Tipo de Artigo: Original

FAKE NEWS DETECTION USING DECISION TREE, SUPPORT VECTOR MACHINE AND K-NEAREST NEIGHBORS ALGORITHMS

DETECÇÃO DE FAKE NEWS USANDO OS ALGORITMOS DECISION TREE, SUPPORT VECTOR MACHINE E K-NEAREST NEIGHBORS

Recebido em: 07/03/2024

Aprovado em: 25/05/2024

Lucas Monteiro Bastos (Orcid: https://orcid.org/0009-0005-1362-209X)
Especialista em Ciência de Dados. Centro Universitário UNDB. São Luís, Maranhão, Brasil.

Allan Kássio Beckman Soares da Cruz (Orcid: https://orcid.org/0000-0002-2631-2032)
Doutorando em Ciência da Computação. Programa de Pós-Graduação Doutorado em Ciência da Computação - Associação UFMA-UFPI. São Luís, Maranhão, Brasil.

Pamela Torres Maia Beckman da Cruz (Orcid: https://orcid.org/0000-0002-9147-6682)
Doutoranda em Ciência da Informação. Faculdade de Letras – Universidade de Coimbra. Coimbra, Portugal.

Mario Meireles Teixeira (Orcid: https://orcid.org/0000-0001-8771-1478)
Doutor em Ciência da Computação. Programa de Pós-Graduação Doutorado em Ciência da Computação - Associação UFMA-UFPI. São Luís, Maranhão, Brasil.

Carlos de Salles Soares Neto (Orcid: https://orcid.org/0000-0002-6800-1881)
Doutor em Ciência da Computação. Programa de Pós-Graduação Doutorado em Ciência da Computação - Associação UFMA-UFPI. São Luís, Maranhão, Brasil.

DOI: https://doi.org/10.5281/zenodo.14057842

Autor para correspondência:

Allan Kássio Beckman Soares da Cruz (Orcid: https://orcid.org/0000-0002-2631-2032)
Doutorando em Ciência da Computação. Programa de Pós-Graduação Doutorado em Ciência da Computação - Associação UFMA-UFPI. São Luís , Maranhão , Brasil.

ABSTRACT
Fake news represents misinformation, typically spread across social networks, and possesses a significant potential for widespread dissemination online, which can lead to substantial societal issues. Therefore, it becomes imperative to explore and develop strategies aimed at minimizing such impacts, including the detection of fake news through the employment of machine learning (ML) techniques and algorithms. The purpose of this study is to examine the

effectiveness and application of ML algorithms in identifying fake news. This research adopts an applied approach, focusing on a descriptive and quantitative analysis. The data for this study were sourced from the Kaggle platform, with data extraction conducted using Python, and analysis performed on the Jupyter Notebook platform.

Keywords: Fake News. Detection. Machine Learning.

RESUMO
Notícias falsas representam desinformação, tipicamente disseminadas pelas redes sociais, e possuem um potencial significativo para disseminação em larga escala online, o que pode levar a problemas substanciais na sociedade. Portanto, torna-se imperativo explorar e desenvolver estratégias voltadas à minimização de tais impactos, incluindo a detecção de notícias falsas por meio do uso de técnicas e algoritmos de aprendizado de máquina (ML). O objetivo deste estudo é examinar a eficácia e aplicação de algoritmos de ML na identificação de notícias falsas. Esta pesquisa adota uma abordagem aplicada, focando em uma análise descritiva e quantitativa. Os dados para este estudo foram obtidos da plataforma Kaggle, com a extração de dados realizada usando Python, e a análise executada na plataforma Jupyter Notebook.

Palavras-chave: Notícias Falsas. Detecção. Aprendizado de Máquina.

1 INTRODUCTION

The harmful effects that news can have on societal risks are profound. The dissemination of adulterated information, widely known as Fake News, can adversely affect many aspects of society, including politics, security, and health domains. Hence, this document will introduce practices aimed at determining the authenticity of shared information. Every kind of false information, no matter how simple or absurd, holds significance, as such misinformation can mislead people into error. This is because the news might mix false information with truths (Delmazo and Valente, 2020).

This backdrop prompts the question: Can the truthfulness of news be verified using an algorithm that employs machine learning techniques? From this inquiry, three hypotheses were developed for the study, suggesting that: Machine Learning algorithms might predict outcomes on specific subjects with pertinent themes; developing an algorithm could achieve a level of accuracy in determining the veracity of facts; and machine learning techniques could help in identifying truthful information, thereby reducing Fake News instances.

This exploration was inspired by an academic endeavor that involved developing a practical project for a Web Programming course focused on detecting Fake News through a Machine Learning algorithm. This project piqued interest in further understanding the technology, as well as in its analysis and application in personal research, such as a thesis project.

The social and professional relevance of this research is evident. Finding potential solutions for fact verification and distinguishing between real news and fake news is aimed at helping society at large. News with significant relevance, be it in economic, political, or social spheres, can have substantial or irreversible effects on a specific social group.

The study conducted is of an applied nature and follows a descriptive and quantitative theoretical model. Data was collected from the Kaggle platform, then extracted using Machine Learning algorithms (Albon, 2018), and analyzed with the Python programming language (Rising and Odegua, 2017). The extracted data were processed and analyzed via the interactive Jupyter Notebook platform.

The importance of this application could bring remarkable benefits. It was discovered that accurately detecting such information has become a complex and challenging issue for several reasons. The spread of Fake News is often unintentional, aimed at sharing information to warn family members or counter someone's argument (Carvalho and Mateus, 2018).

2 THEORETICAL REFERENCE

In this section, the theoretical foundations of the research will be highlighted. It addresses concepts and approaches related to Fake News, libraries for data analysis such as Pandas, NumPy, Matplotlib, Seaborn, and Scikit-learn, the Python programming language, and Machine Learning. Additionally, it discusses algorithm techniques like Decision Trees, Support Vector Machines (SVM), and K-Nearest Neighbors (KNN).

2.1. Types of Machine Learning

The most effectively utilized ML algorithms are those that can automate decision-making processes based on known patterns. In supervised learning, for instance, the user supplies the algorithm with pairs of input and desired output, and the algorithm devises a method to produce the desired output for a new input. Importantly, the algorithm can generate an output for an input it has never encountered before, without human intervention (Müller & Guido, 2017).

2.2. Supervised Learning

Machine Learning algorithms that learn from pairs of input and output are termed supervised because an "instructor" oversees the algorithms by providing them with the desired outcomes for each example they learn from (Müller & Guido, 2017). This model is commonly used in applications like facial and voice recognition, product or movie recommendations, and sales forecasting. Supervised Learning is divided into two types: Regression and Classification (Duda *et al.*, 2000).

Regression is used to train and predict a continuous value response, for example, forecasting real estate prices. Classification aims to identify the appropriate class label, such as distinguishing between positive and negative sentiments, male and female individuals, benign and malignant tumors, and safe and risky loans (Izbicki & Santos, 2020).

2.3. Unsupervised Learning

Unsupervised Learning is employed to detect anomalies, outliers, fraud, or faulty equipment, or to group customers with similar behaviors for a sales campaign. Unlike Supervised Learning, this model does not utilize labeled data (Duda *et al.*, 2000).

An unsupervised learning algorithm attempts to efficiently classify a dataset into a certain number of groups. These algorithms are crucial tools for data analysis, identifying patterns and trends, and are most commonly used to cluster similar entries into logical groups. Examples include Kmeans, Random Forests, Hierarchical Clustering, and others.

2.4. Decision Tree

Decision Trees are guided by algorithms from the Top-Down Induction of Decision Trees family. A decision tree is a data structure classified as either a leaf node that represents a class or a decision node that poses an estimate about some type of property. Each estimation implication has an edge leading to the development of a subtree, which maintains the same structure as the original tree (Monard & Baranauskas, 2003).

Due to their predictive features and the elucidation of difficulties in a practical and direct manner, decision trees are categorical standards demonstrating utility across various learning domains (Vieira *et al.*, 2018).

2.5. Support Vector Machine (SVM)

According to Coutinho (2019), SVM is a widely used supervised learning algorithm in situations requiring data classification into distinct sets, but it also applies to regression. Its aim is to find a line that separates two different classes, examining the positions of specific groups, especially those close to the other class or subsequent neighbors (Coutinho, 2019).

The SVM algorithm is noted for its precision in classification, outperforming Logistic Regression and Decision Trees in this regard. Its applications span various sectors, including facial recognition, intrusion detection, email spam recognition, and the classification of news and web pages (Navlani, 2019).

2.6. K-Nearest Neighbors (KNN)

KNN is a supervised classification algorithm that requires a large number of labeled samples to predict the class of others. To label a new sample, the algorithm relies on the closest labeled ones, which are its k nearest neighbors. The "k" in K-Nearest Neighbors represents the number of neighbors observed and checked for their classification (Beyer *et al.*, 1999).

KNN was one of the first algorithms used and is considered one of the simplest supervised classification methods, which can also be applied to multivariate regression (Fosseng, 2013).

2.7. Confusion Matrix

This tool is commonly used to evaluate classification models in Machine Learning. A confusion matrix provides a visual layout with pertinent information about the model's accuracy (Stehman, 1997). It is a 2x2 matrix where each cell reveals different information about the model's errors and accuracies. Its interpretation is described as follows:

- True Positive (TP): Counts how often the model predicted the condition as positive, and it was indeed positive.
- False Positive (FP): Shows how often the model predicted the condition as negative, but it was actually positive.

- True Negative (TN): Indicates how often the model predicted the condition as negative, and it was genuinely negative.
- False Negative (FN): Represents how often the model predicted the condition as negative, however, it was positive.

2.8. Evaluation Metrics

Various metrics can be employed to determine whether a program is learning to perform its task effectively. For Supervised Learning problems, performance metrics measure the number of prediction errors (Duda et al., 2000; Monard and Baranauskas, 2020). To evaluate models more accurately, certain evaluation metrics such as accuracy, sensitivity, specificity, and precision are used (Ferrari & Silva, 2017).

Accuracy is one of the most intuitive and straightforward metrics for evaluating the performance of a classification model. It calculates the proportion of total correct predictions (both true positives and true negatives) to the total number of predictions made by the model. The formula is:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{FN} + \text{TN}}$$

High accuracy indicates that the model is effective across both classes but doesn't specify how well the model distinguishes between those classes. Therefore, while accuracy is a useful general indicator, it might be misleading in cases of unbalanced class distributions, where one class significantly outnumbers the other.

Sensitivity, also known as recall, measures the model's ability to correctly identify positive outcomes from the actual positives available within the data. In medical testing, for example, sensitivity would refer to the test's ability to correctly identify patients with a disease. The formula is:

$$\text{Sensitivity : } \frac{\text{TP}}{\text{TP} + \text{FP}}$$

High sensitivity is crucial in scenarios where missing out on true positive cases has serious consequences, such as in disease screening where failing to identify a condition could have dire health implications.

Specificity assesses the model's ability to accurately identify negative outcomes from the actual negatives present in the dataset. It's the complement to sensitivity, focusing on the true negative rate. For a medical test, specificity measures how well the test identifies those without the disease when they are truly disease-free. The formula is:

$$\text{Specificity: } \frac{\text{TN}}{\text{TN} + \text{FN}}$$

High specificity is essential in contexts where falsely identifying negative instances as positive could lead to unnecessary interventions, costs, or anxiety, such as in the screening for conditions that require invasive follow-up testing.

Precision focuses on the proportion of positive identifications that were actually correct. It is especially important in situations where the cost of a false positive is high. For instance, in email filtering, precision measures the proportion of emails correctly identified as spam against all emails flagged as spam, whether correctly or not. The formula is:

$$\text{Precision : } \frac{\text{TP}}{\text{TP} + \text{FP}}$$

High precision ensures that when the model predicts a positive result, it can be trusted with a high degree of confidence. This is crucial in fields like finance or law, where false positives can have costly implications.

Each of these metrics offers valuable insights into different aspects of a model's performance. When considered together, they provide a comprehensive picture of a model's effectiveness, strengths, and weaknesses in classification tasks, enabling developers to choose or tune models according to the specific needs and priorities of their application.

3 RELATED WORK

The problem of fake news detection has garnered significant attention from the research community, resulting in a variety of approaches leveraging different machine learning techniques. This section reviews recent studies that have explored the use of Decision Tree, Support Vector Machine (SVM), and K-Nearest Neighbors (KNN) algorithms in the context of fake news detection.

A recent study by Ahmed *et al.* (2020) utilized an ensemble of machine learning techniques to detect fake news. Their approach combined Decision Tree, SVM, and KNN with other algorithms such as Naive Bayes and Random Forest, achieving an accuracy of over 90% on the Fake News Detection Dataset from Kaggle. Their findings indicate that ensemble methods often outperform individual classifiers by leveraging their complementary strengths.

In another study, Shu *et al.* (2020) investigated the use of SVM in fake news detection. They utilized a dataset containing news articles from multiple domains and reported that SVM performed significantly well in classifying fake and real news, achieving an accuracy of 93%. The study highlighted the importance of feature selection and text preprocessing in enhancing the performance of the SVM model.

A comparative analysis conducted by Pérez-Rosas *et al.* (2018) evaluated the effectiveness of various machine learning algorithms, including Decision Tree and KNN, for fake news detection. Their experiments demonstrated that while Decision Tree algorithms provided a good balance between precision and recall, KNN exhibited superior performance in scenarios with balanced datasets. The study emphasized the need for careful algorithm selection based on the characteristics of the dataset.

Prachi *et al.* (2022) focused on the application of Decision Tree algorithms in fake news detection. Their research utilized a large dataset of news articles and employed advanced text feature extraction techniques. The Decision Tree model achieved an accuracy of 88%, showcasing its capability to handle complex decision-making processes inherent in fake news detection.

Furthermore, a study by Madani *et al.* (2024) applied KNN for fake news detection, integrating it with natural language processing techniques to improve feature extraction. Their results showed that KNN could achieve competitive accuracy levels, particularly when combined with feature selection methods that reduce the dimensionality of the data.

Overall, these studies underscore the potential of using machine learning algorithms, including Decision Tree, SVM, and KNN, for effective fake news detection. They also highlight the critical role of data preprocessing, feature selection, and the combination of multiple algorithms in enhancing detection performance.

4 RESULTS

The Decision Tree model shows high performance across all metrics, indicating a strong ability to classify both positive and negative outcomes accurately. The high specificity suggests excellent performance in identifying true negatives, which is crucial for applications where false alarms are particularly undesirable as we can observe in Table 1.

Table 1 – Decision Tree.

Metric	Value (%)
Accuracy	98
Precision	98.5
Sensitivity	99
Specificity	99.4

Font: Created by the authors.

The SVM model outperforms the Decision Tree in accuracy and sensitivity, showing its strength in correctly identifying true positives. This model's high performance underscores its effectiveness in distinguishing between classes with a high degree of accuracy, making it suitable for tasks requiring precise classification as we can observe in Table 2.

Table 2 – Support Vector Machine (SVM).

Metric	Value (%)
Accuracy	99.5
Precision	98.5
Sensitivity	99.4
Specificity	99.3

Font: Created by the authors.

While the KNN model demonstrates exceptional precision and sensitivity, its notably lower accuracy and specificity highlight a significant challenge in correctly identifying true negatives. This discrepancy might suggest that while the model is highly capable of recognizing similar instances, it struggles with diverse or outlier data points, affecting its overall accuracy as we can observe in Table 3.

Table 3 – K-Nearest Neighbors (KNN)

Metric	Value (%)
Accuracy	60.8
Precision	99.6
Sensitivity	99.4
Specificity	17.7

Font: Created by the authors.

The results in Table 4 show a notable variation in the performance of different algorithms on the same dataset. SVM shows the highest accuracy, making it the most reliable model for this particular task. The Decision Tree also performs well, offering a balanced approach between sensitivity and specificity. However, KNN's lower accuracy and specificity indicate it may not be the best choice for datasets with high variability or when the identification of true negatives is as critical as the identification of true positives.

Table 4 – Models Performance Comparated

Model	Accuracy (%)	Precision (%)	Sensitivity (%)	Specificity (%)
Decision Tree	98	98.5	99	99.4
Support Vector Machine	99.5	98.5	99.4	99.3
K-Nearest Neighbors (KNN)	60.8	99.6	99.4	17.7

Font: Created by the authors.

These findings emphasize the importance of choosing the right algorithm based on the specific requirements of the task at hand, including the nature of the data and the importance of different types of errors. Future work could explore combining these models in an ensemble approach to leverage their strengths and mitigate weaknesses, potentially leading to improved overall performance.

5 CONCLUSION

The rise of social media, coupled with the ease of access through smartphones, tablets, and laptops, has led to both benefits and drawbacks, including the increasing spread of Fake News. False news disseminated as true facts can create political, educational, and health issues within a society. These challenges necessitate solutions that make it easier to discern what may be real or false, especially on the internet, where news is rapidly produced and shared without discrimination.

Therefore, the aforementioned factors highlight the importance of conducting research and developing technological applications that assist in addressing issues arising from the digital world, such as the spread of Fake News.

Machine Learning, with its capability to extract knowledge from continuously produced data and assist in research methodologies, can play a significant role in verifying the truthfulness of information shared online. Its application is evolving across various fields, including the detection and verification of articles' authenticity. Developing such applications can offer considerable benefits, as the models used have proven capable of accurately identifying false or true information.

This study employed Machine Learning techniques such as Holdout and various classification algorithms to predict the veracity of articles in the dataset, based on characteristics such as instances and attributes. Five models were applied, four of which achieved accuracy rates over 90%, and only one fell to the 60% range, as detailed in previous chapters.

Furthermore, the search for a standardized database containing both Fake and True news represented a limitation of this work, as it was challenging to find a database with concise information. Additionally, it is crucial to note that there is a scarcity of applications for detecting fake news, particularly in Brazil. Thus, there is a need for further development and continuation of projects in this area that aim to evaluate additional classification algorithms for article and news prediction, such as Naive Bayes, Boosting, K-means, and Gradient Descent. Implementing other Machine Learning techniques like Cross-Validation, K-fold, and Leave-one-out to measure optimal performance, evaluating model performance without the target variable, and expanding the use of Portuguese language (Brazil) databases are also recommended.

Accordingly, the methodology adopted in this study looks forward to future work involving the integration with a web application for real or fake news detection, classifying them and thereby utilizing and demonstrating the significance of Machine Learning's applicability in reducing the dissemination and adverse effects that Fake News can have on society at large.

REFERENCES

AHMED, Hadeer; TRAORE, Issa; SAAD, Sherif. **Detection of online fake news using n-gram analysis and machine learning techniques**. In: Intelligent, Secure, and Dependable Systems in Distributed and Cloud Environments: First International Conference, ISDDC 2017, Vancouver, BC, Canada, October 26-28, 2017, Proceedings 1. Springer International Publishing, 2017. p. 127-138.

ALBON, Chris. **Machine Learning with Python Cookbook**. 1ª ed. O'Reilly Media, Abril de 2018.

BEYER, Kevin *et al.* **When is "nearest neighbor" meaningful?** In: International conference on database theory. Springer, Berlin, Heidelberg, 1999. p. 217-235.

CARVALHO, Mariana Freitas Canielo de; MATEUS, Cristielle Andrade. Fake News e **Desinformação no Meio Digital**: Análise da Produção Científica sobre o Tema na Área de Ciência da Informação. V EREBD, 2018, Universidade Federal de Minas Gerais, Belo Horizonte – MG.

COUTINHO, Bernardo. **Modelos de predição SVM**. Medium, 2019. Disponível em: https://medium.com/turing-talks/turing-talks-12-classifica%C3%A7%C3%A3oporsvm-f4598 094a3f1. Acesso em: 01 out. 2020.

DE ARAUJO VIEIRA, Elamara Marama *et al.* **Avaliação da performance do algoritmo J48 para construção de modelos baseados em árvores de decisão**. Revista Brasileira de Computação Aplicada, v. 10, n. 2, p. 80-90, 2018.

DELMAZO, Caroline; VALENTE, Jonas C.L. **Fake news nas redes sociais online: propagação e reações à desinformação em busca de cliques**. Media & Jornalismo, Lisboa, v. 18, n. 32, p. 155-169, abr. 2018. Disponível em: http://www.scielo.mec.pt/scielo.php?script= sci_arttext&pid=S2183-54622018000100012&lng=pt&nrm=iso . Acesso em: 25 out. 2020.

DUDA, Richard O.; HART, Peter E.; STORK, David G. **Pattern Classification**. 2ª ed. Wiley-Interscience, 2000.

FERRARI, D. G.; DE CASTRO SILVA, L. N. **Introdução a mineração de dados**. Saraiva Educação S.A., 2017.

FOSSENG, S**. Learning Distance Functions in k-Nearest Neighbors**. Dissertação (Mestrado) — Institutt for datateknikk og informasjonsvitenskap, 2013.

IZBICKI, Rafael; SANTOS, Tiago Mendonça dos. **Aprendizado de máquina: uma abordagem estatística**. São Carlos, SP, 2020. [Livro eletrônico].

MADANI, Mirmorsal; MOTAMENI, Homayun; ROSHANI, Reza. **Fake news detection using feature extraction, natural language processing, curriculum learning, and deep learning**. International Journal of Information Technology & Decision Making, 2024, 23.03: 1063-1098.

MONARD, Maria Carolina; BARANAUSKAS, José Augusto. **Conceitos sobre Aprendizado de Máquina**. Universidade de São Paulo, 2003. Disponível em: http://dcm.ffclrp.usp.br /~augusto/publications/2003-sistemas-inteligentes-cap4.pdf. Acesso em: 24 out. 2020.

MONARD, Maria Carolina; BARANAUSKAS, José Augusto. **Indução de Regras e Árvores de Decisão**. Em: Sistemas Inteligentes - Fundamentos e Aplicações. 1 ed. Barueri, SP: Manole Ltda, 2003. p. 115-140. ISBN 85-204-1680.

MÜLLER, Andreas C.; GUIDO, Sarah. **Introduction to Machine Learning with Python:** A Guide for Data Scientists. 1ª ed. Nova York: O'Reilly Media, 2017.

NAVLANI, Avinash. **Support vector machines with Scikit-Learn**. DataCamp, 2019. Disponível em: https://www.datacamp.com/community/tutorials/svm-classification-scikit-learn-python. Acesso em: 01 out. 2020.

PÉREZ-ROSAS, Verónica, *et al.* **Automatic Detection of Fake News**. In: Proceedings of the 27th International Conference on Computational Linguistics. 2018. p. 3391-3401.

PRACHI, Noshin Nirvana, *et al.* **Detection of Fake News Using Machine Learning and Natural Language Processing Algorithms** [J]. Journal of Advances in Information Technology, 2022, 13.6.

RISING, F. O. I.; ODEGUA, O. "**DataSist:** A Python-based library for easy data analysis, visualization and modeling," 2017.

SHU, Kai; WANG, Suhang; LIU, Huan. **Beyond news contents: The role of social context for fake news detection**. In: Proceedings of the twelfth ACM international conference on web search and data mining. 2019. p. 312-320.

STEHMAN, S. V. **Selecting and interpreting measures of thematic classification accuracy**. Remote Sensing of Environment, v. 62, n. 1, p. 77–89, 1997.

Conflict of interest

We declare that there was no conflict of interest.