International Journal of Computer Science & Communication (ISSN: 0973-7391)
Volume 11 • Issue 2 pp. 56-63 April 2020 - Sept 2020 www.csjournals.com

Hindi Fake News Detection Using Machine Learning

Shubhnandan Singh Jamwal
PG Department of Computer Science and IT, University of Jammu, J&K

Abstract: The social media and online platforms are very important sources for the proliferation of the fake news which can spread rapidly and reach a wide audience within a short period. The viral nature of fake news can amplify its impact and influence public opinion, beliefs, and behaviours. The main objective of the paper is to study the performance of Linear Regression (LR), Support Vector Machine (SVM) and Naive Bayes in detection the fake and similar news with the highest accuracy for news in Hindi. The data has been taken from online resources and divided into training and test data set which are inputted into the different models. The testing of the all the three models is executed hundred times for each one. The accuracy of the LR, SVM, and Naive Bayes is measured as 0.917, 0.916 and 0.90 respectively.

Keywords: Fake news, Linear Regression, Naive Bayes, Support Vector Machine, Machine Learning.

Introduction

Fake news refers to false or misleading information which is presented as a fact. It can take various forms, including fabricated stories, hoaxes, misinformation, and disinformation, and is often spread through traditional media outlets, social media platforms, websites, and other online channels. Fake news articles are intentionally deceptive and these articles are not only limited to politics and election but also results in severe injury and death by triggering certain actions. Hence it becomes imperative to classify and identifying fake news article and alerting human readers.

Social media is low cost, easy to use and help in spreading information rapidly. This enables people to consume and spread news whether it is genuine news or fake news. With the proliferation of social media and online platforms, fake news can spread rapidly and reach a wide audience within a short period. The viral nature of fake news can amplify its impact and influence public opinion, beliefs, and behaviours. Nowadays many people use social media to spread rumours, low quality news with intentionally fake or wrong information. A. Bedi and et. al., [1] have measured fake content from 567 fake news websites and approximate 9,500 fake stories on Twitter and Facebook and also proposed a model how we can identify and secure the issues of fake news in social networks. It is not only the fake news, which is posing challenges but the spammers on the social media is also a threat. It has been observed [2] that vast usage of social media makes it a familiar platform for malicious users referred as social spammers to overwhelm usual users with unwanted content. Different ways for detection of social spammer is done by constructing a classifier based on social network and content information. However social spammers are adaptable and sophisticated to game the system with rapidly developing network and content patterns. The rigid anti spam norms have resulted in development of spammers. They look alike legal users who are difficult to recognize. A spammer classification method based on LDA (Latent Dirichlet Allocation) a topic model was also proposed. Fake news typically lacks credible sources, evidence, or factual basis to support its claims. It may contain exaggerated, distorted, or entirely fabricated information designed to mislead readers or viewers.

Fake News Objectives:

Fake news is deliberately created and disseminated with the intent to deceive or manipulate the audience. It is often generated for political, ideological, financial, or sensational purposes. Fake news often employs emotional language, sensational headlines, and provocative imagery to evoke strong emotional reactions from the audience. It aims to elicit fear, outrage, or other intense emotions to garner attention and engagement. Fake news preys on individuals' confirmation bias, the tendency to interpret information in a way that confirms one's existing beliefs or biases. People may be more likely to accept and share fake news that aligns with their preconceived notions or ideological views. The proliferation of fake news erodes public trust in traditional media sources and institutions, undermining the credibility of legitimate journalism and weakening democratic processes. Combating fake news requires critical thinking skills, media literacy education, fact-checking initiatives, and responsible journalism practices. It also involves efforts from technology companies, policymakers, educators, and civil society to address the root causes of misinformation and promote the dissemination of accurate, reliable information in the digital age.

The objective of the paper is to use and analyse the different machine learning models for the identification of the fake and similar news. The machine learning models can then be employed and used to prepare a system for the detection of effective fake news.

Literature Review

Yavary et. al.,[3] considered two main sources for information verification in social networks that include user feedback and news agencies. User feedbacks as the first source can be user conversational tree and some patterns are extracted from this tree. News agencies as the second source are also utilized for verification of information by textual entailment methods. Finally, these two types of features are aggregated to classify the information in one of the three classes of true, false, or unverified. This method was also tested through the experiments with public datasets. The results of experiments show that the hybrid suggested method for information verification could pass the state-of-the-art methods in information verification. N. Snell and et. al., [4] presented dataset which is manually identified and classified containing news stories that can be used for the training and testing of classification systems that identify legitimate versus fake and manipulative news stories. K. Shu and et. al.,[5] constructed real-world datasets measuring users trust level on fake news and select representative groups of both “experienced” users who are able to recognize fake news items as false and “naïve” users who are more likely to believe fake news. They performed a comparative analysis over explicit and implicit profile features between these user groups, which reveal their potential to differentiate fake news. J. Lin and et. al., [6] proposed a framework which extracts 134 features and builds traditional known machine learning models like Random Forest and XGBoost. They also proposed a deep learning based model (LSTM with self-attention mechanism) to see which one performs better in the fake news article detection in both political news and celebrity news domains. The experimental results show that XGBoost model improved 16.4% and 13.1% over the best baseline in terms of accuracy in both political news articles and celebrity news articles. E. Qawasmeh et. al., [7] investigated the automatic identification of fake news over online communication platforms and proposed an automatic identification of fake news using modern machine learning techniques. The proposed model is a bidirectional LSTM concatenated model that is applied on the FNC-1 dataset with 85.3 % accuracy performance. S. Gaonkar and et. al., [8] proposed a model that classifies unreliable news into real and fake news after computing a score and will be able to distinguish between real and fake news based on various parameters obtained from a Uniform Resource Locator (URL). The proposed model is based on various Machine Learning and Natural Language Processing techniques to achieve maximum accuracy.

Martens et. al., [9] studied fake reviews, their providers, characteristics, and how well they can be automatically detected and also conducted disguised questionnaires with 43 fake review providers and studied their review policies to understand their strategies and offers. By comparing 60,000 fake reviews with 62 million reviews from the Apple App Store they found significant differences, e.g., between the corresponding apps, reviewers, rating distribution, and frequency. This inspired the development of a simple classifier to automatically detect fake reviews in app stores. On a labelled and imbalanced dataset including one-tenth of fake reviews, as reported in other domains, the proposed classifier achieved a recall of 91% and an AUC/ROC value of 98%.

Zhang et.al., [10] proposed a deep learning approach for text representation called DCWord (Deep Context representation by Word vectors) to deceptive review identification. The basic idea is that since deceptive reviews and truthful reviews are composed by writers without and with real experience on using the online purchased goods or services, there should be different contextual information of words between them. Unlike state-of-the-art techniques in seeking best linguistic features for representation, they used word vectors to characterize contextual information of words in deceptive and truthful reviews automatically. The average-pooling strategy (called DCWord-A) and max-pooling strategy (called DCWord-M) are used to produce review vectors from word vectors. Experimental results on the Spam dataset and the Deception dataset demonstrate that the DCWord-M representation with LR (Logistic Regression) produces the best performances and outperforms state-of-the-art techniques on deceptive review identification.

The news readers adopt biased views, when the reporting of the news agencies is biased. The social media is much more responsible for the transmission of the biased news. [11–15]. Fake news content is difficult to identify because the term "fake news" covers intentionally false, deceptive stories as well as factual errors, satire, and sometimes, stories that a person just does not like. Jennifer Golbeck et. al., [16] presented a dataset of fake news and satire stories that are hand coded, and in the case of fake news, include rebutting stories. They also included a thematic content analysis of the articles, identifying major themes that include hyperbolic support or condemnation of a gure, conspiracy theories, racist themes, and discrediting of reliable sources. In addition to releasing dataset for research use, they analyzed it and show results based on language that are promising for classification purposes. S. Mo Jang et. al., [17] retrieved 307,738 tweets about 30 fake and 30 real news stories, and examined the root content, producers of original source, and evolution patterns. The findings revealed that root tweets about fake news were mostly generated by accounts from ordinary users, but they often included a link to non-credible news websites. Additionally, they observed significant differences between real and fake news stories in terms of evolution patterns. Their observation on evolution tree analysis, tweets about real news showed wider breadth and shorter depth than tweets about fake news. The results also indicated that tweets about real news spread widely and quickly, but tweets about fake news underwent a greater number of modifications in content over the spreading process.

Implementation

Different machine learning algorithms i.e., Logistic Regression, SVM and Naive Bayes are implemented and used on the data to analyse their performance for the detection of the fake and similar news. The steps followed in the implementation of the experiments are:

1) Dataset: The dataset which have been used to train the different models contains 2080 fake and 1246 true stories. The 1240 and 840 instances from Kaggle and Github respectively, has been used to collect the fake news, where as 945 and 301 instances has been used from the same sources for the creation of the true news dataset. The 80% data is used for training and 20% is used for testing in all the models.

International Journal of Computer Science & Communication (ISSN: 0973-7391)
Volume 11 • Issue 2 pp. 56-63 April 2020 - Sept 2020 www.csjournals.com

2) Pre-processing: The pre-processing of the data involves number of steps. Natural Language Toolkit (NLTK) library of python is used for the tasks of pre-processing. All the punctuations, numerals, blank spaces and stop words (table 1) are removed from the text.

मैं	मुझको	मेरा	अपने	हमने	हमारा	अपना	हम	आप	आपका
तुम्हारा	अपने	स्वयं	वह	इसे	उसके	खुद	कि वह	उसकी	उसका
खुद ही	यह	इसके	उन्होने	अपने	क्या	जो	किसे	किसको	कि
ये	हूँ	होता है	रहे	थी	थे	होना	गया	को	किया
है	पडा	होने	करना	करता है	किया	रही	एक	लेकिन	अगर
या	क्यूंकि	जैसा	जब तक	जबकि	की	पर	द्वारा	के लिए	साथ
के बारे	खिलाफ	बीच	में	माध्यम	दौरान	से पहले	के बाद	ऊपर	नीचे
को	से	तक	से नीचे	करने में	निकल	बंद	से	तहत	दुबारा
आगे	फिर	एक	यहाँ	वहाँ	कब	कहाँ	क्यों	कैसे	सारे
किसी	दोनो	प्रत्येक	ज्यादा	अधिकांश	अन्य	में कुछ	ऐसा	में कोई	मात्र
खुद	समान	इसलिए	बहुत	सकता	जायेंगे	जरा	चाहिए	अभी	और
कर	रखें	का	हैं	इस	होता	करने	ने	बनी	तो
ही	हो	इसका	था	हुआ	वाले	बाद	लिए	सकते	इसमें
दो	वे	करते	कहा	वर्ग	कई	करें	होती	अपनी	उनके
यदि	हुई	जा	कहते	जब	होते	कोई	हुए	व	जैसे
सभी	करता	उनकी	तरह	उस	आदि	इसकी	उनका	इसी	पे
तथा	भी	परंतु	इन	कम	दूर	पूरे	गये	तुम	मै
यहां	हुये	कभी	अथवा	गयी	प्रति	जाता	इन्हें	गई	अब
जिसमें	लिया	बड़ा	जाती	तब	उसे	जाते	लेकर	बड़े	दूसरे

Table 1: Stop Words removed from the data

The words are also converted to their base forms by using WordNetLemmatizer. These lemmatized words are used to classify the documents having similar features.

3) **Feature Extraction:** The lexical and syntactic features are used to classify the text data into features. The average length of words, percentage of characters usage like '?', POS frequency of word in one article, number of words expressing certainty, number of tentative words, number of adjectives and adverbs and count of removed stop words are used as features for training and testing.

TF-IDF is commonly employed in natural language processing and information retrieval like text mining, document classification, information extraction etc. TF-IDF is utilized as a feature extraction technique to help identify important terms or words within documents. Term Frequency is used to measures how many times a term appears in a document. It's calculated as the number of times a term appears in a document divided by the total number of terms in the document. The frequency of the term used in the document shows the context of the similar news. Inverse Document Frequency is calculated as the logarithm of the total number of documents divided by the number of documents containing the term. The TF-IDF score of a term in a document is calculated by multiplying its TF value by its IDF value. Terms with high TF-IDF scores are those that are common within a document but rare across the entire document collection, making them potentially significant in distinguishing the document from others. In the context of fake news detection, TF-IDF can be used in several ways:

Feature Extraction: By computing TF-IDF scores for each term in a document it is converted to vector of TF-IDF scores. This vector can then be used as input to machine learning algorithms for classification or clustering tasks.

Keyword Extraction: TF-IDF helps in identifying keywords or important terms within a document. These keywords can then be used to summarize the content of the document or to identify key features that distinguish between fake and genuine news articles.

Document Similarity: TF-IDF is also used to measure the similarity between documents. By computing the cosine similarity between the TF-IDF vectors of different documents the relationship in terms of their content is compared. Scikit-learn function *TfidfVectorizer* is used for the converting the text data to a matrix which is also known as BoW representation.

4) Training:

In the context of fake news detection, TF-IDF alone may not be sufficient to accurately distinguish between fake and genuine news articles. It is often used in conjunction with other techniques such as natural language processing, sentiment analysis, and machine learning algorithms to build more robust fake news detection systems. Additionally, the effectiveness of TF-IDF depends on the quality of the underlying text data and the specific characteristics of the fake news articles being analyzed.

a) Logistic Regression: Fake news detection is essentially a binary classification problem where the model needs to classify whether a given piece of news is fake or real based on certain features. The Logistic Regression model is trained with the training data set. Scikit-learn is a popular open-source machine learning library for Python which provides simple and efficient tools for data mining and data analysis, built on top of other popular Python libraries like NumPy, SciPy, and matplotlib.

data = pd.read_csv('your_dataset.csv')

```
# Split the dataset into features and labels
X = data['features']
y = data['label']
# Split the data into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2,
random_state=42)
# Create a linear regression model
model = LinearRegression()
# Train the model
model.fit(X_train, y_train)
# Make predictions on the test set
y_pred = model.predict(X_test)
```

d) Naive Bayes Classifier: Naive Bayes classifiers are another popular choice in NLP tasks like fake news detection, text classification etc. due to simplicity, efficiency, and effectiveness with relatively small datasets. Naive Bayes classifiers are known for their computational efficiency and ability to handle large feature spaces, which makes them suitable for text classification tasks like fake news detection. Naive Bayes classifiers are probabilistic classifiers based on Bayes' theorem and assume that features are conditionally independent given the class label. The Scikit-learn library is used for the implementation of Naive Bayes classifier.

```
# Split the dataset into features (text) and labels
X = data['text']
y = data['label']
# Convert text data into numerical features using CountVectorizer
vectorizer = CountVectorizer()
X = vectorizer.fit_transform(X)
# Split the data into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2,
random_state=42)
# Create a Multinomial Naive Bayes classifier
clf = MultinomialNB()
```

International Journal of Computer Science & Communication (ISSN: 0973-7391)
Volume 11 • Issue 2 pp. 56-63 April 2020 - Sept 2020 www.csjournals.com

```
# Train the classifier
clf.fit(X_train, y_train)
# Make predictions on the test set
y_pred = clf.predict(X_test)
# Calculate the accuracy of the classifier
accuracy = accuracy_score(y_test, y_pred)
print("Accuracy:", accuracy)
```

e) Support Vector Machine: Support Vector Machine (SVM) classifiers are commonly used in fake news detection tasks due to their ability to handle high-dimensional feature spaces and their effectiveness in separating data points into different classes through the use of hyperplanes. SVM can be computationally intensive, especially with large datasets, and the choice of kernel function and parameters can significantly impact the performance. Cross-validation is used to optimize the performance of the SVM model by tuning C and sigma parameters. C parameter controls the trade-off between the training error and the margin, while the sigma parameter controls the smoothness of the decision boundary and the influence of individual training examples.

```
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2,
random_state=42)
# Create an SVM classifier
clf = svm.SVC(kernel='linear')
# Train the classifier
clf.fit(X_train, y_train)
# Make predictions on the test set
y_pred = clf.predict(X_test)
# Calculate the accuracy of the classifier
accuracy = accuracy_score(y_test, y_pred)
print("Accuracy:", accuracy)
```

Model Evaluations

The LR, SVM and Naive Bayes machine learning models are implemented and tested in this paper. Multiple evaluation metrics, including the accuracy, precision, F1-score, and recall were adopted to evaluate the performances of standard models. The accuracy is the ratio of the number of samples correctly classified to the total number of samples in a given test dataset.

$$\text{Accuracy} = (TP + TN) / (TP + TN + FP + FN)$$

The precision is the ratio of the true positive samples to the sum of the true positive and false positive samples.

$$\text{Precision} = TP/(TP + FP)$$

The recall indicates to the ratio of the true positive samples to the sum of the true positive and false negative samples. The F1-score value is used to evaluate the success of machine learning algorithms.

$$\text{Recall} = TP/(TP + FN)$$

The F1-score is the weighted average of precision and recall.

$$\text{F1-score} = 2 * \text{Precision} * \text{Recall} / (\text{Precision} + \text{Recall})$$

TP, TN, FP, and FN represent the numbers of true positive, true negative, false positive, and false negative in the confusion matrix, respectively.

Model	Accuracy	Precision	Recall	F1-Score
LR	0.917	0.86	1.00	0.93
SVM	0.916	0.84	1.00	0.91
Naive Bayes	0.90	0.81	1.00	0.90

Table 2: Results of Different Models

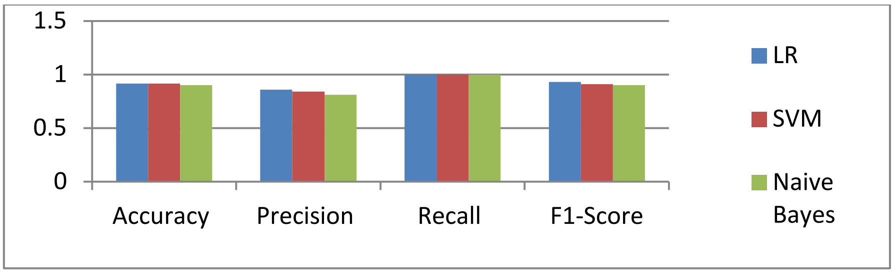

Fig1: Graphical Representation of Table 1

Conclusion and Future Scope

The main objective of the paper is to study the performance of common machine learning models and to identify the optimal model for detecting fake news with the highest accuracy. The entire data was divided into the training and test data set and inputted into the model implementation process. The testing of the all the three models is executed hundred times for each one. The accuracy of the LR, SVM, and Naive Bayes is measured as 0.917, 0.916 and 0.90 respectively. The precision of the LR, SVM, and Naive Bayes is measured as 0.86, 0.84 and 0.81 respectively. The Recall in all the cases is measured as 1. The F1-score of the LR, SVM, and Naive Bayes is measured as 0.93, 0.91 and 0.90 respectively. Logistic regression is a good baseline model and is often used in conjunction with other techniques in ensemble methods. SVMs become slow and inefficient when dealing with very large datasets. Naive Bayes classifiers remain popular due to their simplicity, scalability, and ability to handle high-dimensional data efficiently and are often used as baseline models and can provide good performance in many real-world classification tasks, especially when the independence assumption holds reasonably well.

References

[1] A. Bedi, N. Pandey and S. K. Khatri, "A Framework to Identify and secure the Issues of Fake News and Rumours in Social Networking," 2019 2nd International Conference on Power Energy, Environment and Intelligent Control (PEEIC), Greater Noida, India, 2019, pp. 70-73, doi: 10.1109/PEEIC47157.2019.8976800.

[2] Jose, T., Babu, S.S. Detecting spammers on social network through clustering technique. J Ambient Intell Human Comput (2019). https://doi.org/10.1007/s12652-019-01541-6

[3] Yavary, A., Sajedi, H. & Saniee Abadeh, M. Information verification in social networks based on user feedback and news agencies. Soc. Netw. Anal. Min. 10, 2 (2020). https://doi.org/10.1007/s13278-019-0616-4

[4] N. Snell, W. Fleck, T. Traylor and J. Straub, "Manually Classified Real and Fake News Articles," 2019 International Conference on Computational Science and Computational Intelligence (CSCI), Las Vegas, NV, USA, 2019, pp. 1405-1407, doi: 10.1109/CSCI49370.2019.00262.

[5] K. Shu, S. Wang and H. Liu, "Understanding User Profiles on Social Media for Fake News Detection," 2018 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR), Miami, FL, USA, 2018, pp. 430-435, doi: 10.1109/MIPR.2018.00092.

[6] J. Lin, G. Tremblay-Taylor, G. Mou, D. You and K. Lee, "Detecting Fake News Articles," 2019 IEEE International Conference on Big Data (Big Data), Los Angeles, CA, USA, 2019, pp. 3021-3025, doi: 10.1109/BigData47090.2019.9005980.

[7] E. Qawasmeh, M. Tawalbeh and M. Abdullah, "Automatic Identification of Fake News Using Deep Learning," 2019 Sixth International Conference on Social Networks Analysis, Management and Security (SNAMS), Granada, Spain, 2019, pp. 383-388, doi: 10.1109/SNAMS.2019.8931873.

[8] S. Gaonkar, S. Itagi, R. Chalippatt, A. Gaonkar, S. Aswale and P. Shetgaonkar, "Detection Of Online Fake News : A Survey," 2019 International Conference on Vision Towards Emerging Trends

in Communication and Networking (ViTECoN), Vellore, India, 2019, pp. 1-6, doi: 10.1109/ViTECoN.2019.8899556.
[9] Martens, D., Maalej, W. Towards understanding and detecting fake reviews in app stores. Empir Software Eng 24, 3316–3355 (2019). https://doi.org/10.1007/s10664-019-09706-9
[10] Zhang, W., Wang, Q., Li, X. et al. DCWord: A Novel Deep Learning Approach to Deceptive Review Identification by Word Vectors. J. Syst. Sci. Syst. Eng. 28, 731–746 (2019). https://doi.org/10.1007/s11518-019-5438-4
[11] Groseclose, T., Milyo, J.: A measure of media bias. Q. J. Econ. 120, 1191–1237 (2005)
[12] Sunstein, C.R.: Echo Chambers: Bush v. Gore, Impeachment, and Beyond. Princeton University Press, Princeton (2001)
[13] Frey, D.: Recent research on selective exposure to information. Adv. Exp. Soc. Psychol. 19, 41–80 (1986)
[14] Mutz, D.C.: Facilitating communication across lines of political difference: the role of mass media. Am. Polit. Sci. Assoc. 95(01), 97–114 (2001)
[15] Mullainathan, S., Shleifer, A.: The market for news. Am. Econ. Rev. 95, 1031–1053 (2005)
[16] Jennifer Golbeck, Matthew Mauriello, Brooke Auxier, Keval H. Bhanushali, Christopher Bonk, Mohamed Amine Bouzaghrane, Cody Buntain, Riya Chanduka, Paul Cheakalos, Jennine B. Everett, Waleed Falak, Carl Gieringer, Jack Graney, Kelly M. Hoffman, Lindsay Huth, Zhenya Ma, Mayanka Jha, Misbah Khan, Varsha Kori, Elo Lewis, George Mirano, William T. Mohn IV, Sean Mussenden, Tammie M. Nelson, Sean Mcwillie, Akshat Pant, Priya Shetye, Rusha Shrestha, Alexandra Steinheimer, Aditya Subramanian, and Gina Visnansky. 2018. Fake News vs Satire: A Dataset and Analysis. In Proceedings of the 10th ACM Conference on Web Science (WebSci '18). Association for Computing Machinery, New York, NY, USA, 17–21. https://doi.org/10.1145/3201064.3201100
[17] S. Mo Jang, Tieming Geng, Jo-Yun Queenie Li, Ruofan Xia, Chin-Tser Huang, Hwalbin Kim, Jijun Tang, A computational approach for examining the roots and spreading patterns of fake news: Evolution tree analysis, Computers in Human Behavior, Volume 84, 2018, Pages 103-113, ISSN 0747-5632, https://doi.org/10.1016/j.chb.2018.02.032.