UDC 621.396
DOI: 10.15587/1729-4061.2026.352117

DEVELOPMENT OF CLASSIFICATION ALGORITHM AND BOOSTING METHOD FOR FAKE NEWS DETECTION: FILTRATION AND ORIENTATION

The object of this study is text-based news content distributed through online media and social media platforms, presented as vector objects formed on the basis of unstructured text data and used for subsequent automated analysis. The problem solved in this study is the limited effectiveness, stability, and generalizing ability of traditional machine learning methods in detecting fake news, especially in conditions of heterogeneous datasets, noisy textual characteristics, and dynamically changing linguistic patterns, which negatively affects the quality of classification.

The article proposes a method for improving the efficiency of machine learning based on the combined use of SVM and the AdaBoost algorithm. To form an informative representation of text data, complex preprocessing and feature extraction using the TF-IDF model are used. The experimental verification of the method was performed on four open datasets: ISOT, Kaggle, News Trends and Reuters.

The results show that the proposed SVM ensemble model with AdaBoost is superior to the basic SVM classifier and a number of traditional algorithms. Accuracy increased from 0.8175 for the base model to 0.83 for SVM+AdaBoost, while memorization increased by 4.02%, average memorization accuracy increased by 2.22%, and the F1 index increased by 1.84%, while the stability of the test accuracy decreased only slightly by 0.19%. The improvement is explained by AdaBoost's adaptive enhancement of the contribution of hard-to-classify objects and a reduction in the number of errors with moderate computational complexity.

The developed approach can be effectively applied in automated systems for monitoring news content and social networks in the presence of marked-up text data and limited computing resources

Keywords: fake news, social media, infodemy, machine learning, automatic detection of misinformation

Zhanna Suimenbayeva
Master of Technology, Head
Office of Innovative Projects**
ORCID: https://orcid.org/0000-0002-2347-356X

Pramod Kumar Aylapogu
Corresponding author
Professor
Department of Electrical Communication Engineering
B.V. Raju Institute of Technology
Vishnupur, Narsapur, Taljarampet, Telangana, India, 502313
E-mail: brustembayev5@gmail.com
ORCID: https://orcid.org/0000-0003-2810-2765

Ruslan Kassym
Supervisor Project, Researcher
Department of Information and Communication Technologies**
Department of Electrical Engineering
University of Jaén
Campus Las Lagunillas s/n, Jaén, Spain, 23071
ORCID: https://orcid.org/0000-0001-8024-5224

Arai Tolegenova
PhD, Associate Professor***
ORCID: https://orcid.org/0000-0001-6318-8328

Begaidar Sarsekulov
Researcher***
ORCID: https://orcid.org/0009-0005-8366-5335

Nurlan Suimenbayev
Director
"INSAT Alatau" LLP
Makateva str., 111, Almaty, Republic of Kazakhstan, 050027
ORCID: https://orcid.org/0000-0003-1634-6738

Yerdaulet Beibit
Researcher*
ORCID: https://orcid.org/0000-0003-3360-0419

Akmaral Tlenshiyeva
Senior Lector
Department Information Communication Technologies**
ORCID: https://orcid.org/0000-0001-8105-1632

Ayaulym Kassym
Doctoral Student
School of Sciences and Humanities (SSH)
Nazarbayev University
Kabanbay Batyr ave., 53, Astana, Republic of Kazakhstan, 010000
ORCID: https://orcid.org/0009-0009-0158-7038

Mussapirova Gulzada
Candidate of Technical Sciences, Associate Professor*
ORCID: https://orcid.org/0009-0008-6569-8557
*Department of Radio Engineering and Telecommunication**
**ALT University
Chevchenko str., 97, Almaty, Republic of Kazakhstan, 050013
***Department of Information Communication Technologies
S.Seifullin Kazakh Agrotechnical Research University
Zhenis str., 62, Astana, Republic of Kazakhstan, 010011

Received 19.11.2025
Received in revised form 26.01.2026
Accepted date 03.02.2026
Published date 27.02.2026

How to Cite: *Suimenbaeva, Z., Aylapogu, P. K., Kassym, R., Tolegenova, A., Sarsekulov, B., Suimenbayev, N., Beibit, Y., Tlenshiyeva, A., Kassym, A., Gulzada, M. (2026). Development of classification algorithm and boosting method for fake news detection: filtration and orientation. Eastern-European Journal of Enterprise Technologies, 1 (2 (139)), 25–35. https://doi.org/10.15587/1729-4061.2026.352117*

1. Introduction

Rapid progress in electronic communications and the World Wide Web (WWW) has led to massive changes in the way people consume news. Social networks and online platforms for news content have become the main sources of operational information, providing unprecedented speed and scale of its dissemination. Fake news is false or misleading information disguised as a real and reliable news article. Almost all fake content is created simply to distract people and cause distrust among readers, changing their mentality [1]. In recent years, the spread of fake news has become global due to the widespread use of social media and instant access to information. Modern platforms promote content that causes emotional reactions (surprise, fear, anger) because it increases viewing time, likes, and reposts. Fake news is usually formulated more sensationally and simplistically than verified facts, which is why algorithms more often reinforce false content, regardless of its truth. In conditions of information overload, people often share headlines or short phrases, which makes fake news especially vulnerable to mass dissemination. This was especially evident during the COVID-19 pandemic, when false claims about health and safety spread faster than actual data, which led to significant social consequences. In such circumstances, there is an urgent need to develop effective automated methods for identifying and filtering fake news. As a result of the Internet outage, a 25% increase in social media activity was recorded worldwide [2]. Instagram Facebook, Twitter and messaging apps like WhatsApp have become key sources of information about what is happening in the world in real time. Twitter is a platform that distributes fake news daily to 1.5 million active users [3]. These facts indicate that the problem of fake news is not an episodic one, but a systemic one.

In the Internet age, fake news spreads faster [4]. Stories are often related to topics that are relevant on social media. Stories tend to have outrageous headlines designed to be clicked on. It is very common to see that fake news attracts more views and engagement than real news. Fake news is actively spreading on social media due to a combination of technological, psychological and social factors. The algorithms of the platforms are focused on increasing user engagement, which contributes to the promotion of emotionally colored content, regardless of its authenticity. The low threshold for information dissemination, cognitive distortions of users, the effect of echo chambers and the lack of strict control over sources create a favorable environment for the mass dissemination of disinformation. Automated accounts and crisis social contexts play an additional role, in which society's vulnerability to false information increases. Despite the development of information technology and analytical tools, the problem of the spread of fake news remains unresolved both from a scientific and practical point of view. Existing information verification mechanisms often prove to be insufficiently effective in conditions of high speed, volume and variety of content typical of modern social media. In addition, the development of automated content generation tools and coordinated disinformation campaigns reduces the effectiveness of previously developed methods for detecting false information. In this regard, the scientific community is faced with the need to develop more sustainable, adaptive and scalable approaches to identify and prevent the dissemination of false information.

False news on social networks such as Twitter and Facebook can be detected using various algorithms based on machine learning [5, 6]. Fake news has attracted a lot of attention due to the widespread use of social networks. This shows that research aimed at developing and improving methods for detecting and countering fake news in digital information systems is an urgent and practically significant scientific topic that requires further systematic study.

2. Literature review and problem statement

The paper [7] presents the results of a study of transformer models by BERT, ALBERT and XLNET for recognizing false news in social networks. It is shown that transformer architectures provide high-quality classification due to in-depth analysis of the text context. However, there are unresolved issues related to the high computational complexity of the models and the need for large amounts of data for training, which limits their practical application in resource-limited systems. An attempt to simplify the process of detecting fake news without using complex neural architectures, discussed in [8], which offers an unsupervised method for detecting spammers based on an analysis of user interests and interactions. This approach makes it possible to identify potentially undesirable activity without the need for labeled data, reducing dependence on pre-classification. However, the method is focused on the behavioral and thematic characteristics of users, which limits its direct application for semantic text analysis and identification of complex lexico-semantic patterns of fake information. To improve classification quality within the framework of classical machine learning methods, the study [9] proposed the use of combinations of various algorithms, including SVM, KNN, and logistic regression. The effectiveness of algorithms depends on the structure of a particular data set. In one case, SVM shows the best results, in the other, logistic regression. Nevertheless, the problem of model stability remains unresolved when data characteristics change and there is no universal algorithm that ensures stable quality at high speed.

The paper [10] provides an overview of modern methods for detecting fake news, which analyzes in detail the trends and problems associated with the use of machine learning algorithms and natural language processing approaches. Despite the systematization of recent achievements, the review emphasizes that many modern methods require significant computing resources and large amounts of data for training, which limits their practical application in conditions of limited resources. To eliminate the limitations of individual models, the study [11] considered ensemble machine learning methods for classifying fake news. The authors show that combining several basic algorithms makes it possible to improve the accuracy and completeness of classification compared to single models. However, an increase in the number of basic models leads to an increase in the computational load and complication of system configuration, which reduces its efficiency when processing large data streams in real time. A simpler and more accessible solution is presented in [12], where the use of standard machine learning algorithms from the Scikit-Learn library is considered. It has been demonstrated that such approaches have low computational complexity and ease of implementation, but they are inferior to ensemble and more complex models in terms of generalization quality and resistance to changes in data characteristics.

The study [13] proposed a two-stage automated pipeline based on transformer models for detecting fake news related to COVID-19 and subsequent fact checking. Despite their high accuracy, such systems require significant computing resources and complex configuration, which limits their use in scalable application solutions.

The paper [14] shows the results of combining text analysis methods and controlled machine learning algorithms within a multilevel voting model. Ensembles based on decision trees and simple classifiers provide good average values of accuracy and F-score. However, choosing the optimal ensemble configuration remains a difficult task and requires evaluating a large number of models. To increase the efficiency of feature extraction, the study [15] analyzes various vectorization methods, including TF-IDF, hashing vectorizer, and count vectorizer. It is shown that the use of TF-IDF and ensemble voting schemes makes it possible to increase classification accuracy, but an increase in the number of models again leads to an increase in computational costs. The use of a collective voting classifier (EVC) is considered in [16], where it is shown that ensemble methods can be superior to single classifiers. However, such systems require careful configuration and do not always provide sufficient completeness for detecting fake news.

Studies [17, 18] show deep neural architectures, including CNN-LSTM and models based on gradient boosting. Such approaches are capable of achieving high accuracy, including up to 86% for multi-class tasks. However, their main disadvantage remains their high computational complexity and significant resource requirements, which makes it difficult to use them in real monitoring systems.

Thus, an analysis of existing research shows that, despite the variety of approaches to detecting fake news, common unresolved problems remain related to ensuring high accuracy and completeness of classification with moderate computational complexity and the resilience of models to data changes. This makes it advisable to conduct research aimed at using ensemble methods based on classical machine learning algorithms, in particular, a combination of SVM and AdaBoost in the TF-IDF feature space, which partially eliminates these limitations without using resource-intensive neural architectures.

3. The aim and objectives of the study

The aim of this study is to develop of an automatic fake news detection system based on natural language processing and machine learning algorithms, including linear SVM and the ensemble SVM method with AdaBoost, using a set of open news datasets.

To achieve this aim, the following objectives were solved:

– to generate and prepare a single data set of news texts;

– to implement a text data preprocessing module;

– to get a numerical representation of news texts for subsequent classification;

– to develop and experimentally evaluate fake news classification models.

4. Materials and methods

4. 1. The object and hypothesis of the study

The object of this study is text-based news content distributed through online media and social media platforms, presented as vector objects formed on the basis of unstructured text data and used for subsequent automated analysis.

The subject of the study is machine learning methods and algorithms for classifying news texts, as well as ways to increase their effectiveness and resilience in detecting fake news based on textual features.

The main hypothesis of the study is the assumption that using an ensemble approach based on a combination of the support vector machine (SVM) method and the AdaBoost efficiency enhancement algorithm combined with an expanded feature space generated using TF-IDF can improve the accuracy, completeness and stability of fake news classification compared to basic classical machine learning models without significant increase in computational complexity.

The assumptions made in the study are that the textual content of news contains stable lexico-stylistic and structural patterns characteristic of fake information, which can be identified and formalized using statistical text processing methods. It is also assumed that the training and test samples are representative and reflect the main characteristics of real news streams.

The simplifications adopted in the study include the use of exclusively textual features without taking into account additional metadata, such as the source of the publication, time characteristics and social reactions of users. In addition, the analysis is limited to static datasets and does not take into account the effects of conceptual drift characteristic of dynamically changing news streams in real time.

4. 2. Structure and algorithmic implementation of the fake news classification system

In the proposed work, machine learning ensemble technique is explored for fake news classification. SVM, a supervised machine learning method, is enhanced by combining it with the AdaBoost algorithm, an unsupervised machine learning technique. SVM has been chosen due to its reported high accuracy when considered with other classical machine learning algorithms for text classification. Fig. 1 shows the general architecture of fake news detection system.

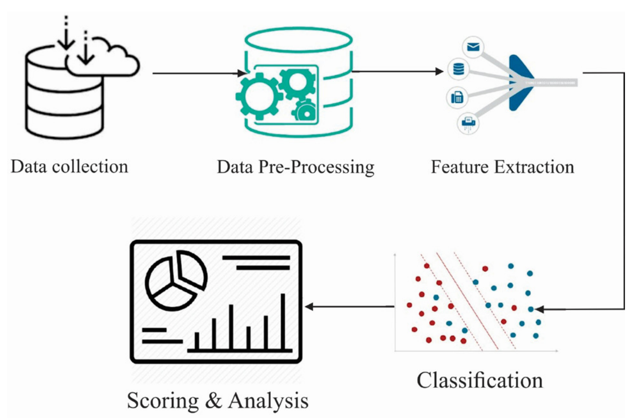

Fig. 1. General architecture of fake news detection

It consists of five basic modules: Data collection, data pre-processing, feature extraction, classification and scoring & analysis. The detailed architecture diagram for proposed system is shown in Fig. 2.

During data collection, four different datasets such as ISOT [11], Kaggle [12], News Trends [13] and Reuters [19–22] are used which contains a labelled collection of fake and real news articles. Data collection is performed on these four different datasets to create a unique dataset of 2000 records, which is to be used for training and testing the models. Pre-processing is done on the data for text normalization. During pre-processing punctuation is removed, text is converted to lowercase and tokenization is done to obtain an array of word tokens.Stop-word removal is performed on the tokens to eliminate the very common words in English which are not relevant as training features.

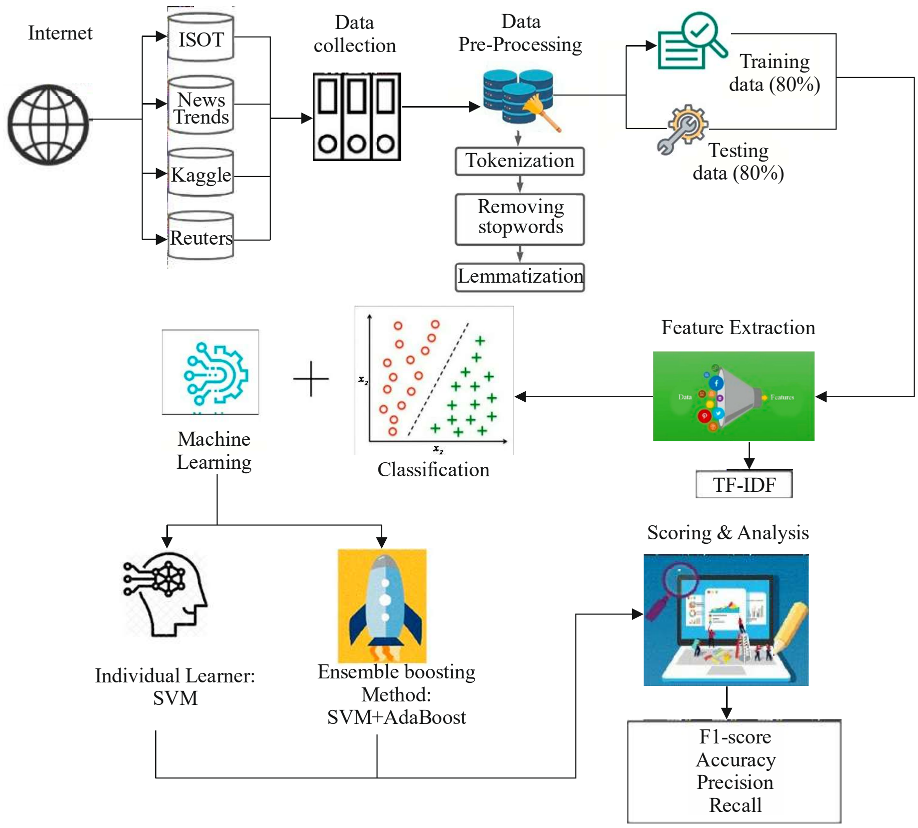

Fig. 2. Detailed architecture of proposed system

After pre-processing, the features are extracted using TF-IDF vectorization. A score is calculated for each token based on its uniqueness, frequency as well as relevance to the particular news article text and to every record on the dataset as a whole and a matrix of feature vectors is obtained. The resulting vectorized dataset acts as coordinates for the training and testing phase. The ratio of data set assigned for training and testing is 80:20. SVM algorithm is used for training. Support vector classifier is used with a "linear" kernel and the features in the vectorized training set are fitted into the model. On training, an optimal hyperplane is obtained which classify the news as "REAL" or "FAKE".

The vectorized testing set is used to assess the F1 score, accuracy, recall, precisionand other performance metrics of the algorithm. The boosting algorithm AdaBoost is used with SVM as the base learner. The number of estimators is setas 50 and the learning rate is 1.0. Similarly, the "linear" kernel is used for the support vector classifier and the model is trained by fitting the training vectors and finding the optimal hyperplane after multiple iterations of SVM occurs. In each iteration, the weight of the misclassified features is increased and that of easily classified features is decreased using AdaBoost classifier and the model is tested using the testing dataset to check the precision, accuracy, F1 score, recall and other performance metrics of the ensemble system. Scoring is done for both models to compare the performance between SVM used with boosting algorithm AdaBoost and SVM based on the metrics such as precision, accuracy, F1 score, recall and average precision-recall score. Hence an assessment of the ensemble model will reveal how it fares against individual learners and whether or not thismodel is suitable for effective detection of fake news and real news.

4. 3. Data collection

The primary sources of fake news article are from Social media network platforms like Twitter, Facebook, Instagram etc., In this work, the datasets such as ISOT, Kaggle, News Trends and Reuters which are available for downloading with columns such as title, text, author, subject and published date. The ISOT, News Trends, Kaggle and Reuters consists of 44,898, 44,676, 6,335 and 19,968 news articles respectively, labelled as either fake or real news. From each of these datasets, 500 news articles are manually collected and then concatenated to produce a unique dataset of about two thousand (2000) datasets. Table 1 shows the statistics of four collected datasets.

Table 1

Statistics of collected datasets

Datasets	Total articles	Real articles	Fake articles
ISOT	500	250	250
Kaggle	500	213	287
News trends	500	255	245
Reuters	500	268	232

An equal number of samples taken from each source ensures a source-level balance, while an almost balanced distribution of real and fake articles ensures reliable supervised learning and an objective assessment of effectiveness.

4. 4. Pre-processing

In the pre-processing phase, the unstructured data will be converted into structured data. Data pre-processing is the crucial first step while creating any machine learning model. It is the process of taking the raw data, cleaning it and making it suitable for processing by a machine learning algorithm. The steps involved in pre-processing are tokenization, removing stop words and lemmatization.

Tokenization is the process in which textual data is divided into meaningful fragments or pieces. These pieces are generally called tokens and they remove every single punctuation from the text [14]. During this process, a distinctive ID of type integer is indexed to each word and then the frequency of occurrence of each token is tallied and then normalization occurs [15]. An array of text tokens is obtained in this process.

Stop-words are language specific words that are not very important and do not carry any information when used on their own. These words are commonly used to make the sentence structure complete and join different phrases. Stop-words include conjunctions, prepositions, and pronouns used in the language. The English language has around 400–500 of such words. Some examples of stop words in the English language are *a, an, and, am, but, does, on, once, until, too, when, where, what, any,* etc. [15]. During this step of pre-processing the English stop-words are removed during this stage and the resulting tokens don't contain these less relevant words.

Lemmatization is the process by which a word is converted to its root form. The context in which the word occurs is observed and suitably it is changed to its basic form. In order to lemmatize, an instance of the WordNetLemmatizer() is created and the lemmatize() function is called on every single word token. Fig. 3 depicts how lemmatization occurs on a word.

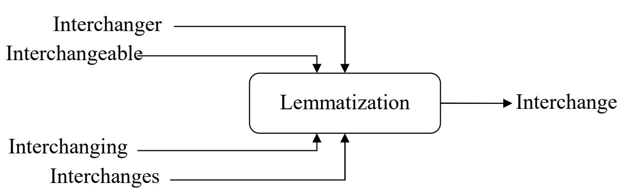

Fig. 3. Lemmatization

After pre-processing, the whole dataset is divided in the ratio of 80:20 for training and testing dataset using *train_test_split* function provided in Python.

4. 5. Feature extraction

TF-IDF is used for extraction of features from the dataset. It is a weighing matrix commonly used for measuring the importance of a word. This is basically the sum of the count and weight. It is used to assess and quantify how relevant a word is to an article in a collection of articles. This is performed by taking the product of the following two metrics: how often a word appears in a news article (*tf*), and in how many articles in the dataset the word appears (*idf*). *TF-IDF* score is represented by (1)

$$TF.IDF(w,d,D) = TF(w,d)\cdot IDF(w,D), \tag{1}$$

where w – word, d – document, D – document set.

TF is calculated by dividing the total number of words in a text or article by the total number of words in the document. In order to get to the *TF*, one must use (2)

$$TF(w,d) = \log(1 + freq(w,d)). \tag{2}$$

The logarithmic value of the dataset N divided by the number of documents containing the word w yields the *IDF*. (3) is used to compute *IDF*

$$IDF(w,d) = \log\left(\frac{N}{count(d \in D : web)}\right), \tag{3}$$

where N – the number of documents.

4. 6. Classification module

Classification of false news items is then carried out using the improved dataset that was acquired during pre-processing and feature extraction. The ML model used here is Support Vector Machine (SVM) along witha boosting algorithm, AdaBoost is used and the classification is performed by ensemble methodology.

Minimization is the goal of SVM, a method for supervised machine learning that uses this idea. Also known as discriminative classifier [15], it is a kind of algorithm. Decision boundaries between vectors belonging to a group, class, or category and vectors that do not belong to it are determined using this method A vector is a collection of numerical numbers that indicate a set of coordinates in a particular region of space.

The feature vectors obtained after TF-IDF feature extraction process are plotted in a two-dimensional space as a set of coordinates. Using support vector machines, the SVM method may then classify the input data and choose the best hyperplane for separating the two classes (fake and genuine) with the widest possible margin. Margin is the term used to describe the separation between the vectors and the hyperplane. It is the goal of SVM to increase this margin distance to the maximum possible value. The hyper-plane which has the maximum margin between the support vectors iscalled the optimal hyper-plane. The plotting of vectors and the relationship between support vectors and the optimal hyperplane are shown in Fig. 4.

During training the optimal hyperplane is found from an infinite number of options which classified the training feature vectors into "real" and "fake" categories and during testing time this hyperplane is used on the testing feature vectors to assess how the result of the SVM algorithm actually performs.

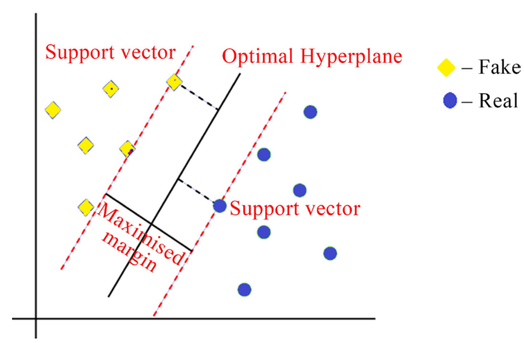

Fig. 4. Graphical representation of support vector machine

AdaBoost is a popular boosting technique that works by combining many weak classifiers or base classifiers to build a single strong classifier. It is selectedto boost the performance of any algorithm. The Adaboost classifier first assigns a higher weightage to feature vectors which are tough to deal with and a lower weightage to the features that may be more effortlessly processed. This process happens recurrently until the classifier is able to more accurately classify the training data. Adaboost can be utilized alongside any classifier to improve it andproduce a more accurate model

AdaBoost may be used to improve SVM in the following ways: First, assign equal weights to all of the findings:

Step 1. To begin, use (4) to give the same weights to each record in the collection

$$Weight(xi)=\frac{1}{N}, \quad (4)$$

where N – the number of records.

Step 2. Classify random samples using stumps.

Replace original data in random samples with the same probability as sample weights and then build a model based on these samples.

Step 3. Calculate total error.

The sum of the weights of misclassified records is the total error. Zero to one mistake is guaranteed at all times. (5) is used to arrive at the answer

$$Total\ Error = Weight\ of\ misclassified\ records. \quad (5)$$

The number 0 denotes a perfect stump, whereas the number 1 denotes a weak stump (misclassification).

Step 4. Calculate performance of stump (α).

The performance of stump is calculated using (6)

$$Performance\ of\ stump(\alpha)=\frac{\frac{1}{2}\ln[1-TE]}{TE}, \quad (6)$$

where ln – natural log, TE – total error.

Step 5. Update weights.

For misclassified samples, the weights are updated as (7)

$$New\ Weight = Weight\ Xe^{(+\alpha)}. \quad (7)$$

For correctly classified samples, the weights are updated as (8)

$$New\ Weight = Weight\ Xe^{(-\alpha)}. \quad (8)$$

Step 6. Update weights in iteration.

Step 7. Final predictions.

Final prediction is made using (9)

$$Final\ prediction/Sign(Weighted\ sum)= =\sum(\alpha_i X(Predicted\ value\ at\ each\ iteration)). \quad (9)$$

The steps are repeated and the number of iterations can be specified in the n_ estimators parameter when using the AdaBoost classifier function provided by sklearn. When SVM with AdaBoost model is run, the specified number of iterations of the base learner occurs, and each time the misclassified records i.e., the vectors which were incorrectly classified by SVM during testing are given higher weightage than ones which were correctly classified.

4. 7. Scoring and analysis

Scoring and analysis is the process of finding the efficacy of the proposed model. Fig. 5 shows the actual and predicted values. Actual values can be either true or false and they are the real and expected outcome for any input given. Predicted values can be either positive or negative and they are the observed outcomes which are the result of a machine learning model. It is very helpful for calculating recall, precision, and accuracy and also crucial for generating the confusion matrix.

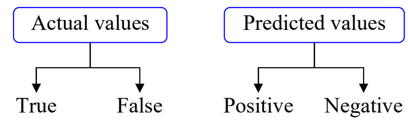

Fig. 5. Actual and predicted values

Confusion matrix is commonly used for measurement of performance ofdifferent machine learning classification models in which the output produced isin the form of two or more number of classes. It is a table which consists of fourdistinct groupings of predicted and actual values as shown in Fig. 6.

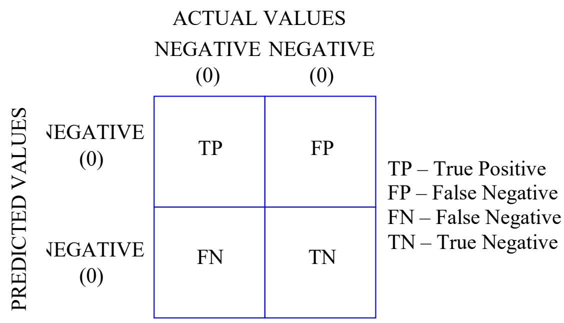

Fig. 6. Different combinations in confusion matrix

To evaluate and analyze the system, accuracy, precision, responsiveness, and F1-score were used as evaluation indicators.

The evaluation of the system took into account such indicators as accuracy, error-free, responsiveness and F1, which are part of the verification process.

The accuracy (A) of a model's predictions or its categorization of test data may be determined by calculating the accuracy (A) as a percentage. Calculated by summing the number of correct forecasts to the number of correct ones, and then dividing the result by the total.

Precision (P) is used to measure the model's ability to accurately forecast future outcomes. To calculate it, it is necessary to divide the total of the true positive forecasts by the ratio of the true positive predictions.

When recall (R) is considered, it refers to the sensitivity of the model to relevant instances.

F1-score indicates about the accuracy of the model for positive and negativeclasses. It is used when the data set is not balanced. The score conveys the extensiveness of the proposed model. F1-score is given using harmonic mean of precision and recall which provides a balance between the two. F1-score is calculated using (10)

$$F1=2\cdot\frac{precision.recall}{precision+recall}. \quad (10)$$

These estimates generally characterize different aspects of the model's operation. Accuracy reflects the overall correctness of forecasts, while accuracy and responsiveness assess the reliability and completeness of fake news detection, respectively. Given the types of data typical of news data and the class imbalance that is often present in the real world, the F1 score serves as an important indicator of the effectiveness of balanced classification. Therefore, the selected indicators provide a reliable basis for comparing models and confirming the effectiveness of the proposed approach.

5. The results of a study of an automatic fake news detection system based on machine learning methods

5. 1. Data collection module

The fake news detection datasets, News Trends, Kaggle, Reuters and ISOT are modified to contain 500 of the original records each to keep the size manageable. Fig. 7 shows the first 5 rows of the dataset obtained after collection

	title	text	label
0	Special election to replace Conyers to be held…	WASHINGTON (Reuters) - Michigan Governor Rick …	REAL
1	Videos on the Pacific Crest Trail Association …	Posted on October 30, 2016 by Graywolf Publish…	REAL
2	Hillary Clinton Tops "Islamist Money in Politi…	Hillary Clinton Tops "Islamist Money in Politi…	FAKE
3	UKIP MEPs Steven Woolfe & Mike Hookem reported…	UKIP MEPs Steven Woolfe & Mike Hookem reported…	FAKE
4	House tax panel chair to urge longer-lasting i…	WASHINGTON (Reuters) - The head of the U.S. Ho…	REAL

Fig. 7. Sample result obtained from data collection module

The CSV input files are loaded using Python package Pandas. The individual datasets are then concatenated into one unique dataset containing 2000 records total. Each news article is labelled as "REAL" or "FAKE". The data is shuffled to reduce bias and unused columns are dropped. There are real and fake news articles.

5. 2. Data preprocessing module

Textual data needs to be refined and encoded to numerical values before feeding them into any machine learning model. The punctuation is removed and the text is converted to lowercase form. As can be seen in Fig. 8, a sample dataset has been processed to remove any unnecessary punctuation, tokenize the text, remove any stop words, lemmatize the tokens, create clean text, and then have those labels converted to the goal values of 0 and 1.

This is followed by tokenization in which the text sentences are converted into an array of word tokens. Lemmatization is performed on the tokens to get root words. Stop-word elimination is then carried out. The label "FAKE" has a goal value of "0," whereas the label "REAL" has a target value of "1. The dataset should then be divided into training and testing sets. It is 80% for training and 20% for testing the algorithm.

5. 3. A module for obtaining a numerical representation of news texts

An instance of TfidfVectorizer is first initialized. The TfidfVectorizer is used to turn a collection of pre-processed text into a set of TF-IDF features. Then fit and transform the vectorizer on the training dataset to obtain a matrix of feature vectors. The importance of each word in the text is calculated. Fig. 9 shows the result of feature extraction on the dataset in the form of a matrix of feature vectors. The number of features is 40195 and the values of each feature act as vector coordinates during classification using SVM.

```
[[0.         0.         0.        ... 0.         0.         0.        ]
 [0.         0.         0.1164288 ... 0.         0.         0.        ]
 [0.         0.         0.        ... 0.         0.         0.        ]
 ...
 [0.         0.         0.        ... 0.         0.         0.        ]
 [0.         0.         0.        ... 0.         0.         0.        ]
 [0.         0.         0.        ... 0.         0.         0.        ]]
(1600, 40195)
[[0. 0. 0. ... 0. 0. 0.]
 [0. 0. 0. ... 0. 0. 0.]
 [0. 0. 0. ... 0. 0. 0.]
 ...
 [0. 0. 0. ... 0. 0. 0.]
 [0. 0. 0. ... 0. 0. 0.]
 [0. 0. 0. ... 0. 0. 0.]]
(400, 40195)
```

Fig. 9. Result obtained from feature extraction module

Despite the fact that the resulting feature space is large, it is well suited for linear classifiers that can efficiently process sparse and multidimensional data. Thus, the extracted features provide an informative and reliable representation for the subsequent classification stage.

5. 4. Classification module

Classification is done by initializing the SVC classifier with kernel set as "linear" and C = 1.0. Then fit this on training set (training feature vectors) and y_train.

	title	text	label	removed_punc	tokens	stop_tokens	lem_words	clean_text	target
0	Special election to replace Conyers to be held…	WASHINGTON (Reuters) - Michigan Governor Rick …	REAL	WASHINGTON Reuters Michigan Governor Rick Sny…	[washington, reuters, michigan,	[washington, reuters, michigan, governor, rick…	[washington, reuters, michigan, governor, rick…	washington reuters michigan governor rick snyd…	0
1	Videos on the Pacific Crest Trail Association …	Posted on October 30, 2016 by Graywolf Publish…	REAL	Posted on October 30 2016 by Graywolf Publishe…	[posted, on, october, 30, 2016, by, graywolf, …	[posted, october, 30, 2016, graywolf, publishe…	[posted, october, 30, 2016, graywolf, publishe…	posted october 30 2016 graywolf published oct …	0
2	Hillary Clinton Tops "Islamist Money in Politi…	Hillary Clinton Tops "Islamist Money in Politi…	FAKE	Hillary Clinton Tops Islamist Money in Politic…	[hillary, clinton, tops, islamist, money, in, …	[hillary, clinton, tops, islamist, money, poli…	[hillary, clinton, top, islamist, money, polit…	hillary clinton top islamist money politics li…	1
3	UKIP MEPs Steven Woolfe & Mike Hookem reported…	UKIP MEPs Steven Woolfe & Mike Hookem reported…	FAKE	UKIP MEPs Steven Woolfe Mike Hookem reported …	[ukip, meps, steven, woolfe, mike, hookem, rep…	[ukip, meps, steven, woolfe, mike, hookem, rep…	[ukip, meps, steven, woolfe, mike, hookem, rep…	ukip meps steven woolfe mike hookem reported f…	1
4	House tax panel chair to urge longer-lasting i…	WASHINGTON (Reuters) - The head of the U.S. Ho…	REAL	WASHINGTON Reuters The head of the US House o…	[washington, reuters, the, head, of, the, us, …	[washington, reuters, head, us, house, represe…	[washington, reuters, head, u, house, represen…	washington reuters head u house representative…	0

Fig. 8. Result obtained from data pre-processing module

Then, the test set is provided, and the accuracy is computed using the function accuracy_score(). The obtained results are accuracy comes with 0.817500, the precision is 0.816901, the recall score is 0.836538, F1 score is 0.826603, and the average precision-recall score is 0.90. The confusion matrix shows 153 true positive results, 174 true negative results, 39 false negative results and 34 false positive results.

For example, the y-axis of a PR curve maps accuracy to recall, whereas the x-axis does the same thing. The curve shows the trade-off involving the precision and recall values for a certain threshold value. If the area under the curve is high then the algorithm is said to have great recall and precision values,in which case the high precision represents having a small false positive rate, and high recall represents a small false negative rate. This indicates that the classifier model is generally providing correct outcomes (high precision), and also giving mostly positive outcomes (high recall). Fig. 10 shows the PR curve for the classifier SVM, which shows that area under the curve is high. This indicates a good recall and precision score.

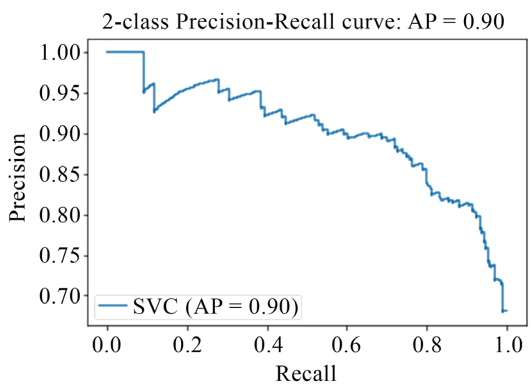

Fig. 10. Precision-recall curve using support vector machine

AdaBoost is initialized using AdaBoost classifier with base learner as SVC with kernel set to “linear”, the n_ estimators = 50 and learning rate as 1.0. This classifier model is fit on training feature vectors and y_ train. Then the model is evaluated on the test set and the accuracy is calculated. Also evaluate the precision, recall values and produce a graph for analysis.

Fig. 11 shows the result of classification using the ensemble boosting model, where AdaBoost is utilized with SVM as base learner. The result obtained using SVM with AdaBoost are accuracy comes out as 0.830000, the precision is 0.815315, the recall score is 0.870192, F1 score is 0.836186, and the average precision-recall score is 0.92. The confusion matrix shows 151 true positive results, 181 true negative results, 41 false negative results and 27 false positive results. Fig. 11 shows the precision-recall curve for the classification algorithm SVM with AdaBoost, which shows a high area under the curve.

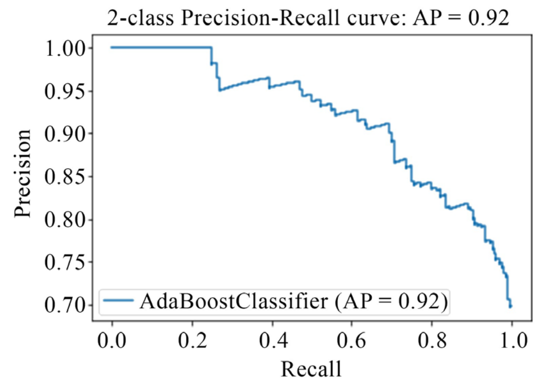

Fig. 11. Precision-recall curve using support vector machine with AdaBoost

The results obtained from the two models are summarized in Table 2 which clearly shows that when the boosting technique, AdaBoost is used with SVM algorithm an improvement is observed in the model’s accuracy, recall, average precision-recall and F1 score.

Fig. 12 shows the graphical representation of the analysis of performance metrics using SVM and SVM+AdaBoost. It clearly shows that the accuracy increases by 1.52%, precision decreases by 0.19%, recall increases by 4.02%, average precision-recall score increases by 2.22% and F1 score increases by 1.84%.

Table 2

Analysis of performance metrics using SVM and SVM+AdaBoost

PERFORMANCE METRICS	SVM	SVM +ADABOOST
Accuracy	0.817500	0.830000
Precision	0.816901	0.815315
Recall	0.836538	0.870192
Average precision-recall	0.90	0.92
F1-score	0.826603	0.841860

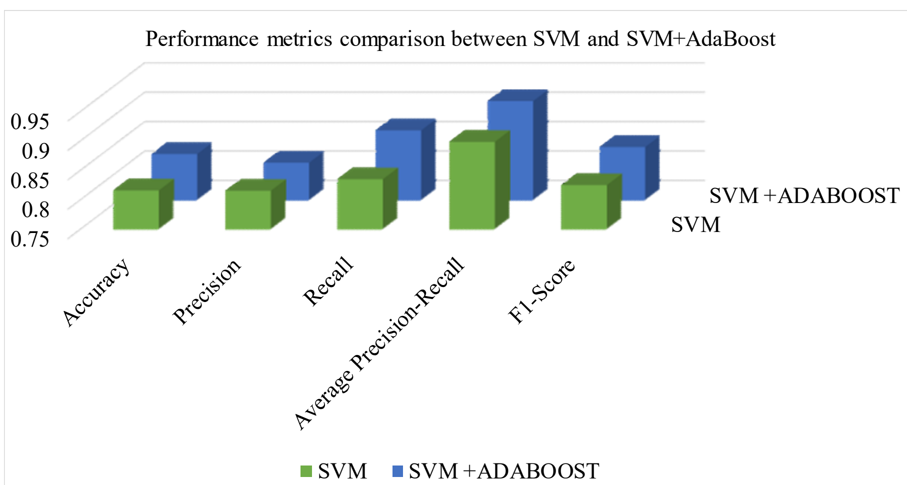

Fig. 12. Graphical representation of comparison of performance metrics using support vector machine and support vector machine + AdaBoost

A slight decrease in accuracy indicates a slight increase in the number of false positives, which is a compromise to achieve higher sensitivity to actual positive cases. Overall, the large area under the accuracy curves confirms that both models effectively distinguish real news from fake news, while the improved model demonstrates superior reliability.

6. Discussion of the results of fake news classification using SVM and the AdaBoost ensemble method

The results of the fake news classification obtained in the work are explained by the cumulative influence of the stages of text preprocessing, feature extraction, and the choice of machine learning algorithms. As shown in Fig. 8, 9, the use of a standard natural language processing pipeline – including text purification, tokenization, lemmatization, and removal of stop words – allowed to form a homogeneous and informative textual representation of data. Application of TF-IDF vectorization (Fig. 9) ensured that both the frequency characteristics of words and their significance in the corpus of documents are taken into account, which is especially important when analyzing news texts with characteristic lexical markers of fake information. The high dimension of the feature space (40,195 features) contributes to better class separability, which is confirmed by the accuracy and completeness indicators in Table 2. In contrast to paper [12], where the use of simple Scikit-Learn tools provided limited classification accuracy due to fewer features and simplified text processing, the presented approach allows for a more stable class separation due to complex preprocessing and expanded feature space. Thus, data preparation and feature extraction create a reliable basis for the correct operation of classifiers.

The experiments in this study were conducted using datasets in English, the proposed approach is inherently language-independent and can be effectively applied to detect fake news in Slavic languages. The TF-IDF representation is based on statistical weighting of terms rather than language-specific grammatical rules, which makes it suitable for texts written in Slavic languages based on Cyrillic and Latin. In addition, linear SVM classifiers are well known for their reliability in a multidimensional and sparse feature space, which is especially important for morphologically rich languages such as Russian, Ukrainian, Polish, and Czech. By adapting the preprocessing stage to include language-specific vocabulary, stop word lists, and lemmatization, the proposed structure can be directly applied to the analysis of fake news in the media in Slavic languages without changing the basic classification model.

From an interlanguage point of view, previous studies have shown that fake news in different languages has common stylistic and semantic features, such as emotionally saturated vocabulary, simplified narrative structures, and repetitive lexical cues, although their specific lexical implementations may vary. Therefore, although the pre-processing stage (for example, tokenization, removal of stop words and lemmatization) must be adapted to each language, the basic TF-IDF–based function representation and the SVM-AdaBoost classification system can be effectively applied to detect fake news in the media in Slavic languages.

The classification results using linear SVM demonstrate balanced values of accuracy and completeness (Table 2), which indicates the stable ability of the model to correctly recognize both fake and real news. The high area under the precision–recall curve (Fig. 10) confirms that the model retains high classification quality when the decision threshold is changed. Improved performance when using the SVM + AdaBoost ensemble approach (Fig. 11, Table 2) is explained by the fact that the boosting method consistently enhances the influence of difficult-to-classify objects, reducing the proportion of false negative results and increasing the completeness of fake news recognition. This is especially noticeable in the increase in recall by 4.02% and F1-score by 1.84% (Fig. 12). Unlike [11, 15], where ensemble methods increased accuracy by increasing the number of basic models, which increased computational costs, this approach provides comparable or better performance with a moderate computing load. The proposed combination of SVM + AdaBoost makes it possible to achieve high accuracy (0.83) and F1-measure (0.8419) without the need for large amounts of data and complex network configuration. The observed improvement in accuracy and F1 performance can be explained by a combination of TF-IDF feature representation, SVM linear classification, and the AdaBoost ensemble. TF-IDF captures the importance of keywords throughout the corpus, allowing the model to recognize subtle patterns in fake news texts. SVM efficiently processes a multidimensional space of sparse objects, while AdaBoost focuses on misclassified instances, increasing the classifier's attention to complex examples. Together, these mechanisms reduce the number of false positives and false negatives, resulting in improved overall performance. This was made possible by integrating a carefully prepared feature space and focusing the boosting mechanism on difficult-to-classify objects, which reduces the proportion of false positive and false negative classifications.

The proposed TF-IDF + SVM + AdaBoost platform not only surpasses classical SVM models in test datasets, but also demonstrates resilience to new forms of disinformation. The method identifies the characteristic lexical and structural patterns characteristic of fake news, which remain relevant even in modern news, supplemented by videos created using artificial intelligence. This highlights the potential applicability of the approach to modern multimodal fake news streams, while maintaining interpretability and low computational costs.

In comparison with existing research in the field of fake news detection, where naive Bayes classifier, logistic regression and basic SVM models are widely used, the proposed SVM + AdaBoost approach demonstrates comparable or higher accuracy (up to 83%) with lower computational costs compared to deep neural networks (CNN, LSTM). High efficiency is ensured by careful preprocessing of the text, extended feature space (TF-IDF) and an amplification mechanism on difficult-to-classify objects, which reduces the number of false positive and false negative classifications.

The practical significance of the study is in the creation of an automatic fake news detection system capable of detecting false texts with high accuracy with limited computing resources. The method can be applied to various languages, including Slavic, and adapted to monitor streaming news in real time. It is able to identify common patterns of fake news, including modern forms of disinformation with AI-generated videos. The use of the system helps to improve the quality of the information space and reduce the spread of misinformation among users.

The main limitations of the work are related to the limited amount of data set (2000 articles) and the use of only English-language sources, which limits the portability of the model to other languages and dynamic news streams. Despite the use of well-known open datasets (ISOT, Kaggle, News Trends, Reuters), new fake news appears daily, and their lexical and stylistic characteristics may change, which creates a problem of concept drift. Also, the current model relies solely on textual attributes and does not take into account metadata (source, time, social reactions), which may reduce accuracy in real conditions.

In the future, it is necessary to expand the data corpus, including multilingual sources (including Slavic languages), the integration of semantic representations of text (word/sentence embeddings) and the development of adaptive model update methods for working with streaming news data in real time. These measures will increase the system's resilience to the rapidly changing characteristics of fake news and improve its practical applicability.

The proposed system shows promising results for a model that uses a common approach to detecting fake news. In the future, its performance can be improved by experimenting with different core learners in an ensemble, applying evolutionary weight optimization techniques, integrating additional feature extraction techniques, and adapting the system to real-time news streams to detect misinformation early.

7. Conclusions

1. The task of forming a single and manageable data set from several sources (News Trends, Kaggle, Reuters, ISOT) has been successfully solved. Each dataset was reduced to 500 records and then combined into a single dataset containing 2,000 news articles. The first 5 lines of the collected dataset illustrate the successful integration and labeling of "REAL" and "FAKE" news. The resulting data set provides a balanced representation of classes, ensuring the statistical reliability of subsequent analysis.

2. The separation of training and testing (80/20) allowed for effective training and assessment. The quantitative results of this stage indirectly affect the final accuracy. This reduction in noise and the elimination of lexical redundancy made it possible to increase the informative value of the features, which indirectly affected the final classification indicators: the accuracy of the basic SVM model was 0.8175, while the use of the SVM + AdaBoost ensemble increased this indicator to 0.83.

3. TF-IDF vectorization transformed the processed text into a multidimensional feature space consisting of 40,195 features, reflecting both the frequency of occurrence of words and their significance. At this stage, a quantitative framework for effective classification was created. Feature extraction provided a detailed representation of lexical markers associated with fake news, which led to an increase in the performance of the proposed system.

4. The classification problem was solved using linear SVM and the SVM + AdaBoost ensemble method. The SVM classifier alone achieved an accuracy of 0.8175, a return of 0.8365, an F1 score of 0.8266 and an average return accuracy of 0.90, while SVM + AdaBoost improved these indicators to an accuracy of 0.83, 0.8702 reviews, 0.8419 F1 points and 0.92 average return accuracy, among others. The large area under the accuracy curves indicates a significant difference between real and fake news. Compared to the classical SVM models described in the literature, which showed an accuracy of 78–82%, the integrated approach demonstrates excellent performance. This improvement is attributed to an efficiency improvement mechanism that focuses on difficult-to-classify cases, reducing the number of false positives and increasing the number of reviews.

Conflict of interest

The authors declare that they have no conflict of interest in relation to this study, whether financial, personal, authorship or otherwise, that could affect the study and its results presented in this paper.

Financing

This research has been funded by the Science Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan Grant № AP 22686357 "Research and construction of an information system for processing nanosatellite orientation data based on filtration methods".

Data availability

Manuscript has data included as electronic supplementary material.

Use of artificial intelligence

The authors confirm that they did not use artificial intelligence technologies when creating the current work.

Acknowledgment

We express our acknowledgement to the Science Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan (grant. No. AP19679190 "Research and optimization of technology of intelligent reflective surfaces using artificial intelligence").

In appreciation for allowing them to utilize the R-Language and Python tool at Malla Reddy Engineering College in Hyderabad, the authors wish to acknowledge the CSE and ECE group at this college.

Authors' contributions

Zhanna Suimenbaeva: Validation, Investigation, Project administration; **Pramod Kumar Aylapogu**: Conceptualization, Methodology, Software, Formal analysis, Writing – original draft, Writing – review & editing, Visualization; **Ruslan Kassym:** Data Curation, Resources, Supervision; **Arai Tolegenova**: Validation, Investigation, Writing – review & editing; **Begaidar Sarsekulov**: Methodology, Resources, Data Curation; **Nurlan Suimenbayev**: Formal analysis, Investigation, Supervision; **Yerdaulet Beibit**: Software, Validation, Visualization; **Akmaral Tlenshiyeva**: Conceptualization, Data Curation, Project administration; **Ayaulym Kassym**: Investigation, Resources, Writing – review & editing; **Mussapirova Gulzada**: Methodology, Formal analysis, Project administration.

References

1. Zhou, X., Zafarani, R. (2020). A Survey of Fake News. ACM Computing Surveys, 53 (5), 1–40. https://doi.org/10.1145/3395046
2. Yang, K.-C., Torres-Lugo, C., Menczer, F. (2020). Prevalence of low-credibility information on twitter during the covid-19 outbreak. Workshop: Workshop on Cyber Social Threats (CySoc 2020). https://doi.org/10.36190/2020.16
3. Jamieson, K. H., Cappella, J. N. (2008). Echo chamber: Rush Limbaugh and the conservative media establishment. Oxford University Press.
4. Holan, A. D. (2016). 2016 Lie of the Year: Fake news. Available at: https://www.politifact.com/article/2016/dec/13/2016-lie-year-fake-news/
5. Vosoughi, S., Roy, D., Aral, S. (2018). The spread of true and false news online. Science, 359 (6380), 1146–1151. https://doi.org/10.1126/science.aap9559
6. Allcott, H., Gentzkow, M. (2017). Social Media and Fake News in the 2016 Election. Journal of Economic Perspectives, 31 (2), 211–236. https://doi.org/10.1257/jep.31.2.211
7. Gundapu, S., Mamidi, R. (2021). Transformer based Automatic COVID-19 Fake News Detection System. arXiv. https://doi.org/10.48550/arXiv.2101.00180
8. Koggalahewa, D., Xu, Y., Foo, E. (2022). An unsupervised method for social network spammer detection based on user information interests. Journal of Big Data, 9 (1). https://doi.org/10.1186/s40537-021-00552-5
9. Kesarwani, A., Chauhan, S. S., Nair, A. R., Verma, G. (2020). Supervised Machine Learning Algorithms for Fake News Detection. Advances in Communication and Computational Technology, 767–778. https://doi.org/10.1007/978-981-15-5341-7_58
10. Thakar, H., Bhatt, B. (2024). Fake news detection: recent trends and challenges. Social Network Analysis and Mining, 14 (1). https://doi.org/10.1007/s13278-024-01344-4
11. Ahmad, I., Yousaf, M., Yousaf, S., Ahmad, M. O. (2020). Fake News Detection Using Machine Learning Ensemble Methods. Complexity, 2020, 1–11. https://doi.org/10.1155/2020/8885861
12. Gao, Q., Lantz, N. (2022). Using Machine Learning Algorithms to Detect Fake News. Journal of Student Research, 11 (4). https://doi.org/10.47611/jsrhs.v11i4.3446
13. Vijjali, R, Potluri, P., Kumar, S., Teki, S (2020). Two Stage Transformer Model for COVID-19 Fake News Detection and Fact Checking. arXiv. https://doi.org/10.48550/arXiv.2011.13253
14. Kaur, S., Kumar, P., Kumaraguru, P. (2019). Automating fake news detection system using multi-level voting model. Soft Computing, 24 (12), 9049–9069. https://doi.org/10.1007/s00500-019-04436-y
15. Mahabub, A. (2020). A robust technique of fake news detection using Ensemble Voting Classifier and comparison with other classifiers. SN Applied Sciences, 2 (4). https://doi.org/10.1007/s42452-020-2326-y
16. Umer, M., Imtiaz, Z., Ullah, S., Mehmood, A., Choi, G. S., On, B.-W. (2020). Fake News Stance Detection Using Deep Learning Architecture (CNN-LSTM). IEEE Access, 8, 156695–156706. https://doi.org/10.1109/access.2020.3019735
17. Kaliyar, R. K., Goswami, A., Narang, P., Sinha, S. (2020). FNDNet – A deep convolutional neural network for fake news detection. Cognitive Systems Research, 61, 32–44. https://doi.org/10.1016/j.cogsys.2019.12.005
18. Fake News. University of Victoria. Available at: https://libguides.uvic.ca/fakenews
19. Ozbay, F. A., Alatas, B. (2020). Fake news detection within online social media using supervised artificial intelligence algorithms. Physica A: Statistical Mechanics and Its Applications, 540, 123174. https://doi.org/10.1016/j.physa.2019.123174
20. Sultan, A., Yermoldina, G., Kassym, R., Serikov, T., Bekbosynov, S., Yernazarov, N. et al. (2024). Research and construction of an adaptive drive with increased efficiency based on a balancing friction clutch. Vibroengineering Procedia, 54, 334–340. https://doi.org/10.21595/vp.2024.23971
21. Nagima, B., Kumar, P., Suman, N., Serikov, T., Temirbekova, M., Sergazin, G. et al. (2024). Development of an innovative patch antenna design and implementation using ENG materials for health care systems and 5G networks. Eastern-European Journal of Enterprise Technologies, 4 (5 (130)), 26–33. https://doi.org/10.15587/1729-4061.2024.309080
22. Yelikbai, K., Kumar, P., Serikov, T., Temirbekova, M., Mussapirova, G., Arai, T. et al. (2025). Improving inteligent reflective surface efficiency using ENG-metamaterials and L-shaped patch antennas in a dynamic radio environment. EUREKA: Physics and Engineering, 4, 99–109. https://doi.org/10.21303/2461-4262.2025.003914