Accuracy	SMO > LMT = Bagging > SGD > BayesNet > SimpleCart > J48 > JRip > Ridor > MLP > Decision Tree > RFC > IBk > ASC > OneR > LWL > OLM > Decision Stump > KLR > CvC > ZeroR > CVPS > WIHW
Precision	Decision Stump > LWL > SimpleCart > SMO > Bagging > LMT > SGD > BayesNet > J48 > MLP > JRip > RFC > Decision Tree > Ridor > ASC > OLM > OneR > CvC > ZeroR > CVPS > WIHW > IBk > KLR
Recall	ZeroR = CVPS = WIHW > OneR > CvC > ASC > JRip > Ridor > SMO > SGD > LMT > Bagging > BayesNet > IBk > J48 > Decision Tree > MLP > SimpleCart > RFC > LWL > OLM > Decision Stump > KLR
F-measure	SMO > LMT > Bagging > SGD > JRip > BayesNet > ZeroR > CVPS > WIHW > Ridor > J48 > ASC > SimpleCart > Decision Tree > OneR > MLP > CvC > RFC > IBk > OLM > KLR > LWL > Decision Stump

Table 12
Performance comparison of the algorithms for ISOT Fake News data set.

Accuracy	Decision Tree > JRip > LMT > KLR > SimpleCart > Bagging > SGD > ASC > BayesNet > LWL > MLP > Decision Stump > OneR > J48 > Ridor > CvC > IBk > SMO > RFC > OLM > ZeroR > CVPS > WIHW
Precision	Decision Tree > JRip > SimpleCart > KLR > LMT > Bagging > ASC > SGD > BayesNet > Decision Stump > LWL > OneR > MLP > Ridor > J48 > CvC > IBk > OLM > SMO > RFC > ZeroR > CVPS > WIHW
Recall	ZeroR = CVPS = WIHW > Decision Tree > LMT > KLR > JRip > BayesNet > SimpleCart > SGD > Bagging > MLP > LWL > Decision Stump > J48 > OneR > Ridor > IBk > RFC > ASC > CvC > OLM > SMO
F-measure	Decision Tree > ZeroR > CVPS > WIHW > LMT > KLR > JRip > SimpleCart > Bagging > BayesNet > SGD > MLP > LWL > ASC > J48 > IBk > Decision Stump > Ridor > OneR > CvC > RFC > SMO > OLM

Clustering (CvC), Weighted Instances Handler Wrapper (WIHW), Ridor, Multi-Layer Perceptron (MLP), Ordinal Learning Model (OLM), Simple Cart, Attribute Selected Classifier (ASC), J48, Sequential Minimal Optimization (SMO), Bagging, Decision Tree, IBk, and Kernel Logistic Regression (KLR) have been for the first time analyzed for the problem of fake news detection in this study.

The order of performance of the algorithms for the BuzzFeed Political News Data set has been shown in Table 10. IBk seems the worst intelligent classification algorithm depending on precision and recall metrics.

Performance ordering of the algorithms for the Random Political News data set has been listed in Table 11. SMO algorithm seems the best algorithm according to accuracy and F-measure values. KLR seems the worst intelligent classification algorithm depending on precision and recall metrics.

Performances of the intelligent classification algorithms for the ISOT Fake News data set have been shown in Table 12. The decision tree algorithm has outperformed all other intelligent classification algorithms according to all evaluation metrics except recall within this data set. JRip algorithm seems the second best algorithm according to accuracy and precision metrics.

Mean performances of all supervised artificial intelligence algorithms with respect to all evaluation metrics for three data sets have been shown in Table 13. Mean performances have also been demonstrated in Fig. 9. According to the obtained results, the best mean values in terms of accuracy, precision, and F-measure have been obtained from the Decision Tree algorithm. The best mean recall value as 1000 has been achieved by ZeroR, CVPS, and WIHW algorithms.

Table 13
Mean performance of supervised artificial intelligence algorithms within all data sets.

	Mean accuracy	Mean precision	Mean recall	Mean F-measure
BayesNet	0,610	0,630	0,617	0,623
JRip	0,615	0,613	0,648	0,629
OneR	0,543	0,543	0,662	0,587
Decision Stump	0,548	0,685	0,548	0,505
ZeroR	0,505	0,505	**1,000**	0,671
SGD	0,621	0,624	0,624	0,624
CVPS	0,505	0,505	**1,000**	0,671
RFC	0,583	0,594	0,563	0,581
LMT	0,633	0,633	0,645	0,640
LWL	0,561	0,646	0,561	0,501
CvC	0,527	0,529	0,696	0,594
WIHW	0,505	0,505	**1,000**	0,671
Ridor	0,587	0,586	0,624	0,599
MLP	0,613	0,616	0,616	0,616
OLM	0,534	0,562	0,546	0,477
SimpleCart	0,636	0,645	0,623	0,634
ASC	0,585	0,600	0,614	0,579
J48	0,622	0,624	0,635	0,629
SMO	0,611	0,613	0,606	0,609
Bagging	0,641	0,646	0,637	0,641
Decision Tree	**0,745**	**0,741**	0,780	**0,759**
IBk	0,557	0,492	0,571	0,525
KLR	0,553	0,516	0,581	0,546

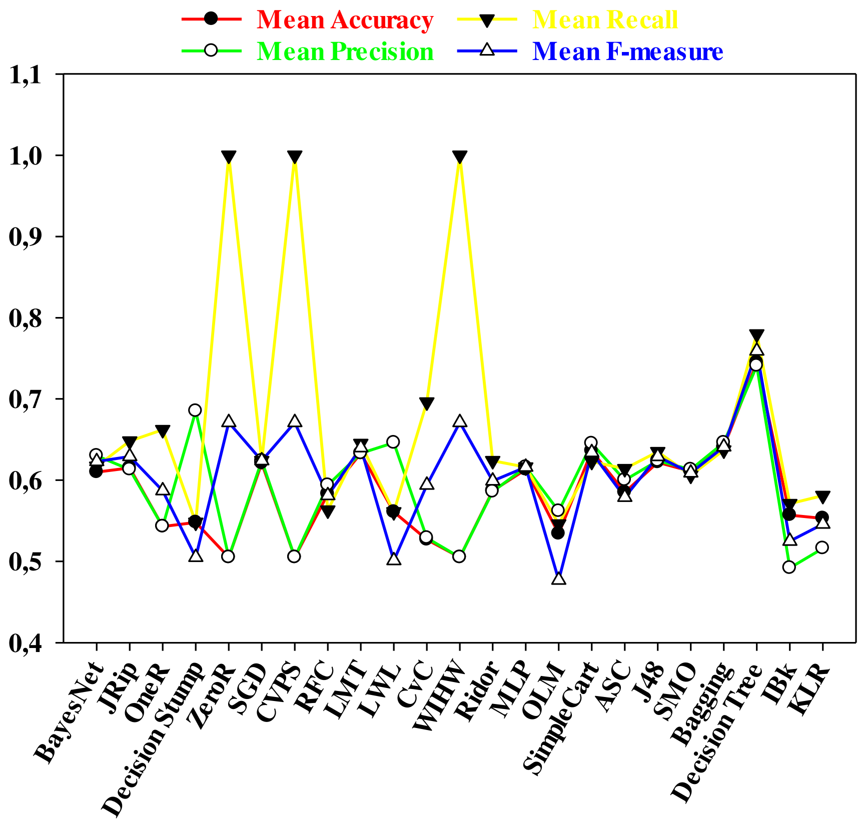


Fig. 9. Demonstration of mean performances of supervised artificial intelligence algorithms.

6. Conclusions and future works

In recent years, it has become difficult for users to access accurate and reliable information because of the increased amount of information on social media. In this study, a model is proposed to detect fake news in social media by combining text mining methods and supervised artificial intelligence algorithms. Text mining analysis and supervised artificial intelligence algorithms have been conducted separately. This combined model has been tested on three different real-world data set and evaluated according to accuracy, recall, precision, and F-measure values. Mean performances of all supervised artificial intelligence algorithms with respect to all evaluation metrics within three data sets have been calculated. According to the obtained results, the best mean values in terms of accuracy, precision, and F-measure have been obtained from the Decision Tree algorithm. ZeroR, CVPS, and WIHW algorithms, with 1000 value, seem the best algorithms in terms of recall metric.

In future works, the current work may be improved by exploring new algorithms, hybridizing the current algorithms, integrating intelligent optimization algorithms for better results. Ensemble methods and different feature extraction methods may also be integrated for improving the performances of the models.

References

[1] H. Allcott, M. Gentzkow, Social media and fake news in the 2016 election, J. Econ. Perspect. 31 (2017) 211–236.
[2] Y. Zhang, Y. Su, L. Weigang, H. Liu, Rumor and authoritative information propagation model considering super spreading in complex social networks, Physica A 506 (2018) 395–411, http://dx.doi.org/10.1016/j.physa.2018.04.082.
[3] V. L. Rubin, Y. Chen, N. J. Conroy, Deception detection for news: three types of fakes, in: Proceedings of the 78th ASIS & T Annual Meeting: Information Science with Impact: Research in and for the Community. Missouri, 2015.
[4] N. J. Conroy, V. L. Rubin, Y. Chen, Automatic deception detection: Methods for finding fake news, in: Proceedings of the Association for Information Science and Technology, Vol. 52, 2015, pp. 1-4.
[5] V. L. Rubin, N. J. Conroy, Y. Chen, Towards news verification: Deception detection methods for news discourse, in: Proceedings of Hawaii International Conference on System Sciences, USA, 2015.
[6] S. M. H. Dadgar, M. S. Araghi, M.M. Farahani, A novel text mining approach based on TF-IDF and Support Vector Machine for news classification, in: Proceedings of IEEE International Conference on Engineering and Technology (ICETECH), India, 2016, pp. 112-116.
[7] V. L. Rubin, N.J. Conroy, Y. Chen, S. Cornwell, Fake news or truth? using satirical cues to detect potentially misleading news, in: Proceedings of the Second Workshop on Computational Approaches to Deception Detection, California, 2016, pp. 7-17.
[8] Z. Jin, J. Cao, Y. Zhang, J. Luo, News verification by exploiting conflicting social viewpoints in microblogs, in: Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, Arizona, 2016, pp. 2972-2978.
[9] E. Tacchini, G. Ballarin, M. L. Della Vedova, S. Moret, L. de Alfaro, Some like it hoax: Automated fake news detection in social networks, in: Proceedings of the 2nd Workshop on Data Science for Social Good, Macedonia, 2017.
[10] K. Shu, A. Sliva, S. Wang, J. Tang, H. Liu, Fake news detection on social media: A data mining perspective, ACM SIGKDD Explor. Newsl. 19 (2017) 22–36, http://dx.doi.org/10.1145/3137597.3137600.
[11] H. Ahmed, I. Traore, S. Saad, Detecting opinion spams and fake news using text classification, Secur. Priv. 1 (2017) http://dx.doi.org/10.1002/spy2.9.
[12] S. Gilda, Evaluating machine learning algorithms for fake news detection, in: Proceedings of the IEEE 15th Student Conference on Research and Development, Malaysia, 2017, pp. 110-115.
[13] N. Ruchansky, S. Seo, Y. Liu, CSI: A hybrid deep model for fake news detection, in: Proceedings of the 2017 ACM on Conference on Information and Knowledge Management, Singapore, 2017, pp. 797-806.
[14] K. Shu, S. Wang, H. Liu,
[15] Y. Long, Q. Lu, R. Xiang, M. Li, C. R. Huang, Fake news detection through multi-perspective speaker profiles, in: Proceedings of the Eighth International Joint Conference on Natural Language Processing, Taiwan, 2017, pp. 252-256.
[16] Á. Figueira, L. Oliveira, The current state of fake news: challenges and opportunities, Procedia Comput. Sci. 121 (2017) 817–825, http://dx.doi.org/10.1016/j.procs.2017.11.106.
[17] C. Janze, M. Risius, Automatic detection of fake news on social media platforms, in: Proceedings of 21st Pacific-Asia Conference on Information Systems, Malaysia, 2017.
[18] V. Pérez-Rosas, B. Kleinberg, A. Lefevre, R. Mihalcea, Automatic detection of fake news, 2017, arXiv preprint arXiv:1708.07104.
[19] C. Buntain, J. Golbeck, Automatically identifying fake news in popular twitter threads, in: Proceedings of the IEEE International Conference on Smart Cloud, New York, 2017, pp. 208-215.
[20] A. Bessi, On the statistical properties of viral misinformation in online social media, Physica A 469 (2017) 459–470, http://dx.doi.org/10.1016/j.physa.2016.11.012.
[21] H. Zhu, H. Wu, J. Cao, G. Fu, H. Li, Information dissemination model for social media with constant updates, Physica A 502 (2018) 469–482, http://dx.doi.org/10.1016/j.physa.2018.02.142.
[22] K. Shu, S. Wang, H. Liu, Understanding user profiles on social media for fake news detection, in: Proceedings of the 1st IEEE Inernational Workshop on Fake MultiMedia, USA, 2018.
[23] S. Tschiatschek, A. Singla, M. Gomez Rodriguez, A. Merchant, A. Krause, Fake news detection in social networks via crowd signals, in: Proceedings of the World Wide Web Conferences, France, 2018, pp. 517-524.
[24] G. B. Guacho, S. Abdali, E. E. Papalexakis, Semi-supervised content-based fake news detection using tensor embeddings and label propagation, in: Proceedings of the 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), Spain, 2018, pp. 322-325.
[25] K. Shu, H.R. Bernard, H. Liu, Studying fake news via network analysis: detection and mitigation, in: Emerging Research Challenges and Opportunities in Computational Social Network Analysis and Mining, Springer, Cham, 2019, pp. 43–65.
[26] F. Monti, F. Frasca, D. Eynard, D. Mannion, M.M. Bronstein, Fake news detection on social media using geometric deep learning, 2019, arXiv preprint arXiv:1902.06673.
[27] A. Olivieri, S. Shabani, M. Sokhn, P. Cudré-Mauroux, Creating task-generic features for fake news detection, in: Proceedings of the 52nd Hawaii International Conference on System Sciences, Maui, 2019.
[28] C. Guo, J. Cao, X. Zhang, K. Shu, M. Yu, Exploiting emotions for fake news detection on social media, 2019, arXiv preprint arXiv:1903.01728.
[29] L. Kumar, P.K. Bhatia, Text Mining: concepts, process and applications, Int. J. Glob. Res. Comput. Sci. 4 (2013) 36–39.
[30] L.A. Mullen, K. Benoit, O. Keyes, D. Selivanov, J. Arnold, Fast, consistent tokenization of natural language text, J. Open Source Softw. 3 (2018) http://dx.doi.org/10.21105/joss.00655.
[31] S. Dharmendra, J. Suresh, Evaluation of stemming and stop word techniques on text classification problem, Int. J. Sci. Res. Comput. Sci. Eng. 3 (2015) 1–4.

[32] N. Azam, J. Yao, Comparison of term frequency and document frequency based feature selection metrics in text categorization, Expert Syst. Appl. 39 (2012) 4760–4768, http://dx.doi.org/10.1016/j.eswa.2011.09.160.
[33] P. Langley, W. Iba, K. Thompson, An analysis of bayesian classifiers, in: Proceedings of the Tenth National Conference of Artificial Intelligence, California, 1992, pp. 223-228.
[34] W. W. Cohen, Fast effective rule induction, in: Proceeding of the Twelfth International Conference on Machine Learning, California, 1995, pp. 115-123.
[35] R.C. Holte, Very simple classification rules perform well on most commonly used data sets, Mach. Learn. 11 (1993) 63–90.
[36] E. Frank, I. Witten, Generating Accurate Rule Sets Without Global Optimization, in: Proceedings the Fifteenth International Conference on Machine Learning, San Francisco, 1998.
[37] H. B. Mirza, V. R. Ratnaparkhe, Classifier tools: A comparative study, in: Proceedings of the Second International Conference on Intelligent Computing and Control Systems (ICICCS), India, 2018, pp. 1543-1547.
[38] F.S. Esmail, M.B. Senousy, M. Ragaie, Predication model for leukemia diseases based on data mining classification algorithms with best accuracy, Int. J. Comput. Electr. Autom. Control Inf. Eng. 10 (2016) 842–851, http://dx.doi.org/10.5281/zenodo.1124005.
[39] S. Ruder, An overview of gradient descent optimization algorithms, 2016, arXiv preprint arXiv:1609.04747.
[40] A. Joshuva, V. Sugumaran, Selection of a meta classifier-data model for classifying wind turbine blade fault conditions using histogram features and vibration signals: a data-mining study, Prog. Ind. Ecol. Int. J. 13 (2019) 232–251.
[41] M. Karabatak, T. Mustafa, Performance comparison of classifiers on reduced phishing website data set, in: Proceedings of the 6th International Symposium on Digital Forensic and Security (ISDFS), Turkey, 2018, pp. 1-5.
[42] W. Chen, X. Xie, J. Wang, B. Pradhan, H. Hong, D.T. Bui, Z. Duan, J. Ma, A comparative study of logistic model tree, random forest, and classification and regression tree models for spatial prediction of landslide susceptibility, Catena 151 (2017) 147–160, http://dx.doi.org/10.1016/j.catena.2016.11.032.
[43] V. Mahesh, A. Kandaswamy, C. Vimal, B. Sathish, ECG arrhythmia classification based on logistic model tree, J. Biomed. Sci. Eng. 2 (2009) http://dx.doi.org/10.4236/jbise.2009.26058.
[44] I. Furxhi, F. Murphy, M. Mullins, C.A. Poland, Machine learning prediction of nanoparticle in vitro toxicity: A comparative study of classifiers and ensemble-classifiers using the Copeland Index, Toxicol. Lett. 312 (2019) 157–166, http://dx.doi.org/10.1016/j.toxlet.2019.05.016.
[45] J. Soni, U. Ansari, D. Sharma, S. Soni, Predictive data mining for medical diagnosis: An overview of heart disease prediction, Int. J. Comput. Appl. 17 (2011) 43–48, http://dx.doi.org/10.5120/2237-2860.
[46] K. B. N. Lakmali, P. S. Haddela, Effectiveness of rule-based classifiers in Sinhala text categorization, in: Proceeding of the National Information Technology Conference (NITC), Sri Lanka, 2017, pp. 153-158.
[47] H. Ramchoun, M.A.J. Idrissi, Y. Ghanou, M. Ettaouil, Multilayer perceptron: architecture optimization and training, Int. J. Interact. Multimedia Artif. Intell. 4 (2016) 26–30, http://dx.doi.org/10.9781/ijimai.2016.415.
[48] A. Ben-David, L. Sterling, Y.H. Pao, Learning and classification of monotonic ordinal concepts, Comput. Intell. 5 (1989) 45–49, http://dx.doi.org/10.1111/j.1467-8640.1989.tb00314.x.
[49] L. Breiman, J.H. Friedman, R.A. Olshen, C. Stone, Classification and Regression Trees, Wadsworth Books, 1984.
[50] S. Kalmegh, Analysis of weka data mining algorithm reptree, simple cart and random tree for classification of indian news, Int. J. Innov. Sci. Eng. Technol. 2 (2015) 438–446.
[51] S. Shafi, S. M. Hassan, A. Arshaq, M. J. Khan, S. Shamail, Software quality prediction techniques: A comparative analysis, in: Proceedings 4th International Conference on Emerging Technologies, 2008, pp. 242-246.
[52] C. Coşkun, A. Baykal, An Application for Comparison of Data Mining Classification Algorithms, Akademik Bilişim, 2011, pp. 1–8.
[53] J.C. Platt, Sequential minimal optimization: A fast algorithm for training support vector machines, in: Scholkopf, Burges, Smola (Eds.), Advances in Kernel Method: Support Vector Learning, MIT Press, Cambridge, MA, 1998, pp. 185–208.
[54] L. Breiman, Bagging predictors, Mach. Learn. 24 (1996) 123–140.
[55] M. Kantardzic, Data Mining: Concepts, Models, Methods, and Algorithms, John Wiley & Sons, 2011.
[56] D.S.V.G.K. Kaladhar, B.K. Pottumuthu, P.V.N. Rao, V. Vadlamudi, A.K. Chaitanya, R.H. Reddy, The elements of statistical learning in colon cancer data sets: data mining, inference and prediction, Algorithms Res. 2 (2013) 8–17.
[57] M.K. Elbashir, J. Wang, Kernel logistic regression algorithm for large-scale data classification, Int. Arab J. Inf. Technol. 12 (2015) 465–472.
[58] M. Sokolova, N. Japkowicz, S. Szpakowicz, Beyond accuracy, F-score and ROC: a family of discriminant measures for performance evaluation, in: Australasian Joint Conference on Artificial Intelligence, Springer, Berlin, Heidelberg, 2006, pp. 1015–1021.
[59] C. Silverman, This analysis shows how fake election news stories outperformed real news on facebook, https://www.BuzzFeed.com/craigsilverman/viral-fake-election-news-outperformed-real-news-on-facebook?utm_term=.gl5qYwrP#.gt1ygzDN, 2016 (accessed 6 2019).
[60] B. D. Horne, S. Adali, Just in: fake news packs a lot in title, uses simpler, repetitive content in text body, more similar to satire than real news, in: Proceedings of Eleventh International AAAI Conference on Web and Social Media, USA, 2017.
[61] S. Yang, K. Shu, S. Wang, R. Gu, F. Wu, H. Liu, Unsupervised fake news detection on social media: A generative approach, in: Proceedings of 33rd AAAI Conference on Artificial Intelligence, Hawaii, USA 2019.
[62] M. Zimdars, False, Misleading, Clickbait-Y, and Satirical 'News' Sources, Google Docs, 2016.
[63] P. Engel, Here are the Most-and Least-Trusted News Outlets in America, Vol. 21, Business Insider. Business Insider. Business Insider, Inc, 2014.
[64] H. Ahmed, I. Traore, S. Saad, Detecting opinion spams and fake news using text classification, Secur. Priv. 1 (2018) http://dx.doi.org/10.1002/spy2.9.