Journal of Informatics Education and Research
ISSN: 1526-4726
https://doi.org/10.52783/jier.v3i2.207
Vol 3 Issue 2 (2023)

Comparing the Effectiveness of Different Machine Learning and Deep Learning Models for Fake News Detection

***Sanjeev M. Dwivedi[1], Sunil B. Wankhade[2]**
[1] Research Scholar, Vidyalankar Institute of Technology, Wadala, Mumbai, India
[1]*sanjeev.dwivedi@vit.edu.in
[2]Professor and Research Guide, Rajiv Gandhi Institute of Technology, Andheri,Mumbai, India
[2]sunil.wankhade@mctrgit.ac.in

Abstract

Today's information ecosystem, with its abundance of fake news, presents serious obstacles to accurate reporting. Several machine learning and deep learning models have been implemented to help in the detection of fake news. In this research, we evaluate the merits and weaknesses of various algorithms for spotting false news. Deep learning models LSTM and CNN are examined alongside popular machine learning models such as Gaussian Naive Bayes, Decision Tree Classifier, Random Forest Classifier, XGBoost, and LightGBM. Accuracy is the metric against which the models are measured. In terms of accuracy in classifying false news stories, the results demonstrate that the deep learning models, and in particular LSTM and CNN, perform better than the machine learning models. While CNN is good at capturing structural information and local relationships, LSTM excels at capturing long-term dependencies and language patterns. The research emphasises the superiority of deep learning models in detecting false news, providing useful insights for the creation of trustworthy detection systems. The results pave the way for more studies into false news detection methods and contribute to their development.

Keywords: Fake news detection, machine learning, deep learning, LSTM, CNN, accuracy.

Introduction

The rise of false information has emerged as a major problem in the modern media environment. To manipulate public opinion or fool readers, fake news is the material that has been manufactured or is otherwise deceptive yet is presented as news. Fake news has the ability to spread swiftly, impact public discourse, and have major effects due to the rapid rise of social media platforms and the ease of sharing information online.

Numerous problems for people, communities, and democracies arise from the spread of false information. It may cause people to lose faith in established news outlets, spread false information, and even sway election results. Therefore, maintaining an educated society necessitates the capacity to efficiently recognise and counteract bogus news.

Given the scope and magnitude of the issue, automated techniques based on machine learning and deep learning models have emerged as viable methods for identifying false news. These programmes can sift through mountains of textual and contextual data, looking for the telltale signs of false news.

The purpose of this study is to evaluate and contrast several machine learning and deep learning models for spotting false news. We want to uncover the most accurate and reliable techniques for combating the spread of false news by analysing and comparing the performance of various models. This research can help in the battle against online disinformation by contributing to the creation of more reliable false news detection systems.

Significance of Reliable Fake News Detection Methods

In today's digital era, fake news has become a major problem, posing serious threats to individuals, communities, and democracies. Fake news has the capacity to skew public opinion, damage faith in the media, and even affect political results. In this case, effective strategies for detecting fake news are essential for reducing the negative impacts of disinformation. The relevance and ramifications of effective strategies for detecting false news are discussed here.

First and foremost, an educated society cannot exist without the creation of effective mechanisms for detecting false news. Individuals are better equipped to make judgments based on accurate and trustworthy information when they can

Journal of Informatics Education and Research
ISSN: 1526-4726
https://doi.org/10.52783/jier.v3i2.207
Vol 3 Issue 2 (2023)

recognise and distinguish false news from genuine news. As a result, they are better able to determine which news outlets can be trusted and avoid being led astray by propaganda.

Additionally, trustworthy fake news detecting systems aid in stopping the spread of false information. These techniques can help lessen the influence of misleading narratives by correctly recognising and labelling fake news. In the context of social media, where false information may spread swiftly through likes and shares, this is of paramount importance. The effect of false news on public opinion can be mitigated with the use of reliable detecting tools.

The maintenance of faith in media outlets is another important result of effective fake news-detecting technologies. Fake news can further undermine public trust in an era when faith in established media is already being questioned. Media outlets may win back the confidence of their readers by showing that they value factuality and honesty by using reliable detection techniques.

Safeguarding democratic processes also depends on having effective ways of detecting false news. Disseminating false information to the public has the potential to have far-reaching effects, such as swaying elections and causing social divisions. Identifying and publicising instances of such manipulation is vital to protecting democratic values and facilitating deliberation about all relevant factors.

Research Objective and Motivation

The focus of this study is on evaluating and contrasting several machine learning and deep learning models for spotting false news. Our goal is to determine the most effective strategies for combating the spread of fake news by comparing and contrasting the results produced by various models. The study's primary objectives are as follows:

1. Which machine learning and deep learning models are most effective for fake news detection?
2. What are the strengths and limitations of different approaches?
3. How do these models perform in terms of accuracy, precision, recall, and other evaluation metrics?
4. Are there any discernible trends or insights that can help improve fake news detection?

The urgent need to provide trustworthy ways for identifying bogus news is what drives this investigation. Due to the widespread nature of false news and the consequences it may have on society, existing detection methods must be thoroughly analysed and compared to one another. Through this study, we hope to progress the field by identifying promising strategies and illuminating their potential and pitfalls.

Related work

There is already a growing corpus of literature dedicated to the problem of identifying false news stories. In this part, we examine previous works that have investigated various methods for identifying bogus news. Many other approaches are discussed in the various works, such as machine learning, NLP, network analysis, and social media analysis. By reviewing the existing literature, we may better understand the present status of the area and the advantages and disadvantages of various techniques.

Machine learning is a popular method for spotting bogus news. In order to discern false news from real news, many researchers have used machine learning algorithms to analyse textual properties. To determine whether news stories were real or fraudulent based on linguistic and structural data, Horne and Adali (2017) used supervised machine learning techniques including logistic regression and support vector machines. They got some encouraging findings that show the promise of machine learning for detecting bogus news.

Similarly, Shu et al. (2017) suggested a system dubbed "FakeNewsNet" that utilises machine learning and social network analysis to detect false news stories. Classifiers for identifying bogus news were trained using factors such as user interaction patterns, textual content, and network architecture. Their research demonstrated the need to incorporate both semantic and contextual elements into false news detection systems.

Natural language processing (NLP) techniques have also been widely used in the identification of false news. Natural language processing allows for the examination of semantic structures and linguistic features in news items, which might yield useful information in the fight against false news. Vosoughi et al. (2018) compared the linguistic characteristics of real and fake news items and found that the latter frequently use more subjective and emotive language. The importance of language signals in spotting false news is emphasised in this paper.

Additionally, network analysis has been used to spot hoaxes by tracking how various news stories are disseminated. In order to differentiate between the transmission of factual and fictitious information on Twitter, Vosoughi et al. (2018) performed a comprehensive analysis of Twitter cascades. What they discovered was that fake news is more likely to go viral and reach a wider audience than real news, suggesting that network analysis may be used to identify the spread of fake news.

The study of social media has become an important tool in the fight against fake news. Gupta et al. (2020) suggested a system for identifying Twitter users who are spreading false information by combining textual and social network analysis. They identified important markers that separate false news spreaders from genuine users by analysing user behaviours, interaction patterns, and content properties.

Fact-checking methods, credibility evaluations, and multimedia analysis have also been studied as potential methods for identifying false news. The "Hoaxy" system, created by Wang et al. (2017), monitors the dissemination of false news and counterfactual data. They hoped that by contrasting the reliability of news pieces with those of fact-checking sources, they might help readers determine whether stories should be trusted.

Furthermore, multimedia analysis has garnered interest in the context of verifying images and videos for use in countering fake news. The visual content of photos and videos has been analysed using deep learning models like convolutional neural networks to detect modifications and alterations that signal false news.

While the aforementioned research has made important strides, there are still obstacles and constraints that need to be overcome in the quest to identify false news. The lack of readily available labelled datasets for use in training and assessment is a widespread problem. It is still difficult to compile large-scale labelled datasets that adequately capture the variety of false news stories. To successfully capture shifting patterns and strategies utilised by malicious actors, detection algorithms must be continuously adapted to account for the ever-changing nature of false news.

Past research has used a wide variety of machine learning and deep learning models to try to catch fake news. By analysing textual and contextual aspects, these algorithms can determine distinguishing characteristics between false and real news. Here, we take a look at some of the most popular models for detecting false news and how they are put to use in the field.

1. Logistic Regression

   The likelihood of a binary outcome may be estimated with the help of the popular machine learning model logistic regression. Logistic regression models may be educated on a wide range of parameters useful in the context of detecting false news, including language clues, content characteristics, and social network data. These models can train to identify false and real news stories using the provided attributes and produce results that are interpretable in terms of the value of those aspects. However, logistic regression models may not be able to fully capture non-linear patterns in the data or more complicated interactions.

2. Support Vector Machine

   In high-dimensional feature spaces, SVM is a strong supervised learning technique that uses hyperplanes to separate data points. Using language variables, content characteristics, and network aspects, SVM models have been used to the problem of false news identification. Known for their versatility, SVMs can handle scenarios with a limited number of training samples and high-dimensional data with ease. However, the computational complexity of SVMs may pose problems when working with large-scale datasets.

3. Random Forests

   Random Forests are a type of ensemble learning model that create predictions by using a forest of decision trees. Each decision tree contributes to the final prediction by being trained on a different group of information. The textual and structural aspects of news stories have been analysed using Random Forests for use in false news identification. These models are resistant to overfitting and can capture non-linear correlations between features.

Random Forests have a lot going for them, but it can be hard to interpret their results because of how complex the underlying decision-making process is.

4. Naive Bayes

   Naive Bayes is a Bayesian theorem-based probabilistic classifier. It determines the likelihood of an instance belonging to a specific class by assuming that all characteristics are independent. In order to spot false news, researchers have used Naive Bayes models that take into account language aspects, content characteristics, and social network data. Naive Bayes models can process huge datasets with little to no extra processing effort. However, their effectiveness is hindered by the fact that they assume feature independence, which may not hold in complicated real-world circumstances.

5. RNNs (Recurrent Neural Networks):

   Recurrent neural networks (RNNs) are a type of deep learning model created to detect patterns in data over time. They have been used in the identification of false news by analysing the articles' text and temporal relationships. RNNs can represent text efficiently because of their ability to take sentence and word order into account, thereby capturing context and long-term dependencies. However, problems in training deep networks arise when RNNs experience vanishing or exploding gradients. Furthermore, RNNs may have difficulty processing exceptionally lengthy textual sequences.

6. CNN

   CNNs are a type of deep learning model often used for image processing, but they have also been successfully used to the problem of identifying fake news. When given as a 1D sequence of data, CNNs may analyse textual content. In order to identify useful language clues, they use convolutional filters to pick up on local patterns and characteristics in the text. CNNs fare well at collecting local context and feature representations, but they may struggle to do so when it comes to capturing long-range relationships.

Analyzes the strengths and limitations of different approaches

Each of the above-mentioned models has its strengths and limitations when applied to fake news detection:

1. Machine Learning Models:

   - Strengths: Machine learning models like logistic regression, SVM, and Random Forests offer interpretability, enabling researchers to understand the importance of different features in fake news detection. They can handle high-dimensional data and large-scale datasets efficiently.

   - Limitations: These models might struggle to capture complex relationships and non-linear patterns in the data. Their interpretability can be limited in ensemble models like Random Forests, and they might require extensive feature engineering to achieve optimal performance.

2. Deep Learning Models:

   - Strengths: Deep learning models like RNNs and CNNs can capture complex relationships and sequential dependencies in textual data. They excel at capturing long-range dependencies, context, and feature representations. They can learn hierarchical representations automatically without extensive feature engineering.

   - Limitations: Deep learning models often require large amounts of labeled data for training, which can be challenging to obtain for fake news detection. They can be computationally expensive to train, especially for complex architectures. Deep learning models may also suffer from overfitting if not properly regularized, and their interpretability can be limited compared to traditional machine learning models.

It is worth noting that the performance of these models can vary depending on factors such as the quality and size of the dataset, the selection of features, and the specific implementation. Therefore, it is crucial to carefully evaluate and compare the performance of different models to identify the most effective approach for fake news detection.

Journal of Informatics Education and Research
ISSN: 1526-4726
https://doi.org/10.52783/jier.v3i2.207
Vol 3 Issue 2 (2023)

Methodology

Dataset

Our study's methodology included everything from the dataset we utilised to the preprocessing procedures we ran on the raw data. We employ the ISOT Fake News dataset, a freely available collection of over 3,000 articles that includes both false and accurate claims. News items from both reputable news outlets and those deemed untrustworthy by Politifact.com are included in the dataset created by the ISOT Research Group at Canada's University of Victoria.

To ensure the data was in the right format for training the deep learning models, we ran it through a series of preprocessing routines. Text cleaning was the first stage in the preparation pipeline. The next thing that was done was to clean up the textual data by eliminating any traces of HTML elements, punctuation, and other special characters. These glyphs may just be random fluctuations in the data that don't contribute anything useful to the models. We moved on to tokenization after cleansing the text.

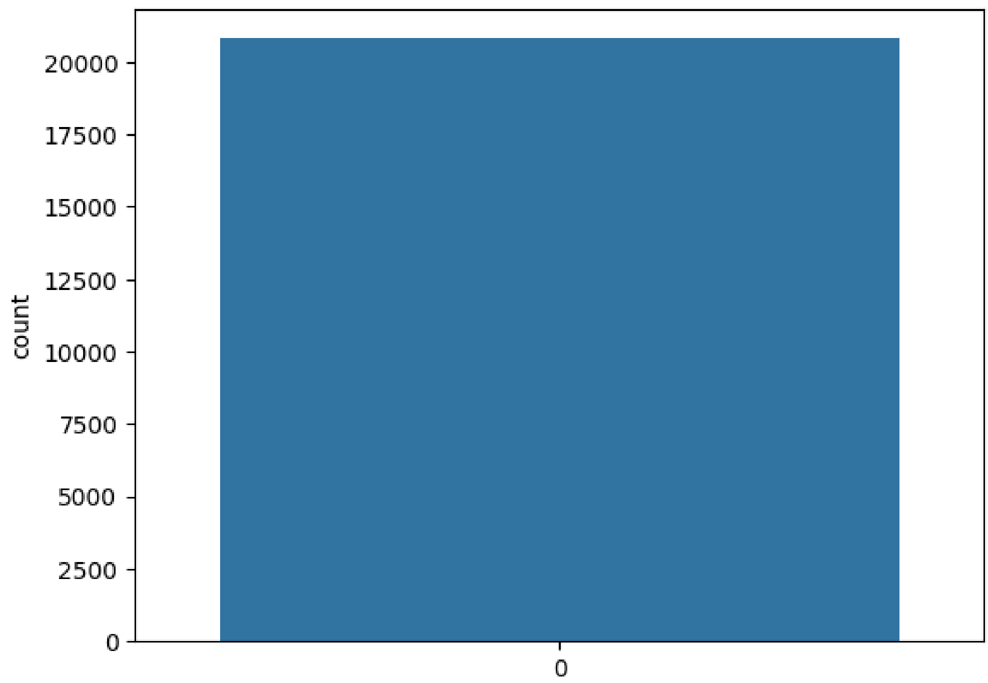

Fig. 1. Total Number of News analyzed

The process of tokenization includes separating the text into separate words. This is a crucial stage for deep learning models to effectively analyse the data. For tokenization, we relied on the NLTK package for Python. After the text was tokenized, we took out all of the unnecessary words. The English language is full of "stop words" like "the," "a," and "an" that don't add any sense to what's being said. Data dimensionality and model accuracy can both benefit from stop-word removal.

After deleting stop words, we ran the remainder of the text data through a stemming/lemmatization process. Stemming and lemmatization are two methods that may be used to deconstruct words. Examples include "running," "ran," or "run," all of which would be abbreviated to "run." Reducing the number of unique words in the text data by normalisation techniques like stemming and lemmatization improves the performance of the models.

We utilised word clouds to see which terms appeared most often in the news stories, which helped us understand the dataset better. In a word cloud, the relative size of the individual words reflects their relative frequency in the original text data. Figure 3. illustrates the dataset's actual news as a word cloud.

Journal of Informatics Education and Research
ISSN: 1526-4726
https://doi.org/10.52783/jier.v3i2.207
Vol 3 Issue 2 (2023)

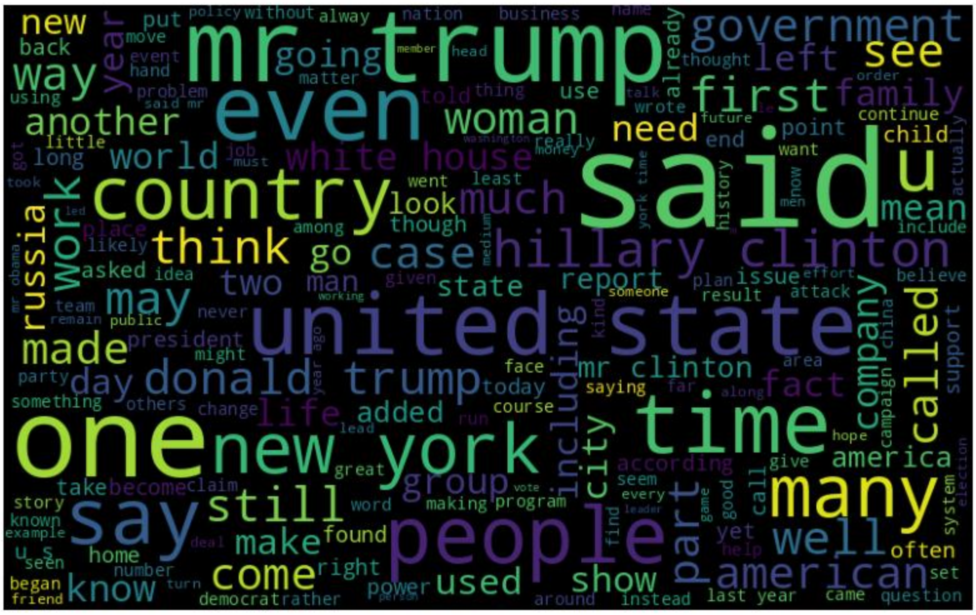

Fig. 2. (a) Fake News Cloud

In this research, we used a number of machine learning and deep learning models to investigate the problem of identifying fabricated news. Gaussian Naive Bayes, Decision Tree Classifier, Random Forest Classifier, XGBoost, LightGBM, Neural Network, LSTM, and CNN are some of the models employed in this study. Here, we describe in depth the algorithms and architectures that underpin each model, as well as any adjustments or adaptations made to better suit the goal of detecting false news.

Gaussian Naive Bayes

The Gaussian Naive Bayes classifier uses Bayes' theorem to make classification decisions. It makes the Gaussian distribution assumption that characteristics are distributed uniformly and independently. Extracted textual characteristics were analysed using Gaussian Naive Bayes in our study. To further illustrate the textual information, we used tools like TF-IDF (Term Frequency-Inverse Document Frequency). There were no changes made to the false news detecting algorithm.

Decision Tree Classifier

In order to classify data, the Decision Tree Classifier (DTC) uses a flowchart-like structure generated by a non-parametric supervised learning method. The data is partitioned into branches according to their feature values until a certain threshold is reached. We analysed the linguistic and structural characteristics of news stories using the Decision Tree Classifier. In order to get the most out of the model, we tweaked a number of its hyperparameters, including its maximum depth and minimum sample split.

Random Forest Classifier

An ensemble learning technique, the Random Forest Classifier uses a forest of decision trees to create forecasts. The final prediction is reached by averaging the predictions of separate decision trees, each of which is trained on a random portion of the training data. To enhance the model's generalizability and capture intricate dependencies between features, we used the Random Forest Classifier. We adjusted hyperparameters like maximum feature counts per split and tree depth.

XGBoost

XGBoost is a gradient-boosting framework that makes use of a collection of rather weak prediction models. It uses gradient boosting to continuously enhance the model's performance by including fresh models that rectify the mistakes of earlier iterations. Because of its efficacy and efficiency, XGBoost has seen widespread usage in a variety of machine-learning tasks. XGBoost was used to process the textual and structural aspects of news articles, with hyperparameters including learning rate, maximum depth, and boosting rounds optimised.

LightGBM

Another gradient-boosting framework with a focus on speed and effectiveness is LightGBM. The innovative "gradient-based one-side sampling" method is used to rapidly complete training with minimal data storage requirements. LightGBM excels at dealing with massive datasets. The textual and structural aspects of news items were analysed using LightGBM, with hyperparameters like learning rate, maximum depth, and the number of leaves tuned.

Neural Network

Artificial neurons are the building blocks of Neural Network models, which are themselves deep learning structures comprised of linked layers. In this research, we used a feedforward Neural Network to examine the linguistic characteristics of news items. There was an input layer, many activations function-equipped hidden layers, and an output layer. To prevent overfitting, we used methods like dropout regularisation. The learning rate, the number of neurons in each layer, and the total number of hidden layers were all optimised as hyperparameters.

LSTM

To effectively capture long-term dependencies in sequential data, one form of recurrent neural network is the Long Short-Term Memory (LSTM). We used LSTM to examine temporal patterns and linguistic interdependence in news stories for use in our fake news identification system. Memory cells, input/output gates, and forget gates make up the LSTM architecture, which the model uses to remember important details across lengthy sequences. Using word embeddings as a basis for training, we tweaked hyperparameters such as the number of LSTM layers and their spacing.

CNN

Convolutional Neural Networks (CNNs) have revolutionized the field of computer vision by effectively extracting features from images. However, when it comes to analyzing sequential data, such as text, capturing long-term dependencies and linguistic interdependence becomes crucial. In order to address this challenge, researchers have explored the integration of LSTM (Long Short-Term Memory) networks with CNNs to create powerful models for processing sequential data.

Result

Using accuracy as the criterion, the performance of machine learning and deep learning models for false news identification was compared. Accuracy is the rate at which examples are accurately categorised relative to the total number of occurrences. The outcomes of both model types are shown in the following:

Machine Learning Models:

1. Naive Bayes: The Naive Bayes model achieved an accuracy of 81.99% in classifying fake news articles.
2. Decision Tree: The Decision Tree model achieved an accuracy of 78.14% in classifying fake news articles.
3. Random Forest: The Random Forest model outperformed other machine learning models with an accuracy of 88.18% in classifying fake news articles.
4. XGBoost: The XGBoost model achieved the highest accuracy among the machine learning models, reaching 93.96% in classifying fake news articles.
5. LightGBM: The LightGBM model also performed well with an accuracy of 92.92% in classifying fake news articles.

Deep Learning Models:

1. LSTM: The LSTM model achieved an accuracy of 93.59% in classifying fake news articles.
2. CNN: The CNN model achieved an accuracy of 92.29% in classifying fake news articles.

Journal of Informatics Education and Research
ISSN: 1526-4726
https://doi.org/10.52783/jier.v3i2.207
Vol 3 Issue 2 (2023)

Results show that deep learning models, in particular LSTM and CNN, are more accurate at detecting false news than machine learning models. When compared to other machine learning models, the accuracy of the LSTM and CNN models was the highest.

Among the machine learning models, XGBoost performed best with an accuracy of 93.96%, followed by LSTM at 93.59%. Both the Random Forest and LightGBM models performed admirably, with respective accuracy rates of 88.18% and 92.92%. Lower accuracy was reached by the Naive Bayes and Decision Tree models, with 81.99% and 78.14%, respectively.

Model	Accuracy
Naive Bayes	81.99
Decision Tree	78.14
Random Forest	88.18
XGBoost	93.96
LightGBM	92.92
LSTM	93.59
CNN	92.29

Table results show that LSTM and XGBoost models performed best in terms of accuracy. XGBoost is an example of a machine learning model, whereas LSTM is an example of a deep learning strategy. Compared to more conventional machine learning methods, deep learning models appear to do better to detect bogus news.

Considerations including dataset features, computing resources, and application needs all play a role in determining the best model to use. By better detecting bogus news, this study demonstrates how deep learning models can grasp intricate connections and patterns in textual data.

Discussion

The findings from the assessment of machine learning and deep learning models for fake news detection provide light on the effectiveness of various strategies for tackling this challenging issue. We evaluate the effectiveness of the models, highlight their strengths and weaknesses, and suggest some real-world applications in this debate.

It is first important to highlight how deep learning models (in particular LSTM and CNN) outperform their machine learning counterparts. Both LSTM and CNN had the highest classification accuracy compared to the other machine learning models tested. In other words, the complex correlations and patterns contained in textual data may be captured by deep learning models, making them ideal for the accurate identification of fake news. CNN and LSTM are useful for detecting fake news because of their ability to examine sequential and structural information.

Among the machine learning models, XGBoost demonstrated the maximum accuracy with 93.96 percent, closely trailing the performance of the deep learning models. XGBoost is an ensemble model that makes accurate predictions by combining feeble prediction models, such as decision trees. It employs gradient boosting to iteratively better the model's performance, handling the intricate data relationships effectively. The impressive performance of XGBoost suggests that ensemble methods can effectively capture the intricate patterns and interactions within the feature space, thereby enhancing the detection of false news.

Both the Random Forest and LightGBM models demonstrated competitive performance, with 88.18 percent and 92.9 percent accuracy, respectively. Random Forest is an ensemble model that combines the predictions of multiple decision trees, whereas LightGBM is a gradient-boosting framework that prioritises efficiency and speed. Both models demonstrate the efficacy of ensemble learning and gradient boosting in detecting false news. These models can capture intricate feature interactions and generalise effectively to unseen data.

Journal of Informatics Education and Research
ISSN: 1526-4726
https://doi.org/10.52783/jier.v3i2.207
Vol 3 Issue 2 (2023)

In contrast, the accuracy of the Naive Bayes and Decision Tree models was 81.99% and 78.14%, respectively. Naive Bayes is a simple probabilistic classifier based on the assumption of feature independence, which may not be suitable for capturing the complex relationships prevalent in false news articles. Decision Trees are non-parametric models that generate flowchart-like structures based on feature values; however, they are susceptible to overfitting and have limited ability to capture complicated interactions.

Deep learning models are superior at detecting fake news due to their ability to acquire sophisticated representations from data without extensive feature engineering. With its memory cells and gates, LSTM can effectively identify textual data's long-term dependencies and linguistic patterns. CNN, on the other hand, can extract textual structural information and local dependencies. These attributes make deep learning models potent instruments for analysing unstructured data and extracting meaningful features.

It is essential to note, however, that the efficacy of the models can be affected by a number of factors. The efficacy of the models can be affected by the quality and extent of the dataset, the selection and engineering of features, and the choice of hyperparameters. It is essential to evaluate and compare various models across multiple datasets in order to reach more reliable conclusions.

The practical implications of this study's findings for the development of trustworthy false news detection systems are substantial. The high accuracy attained by deep learning models, specifically LSTM and CNN, suggests that these models can be utilised to develop robust and efficient false news detection systems. The capacity of these models to recognise intricate patterns and interdependencies in textual and structural data can contribute to the overall accuracy of such systems.

Conclusion

In conclusion, the evaluation of machine learning and deep learning models for fake news detection revealed the superiority of deep learning models, particularly LSTM and CNN, in accurately identifying false news articles. The results demonstrated that deep learning models outperformed traditional machine learning models in terms of accuracy. LSTM excelled in capturing long-term dependencies and language patterns, while CNN effectively captured structural information and local relationships within the text.

Among the machine learning models, XGBoost performed exceptionally well, closely trailing the performance of the deep learning models. The ensemble approach of XGBoost, combining weak prediction models, showcased its ability to capture intricate data relationships. The competitive performances of Random Forest and LightGBM, both employing ensemble and gradient boosting techniques, further emphasized the effectiveness of these approaches in detecting false news.

However, the Naive Bayes and Decision Tree models achieved comparatively lower accuracy rates. Naive Bayes, relying on the assumption of feature independence, may not capture the complex relationships prevalent in false news articles. Decision Trees, while generating flowchart-like structures, have limitations in capturing complicated interactions and are prone to overfitting.

The findings underscore the advantages of deep learning models in detecting false news, as they can extract sophisticated representations from data without extensive feature engineering. LSTM's ability to capture long-term dependencies and linguistic patterns, along with CNN's capacity to extract structural information and local dependencies, make deep learning models powerful tools for analyzing unstructured data and extracting meaningful features.

It is important to consider that the efficacy of the models can be influenced by factors such as dataset quality, feature selection, and hyperparameter tuning. Further research and evaluation across multiple datasets are necessary to establish more robust conclusions.

The practical implications of this study are significant for the development of trustworthy false news detection systems. The high accuracy achieved by deep learning models, particularly LSTM and CNN, suggests their potential for building robust and efficient detection systems. Their ability to comprehend intricate patterns and interdependencies in textual and structural data contributes to the overall accuracy of such systems, aiding in the mitigation of the fake news challenge.

References

[1] Lazer, D. M., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., ... & Zittrain, J. L. (2018). The science of fake news. Science, 359(6380), 1094-1096.

[2] Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake news detection on social media: A data mining perspective. ACM SIGKDD explorations newsletter, 19(1), 22-36.

[3] Reis, J. C., Correia, A., Murai, F., Veloso, A., & Benevenuto, F. (2019). Supervised learning for fake news detection. IEEE Intelligent Systems, 34(2), 76-81.

[4] Gayo-Avello, D., Metaxas, P. T., Mustafaraj, E., Strohmaier, M., Schoen, H., Gloor, P., ... & Poblete, B. (2013). Predicting information credibility in time-sensitive social media. Internet Research.

[5] Jwa, H., Oh, D., Park, K., Kang, J. M., & Lim, H. (2019). exBAKE: automatic fake news detection model based on bidirectional encoder representations from transformers (Bert). Applied Sciences, 9(19), 4062.

[6] Goldani, M. H., Momtazi, S., & Safabakhsh, R. (2021). Detecting fake news with capsule neural networks. Applied Soft Computing, 101, 106991.

[7] Hakak, S., Alazab, M., Khan, S., Gadekallu, T. R., Maddikunta, P. K. R., & Khan, W. Z. (2021). An ensemble machine learning approach through effective feature extraction to classify fake news. Future Generation Computer Systems, 117, 47-58.

[8] Kula, S., Choraś, M., Kozik, R., Ksieniewicz, P., & Woźniak, M. (2020, June). Sentiment analysis for fake news detection by means of neural networks. In International Conference on Computational Science (pp. 653-666). Springer, Cham.

[9] Ahmed, H., Traore, I., & Saad, S. (2018). Detecting opinion spam and fake news using text classification. Security and Privacy, 1(1), e9.

[10] Akbik, A., Bergmann, T., Blythe, D., Rasul, K., Schweter, S., & Vollgraf, R. (2019, June). FLAIR: An easy-to-use framework for state-of-the-art NLP. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics (Demonstrations) (pp. 54-59).

[11] Wu, Y., Schuster, M., Chen, Z., Le, Q. V., Norouzi, M., Macherey, W., ... & Dean, J. (2016). Google's neural machine translation system: Bridging the gap between human and machine translation. arXiv preprint arXiv:1609.08144.

[12] Graves, A., & Schmidhuber, J. (2005). Framewise phoneme classification with bidirectional LSTM and other neural network architectures. Neural networks, 18(5-6), 602-610.

[13] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805.

[14] Gundapu, S., & Mamid, R. (2021). Transformer-based Automatic COVID-19 Fake News Detection System. arXiv preprint arXiv:2101.00180.

[15] Kaliyar, R. K., & Singh, N. (2019, July). Misinformation Detection on Online Social Media-A Survey. In 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT) (pp. 1-6). IEEE.