KAMPALA INTERNATIONAL UNIVERSITY
KIU
Journal of Science, Engineering and Technology
KJSET
Research Article

Machine Learning Base Algorithms Using Truncated Singular Value Decomposition as a Novel Solution for Fake News Analysis and Detection

Hashim Ibrahim Bisallah[1] and Umoru Yahaya Ibrahim[2]

[1]Department of Computer Science, Kampala International University, Uganda
[2]Department of Computer Science, University of Abuja, Nigeria, umoru71@yahoo.ca

Corresponding Author: hashim.bisallah@kiu.ac.ug

Paper history:
Received 12 March 2024
Accepted in revised form
30 April 2024

Keywords
Fake news detection, machine learning, ISOT Dataset, TSVD, social media.

Abstract

Fake news analysis and detection is a method that requires using machine learning classifiers to recognize and detect news content as fake or real. No research work has been carried out on fake news analysis and detection using the machine learning algorithms using Truncated singular value decomposition (TSVD) on the 'ISOT Fake News Dataset' but there is a need for a better design that can provide better accuracy. This research aims to develop and use systematic progression in identifying fake news analysis using machine learning algorithms with TSVD. The machine learning algorithms which include the light Gradient Boosting Machine, Cat boosting classifier, and support vector machine with TSVD were used on 'The ISOT Fake News Dataset' to recognize the fake news from real news in the dataset, thereby improving the performance of these classifiers by cleaning the data to manage the data imbalances more efficiently. The recommended design yields an exactness of 97.33%. The postulated technique out classifies the current classifiers which possess an exactness of 95.05%. We recommend that Machine learning classifiers using TSVD should be applied in Fake News analysis and detection since it gives a better performance accuracy in fake news analysis and detection.

1.0 Introduction

Any inaccurate information or misrepresentation of reality that circulates online through traditional news outlets or social media platforms such as WhatsApp, YouTube, Twitter, and the like is considered fake news. It can be challenging to tell fake news from actual news when they occasionally follow the same pattern. Since Facebook is one of the many ways that fake news may be distributed to a wide audience on social media platforms, it is also inclusive. These platforms offer a lot of benefits, one of which is that they play a big part in making people's lives better. The fact that fake news spreads similarly to how wildfire does in the woods is among its worst drawbacks. One of the reasons bogus material is being shared could be to harm someone's reputation or to gain financial or political advantage over rivals (Alonso et al., 2021). Sentiment analysis is the practice of examining social media comments to determine whether the sentimental tone of the language is genuine or fabricated and is used concerning fake news (Alam et al., 2018; Kesarwani et al., 2021).

Previously, several computer experts have taken advantage of the outcomes that arise from our routine tasks. Professionals have developed a variety of proportionate approaches and strategies to help us solve our day-to-day issues while considering improved performance. Numerous researchers and system designers have worked with deep learning networks, neural networks, and other technologies to provide several explanations. One should confirm the news's source before spreading any information, such as whether it is coming from a different news outlet or an internet source. The reason for this is that news organizations, publications, and social media platforms periodically spread a significant amount of false information to their audience, and when this happens, they realize that they have spread false information.

Spreading false information to make people laugh is a criminal offence. The news of the coronavirus's proliferation serves as one example. People started disseminating false information when the coronavirus started to spread over the world, claiming that scientists had discovered a way to make sure that during the hot season, the entire planet would be free of the coronavirus. Rather, we are witnessing the exact opposite; it has gotten worse than what we saw in the harmattan. The general public shouldn't be exposed to this kind of news because when it turns out to be false, it depresses and discourages individuals. People who are exposed to false information are more likely to have psychiatric illness or despair as a result.

The Internet has increased people's self-assurance in terms of how they get information, make decisions, and adhere to culturally significant protocols (Hamborg et al., 2018). According to research by the Journalism Project at the Pew Research Centre, roughly 53% of American adults will be adults in 2020. According to Jang et al. (2019), social media users "regularly" and "here and there" consume news from these platforms. Of these users, 59% and 54% of Twitter and Facebook clients, respectively, regularly engross themselves in information on these platforms.

The amount of bogus news that is being spread is increasing at an alarming rate due to technological advancements. Due to the rise of enormous tech businesses like Facebook, YouTube, Twitter, and so on, this growth rate can be linked to earlier years. Fake news was strikingly visible in the 2016 United States general election. Large-scale, unverified dissemination of false or misleading information has a profound effect on the reputations of public figures, heads of state, mission chiefs, and professional envoys in a variety of sectors, including but not limited to science, health care, sports, and agriculture (Lazer et al., 2018). The stock exchange sector is another area that is significantly impacted, as bogus news can

One of the main causes of the rapid global dissemination of false information is reliance on information obtained from online media or any news podium. Research has demonstrated that the news that we take to be very important is often later proved to be false (Lai et al., 2020; Wang et al., 2020). The dissemination of false and inaccurate information around the world during the coronavirus outbreak is one such instance (Hua et al., 2020)

According to Ali et al. (2019), Anam et al. (2021), Gupta et al. (2021), Jwa et al. (2019), and Nagi et al. (2021), machine learning algorithms have now improved accuracy across all domains. Some machine learning techniques that are useful in identifying bogus content that is being mislabeled as real or fake news include the gradient boosting classifier technique, the cat boosting classifier approach, and the additional tree classifier methodology.

Advances in technology have made data the engine of the twenty-first century. The relevance of social media platforms as a source of news has increased recently. Additionally, 63%

of news is read in print media, 18% is shared on social media, and the precise figures won't be known until April 2020. This information may be seen by looking back at the volume of information shared on social media in 2013. As a result, contributions to press articles drop by 26%, but the online community expands by 39%. Fake news is becoming a serious problem as a result of social media platforms' ongoing user growth. Fake news is described as a series of real or imagined causal events that are used to mislead people into believing something is factual or to harm someone's reputation (Alonso et al., 2021). The popularity of fake news on social media platforms has been proven by user involvement on "Facebook," where 20 genuine news stories are followed more often than 20 fraudulent ones. It's thought that the widespread distribution of these phony products on social media sites is due to functions like leaving comments, sharing, and tagging our loved ones.

Right now, both machine learning classifiers' tools—which also incorporate natural language processing—help identify fake content. In order to determine if a certain news item is accurate or inaccurate, classifiers are created. This study centres on the necessity of gathering vast amounts of data in order to construct an effective classifier. Various strategies have been implemented to mitigate this issue, and one way forward is to distinguish them and stop their spread. The concept of machine learning techniques in handling fake news articles was introduced in a number of earlier research, including M. Senthil Raja et al. (2022). One technique to determine if news stories are real or fraudulent is the k-Nearest Neighbors (KNN) classifier. However, due to the nature of the text data available on the Internet, this classifier was unable to produce a more accurate conclusion. The machine learning algorithms for fake news identification are presented in this research. This research uses machine learning since it enables the machine to learn without requiring special programming. It is also easy to visualize and has the potential to improve data noise management. However, spreading false information in online forums or WhatsApp groups is a common strategy to attract readers, especially during a difficult time like this one. However, who gains from bogus news and why? There is still much work to be done in order to identify the cause of fake news dissemination. Politics-related disputes are typically what escalates a non-complex situation into a violent war. This discusses the impact of computer-era harm and validates and rejects fake news. In addition to COVID-19, the amount of fake news has increased significantly, creating circumstances that negatively impact people's mental health everywhere in the world. As a result, in addition to the infection, we are dealing with a serious lie that stems from fundamental ignorance and misinformation.

1.1 Research questions

The following exploration questions used in this research:

i. How can machine learning models using TSVD for the analysis and detection of false news be made more resilient and accurate so that they can be used on a variety of datasets?

ii. What strategies may be used to improve machine learning models using TSVD interpretability and explainable for the purpose of detecting and analyzing fake news while maintaining system transparency and user confidence?

iii. How much does the addition of human-in-the-loop methods, like domain experts, enhance the dependability and accuracy of machine learning models using TSVD when it comes to detecting false information?

1.2 Aim and Objectives

This research work aims to develop automated tools that can accurately and efficiently identify and categorize news articles and other sources of information as either real or fake by using machine learning algorithms with TSVD.

The major objectives are:

i. To create a dataset of labeled articles that includes both fake and authentic news articles, which can be used to train the machine learning algorithms with TSVD.

ii. To design and develop an intelligent algorithm that can efficiently categorize fake and authentic news and articles using machine learning with TSVD.

iii. To test and validate the algorithms in (ii) using online and real-time datasets.

iv. To evaluate the algorithms in (ii) with the existing ones and measure their accuracy and performance differences.

iv. To devise and implement procedures and techniques that integrate algorithms in (ii) with real-world applications.

2.0 Related Work and Review of Literature

Numerous studies have been conducted on the analysis and identification of fake news using machine learning algorithms; each study institutes or suggests a particular methodology for doing so. Using a variety of machine learning classifiers, such as the Nave Bayes classifier, the k-Nearest Neighbors (KNN) algorithm, the decision tree, the random forest, the support vector machine (SVM), the deep learning network, and others, this work highlights relevant and existing materials in the field of fake news analysis and detection.

2.1 Truncated Singular Value Decomposition

Truncation singular value decomposition (TSVD) frequently yields a reduced-dimensional space that is more informative and concise than the original article's content, which facilitates machine learning models' ability to learn and distinguish between fake and real news.

TSVD is not just about word frequency. It reveals latent thematic patterns in the data, exposing how fake news stories differ from real news in terms of language styles or subject matter manipulation. We can learn more about the distinctive linguistic and thematic elements of fake news by examining the abbreviated U and Vh matrices. Developing more potent detection techniques and comprehending the strategies employed by those who fabricate fake news can both benefit from this information (Vu et al., 2021).

2.2 Cat boosting Classifier

Cat Boosting is a family collection of boosting hybrid machine teaching techniques that combine the gradient boosting algorithm with several feeble amateurs to become a powerful professional. This type of classifier executes a random transposition of the given dataset and computes the mean categorical value for every article, and it does not categorically make use of binary substitution of the label values (Hancock et al., 2020).

Cat-boosting thereby reduces overfitting that precedes a greater degree of generalizing algorithm (Dong and Qian, 2022). Cat Boosting utilizes ordered target statistics (OTS) to overcome unconditional values for a given feature of the unconditional value, that is to say, the unconditional value is ranked before the specimen is modified with the objective of the initial feature value (Mingyu et al., 2021).

2.3 LGBM Classifier

The Light Gradient Boosting Machine (LGBM) classifier can handle big datasets, quick training times, and a good performance accuracy that have made it a potent instrument for the analysis and identification of fake news. Various research has depicted how well LGBM works to differentiate between authentic and fake news articles (Gupta et al., 2022).

Using the histogram technique and other algorithms, the Light Gradient Boosting Machine (LGBM) classifier is an adaptive gradient boosting model, an effective gradient boosting tree implementation that increases the algorithm's processing capacity and prediction accuracy (Wang et al., 2022).

2.4 Review of Literature

Nasir et al., (2021), reviewed an article to ascertain an automatic fake news detection technique which is the foundation of deep learning network classifier, and machine learning algorithms were explored to fight the wide spread of incorrect news. A novel hybrid deep learning model was recommended by a current study in the categorization of fake news. Two fake news datasets were essentially verified by the algorithm, producing a better performance detection accuracy that is far better than the non-ensemble baseline techniques. The ISOT dataset machine learning technique nevertheless, produce an accuracy that is lower than 90 %.

Fung, Y. et al., (2021), have proposed a new standard for fake news detection at the elementary level of knowledge which also include answer for the process which includes cross-media consistency checking that detect fine-grained knowledge elements that makes news contents misleading.

Lavanya et al., (2021), have proposed that the traditional analytics model was not able to produce an effective and efficient algorithm for the analysis of data by uncovering the masked pattern within the required time like the way Deep learning network does. Deep learning network also discover significant information from the important features of health data which Deep Learning (DL) network stipulates. In particularly, Deep Learning (DL) classifiers comprises of producing better accuracy for social healthcare networks by using the models of pattern recognition.

Nagashri and Sangeetha (2021), research focus on an attempt to find the connection between the news contents and the

circumstances in which they occur inside the text, it can also be used to detect if a texts is real (negative cases) or fake (positive cases). Count Vectorize model was use to transform character-based texts into numeric substitutions and examined which algorithm have sufficient capacity to detect which of the text that is fake or genuine.

Shaikh et al., (2020), research evaluated the prospects of features are taking out of the TF-IDF to detect fake news contents from datasets, and these datasets are limited and at the same time the samples of the dataset are small. With two classifiers; passive-aggressive classifier and Support vector machine model obtaining an accuracy of 95%.

Salman Bin Naeem et al., (2021), research article shows that the manipulations of fake news contents is not only restricted to Politics, various survey conducted depicts that around 1225 fake news tales was disseminated during COVID-19 pandemic, with majority of the news emanating from the website, thereby placing the general public health in serious jeopardy, this important result of spreading fake news to the general public have encouraged rigorous exploration in the field of fake news, especially on the website and Twitter which is the main area of dissemination of fake news in social media network.

Zhang, J. et al., (2020), have proposed FAKEDETECTOR, as a new automatic fake news detection technique, it is basically deduced from a group of very specific and inactive category sunder out of, or pertaining to text information. It creates a meaningful diffusive network technique to learn the exhibitions of news contents, texts and subjects.

Without using cross-validation, Jiang et al. (2021) assessed the effectiveness of three deep learning models and five machine learning models on two datasets of real and fake news of varying sizes. Moreover, several machine learning and deep learning models need the ability to detect fakes via sentiment analysis.

This study aims to fill this research gap and contribute to the discussions by introducing a new multiple imputation method for fake news detection research that takes into account social media and news content, including both unstructured opinions in the form of text and structured opinions shared via sub-ratings like user opinions in comments on particular content attributes. Sentiment analysis is also used in this imputation strategy. Regarding English word segmentation, it is challenging to evaluate the data in the problem holistically and determine the user's genuine intention because existing machine learning techniques are unable to resolve the long-distance connections of texts. As a result, several models based on machine learning are used to identify and categorise false information. The capacity of each model to accurately classify different news items is measured, allowing for the improvement of each model's accuracy in identifying false information.

A mechanism has been presented by Preston, S. et al. (2021) to ascertain whether people who have a greater level of "emotional intelligence" are less likely to fall for fake news. Pandey, S. et al. (2022) have employed a range of methods, such as K-Nearest Neighbour, Support Vector Machine, Decision Tree, Naïve Bayes, and Logistic Regression techniques, to distinguish between authentic and fake news within a given dataset. Additionally, they have enhanced the efficacy of these methods by employing data cleaning to address improperly balanced data. This study offers TSVD-based machine learning techniques for the detection of false news. In this study, machine learning with TSVD is employed since it enables the computer to learn without the need for explicit programming. It is also easy to visualise and has the potential to improve data noise management. Lastly, compared to the existing classifiers, the Cat Boost classifier, Light Gradient Boosting Machine classifier, and Support Vector machine model with TSVD show superior accuracy.

Tables 1 represent the review of related literature connected to fake news analysis and detection.

Table 1: Review of related literature

S/N	Authors	Method used/description	Refs no.
1	Nasir, Khan, and Varlamis	Fake news classification bimodal using convolutional neural network and long short-term memory.	(Nasir et al., 2021)
2	Fung, Thomas, Reddy, Polisetty, Chang, and Sil	Machine Learning Ensemble Methods for detecting fake news.	(Fung, Y. et al., 2021)
3	Lavanya, and Sasikala	Detecting an opinion spams and fakes news using text classification.	(Lavanya et al., 2021)
4	Nagashri, and Sangeetha	Machine learning and data augmentation for fake news identification.	(Nagashri et al., 2021)
5	Shaikh, and Patil	A deep learning ensemble model called 'fake detect' is used for identifying bogus news.	(Shaikh et al., 2020)
6	Salman Bin Naeem, Rubina Bhatti, and Aqsa Khan	Comparative performance of machine learning algorithms for fake news identification.	(Salman Bin Naeem et al., 2021)
7	Zhang, Dong, and Philip	A study of methods for spotting false information and rumours.	(Zhang, J. et al., 2020)
8	Jiang, Li , Haq , Saboor and Ali	Deep learning techniques are optimized and improved for the benefit of society in fake news identification.	(Jiang et al.,2021)
9	Preston, Anderson, Robertson, Shephard, and Huhe	An assessment of the latest techniques for thwarting fake news on social media.	(Preston et al., 2021)
10	Pandey, Prabhakaran, Reddy, and Acharya	Finding bogus news using automatic deception detection techniques.	(Pandey et al., 2022)

3.0 Methodology

This part represents the different methodologies combined to accomplish the preferred objective based on the recommended network. The recommended methodology is further introduced to comprehend this whole task based on the network design. The process of the collection of data is deliberated, and afterward, the resources used in the analyzed data coupled with the checking for robustness are elaborated.

3.1 Network construction

The network construction used in this paper is provided in this part. The network design's serial connection is shown in a flowchart in Figure 1. The abstract of the dataset, which comprises both fake and actual content, is accessible from https://www.uvic.ca/engineering/ece/isot/datasets/fake-news/index.php. It is commonly referred to as the "ISOT Fake News Dataset" (Ahmed et al., 2018). Afterwards, it undergoes pre-processing to address the issue of missing values, discretization, and label encoding. After that, this methodology is developed, trained, and evaluated.

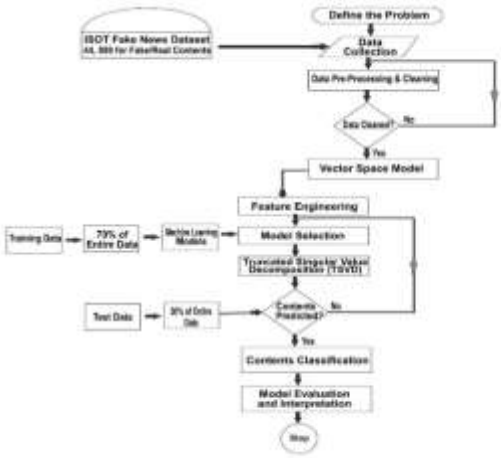

Figure 1: Designed Network flow chart

A bar chart representing the counts of authentic and fraudulent news items relative to the target in a flip chart format is shown in Figure 2. The comprehensive description of the dataset containing true and fraudulent news is shown in Table 2, including the eight different subjects and their recurrence rates are also shown in Table 2.

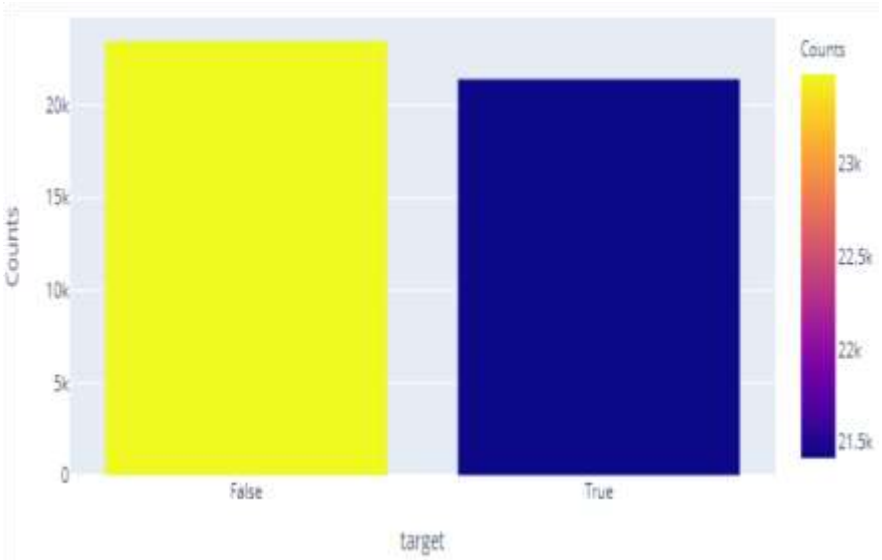

Figure 2. Fake and Real news content counts

Table 2. Description of Fake Contents and Real Contents dataset and their rate of occurrence of each topic.

News	Size (Number of contents)	Topics	
Real-News	21417	**Type**	**Contents size**
		World-News	10145
		Politics-News	11272
Fake-News	23481	**Type**	**Contents size**
		Government News	1570
		Middle-east	778
		US News	783
		left-news	4459
		politics	6841
		News	9050

3.2 Performance Metrics

To calculate how different classifiers are performed based on different performance requirements, such as accuracy, precision, recall, and so on, performance metrics are utilized. The following standards of measurement are discussed:

3.2.1 Confusion Matrix

The ability of three different algorithms to classify these tactics was tested, and the algorithm that produced the highest level of accuracy was found. A confusion matrix was then used to finish the assignment. The confusion matrix displays both the number of false positives and the true positives generated by the classifiers. The confusion metrics found matched the accuracy. The table is a confusion matrix for the binary categorization.

Where the number for accurately identifying phoney content is called True Positive (TP). False Negative (FN): Is the number used to identify fraudulent news that is presented as authentic content? False Positive (FP): A code for phoney information that was assumed to be phoney news. True Negative (TN): Is the amount of authentic content that has been found authentic?

A performance metric called the confusion matrix table is employed to ascertain the unique accomplishment of the matrix as conveyed.

Table 3. Confusion Matrix

		Predicted	
		Negative	Positive
Actual	Negative	TN	FP
	Positive	FN	TP

TN = True Negative
FP = False Positive
FN = False Negative
TP =True Positive

F1 Score

Equation 1 illustrates the equilibrium between precision and recall, which is the goal of the F1 Score measure, which combines the two metrics together.

$$F1Score=\frac{2*precision*Recall}{Precision + Recall} \quad (1)$$

Recall

The equality of real events that are correctly classified as real is known as recall. Another common name for recall is sensitivity, as Equation 2 illustrates.

$$Recall= \frac{TP}{TP+FP} \quad (2)$$

Precision

Equation 3 illustrates precision as the ratio of precisely detected genuine instances to the total number of detected real instances.

$$Precision = \frac{TP}{TP+FN} \quad (3)$$

Accuracy

The ratio of the number of correctly classified occurrences to the total number of real events is

known as accuracy. Equation 4 illustrates how it is comparable to adding TP and TN divided by the total of the occurrence's numeral.

$$Accuracy=\frac{(TP+TN)}{(TP+TN+FP+FN)} \quad (4)$$

4.0 EXPERIMENTAL RESULTS AND DISCUSSION

The experiments were performed on Python language 3.9 for the algorithm development and the result was generated. At the same time, Jupiter notebook will be used as the text editor, with software on HP core i9 computer with 500GG SSD,64Gb RAM 16GB.

4.1 Experimental results

Numerous efficacy tests were performed in the process of estimating the suggested design. In this part, we deliberate on the various operations that had been completed and the accuracy of these applications.

The entire set of three classifiers is executed in Google Colab, which stipulates an environment in the cloud. Due to this reason, Python 3.9 and higher versions of it can be used. The Python packages that were used for training and testing data include Natural Language Toolkit (NLTK), Pandas, Numpy, Matplotlib, Seaborn, and Scikit Learning. The dataset is shared among binary sections (training and testing) in the ratio of 70:30

4.2. Performance of various Classifiers

Table 5 presents an overview of the overall accuracy obtained following the cleaning process. It includes the total quantity of false information detected, the classification of real news, and the results (accuracies), presented in numerical form.

4.3. Performance of Different Classifiers on Fake News Dataset

A comparison Table for the five various Performance Classifiers is represented in Table 6

The comparison of the various research work accuracy in the previous three years is depicted in Table 6. Table 6 reveals further noteworthy findings from other studies, albeit with smaller sample sizes than our paper's accuracy. This illustrates the superiority of our solution. The fact that the ISOT datasets are rarely large enough is one of the primary challenges in the current research. Three machine-learning classifiers using TSVD were used in this study, which was carried out in a unique way. The machine learning algorithms using TSVD outperformed the current classifiers with a higher performance accuracy in the analysis of fake news identification.

4.4. Discussion

The study's findings from the most recent algorithms demonstrate that the classifiers produced a rare outcome, ensuring the success of the investigated research. The present research also yields effectiveness rates of 90%; this study illustrates that misinformation can be promptly identified and effectively addressed. If this paper achieves an accuracy of more than 80%, it is considered effective. The outcome of this latest article, which yielded the finest quality conceivable, is quite remarkable. According to the paper, fake news won't be a significant issue going forward.

Additionally, this work has shown that the most important task for traditional articles is data cleaning. The spread of false information is caused by a number of causes, and this study has identified some potential responses to false information. We believe that any work to continue with this article will require a graphic user interface (GUI). Both the design and the appearance of a software product depend on a

good graphical user interface (GUI). Anyone may flawlessly duplicate and patch any contents into the Graphical User Interface (GUI) platform and acquire the desired results with the help of GUI. GUI illustrates how computers are both entertaining and simplifying our daily lives. Rules governing user processes, computer settings, and decision support: based on our analysis of user requirements, it is clear that our primary user procedure would be to differentiate between real and fake news. End consumers will therefore be able to distinguish between false and accurate news. The technologies employed in this paper included machine learning classifiers using TSVD. After loading the necessary libraries, classifiers such as Cat Boost classifier, Light Gradient Boosting Machine, and support vector classifiers could be used using TSVD. This model was selected for this work due to its ability to produce accurate evaluations with higher performance accuracy.

5.0 Contributions to Literature and implications for practice

The following are the contributions to Literature and implications for practice

5.1 Contributions to Literature:

1. Creation of new techniques: The investigation of machine learning in the analysis and identification of false news has led to the creation of new approaches and procedures for recognizing and detecting fake news, which can aid in the creation of more reliable and efficient models.

2. Understanding the traits of false news, such as the language, tone, and material used to produce it, has been made easier thanks to this study.

3. An understanding of social media's function: The study has brought attention to how social media plays a part in the spread of false information and how social media companies must act to stop the spread of false information.

5.2 Implications for Practice:

1. Development of tools: The study can lead to the development of tools and software that can detect fake news in real-time, which can help organizations and individuals prevent the spread of misinformation.
2. Identification of potential areas for improvement: The study can help organizations identify areas where they need to improve their data collection and analysis practices to detect and prevent the spread of fake news.
3. Promotion of media literacy: The study can promote media literacy and critical thinking skills, which can help individuals identify and avoid fake news.
4. Cooperation with social media platforms: In order to create tactics and resources to counter fake news, the study can promote cooperation between academics, legislators, and social media platforms.

5.3 Limitations of the Study

Machine learning algorithms using TSVD have the potential to identify false news, but they still have several issues that need to be resolved before they can be used as a dependable method. These issues include a lack of labeled data, bias in the data, the dynamic nature of fake news, adversarial attacks, and a limited comprehension of human language.

6. CONCLUSION

Fake news has become one of the most interesting areas of exploratory topics as a result of an increase in cybercrime and rising cases of fake news content. Various researchers have been using distinct techniques in experimenting on distinct datasets to detect Fake news. Approaches in machine learning using TSVD were used for the analysis of Fake news detection.

Experimentation was performed on “ISOT Fake News Dataset” making use of estimation parameters of accuracy, recall, F1-score and precision . Looking at the estimated accuracies, it could be confirmed that the recommended machine learning algorithms using TSVD technique with an accuracy of 97.33% stand as a standard model in comparison with the current classifiers which possess an exactness of 95.05%.

Table 4. Results after administering machine learning classifiers after preprocessing.

S/ N.	Machine Learning Algorithms	Fake News Classified	Real News Classified	Accuracies (Results)
1	LightGBM Boost	2904	2955	98.00 %
2	Cat Boost	2894	2949	97.00 %
3	Support vector classifier	2910	2940	97.00 %

Table 5. Comparison table for the various Performance Classifiers.

Algorithms	Accuracy (%)	Precision (%)	Recall (%)	F1_score (%)
Cat Boost	97	97	97	97
LightGBM Boost	98	98	98	98
Support vector machine (SVM)	97	97	97	97

Table 6. Comparing this research paper with other works.

S/N	**Machine Learning Algorithms**	**Accuracies (Results)**	**Studies**
1	Cat Boost, LightGBM Boost, SVM algorithm with TSVD	97.33%	Proposed study
2	LSTM and BI-LSTM Classifier	91.51%	(Hancock et al.,2020)
3	Random forest algorithm, Perez-LSVM, Linear SVM, multilayer perceptron, bagging classifiers, boosting classifiers, KNN	88.86 %	(Ahmad et al.,2020)
4	Support Vector Machine (SVM) and Term Frequency-Inverted Document Frequency (TF-IDF)	95.05%	(Hamborg et al.,2018)

REFERENCES

Abdullah, A.; Awan, M.; Shehzad, M.; Ashraf, M. (2020). Fake news classification bimodal using convolutional neural network and long short-term memory. Int. J. Emerg. Technol. 2020, 11, 209–212.

Ahmad, I.; Yousaf, M.; Yousaf, S.; Ahmad, M.O. (2020) Fake News Detection Using Machine Learning Ensemble Methods. Complexity 2020, 2020, 1–11.

Ahmed, H.; Traore, I.and Saad, S. (2020). Detection of online fake news using n-gram analysis and machine learning techniques. In Proceedings of the International Conference on Intelligent, Secure, and Dependable Systems in Distributed and Cloud Environments,

Vancouver, BC, Canada, 26–28 October 2017; Springer: Cham, Switzerland, 2017.

Alam, T.M.; Awan, M.J. (2018). Domain analysis of information extraction techniques. Int. J. Multidiscip. Sci.Eng. 2018, *9*, 1 – 9.

Ali, Y.; Farooq, A.; Alam, T.M.; Farooq, M.S.; Awan, M.J.; Baig, T.I.(2019). Detection of Schistosomiasis Factors Using

Association Rule Mining. IEEE Access 2019, *7*, 186108–186114.

Anam, M.; Ponnusamy, V.A.; Hussain, M.; Nadeem, M.W.; Javed, M.; Goh, H.G.; Qadeer, S. 2021). Osteoporosis Prediction for Trabecular Bone using Machine Learning: A Review. Comput Mater. Contin. 2021, 67, 89–105.

Alonso, M.; Vilares, D.; Gómez-Rodríguez, C.; and Vilares, J. (2021). Sentiment Analysis for Fake News Detection. Electronics 2021, 10, 1348.

Ashraf, N., Butt, S., Sidorov, G., and Gelbukh, A. (2021). CIC at check that! 2021: fake news detection using machine learning and data augmentation. In CLEF, 2021–Conference and Labs of the Evaluation Forum.

Aslam, N., Ullah Khan, I., Alotaibi, F. S., Aldaej, L. A., and Aldubaikil, A. K. (2021). Fake detect: A deep learning ensemble model for fake news detection. Complexity, 2021.

Bali, A.P.S.; Fernandes, M.; Choubey, S.; Goel, M. (2019). Comparative performance of machine Learning algorithms for fake news detection. In Proceedings of the International Conference on Advances in Computing and Data Sciences, Ghazibad, India, 12 – 13April 2019; pp. 420–430.

Chauhan, T., and Palivela, H. (2021). Optimization and improvement of fake news detection using deep learning approaches for societal benefit. International Journal of Information Management Data Insights, 1(2), 100051.

Collins, B., Hoang, D. T., Nguyen, N. T., and Hwang, D. (2021). Trends in combating fake news on social media–a survey. Journal of Information and Telecommunication, 5(2), 247-266.

Desai, Meha, and Manan Shah. (2021).An anatomization on breast cancer detection and diagnosis employing multi-layer perceptron neural network (MLP) and Convolutional neural network (CNN). Clinical eHealth 4: 1–11.

Gupta, M.; Jain, R.; Arora, S.; Gupta, A.; Awan, M.J.; Chaudhary, G.; Nobanee, H.(2021). AI- enabled COVID-19 outbreak analysis and prediction: Indian states vs. union territories. Comput.Mater. Contin. 2021, 67, 1–18.

Hamborg, F.; Donnay, K.; Gipp, B. (2018). Automated identification of media bias in news articles: An interdisciplinary literature review. Int. J. Digit. Libr. 2018, 20, 391–415.

Hannah Nithya S, and Sahayadhas A (2022). Automated fake news detection by LSTM enabled with optimal feature selection. J Inf Knowl Manag 21(03). https://doi.org/10.1142/s0219649222500368.

Hua, J.; Shaw, R.J.I. Corona virus (Covid-19) "infodemic" and emerging issues through a data lens: The case of china. Int. J. Environ. Res. Public Health 2020, 17, 2309.

Hancock, J.T.; Khoshgoftaar, T.M. (2020). Cat Boost for big data: An interdisciplinary review. J. Big Data 2020, 7, 94.

Jang, Y.; Park, C.-H.; Seo, Y.-S. (2019).Fake News Analysis Modeling Using Quote Retweet.

Electronics 2019, *8*, 1377. Javed, R.; Saba, T.; Humdullah, S.; Jamail, N.S.M.; Awan, M.J. (2021). An Efficient Pattern Recognition Based Method for Drug-Drug Interaction Diagnosis. In Proceedings of the 2021 1st International Conference on Artificial Intelligence and Data Analytics (CAIDA), Riyadh, Saudi Arabia, 6–7. April 2021; pp. 221–226.

Jiawei Zhang, Bowen Dong, and S Yu Philip. (2020). "Fake detector: Effective fake news detection with deep diffusive neural network". In: 2020IEEE 36th International Conference on Data Engineering (ICDE). IEEE. 2020, pp. 1826–1829.

Jwa, H., Oh, D., Park, K., Kang, J.M., Lim, and H. (2019): exBAKE: automatic fake news Detection Model based on Bidirectional Encoder Representations from Transformer (BERT). Appl. Sci.9, 4062 (2019). https://doi.org/10.3390/app9194062.

Kesarwani, A., Chauhan, S.S., Nair, A.R. and Verma, G. (2021). Supervised Machine Learning Algorithms for Fake News Detection. In Advances in Communication and Computational Technology (pp. 767-778). Springer, Singapore.

Kogan, S.; Moskowitz, T.J.; Niessner, M. (2017). Fake News in Financial Markets; Working Paper; Yale University: New Haven, CT, USA, 2017.

Lai, C.M.; Shiu, H.-J.; Chapman, J. (2020). Quantifiable Interactivity of Malicious URLs and the Social Media Ecosystem. Electronics 2020, *9*, 2020.

Lavanya, P. M., and Sasikala, E. (2021). Deep Learning Techniques on Text Classification Using Natural Language Processing (NLP) In Social Healthcare Network: A Comprehensive Survey. In 2021 3rd International Conference on Signal Processing and Communication *(*ICPSC*)* (pp. 603- 609).

Lazer, D.M.J.; Baum, M.A.; Benkler, Y.; Berinsky, A.J.; Greenhill, K.M.; Menczer, F.; Metzger,

M.J.; Nyhan, B.; Pennycook, G.; Rothschild, D. (20180. The science of fake news. Science 2018, 359, 1094–1096.

Meesad, P. (2021). Thai Fake News Detection Based on Information Retrieval, Natural Language Processing and Machine Learning. SN Computer Science, 2(6), 1-17.

Mostafa, K. R., Khan, M. R., Shifat, U. M., and Mahmud, M. T. (2023). Deep Learning Algorithms: Benefits and Drawbacks. In Emerging Trends in Smart Technologies for Security and Surveillance (pp. 43-60). Springer, Singapore.

M.Senthil Raja, L.Arun and Raj, (2022).Fake news detection on social networks using Machine learning techniques, Materials Today: Proceedings, Volume 62, Part 7, 2022, Pages 4821-4827, ISSN 2214-7853, https://doi.org/10.1016/j.matpr.2022.03.351. (https://www.sciencedirect.com/science/article/pii/S2214785322016947)

Mousavirad SJ, Schaefer G, Korovin I, Oliva D. (2021). Rde-op: A region-based differential evolution algorithm incorporation opposition-based learning for optimizing the learning process of multi-layer neural networks. In: International Conference on the Applications of Evolutionary Computation (Part of EvoStar). Springer, pp 407–420.

Nagi, A.T.; Awan, M.J.; Javed, R.; Ayesha, N. (2021). A Comparison of Two-Stage Classifier Algorithm with Ensemble Techniques on Detection of Diabetic Retinopathy. In Proceedings of the 2021 1st International Conference on Artificial Intelligence and Data Analytics (CAIDA), Riyadh, Saudi Arabia, 6–7 April 2021; pp. 212–215.

Namoun, A.; Alshanqiti, A. (2021). Predicting student performance using data mining an learning analytics techniques: A systematic literature review. Appl. Sci. 2021, 11, 237.

Sharma, N.Sharma, R. and Jindal. (2021).'Machine learning and deep learning applications-a vision, Global Transitions proceedings, vol.2, no.1, pp.24-28, 2021.

Pulido, C.M.; Ruiz-Eugenio, L.; Redondo-Sama, G.; Villarejo-Carballido, B. (2020). A New Application of Social Impact in Social Media for Overcoming Fake News in Health. Int. J. Environ.Res. Public Health 2020, 17, 2430.

Rana, Arti, Arvind Singh Rawat, Anchit Bijalwan, and Himanshu Bahuguna. (2018). Application of multi-layer (perceptron) artificial neural network in the diagnosis system: A systematic review. Paper presented at 2018 International Conference on Research157 in Intelligent and Computing in Engineering (RICE), San Salvador, El Salvador, August 22–24; pp. 1–6.

R Seetharaan, M Tharun, S.S. Sreeja Mole, and K Anandan, (2022). Analysis of fake news detection using machine learning technique, Materials Today: Proceedings, Volume 51, Part 8, 2022,

Shaikh, J.; Patil, R. (2020). Fake News Detection using Machine Learning. In Proceedings of the 2020 IEEE International Symposium on Sustainable Energy, Signal Processing and Cyber Security (iSSSC), San Francisco, CA, USA, 16–17 December 2020; pp. 1 – 5.

Sharma, D.K.; Garg, S.; Shrivastava, P. (2021). Evaluation of Tools and Extension for Fake News Detection. In Proceedings of the 2021 International Conference on Innovative Practices in Technology and Management (ICIPTM), Gautam Buddh Nagar, India, 17 – 19 February 2021; pp. 227–232.

Vu, T., Chunikhina, E., and Raich, R. (2021). Perturbation expansions and error bounds for the truncated singular value decomposition. Linear Algebra and its Applications, 627, 94-139.

Wang, Y., Yang, W., Ma, F., Xu, J., Zhong, B., Deng, Q, Gao, and J. (2020) Weak supervision for fake news detection via reinforcement learning. In: AAAI 2020—34th AAAI Conference on Artificial Intelligence, pp. 516–523 (2020).

Xu, Yuanjin, Fei Li, and Armin Asgari. (2022). Prediction and optimization of heating and cooling loads in a residential building based on multi-layer perceptron neural network and different optimization algorithms. Energy 240: 122692.