A Study of Algorithm-Based Detection of Fake News in Brazilian Election: Is BERT the Best?

Lara Souto Moreira, Gabriel Machado Lunardi, Matheus de Oliveira Ribeiro,
Williamson Silva, Fábio Paulo Basso

***Abstract*—The recent Brazilian election was plagued by the proliferation of false news on the internet. Many people turned to social media to fact-check information and verify its authenticity. In today's digital and data-driven world, fake news can spread rapidly, causing detrimental effects, such as potentially influencing the outcome of an election. In light of this, verifying information has become increasingly reliant on software. While intelligent software can be used to detect and mitigate the spread of fake news, there is a lack of research on the use of such technology in the Portuguese language, particularly when it comes to the implementation of newer strategies such as the Representation of a Bidirectional Transformer Encoder (BERT). Our study evaluated BERT's ability to detect fake news compared to traditional machine learning algorithms, using text classification to identify false news. The results demonstrate BERT's superiority over other algorithms, with a statistically significant difference in all cases. BERT can be considered a viable option for detecting fake news.**

***Index Terms*—BERT, Brazil, Fake news, Machine Learning, Natural Language Processing.**

I. Introduction

Software solutions from companies like Google (when it auto-completes or comes up with a search string), Netflix (when recommending a movie or series according to user preferences), and Amazon (when Alexa personal assistant interacts with people indoors) are applications of Natural Language Processing (NLP).

NLP supports the development of applications based on textual data, whether written or spoken [1]. One of the tasks within the NLP area is textual classification/prediction, which consists of automatically allocating a piece of text to a particular category based on machine learning algorithms [2]. One of the emerging applications of the text classification task is detecting fake news, which aims to classify news text as "false" or "true". Fake news happens in all idioms and causes different emotional reactions in people, such as anger, indignation, and frustration. As a result, the reader finds it difficult to critically analyze the content, which generates a desire to share the news immediately [3], [4].

The spread of fake news reached alarming levels in the 2016 US presidential elections [5] and the Brazilian elections of 2022. In 2018, the same effect occurred in Brazil in presidential elections and, more recently, on a global scale, involving the SARS-CoV-2 virus, which causes COVID-19, and the vaccines against it [6]. Scenarios like these strongly motivated the design of NLP solutions in different languages to minimize the misinformation caused by the sharing of fake news. In Portuguese, however, few efforts are related to the automatic detection of fake news. This is because there are few datasets with previously labeled news, essential for detecting fake news based on text classification [7]. A reference effort is the dataset Fake.Br, the result of work developed at the Interinstitutional Nucleus of Computational Linguistics (NILC-USP) using traditional NLP approaches to detect fake news.

Another gap that guides this work is BERT (*Bidirectional Transformers for Language Understanding*), a deep learning algorithm launched by Google, which has revolutionized the field of NLP [8]. However, little effort has been devoted to analyzing its performance in the Portuguese language, using its variant BERTimbau [9]. Therefore, this article aims to evaluate the ability of the BERTimbau algorithm to detect fake news against algorithms considered traditional in text classification.

In this direction, we have the following hypothesis to test: *BERTimbau presents greater accuracy when compared to traditional machine learning algorithms*. Besides comparing BERTimbau and BERT, our study also evaluated the performance of traditional machine learning algorithms commonly used for fake news detection. This approach provides a comprehensive analysis of the effectiveness of BERTimbau and BERT compared to established methods in fake news detection. Overall, our study contributes to advancing research on identifying fake news, particularly in under-researched languages such as Portuguese, and highlights the potential for cross-lingual applications of machine learning algorithms.

The remainder of this paper is organized as follows. Section II presents the core concepts behind the algorithms. Section III presents related works. Section IV describes the methodology that includes the dataset, algorithms, and metrics that we used to evaluate the algorithms' performance. Section V shows the results and discussion of our findings. Finally, Section VI highlights the main contributions and future work.

II. Background

Text classification is one of the most important tasks of supervised machine learning. It consists of a machine learning technique that assigns predefined categories to open text. Text classifiers can be used to organize, structure, and categorize virtually any type of text. Most text classification and document categorization systems can be broken down into the following four phases: features extraction, dimension reductions, classifier selection, and evaluations [10], [11].

The text can also be analyzed using other methods, such as exposing its entities to a classifier using a word cloud-based data visualization technique. In addition, part-of-speech markup distribution can provide hints in text style and therefore be helpful as a resource in classifying fake news [12].

Two models of textual representation were used in this work, namely, the Bag of Words (BoW) and the Term frequency–inverse document frequency (Tf-idf), both are exemplary models for the NLP area because through them it is possible to perform feature extraction techniques. This approach allows us to transform unstructured data into structured data, making it possible to process it systematically using algorithms. [13].

The BoW technique indicates the number of occurrences of each word chosen in the training corpus, the Tf-idf follows the same concept as the BoW in counting the words present in a corpus, but it has the difference in that it performs a weighting of these words, or that is, its meaning increases in proportion to the number of times the same word is repeated, being compensated by its frequency in the set. The use of Tf-idf becomes interesting when one wants to measure the number of times that a given word appeared in a set of texts.

After converting text into feature vectors, classification algorithms can be applied to predict text classes. In this article, both traditional machine learning algorithms and BERT will be used to classify news. BERT is a newer model in natural language processing that is different from many other models because it uses bidirectional training. [14].

III. Related Works

Some of the main works related to this research will be presented below.

In the article [15] the authors used a corpus of more than 8 million tweets about the general election in the U.S and proposed a bidirectional training approach that improves the classification of fake news by capturing semantic and long-distance dependencies in sentences. The results show that the BERT-based machine learning approach achieved improved classification results compared to other models, with an accuracy of 98.90%.

The authors of [16] conducted a comparison of four machine learning algorithms - Logistic Regression (LR), Stochastic Gradient Descent (SGD), Support Vector Machine (SVM), and Multilayer Perceptron (MLP) - for detecting fake news on the internet. The study utilized the Fake.BR and Sirene News datasets, which totaled 11,942 news items, employed a cross-validation technique with 10 iterations to evaluate the accuracy of the algorithms. The results indicated that all four models achieved precision greater than 90%, with SVM demonstrating the highest performance (96.39%), followed by MLP (95.14%), LR (94.30%), and SGD (92.90%).

The study presented in [17] analyzed the accuracy of thirty-eight algorithms for identifying fake news and identified sixteen datasets used for this purpose. The three algorithms with the highest accuracy were the Stacking Method, Bidirectional Recurrent Neural Network (BiRNN), and Convolutional Neural Network (CNN), with accuracy rates of 99.94%, 99.82%, and 99.80%, respectively. All other algorithms analyzed had accuracy rates greater than 90%. The most commonly used datasets were Kaggle, Weibo, and Fake News Challenges (FNC).

In [18], a comparison is made between SVM and Naive Bayes algorithms to identify which is more effective in classifying political news as true or false. The authors used BoW and Tf-idf techniques for natural language processing and defined metrics to evaluate the algorithms' accuracy. The results showed that SVM outperformed Naive Bayes, with the BoW technique achieving the highest precision of 80.4%.

The article [19] compares different approaches for fake news classification using machine learning algorithms. The author used Logistic Regression, Random Forest, XGBoost, and word embeddings from Multilingual BERT and BERT. The results showed that the XGBoost and BERTimbau models achieved F1-scores of 96% and 92%, respectively. These results suggest that using more advanced machine learning algorithms can significantly improve the accuracy of fake news detection.

The article [20] presents an analysis of Fake News using machine learning algorithms. The researchers employed the FakeRecogna Anomaly, consisting of 101.000 news articles. They evaluated the performance of 11 machine learning algorithms. The results showed that the algorithms did not achieve satisfactory outcomes in accurately detecting fake news.

The study [21] aimed to use text mining techniques to classify news articles as genuine or fake. The WEKA data mining tool was utilized, employing decision tree analysis with Bag-of-Words (BoW) and Term Frequency-Inverse Document Frequencyigura1. (Tf-idf) approaches. The evaluation of the experiment included accuracy, precision, recall, and F-score as metrics. The models were validated using the Fake.br Corpus dataset. The results showed that the Tf-idf model outperformed the BoW model, achieving an accuracy of 89.82%. These findings highlight the effectiveness of Tf-idf in accurately distinguishing between genuine and fake news articles.

Based on the related works presented, our work differs, as far as we know, in the use of BERTimbau applied to a well-known Brazilian dataset. Also, we compare the performance, side by side, of the BERT original English implementation and BERTimbau, the BERT's variant for Portuguese. Lastly, we carefully applied statistical tests to compare BERTimbau against traditional machine learning algorithms, an analysis overlooked in the literature reviewed.

IV. Methodology

Machine learning algorithms were employed to classify news. The process of conducting this work involved several steps, as depicted in Fig. 1 selecting a suitable database, data transformation, training and validation, and model comparison.

A. Dataset

To perform algorithms tests and evaluations, a database was necessary to serve as input for such algorithms. In this work, we adopted the database Fake.br, one of the few annotated databases in Portuguese. The dataset contains 3,600 true and 3,600 false news, totaling 7,200 news. According to [22], the news is divided into six categories: 4,180 on "politics", 1,544 on "TV & celebrities", 1,276 on "society & daily news", 112 on ' 'science & technology", 44 on "economics" and 44 news

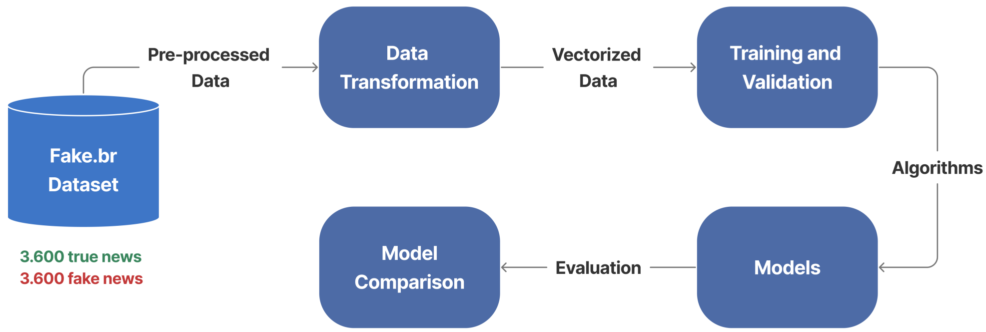


Fig. 1. Methodology adopted in this research.

on “religion”. They all date between January 2016 and January 2018, with free text formatting.

The availability of a large number of Portuguese news articles and annotated data in the Fake.br database were key factors that led to the decision to use it instead of creating a new database. Moreover, the Fake.br dataset is a diverse dataset that contains a variety of fake news articles, covering different topics and coming from different sources. This diversity is essential in ensuring that the algorithms are tested on a range of different types of fake news and are robust enough to identify fake news across different contexts.

The development of a new database would require, in addition to collecting a large number of news, the manual annotation of each piece of news as true or false. This task, therefore, would require a human evaluation to ensure its validation and use in the training of algorithms. Therefore, it would take a lot of time to design, develop and evaluate this new base.

To summarize the database content, we employed the word cloud data visualization technique. The Fig.2 presents the most frequently named entities in real news (true), and the Fig. 3 with the most frequently named entities in fake news. Named entities are classifications of some text elements, such as people, places, and dates. This type of strategy is useful for extracting relevant information from textual data [23].

Fig. 2. *True News.*

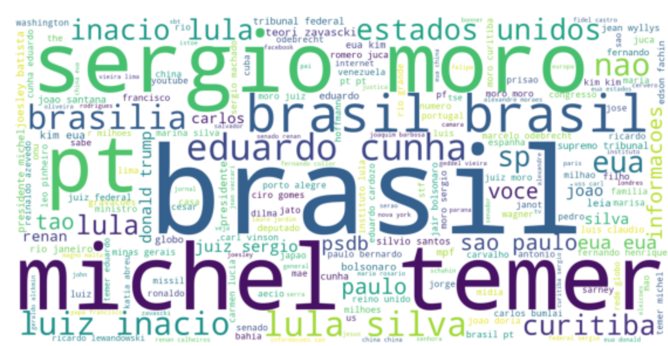


Fig. 3. *Fake News.*

B. Pre-Processing and Data Transformation

Fake.br provides a CSV file with the pre-processed text, including the stopwords removal, accents, and special characters that do not add semantics to the machine learning models. Then, the news was submitted to a data transformation step, a process in which the texts are transformed into a vector of numerical values interpreted by the classifier. The *Scikit-learn* library provides a set of tools that allow you to perform this vectorization.

We employed two methods of vectorization in the analysis. The first method was the *CountVectorizer* class, which applies the “Bag of Words” technique that simply counts the frequency of each word in the text. The second method was the *TfidfVectorizer* class, which applies the Term Frequency-Inverse Document Frequency (Tf-idf) technique. This method assigns a weight to each word based on how often it appears in the text but also takes into account how frequently it appears across the entire corpus.

The Transformers architecture was used to process the data for the BERT algorithm, as it provided methods to instantiate the model and transform the texts into tokens.

In addition to evaluating the effectiveness of BERTimbau in identifying fake news in Portuguese, the study also compared its performance to that of the original BERT model using a translated version of the dataset in English. We utilized the Google Cloud Translation API integrated into a Python environment for translating the “Fake.br” dataset. The Python script for integrating the API and performing the translation and the translated database can be found in the Machine Learning for Fake News Prediction Repository associated with the article.

This approach allowed for a cross-lingual comparison of the two models, highlighting their potential for use in multilingual applications of fake news detection.

C. Model Training and Evaluation

Numerous machine learning algorithms can be used for predictive modeling. Among them, six traditional algorithms were chosen, and BERT, whose main technical innovation is applying bidirectional training of Transformer, a popular attention model, to modeling the language.

BERT consists of a model that establishes a word representation structure to be pre-trained for use in a wide range of [8] tasks. Several pre-models based on BERT were trained for a specific language, such as French [24] and Spanish [25], to try to outperform the results of BERT, which is a multilingual model with support for 104 languages. BERTimbau consists of a pre-trained model for the Portuguese language that uses data from brWaC [9]. Also, two versions are available, namely *Base* and *Large*.

Regarding traditional algorithms, all algorithms chosen were supervised machine learning algorithms that can be used for classification tasks. The specific algorithms we utilized were:

- **SGDClassifier (SGD)**: a linear classifier that uses Stochastic Gradient Descent to learn the model's weights. It is often used for large-scale and sparse data sets.
- **Logistic Regression (LR)**: a linear classifier that is used to predict binary outcomes. It uses logistic regression to model the probability of a binary outcome as a function of the input features.
- **MLPClassifier (MLP)**: a neural network classifier that consists of multiple layers of artificial neurons (also called perceptrons). It is often used for complex classification tasks.
- **Support Vector Machine (SVM)**: a linear classifier that tries to find the hyperplane in the feature space that maximally separates the different classes. It is often used for high-dimensional data sets.
- **AdaBoostClassifier (ADA)**: an ensemble classifier that combines the predictions of multiple weak classifiers to make a strong classifier. The weak classifiers are trained iteratively, with each classifier focusing on the examples that were misclassified by the previous classifiers.
- **Naive Bayes Classifier (NB)**: is a probabilistic classifier based on Bayes' theorem with the assumption of independence between the features. Despite its simplicity, it has been proven to be effective in many text classification and sentiment analysis tasks.

Many factors influenced the choice of machine learning algorithms for a particular problem. Some of the reasons why these algorithms were chosen were because they are widely used and have proven to be effective in many different contexts. It is common for researchers to compare the performance of different algorithms to choose the best one for a particular task or dataset.

For each machine learning algorithm, we set various combinations of hyperparameters values, employed to control the algorithm learning process. We adjusted the hyperparameters according to the algorithm settings presented in the related works and also based on the *Scikit-learn* documentation. All the implemented machine learning algorithms, along with their respective hyperparameters, are also available in the repository. After developing these models for each combination of hyperparameters, we tested the performance of each of these combinations using a 10-fold cross-validation technique. Cross-validation is a method widely used in prediction to assess the generalization capability of predictive models [26].

We adopted the following metrics to compare the algorithms, Precision, Recall, and F-Measure. We generated a ranking report with the metrics used to measure the quality of predictions from all ranking algorithms [2]: Accuracy represents the percentage of predictions matching expected responses, Precision is a measure of how consistent the model's predictions are; Recall is defined as the measure of its completeness; and, finally, F-measure is used to combine precision and recall and can provide overall performance in predicting news.

V. Results and Discussion

As mentioned, we tested the algorithms' performance using a cross-validation technique. This technique divides the dataset into 10 equal parts and alternates these subdivisions between the training and test sets, ensuring that the algorithms are tested on diverse data. Table I shows that most of the algorithms (except Naive Bayes) had an average accuracy of more than 93%. The vectorization used was the *Bag of Words* method. The algorithm that obtained the best performance, with an average of 96.63% of accuracy was the LR, which calculates the sum of the input resources and the logistics of the result. The algorithm with the worst performance was Naive Bayes, with an average of 83.21% accuracy.

Table II shows the results obtained using the Tf-idf vectorization method. The SVM algorithm had the best performance, with an accuracy of 96.50%, and could classify different classes effectively. From these data, the algorithm groups into classes that have the same similarity. Again, we observed that the NB algorithm had the worst performance, with accuracy even lower than that presented previously with *Bag of Words*. Part of this comes from the fact that NB is not an algorithm that behaves well with continuous data. Upon examining Table I and Table II, it can be observed that BERTimbau outperforms all traditional algorithms, regardless of the text representation method used. BERTimbau also outperforms its original implementation (BERT) in English. It was expected because the first one is tailored for Portuguese while the second needs a translation from Portuguese to English. Even a common technique in NLP tasks, the automatic translation implicates a lack of semantics that, in turn, harms the model performance.

After performing the cross-validation, the fold results average was calculated to create a classification report with all the metrics for comparing the algorithms. This is shown in

TABLE I
ACCURACY RESULTS OF EACH FOLD IN BOW

				BoW				
Algorithms	**SGD**	**LR**	**MLP**	**SVM**	**ADA**	**NB**	**BERT**	**BERTimbau**
Iteration 1	96.14%	96.25%	94.44%	96.39%	94.58%	82.08%	91.94%	94.72%
Iteration 2	96.39%	95.97%	93.89%	96.39%	94.86%	81.94%	95.83%	98.47%
Iteration 3	94.44%	95.69%	92.22%	95.14%	93.33%	82.22%	96.39%	97.64%
Iteration 4	96.53%	97.50%	93.61%	96.25%	94.86%	85.56%	97.78%	98.75%
Iteration 5	95.69%	96.67%	93.06%	95.83%	95.83%	85.69%	99.31%	99.86%
Iteration 6	95.56%	97.22%	95.00%	96.81%	95.28%	84.31%	99.03%	99.86%
Iteration 7	97.64%	97.08%	94.44%	97.64%	96.25%	82.08%	98.89%	99.72%
Iteration 8	96.11%	97.08%	94.58%	96.67%	92.36%	78.61%	99.03%	99.72%
Iteration 9	93.47%	96.53%	93.33%	96.11%	92.78%	84.58%	98.61%	100.00%
Iteration 10	95.56%	96.25%	94.58%	96.53%	93.89%	85.00%	98.47%	99.86%
Average	**95.65%**	**96.63%**	**93.92%**	**96.37%**	**94.40%**	**83.21%**	**97.53%**	**98.86%**

TABLE II
ACCURACY RESULTS OF EACH FOLD IN TF-IDF

				Tf-idf				
Algorithms	**SGD**	**LR**	**MLP**	**SVM**	**ADA**	**NB**	**BERT**	**BERTimbau**
Iteration 1	97.50%	95.28%	95.69%	97.36%	95.14%	58.61%	91.94%	94.72%
Iteration 2	96.39%	96.11%	96.67%	96.94%	95.69%	58.75%	95.83%	98.47%
Iteration 3	95.00%	93.33%	95.56%	95.28%	92.64%	61.11%	96.39%	97.64%
Iteration 4	97.50%	95.97%	97.36%	97.64%	94.44%	64.72%	97.78%	98.75%
Iteration 5	96.67%	95.42%	96.25%	96.81%	96.25%	61.11%	99.31%	99.86%
Iteration 6	96.25%	95.42%	96.94%	96.39%	95.28%	63.06%	99.03%	99.86%
Iteration 7	96.39%	95.56%	96.11%	96.53%	95.69%	60.14%	98.89%	99.72%
Iteration 8	95.28%	94.17%	93.47%	95.56%	94.72%	57.92%	99.03%	99.72%
Iteration 9	95.42%	93.75%	95.69%	95.69%	92.92%	61.67%	98.61%	100.00%
Iteration 10	96.53%	96.39%	95.69%	96.81%	93.47%	60.42%	98.47%	99.86%
Average	**96.29%**	**95.14%**	**95.94%**	**96.50%**	**94.63%**	**60.75%**	**97.53%**	**98.86%**

Table III for algorithms with *Bag of Words vectorization* and the Table IV for the *Tf-idf vectorization*.

The BERTimbau model's performance on the classification of false news was evaluated using precision, recall, and, f-measure, with average scores of 99.19%, 98.71%, and 98.95% respectively. The model's performance on the classification of true news was evaluated using the same metrics, with average scores of 98.70%, 98.19%, and 98.74% respectively. The overall accuracy of the model was 98.86%.

The average fold results for true and false news showed that the BERTimbau model outperformed traditional machine learning algorithms in all of the evaluated metrics.

After comparing the algorithms according to their accuracy, recall, and, F-measure, statistical tests were performed comparing the accuracy of all traditional algorithms with BERTimbau. This type of test was carried out to evaluate the hypothesis presented in the introduction of this article: BERTimbau is more accurate than traditional algorithms. To evaluate the accuracy factor of each traditional algorithm compared with BERTimbau, the non-parametric *Mann-Whitney U* test was used, considering that the cross-validation data of the BERTimbau accuracy did not show a normal distribution [27].

Table V presents the *Mann-Whitney U* test results aiming to statistically evaluate the superiority of BERTimbau over traditional classification algorithms. According to the result of Table V, we realized that the differences among the medians of BERTimbau and the other algorithms have statistically significant differences (p-value <0.05). Therefore, we confirm our initial hypothesis that BERTimbau is superior to traditional machine learning algorithms when detecting fake news.

However, this result must be considered with a grain of salt because there are some threats to validity, and we need to present them. The first one regards the context of data. Even the Fake.Br dataset contains many categories (e.g. politics, religion, etc.), word clouds (see Fig. 2 and Fig. 3) denote a predominance of politics. We can conclude it because "Michel Temer", "Sergio Moro", "PT", "Luiz Inacio Lula da Silva", and so on are all entities linked to politics. "Sergio Moro", for example, was highly cited in the 2018 presidential election.

Therefore, information related to politics is ephemeral, i.e., news often has a medium to a short life, as the example of "Sergio Moro" presented previously. Other subjects, like seasonal disasters such as COVID-19 and extreme weather conditions, are also ephemerally affected. Thus, any machine learning model trained for the problem of detecting fake news in a given subject must be robust about the publication time. Then, if we carry out the training in outdated data, it could harm the performance detection of fake news in current news. It is especially valid for traditional machine learning algorithms because they require a training phase and some strategy to keep the model up to date to mitigate the news ephemerality problem.

On the other hand, BERT and BERTimbau are fine with that because these algorithms do not require explicit training once they are pre-trained over large corpora like Wikipedia and brWaC respectively. After all, the more data you have, the more powerful the detection model becomes, and BERT and its variations benefit from the fact the large bases are continually updated. Other characteristics that help to explain BERTimbau superiority are its ability to capture contextual information,

TABLE III
BOW RESULTS TABLE

Algorithms	Average accuracy	Precision		Recall		F-Measure	
		True	Fake	True	Fake	True	Fake
SGD	95.65%	96.78%	94.58%	94.44%	96.86%	95.60%	96.66%
LR	96,63%	97,62%	95,67%	95,58%	97,67%	96,59%	97%
MLP	93.92%	91.21%	97.00%	90.64%	97.19%	93.71%	94.11%
SVM	96.37%	97.42%	95.38%	95.28%	97.47%	96.33%	96.41%
ADA	94.40%	94.91%	93.90%	93.83%	94.97%	94.37%	94.43%
NB	83.21%	76.64%	94.07%	95.53%	70.89%	85.05%	80.85%

TABLE IV
TF-IDF RESULTS TABLE

Algorithms	Average Accuracy	Precision		Recall		F-Measure	
		True	Fake	True	Fake	True	Fake
SGD	96.29%	96.77%	95.82%	95.78%	96.81%	96.27%	96.31%
LR	95.14%	96.75%	93.64%	93.42%	96.86%	95.05%	95.22%
MLP	95.94%	94.99%	96.94%	97.00%	94.89%	95.99%	95.90%
SVM	96.50%	96.89%	96.12%	96.08%	96.92%	96.49%	96.51%
ADA	94.63%	95.11%	94.15%	94.08%	95.17%	94.60%	94.65%
NB	60.75%	56.03%	99.36%	99.86%	21.64%	71.79%	35.54%

transfer learning capabilities, and attention mechanism making it a powerful tool for NLP tasks. Putting it all together we explain why BERTimbau outperforms traditional machine learning algorithms.

Despite its many benefits, the BERT family has some limitations. One of these is its high computational cost, because it requires a large amount of computational resources to train and fine-tune, making it difficult for researchers and practitioners with limited resources to use it. Another limitation is that it is pre-trained on a large corpus of text data, this means that it has been exposed to a wide range of language and may have learned certain biases or stereotypes present in the data. This could lead to the model making unfair or inaccurate predictions about certain groups of people or specific topics.

BERT family also has a limitation in dealing with rare or out-of-vocabulary words, as it has a fixed vocabulary size, it may not be able to accurately represent words that are not present in its training data. Lastly, it has been trained on a large corpus of text data, and therefore it's not optimized for specific tasks or domains. For example, fine-tuning a BERT model on a specific domain or task may not be as effective as using a task-specific model that has been trained on a smaller and more relevant dataset.

TABLE V
MANN-WHITNEY TEST COMPARISON ACCURACY

Algorithm		BoW p Value	Tf-idf p Value
BERTimbau	SGD	0.0019	0.0028
BERTimbau	LR	0.0028	0.0013
BERTimbau	MLP	0.0002	0.0022
BERTimbau	SVM	0.0032	0.0032
BERTimbau	ADA	0.0008	0.0009
BERTimbau	NB	0.0002	0.0002

VI. CONCLUDING REMARKS

In this research, we conducted a statistical analysis of seven machine learning algorithms. We discovered that the Portuguese variation BERTimbau exhibited the highest accuracy compared to the other machine learning algorithms, including when compared to its original implementation in English. However, it is important to note that these results should be viewed with caution. While the BERT family is a highly effective language model with exceptional results in various NLP tasks, it also has limitations such as high computational cost, the potential for bias, challenges with uncommon words, and lack of optimization for specific tasks or domains. These limitations indicate that the BERT family may not be ideal in all situations.

In future work, we intend to extend this contribution to mapping collaboration networks on fake news dissemination. Currently, we are contributing to the Association of Victims and Families of Victims of Covid-19 "Avico Brasil'to monitor fake news regarding vaccines. Mapping the collaboration can be the best way to avoid fake news shortly, thus opening research gaps for crawling collaborations over social networks.

VII. ACKNOWLEDGEMENTS

The authors thank the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código Financeiro 001 and the Universidade Federal do Pampa (UNIPAMPA - Alegrete). Williamson Silva thanks FAPERGS (Projeto ARD/ARC through process number 22/2551-0000606-0). Fabio Paulo Basso thanks to the Pró-Reitoria de Pesquisa (PROPESQ) and AGP, and FAPERGS through the project N. 22/2551-0000841-0. Gabriel M. Lunardi thanks CNPq (Projeto Universal through process number 405973/2021-7).

REFERENCES

[1] C. Manning and H. Schutze, *Foundations of Statistical Natural Language Processing*. MIT Press, 1999.

[2] K. Faceli, *Inteligência artificial: uma abordagem de aprendizado de máquina*. Grupo Gen - LTC, 2018.

[3] K. Shu, A. Sliva, S. Wang, J. Tang, and H. Liu, "Fake news detection on social media: A data mining perspective," *ACM SIGKDD explorations newsletter*, vol. 19, no. 1, pp. 22–36, 2017.

[4] G. M. Lunardi, "Representing the filter bubble: Towards a model to diversification in news," in *Advances in Conceptual Modeling: ER 2019 Workshops FAIR, MREBA, EmpER, MoBiD, OntoCom, and ER Doctoral Symposium Papers, Salvador, Brazil, November 4–7, 2019, Proceedings 38*, pp. 239–246, Springer, 2019.

[5] I. Moraes, "Notícias falsas e pós-verdade: o mundo das fake news e da (des)informação | politize!." https://www.politize.com.br/noticias-falsas-pos-verdade/, 2017. (Accessed on 08/28/2022).

[6] P. Falcão, A. B. d. Souza, *et al.*, "Pandemia de desinformação: as fake news no contexto da covid-19 no brasil," 2021.

[7] R. A. Monteiro, R. L. Santos, T. A. Pardo, T. A. d. Almeida, E. E. Ruiz, and O. A. Vale, "Contributions to the study of fake news in portuguese: New corpus and automatic detection results," in *International Conference on Computational Processing of the Portuguese Language*, pp. 324–334, Springer, 2018.

[8] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "Bert: Pre-training of deep bidirectional transformers for language understanding," 2018.

[9] F. Souza, R. Nogueira, and R. Lotufo, "Bertimbau: Pretrained bert models for brazilian portuguese," in *Intelligent Systems* (R. Cerri and R. C. Prati, eds.), (Cham), pp. 403–417, Springer International Publishing, 2020.

[10] K. Kowsari, K. Jafari Meimandi, M. Heidarysafa, S. Mendu, L. Barnes, and D. Brown, "Text classification algorithms: A survey," *Information*, vol. 10, no. 4, 2019.

[11] G. M. Lunardi, G. M. Machado, F. Al Machot, V. Maran, A. Machado, H. C. Mayr, V. A. Shekhovtsov, and J. P. M. de Oliveira, "Probabilistic ontology reasoning in ambient assistance: predicting human actions," in *2018 IEEE 32nd International Conference on Advanced Information Networking and Applications (AINA)*, pp. 593–600, IEEE, 2018.

[12] d. A. C. Igor Bichara, *Evaluation of Machine Learning Classifiers in Ordinal Multiclass Fake News Detection Scenario*. PhD thesis, Universidade Federal do Rio de Janeiro, 2019.

[13] M. K. Jain, D. Gopalani, Y. K. Meena, and R. Kumar, "Machine learning based fake news detection using linguistic features and word vector features," in *7th Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON)*, pp. 1–6, IEEE, 2020.

[14] J. Devlin, M. Chang, K. Lee, and K. Toutanova, "BERT: pre-training of deep bidirectional transformers for language understanding," *CoRR*, vol. abs/1810.04805, 2018.

[15] R. K. Kaliyar, A. Goswami, and P. Narang, "Fakebert: Fake news detection in social media with a bert-based deep learning approach," *Multimedia Tools and Applications*, vol. 80, no. 8, pp. 11765–11788, 2021. Epub 2021 Jan 7.

[16] V. N. Barbosa, F. M. Mendes Neto, S. A. Filho, and L. Silva, "A comparative study of machine learning algorithms for the detection of fake news on the internet," in *XVIII Brazilian Symposium on Information Systems*, SBSI, p. 8, 2022.

[17] H. F. Villela, F. Corrêa, J. S. de Araújo Nery Ribeiro, A. Rabelo, and E. E. Costa, "Uma análise da acurácia obtida e datasets utilizados em algoritmos de identificação de fake news," in *ISLA 2022*, Lacais, 2022.

[18] L. D. D. Almeida, V. Fuzaro, F. V. Nieto, and A. L. M. Santana, "Identificação de "fake news" no contexto político brasileiro: uma abordagem computacional," *Anais do II Workshop sobre as Implicações da Computação na Sociedade (WICS 2021)*, 2021.

[19] L. Cordeiro, "Comparação entre modelos com diferentes abordagens para classificação de fake news," 2021.

[20] G. L. Garcia, "Detecção de fake news utilizando aprendizado de máquina," 2023.

[21] I. T. Oshita, "Classificação de fake news por mineração de texto," trabalho de conclusão de curso, Londrina, Brazil, 2021.

[22] R. L. S. Santos, R. A. Monteiro, and T. A. S. Pardo, "The fake . br corpus-a corpus of fake news for brazilian portuguese," 2018.

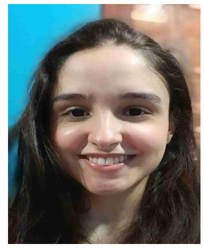
[23] S. Sekine and E. Ranchhod, *Named entities: recognition, classification and use*, vol. 19. John Benjamins Publishing, 2009.

[24] L. Martin, B. Muller, P. J. Ortiz Suárez, Y. Dupont, L. Romary, É. de la Clergerie, D. Seddah, and B. Sagot, "CamemBERT: a tasty French language model," in *58th Annual Meeting of the Association for Computational Linguistics*, pp. 7203–7219, 2020.

[25] J. Canete, G. Chaperon, R. Fuentes, J.-H. Ho, H. Kang, and J. Pérez, "Spanish pre-trained bert model and evaluation data," *Pml4dc at iclr*, vol. 2020, pp. 1–10, 2020.

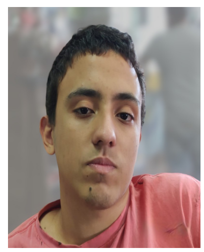
[26] D. Berrar, "Cross-validation," in *Encyclopedia of Bioinformatics and Computational Biology - Volume 1*, pp. 542–545, 2019.

[27] C. Wohlin, P. Runeson, M. Höst, M. C. Ohlsson, B. Regnell, and A. Wesslén, *Experimentation in software engineering*. Springer Science & Business Media, 2012.



Lara Souto Moreira is currently a Software Engineering student at the Federal University of Pampa (UNIPAMPA), Brazil. Her research interests include Data Analysis, Machine Learning, Databases, and Information Systems.

Gabriel Machado Lunardi has a Ph.D. in Computer Science from the Federal University of Rio Grande do Sul (UFRGS), Brazil, and is currently an Adjunct Computer Science Professor at the Federal University of Santa Maria (UFSM), Santa Maria, Brazil. Gabriel has experience in Artificial Intelligence, Natural Language Processing (NLP), Recommender Systems, Machine Learning, and Knowledge Discovery in Databases.



Matheus de Oliveira Ribeiro has a bachelor's degree in Computer Science from the State University of Paraná. He is currently working towards a Master's degree in Software Engineering at the Federal University of Pampa (UNIPAMPA). His research is focused on User Experience and Machine Learning.

Williamson Silva received a Ph.D. in Informatics from the Institute of Computing of the Federal University of Amazonas (UFAM). He is currently an Adjunct Professor of Software Engineering at the Federal University of Pampa (UNIPAMPA). His research interests include Software Engineering, Empirical Software Engineering, Software Quality, Computing Education Research, Usability, User Experience, Machine Learning, and Human-Centered Machine Learning.

Fábio Paulo Basso is currently an Adjunct Professor at the Federal University of Pampa (UNIPAMPA), Brazil. His research interests include software reuse, software architecture, coopetition business models performed through services, technology transfer of computer science invents, precision agriculture and precision animal husbandry.