



# Automated Fake News Detection Using Computational Forensic Linguistics

Ricardo Moura<sup>1</sup>, Rui Sousa-Silva<sup>2,3</sup> , and Henrique Lopes Cardoso<sup>1,4</sup> 

<sup>1</sup> Faculdade de Engenharia, Universidade do Porto, Porto, Portugal  
{up201604912,hlc}@fe.up.pt

<sup>2</sup> Faculdade de Letras, Universidade do Porto, Porto, Portugal  
rsslva@letras.up.pt

<sup>3</sup> Centro de Linguística da Universidade do Porto (CLUP), Porto, Portugal

<sup>4</sup> Laboratório de Inteligência Artificial e Ciência de Computadores (LIACC),  
Porto, Portugal

**Abstract.** Fake news is news-like content that has been produced without following journalism principles. Fake news try to mimic the look and feel of real news to intentionally disinform the reader. This phenomenon can have a strong influence on society, thus being potentially a severe problem. To address this phenomenon, systems to detect fake news have been developed, but most of them build upon fact-checking approaches, which are unfit to detect misinformation when a news piece, rather than completely false, is distorted, exaggerated, or even decontextualized. We aim to detect Portuguese fake news by following a forensic linguistics approach. Contrary to previous approaches, we build upon methods of linguistic and stylistic analysis that have been tried and tested in forensic linguists. After collecting corpora from multiple fake news outlets and from a genuine news source, we formulate the task as a text classification problem and demonstrate the effectiveness of the proposed features when training different classifiers for telling fake from genuine news. Furthermore, we perform an ablation study with subsets of features and find that the proposed feature sets are complementary. The highest results reported are very promising, achieving 97% of accuracy and a macro F1-score of 91%.

**Keywords:** Fake news detection · Forensic linguistics · Natural language processing · Text classification · Disinformation · Misinformation

## 1 Introduction

Technology has evolved significantly in recent years, and its development and adoption have become increasingly fast and easy. One of the technologies that came to define and influence the next generations is new computer-mediated communication channels, such as social media, messaging services, and blogs. These channels made it possible for anyone to share anything about any topic at

any time, instantly and effortlessly. As a result, people are more connected than ever. Companies are aware of this phenomenon and try to use it for their own advantage, e.g. the media now share news on social media. In fact, studies report that people are shifting away from traditional news sources to social media and messaging services to find their news [25]. Even though these platforms have many advantages, they raise a serious problem: the so-called fake news. Because those platforms give all users the freedom to share everything they want at any time, fake news can emerge very easily and rapidly spread disinformation.

The fake news phenomenon can be defined in several different ways and be of multiple types, from satire to fabrication [20], and some of them are even permissible (i.e., satire). The definition of fake news has mutated throughout the years and began to be applied under wrong circumstances [23]. In the context of this paper, fake news is news that does not follow the journalism principles of factuality, objectivity, and neutrality [3, 13]. Instead, fake news pieces try to mimic the look and feel of real news [24] with the intent to mislead the reader. Here lies the distinction between mis- and disinformation: unlike the latter, the former does not intend to mislead.

Although untruthful news accounts have always existed, their use as a way of manipulation and control has recently gained more attention, due to their fast and immediate propagation through social media, without any kind of curation or filtering. Lay people are attracted to this kind of news because of their alluring headlines (used as *clickbait*) and often give more attention to this kind of news than to truthful accounts [4].

Currently, there are two widely used methods to detect fake news: a manual alternative with human intervention and an automatic alternative with Machine Learning methods [8]. The former places the responsibility to assess the news' veracity and accuracy entirely on humans, who then have to flag it depending on their judgment. However, this is not the best option because it has a limited scalability and humans (frequently non-experts) are not sufficiently skilled to distinguish fake from genuine news. The latter alternative to detect fake news consists of using sophisticated computer systems. However, most existing systems are based on fact-checking methods, which fall short of the desired effectiveness, as these systems still lack the robustness to perform a reliable verification of which information is falsely presented [8]. Additionally, detecting fake news goes beyond identifying false information; fact-checking methods are useful when facts are manipulated, but less so when the truth in the news is distorted, exaggerated, or even decontextualized.

This paper presents a system that, contrary to fact-checking, does not depend on the veracity of the facts. Instead, we focus on how the author communicates and how the news is written. In light of this, we address the fake news phenomenon using an approach based on forensic linguistic analysis, i.e. an analysis that considers linguistic and stylistic methods which have been tried and tested in forensic contexts, e.g. to attribute authorship or detect bias in texts [22]. These include, but are not limited to: text statistics (e.g., average text, paragraph, sentence and word length, and  $n$ -gram sequences); spelling; and lexical

choices (e.g., Part-of-Speech). We claim that these approaches have a significant potential to also detect fake news.

Using two corpora collected from multiple sources, we conducted a series of experiments to understand what linguistic characteristics are intrinsic of fake news. Our experiments show promising results with an accuracy of up to 97% and a macro average of F1-score of 90%.

This paper is structured as follows. Section 2 briefly presents previous work on fake news detection using methods similar to the ones applied in this paper. Section 3 introduces the resources used in our experiments, specifically the corpora (Sect. 3.1) and external resources (Sect. 3.2). Section 4 describes the process, from extracting the features to building the model. Next, in Sect. 5, we share, evaluate, and discuss our results. Finally, in Sect. 6 we draw some conclusions, give a perspective into the project's current stage, and discuss what could be the next steps and future work.

## 2 Related Work

Fact-checking is the predominant approach to detect fake news. Notwithstanding, there are alternative methods that seek to make a decision based on linguistic patterns present in the text. The reasoning being that, when someone writes a lie or a deceiving text, they strategically write the text in a way to avoid suspicion [12]. However, not all traces and patterns can be hidden, and hence linguistics-based approaches are often employed for detecting lies, despite being somewhat understudied in the literature.

Ahmed et al. (2017) [1] propose fake news detection using only  $n$ -gram analysis. The authors reached the best performance when using Term Frequency-Inverse Document Frequency (TF-IDF) as a feature extraction technique and a Linear Support Vector Machine (LSVM) as a classifier, with an accuracy of 92%. This accuracy is better than the results obtained by Horne and Adali (2017) [14] (see below). However, this high accuracy score can represent a Population Bias or Representation Bias [19]: as Cruz et al. (2019) [6] highlight, relying only on  $n$ -gram analysis could present a problem because the results of this feature extraction method may vary depending on media content throughout the years.

Perez et al. (2017) [21] made a set of experiments to identify linguistic properties predominating in fake content. The authors constructed two datasets: one was collected via crowd-sourcing covering six news domains; the other was obtained by scraping data from the web, and covers celebrity fake news. They built a fake news detector that achieved the best performance (78% accuracy) using LSVM. The features used were:  $n$ -grams encoded as TF-IDF values; count of punctuation characters; psycho-linguistic features, such as summary categories (e.g. analytical thinking or emotional tone), linguistic processes (e.g. function words or pronouns) and psychological processes (e.g. affective processes or social processes); and features related to readability, such as the number of characters, complex words, long words, number of syllables, word types and paragraphs, among other content features.

Differently from works that focus on the main text, Horne and Adali (2017) [14] consider solely news headlines for detecting fake news. The authors build on the assumption that fake news are targeted at audiences that are not likely to read beyond headlines. They extracted different features and arranged them into three categories: Stylistic Features (e.g. number of stopwords, number of all capital letter words, PoS tagger count on each tag, etc.); complexity features (e.g. readability scores); and psychological features (e.g. number of emotion or informal/swear words). With this set of features extracted from a corpus from 2016 US Election news (retrieved from BuzzFeed) and other scraped news websites related to US politics, the authors have built a LSVM classifier, achieving 71% accuracy.

Overall, these findings show that linguistic-based approaches are understudied. These approaches are, in fact, used but mostly in other contexts and with different goals, such as rumor detection [2], deception detection [18], or hyper-partisanship detection [6]. Such lack of research into fake news detection using approaches other than fact-checking is also evident in Portuguese. Comparing the performance between the works studied is non-trivial, because the authors target different datasets.

### 3 Resources

In this section, we introduce the corpora used in our experiments, as well as the external resources used to build the classifier models used to detect fake news. This project focuses on detecting fake news written in Portuguese. Although Portuguese is one of the most widely spoken languages [26], it still has limited linguistic resources available when compared to English. Due to this limitation, most tools supporting NLP show sub-optimal performance. Nevertheless, we will use tools that already have features and offer support of Portuguese to train the model.

#### 3.1 Corpora

Given the nonexistence of an annotated dataset distinguishing fake from genuine news, we follow a silver standard approach [11] with automatically annotated data [5] when collecting news items for both classes. By using this approach, each news article is labeled (fake or not) according to the category associated with the website where it is published. URLs of the news, which were collected between November and December 2020 and included in the dataset, are made available<sup>1</sup>.

#### Fake News Corpus

Although there are several online corpora of fake news<sup>2</sup>, to the best of our knowledge none is based on Portuguese. We create a corpus by scraping websites that

<sup>1</sup> [drive.google.com/file/d/1jqIMxhcH6H4ozA3zbTnxphriQx1fKi4G/view](https://drive.google.com/file/d/1jqIMxhcH6H4ozA3zbTnxphriQx1fKi4G/view).

<sup>2</sup> <https://github.com/sumeetkr/AwesomeFakeNews>.

are known to publish fake news contents<sup>3</sup>. From those available, we have chosen five: *Bombeiros24*, *JornalDiario*, *MagazineLusa*, *NoticiasViriato*, and *SemanarioExtra*. Some scraped news articles were deemed unusable since they were tagged, by the source, as opinion articles, which have a status that differs from regular news. Our fake news corpus contains 10 343 news pieces posted between 2017 and 2020.

### ***Público* News Corpus**

We build the genuine news corpora by scraping news articles from *Público*, one of the most reputable news outlets in Portugal. Some scraped articles were deemed unusable since the authors categorized them as parody; hence, they should not be considered fake news. Thus, 110 066 news in total were collected from the same period as part of the fake news corpus.

## **3.2 Natural Language Processing Resources**

We explored multiple resources to get the best results for processing the news articles and ended up using a mix between NLTK<sup>4</sup> for the Portuguese stopwords list, the pySpellChecker<sup>5</sup> library for spell checking, and spaCy models for Portuguese<sup>6</sup> for the other tasks (specifically tokenization, part-of-speech tagging, named entity recognition, and lemmatization). We also use Scikit-Learn<sup>7</sup> implementations of the classifiers we have trained and the function CountVectorizer, from the same library to calculate the  $n$ -grams.

## **4 System Description**

Our fake news detection approach includes two phases. The first is a feature extraction phase, where we convert the news articles into a feature-based representation. Subsequently, we train several machine learning models using the representations obtained.

### **4.1 Feature Extraction**

The main text of the news articles is converted into a set of linguistic features. These features (described in more detail in Table 1) can be divided into four categories:

<sup>3</sup> a) [sabado.pt/portugal/detalhe/be-pede-audicao-da-erc-para-esclarecer-registo-de-sites-de-fake-news](http://sabado.pt/portugal/detalhe/be-pede-audicao-da-erc-para-esclarecer-registo-de-sites-de-fake-news)

b) [dn.pt/edicao-do-dia/11-nov-2018/fake-news-sites-portugueses-com-mais-de-fois-milhoes-de-seguidores-10160885.html](http://dn.pt/edicao-do-dia/11-nov-2018/fake-news-sites-portugueses-com-mais-de-fois-milhoes-de-seguidores-10160885.html).

<sup>4</sup> [www.nltk.org/howto/portuguese\\_en](http://www.nltk.org/howto/portuguese_en).

<sup>5</sup> [www.github.com/barrust/pyspellchecker](https://www.github.com/barrust/pyspellchecker).

<sup>6</sup> [www.spacy.io/models/pt](http://www.spacy.io/models/pt).

<sup>7</sup> [www.scikit-learn.org](http://www.scikit-learn.org).

***n*-grams:** We calculate the vocabulary composed of all lemmatized tokens in the documents and subsequently extract a set of unigrams, bigrams, and trigrams, encoded as normalized counts and with TF-IDF. In order to avoid the influence of named entities, we adopt an approach that obfuscates them and focuses on an approach used in forensic linguistic analysis. We use spaCy’s named-entity recognition to replace classified entities with their respective label – person, organization, and location (e.g. “*Cristiano Ronaldo*” becomes “[PERSON]”).

**Frequencies:** We extract a collection of relative frequencies, including the frequency for each punctuation character, the frequency for each Part-of-Speech tag, and the frequency of each type of adverb.

**Text Statistics:** We also obtain a set of statistical features: the number of paragraphs, sentences, tokens, stopwords, characters and syllables. From these, we also generate some average counts: average number of sentences per paragraph, words per paragraph, words per sentence and characters per word.

**Readability:** We compute a set of features that measure how easy it is to read a text. These include vocabulary richness (i.e., how diverse the vocabulary used by an author is), readability indices (e.g. Flesch [9], Flesch-Kincaid [15], Gunning Fog [10] and SMOG [17]), and ratios such as the percentage of long words (>12 characters), obfuscated words [16] (words with numbers or special characters, e.g. “*cr1me*”), misspelled words, and polysyllable words (>2 syllables).

## 4.2 Dataset Description

Figure 1 shows the distribution of the features that seem to differ the most between fake and genuine news. Feature values were normalized and outliers were hidden to facilitate understanding.

As far as *n*-gram features are concerned, (lemmatized) word sequences such as “*primeiro ministro*” (prime minister), “*presidente*” (president), “*empresa*” (company), or “*milhão*” (million), are far more frequent in genuine than in fake news. Conversely, words such as “*rede social*” (social media), “*mostrar*” (show), “*mulher*” (woman), or “*vida*” (life) are more frequent in fake news than in genuine news. The dataset also shows that genuine news tend to reference entities more often than fake news, which results in a higher count of entity-related *n*-grams.

**Table 1.** Features used to build the model for Fake News detection. A star (\*) indicates that the feature is a feature set.

Feature	Description
<i>Text statistics</i>	
Num paragraphs	Number of paragraphs
Num sentences	Number of sentences [spacy]
Num tokens	Number of tokens
Num stopwords	Number of stopwords [nltk]
Num chars	Number of chars
Num of syllables	Number of syllables
Avg sents per para	Average number of sentences per paragraph
Avg words per para	Average number of words per paragraph
Avg words per sents	Average number of tokens per sentence
Avg chars per sents	Average number of characters per sentence
Avg chars per word	Average number of characters per word
<i>Frequencies</i>	
Freq punctuation *	Relative frequency of each punctuation character
Freq PoS tags *	Relative frequency of each PoS tag
Freq type of adverbs *	Relative frequency of each type of adverb
<i>Readability</i>	
Vocabulary richness *	Measures of vocabulary diversity: ratio between the total number of words and the number of unique words – with or without stopwords
Readability indices *	Measures of text reading/understanding difficulty – <i>flesch</i> , <i>fleschkincaid</i> , <i>gunningfog</i> , and <i>smog</i>
% long words	Fraction of words with 12 or more characters
% obfuscated words	Fraction of words containing punctuation or numbers
% misspelled words	Fraction of words with spelling errors
% uppercase words	Fraction of uppercase words
% polysybl words	Fraction of words eith three or more syllables
<i>N -grams</i>	
<i>n</i> -grams *	TF (counts) and TF-IDF of unigram, bigrams, trigrams. In total 600 <i>n</i> -grams

### 4.3 Classification Process

We conduct several experiments with each feature category and with multiple Machine Learning algorithms, specifically: Logistic Regression (LR), Linear Support Vector Machines (LSVM), Random Forest (RF), Decision Tree (DT), Gradient Descent (SGD), Naive Bayes (NB), and Gradient Boosting Classifier (GBC). We use Scikit-Learn’s implementations of these algorithms and resort

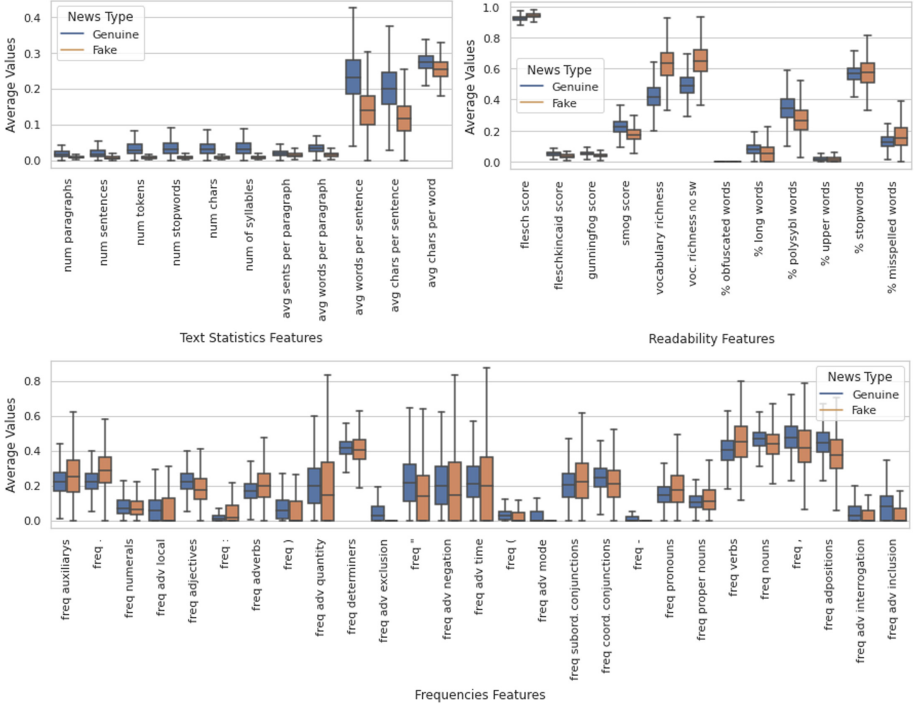


Fig. 1. Distribution values per class for each feature set.

to the default values of the hyperparameters as defined by the library, only specifying (when possible) the class\_weight property to “balanced” to make the algorithms handle both classes with equal importance, and for LR the Lasso penalty (11).

To better assess the performance of each model, we use 5-fold stratified cross-validation. In each fold, we return the following metrics: Accuracy, Precision, Recall, and F1-score. Although we pay attention to all these metrics, we mainly focus on two. The first is Accuracy, which is the metric consistently presented in the related works section (see Sect. 2). However, due to the imbalanced nature of our dataset, the second metric we focus on is the macro average F1-score. Furthermore, we collect the feature importance for every model to understand the features that each model deems more important to choose between the fake and genuine news classes.

## 5 Experimental Results

The results shown in Table 2 are the average performance rates for each model in the 5-fold stratified cross-validation setup. We can observe that Logistic Regression and Random Forest achieve the best results.



**Table 2.** Average results from 5-fold stratified cross-validation.

Model	Acc.	Weighted average			Macro average		
		P	R	F1	P	R	F1
Naive Bayes	0.77	0.93	0.77	0.82	0.63	0.85	0.64
Linear SVM	0.78	0.95	0.78	0.77	0.82	0.77	0.69
SGD	0.87	0.95	0.87	0.90	0.72	0.90	0.76
Gradient boosting	0.92	0.93	0.92	0.92	0.75	0.83	0.78
Decision tree	0.95	0.95	0.95	0.95	0.85	0.85	0.85
Logistic regression	0.95	0.96	0.95	0.96	0.82	<b>0.95</b>	0.87
<b>Random forest</b>	<b>0.97</b>	<b>0.97</b>	<b>0.97</b>	<b>0.97</b>	<b>0.96</b>	0.87	<b>0.91</b>

Tables 3 and 4 show, in more detail, the results obtained by the Logistic Regression and Random Forest models, respectively; we also report the results obtained when using each group of features individually.

**Table 3.** Scores of each feature’s category fitted in a Logistic Regression model.

Features (number of features)	Acc.	Genuine news			Fake news		
		P	R	F1	P	R	F1
<i>N</i> -grams (600)	<b>0.96</b>	0.98	<b>0.98</b>	<b>0.98</b>	0.76	0.80	0.78
Frequencies (64)	0.88	0.98	0.88	0.93	0.39	0.81	0.53
Text Statistics (11)	0.90	<b>0.99</b>	0.90	0.94	0.46	0.90	0.61
Readability (12)	0.89	<b>0.99</b>	0.89	0.94	0.43	0.86	0.57
All Features (687)	0.95	0.65	0.94	0.77	<b>0.99</b>	<b>0.95</b>	<b>0.97</b>

Logistic Regression obtains high accuracy scores regardless of the set of features used, especially the model trained with *n*-grams or the one trained with all the features. The accuracy is even slightly higher when the model trained only with *n*-grams is used, compared to the all-features model. However, if we examine the F1-score, we can see that although the *n*-grams model performs well in finding genuine news, it shows a poor performance when detecting fake news. Since this a fake news detection problem, it makes sense to consider the best model trained with all the features, which achieves a macro-F1 score of 0.87 (as shown in Table 2).

The models trained with Random Forest also present very high accuracy scores, even outperforming Logistic Regression. Nevertheless, we will use the F1-score once more. The best model, in this case, is the one where all the features are used for training. Although the Random Forest model is almost perfect at identifying genuine news, the same cannot be said about fake news. Comparing the best model of each algorithm, we notice that the F1-score for fake news is

**Table 4.** Scores of each feature’s category fitted in a Random Forest model.

Features (number of features)	Acc.	Genuine news			Fake news		
		P	R	F1	P	R	F1
<i>N</i> -grams (600)	0.96	0.97	0.99	0.98	0.89	0.64	0.75
Frequencies (64)	0.96	0.96	0.99	0.98	0.91	0.55	0.69
Text Statistics (11)	0.96	0.97	0.99	0.98	0.81	0.65	0.72
Readability (12)	0.95	0.96	0.99	0.97	0.77	0.53	0.62
All features (687)	<b>0.97</b>	<b>0.98</b>	<b>1.00</b>	<b>0.99</b>	<b>0.94</b>	<b>0.75</b>	<b>0.83</b>

lower in the Random Forest model. Nevertheless, the model trained with Random Forest yields the best results, achieving the highest macro-F1 score among all models (as per Table 2).

In both learning algorithms, we can also notice that the models trained using frequencies or readability properties alone result in comparatively poorer performance. Nevertheless, when combining with the remaining feature sets, the overall performance is improved. Among all feature sets, we can see that the *n*-grams always return the best results for both algorithms. Even though entities were obfuscated, these results may still exhibit some overfitting, as *n*-grams are highly reliant on the vocabulary used.

Results with Logistic Regression also indicate that with the exception of *n*-grams, none of the feature sets can distinguish fake news with a precision higher than 0.5. However, when all of the features are used simultaneously, the model yields an excellent precision score for the fake news class. Additionally, although each feature set performs rather well at distinguishing genuine news when all features are used, precision drops significantly.

## 5.1 Feature Analysis

We analyze the main features used by each model to predict the class label. For Random Forest, we use the *feature\_importance\_* property<sup>8</sup>, while for Logistic Regression we use the *coef\_* property<sup>9</sup>. Since each model has its own way of calculating feature importance, we cannot directly compare the values. Furthermore, the two classifiers make predictions in very different ways. Random Forest is a non-linear classifier composed of a multitude of decision trees, whilst Logistic Regression is based on a linear decision boundary and uses a weighted sum of the features to make predictions. This makes comparing feature importance between the models non-trivial. Nevertheless, what we can do is compare which are the top ten features each model considers the most important:

<sup>8</sup> [scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier](https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html).

<sup>9</sup> [scikit-learn.org/stable/modules/generated/sklearn.linear\\_model.LogisticRegression](https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LogisticRegression.html).

**Logistic Regression**

1. num stopwords
2. num syllables
3. avg words per paragraph
4. avg sents per paragraph
5. 1-gram counts ‘milhão’
6. 2-gram counts ‘milhão euro’
7. freq !
8. freq [
9. smog score
10. freq <<

**Random Forest**

1. num syllables
2. num chars
3. num tokens
4. vocabulary richness
5. avg words per paragraph
6. num stop words
7. vocabulary richness without sw
8. avg chars per sentence
9. avg words per sentence
10. 1-gram counts [ORG]

The feature analysis suggests noticeable differences in fake news articles as compared to genuine news. While Random Forest relies mainly on features from the text statistics category, the Logistic Regression model considers that all feature sets are important.

Similar to Random Forest, the Logistic Regression model places more importance on text statistics, when compared to the other categories. However, Logistic Regression also places some importance on other feature sets: first, the  $n$ -grams “milhão” and “milhão euro”, which are more frequent in fake news, as mentioned in Sect. 4.2. Next, the model uses punctuation frequencies, such as “!”. This frequency can represent the author’s emotions, which are expected to occur more often in fake news. The other two frequencies are more related to the style chosen by the authors, which may represent overfitting. Lastly, the model uses a readability score – SMOG. This metric performs a calculation based on the number of sentences and the number of polysyllable words (both metrics are higher in genuine news) to grant a final score estimating the years of education needed to understand a text.

In addition to the features related to text statistics, the Random Forest model also uses unigram counts [ORG] and vocabulary richness features. The former means that it gives importance to the number of entities identified as organizations. The latter measures language diversity, which is unexpectedly higher in fake news, as mentioned in Sect. 4.2.

## 6 Conclusions

Fake news is news that does not follow the principles of journalism. Instead, the authors of such news try to mimic the look and feel of real news, and have a hidden agenda to disinform the reader. This phenomenon is a severe problem in our society, and the topic has become increasingly relevant in recent years.

For this paper, we collected a corpus of fake news and a corpus of genuine news from the same time frame using a silver standard approach. We then performed feature engineering inspired on approaches used by forensic linguistic analyses.

Although this remains understudied, we conclude that a forensic linguistics-grounded approach for classifying fake news can be applied with great success.

To the best of our knowledge, this is the first work that applies this kind of approach to solve the problem of fake news detection to Portuguese texts.

For future work, we intend to further analyze the robustness of this approach. To do so, we will investigate how our model performs on other corpora and possibly with manually annotated datasets. Furthermore, we will consider exploring the problem in a multi-class formulation exploring different text genres (e.g. fake, genuine, sensationalist news, and so on). We also believe that using neural language models, such as BERT [7], can be a promising direction, and is thus worth exploring.

**Acknowledgments.** This research is supported by project DARGMINTS (POCI/01/0145/FEDER/031460), CLUP (UIDB/00022/2020), and LIACC (FCT/UID/CEC/0027/2020), funded by Fundação para a Ciência e a Tecnologia (FCT).

## References

1. Ahmed, H., Traore, I., Saad, S.: Detection of online fake news using n-gram analysis and machine learning techniques. In: Traore, I., Woungang, I., Awad, A. (eds.) ISDDC 2017. LNCS, vol. 10618, pp. 127–138. Springer, Cham (2017). [https://doi.org/10.1007/978-3-319-69155-8\\_9](https://doi.org/10.1007/978-3-319-69155-8_9)
2. Alkhodair, S.A., Ding, S.H., Fung, B.C., Liu, J.: Detecting breaking news rumors of emerging topics in social media. *Inf. Process. Manage.* **57**, 102018 (2020)
3. Bender, J., Davenport, L., Fedler, F., Drager, M.: *Reporting for the Media*. Oxford University Press, Oxford (2012)
4. Browne, R.: ‘Junk news’ gets massive engagement on Facebook ahead of EU elections, study finds. CNBC (2019). <https://www.cnn.com/2019/05/21/junk-news-gets-higher-engagement-on-facebook-ahead-of-eu-elections.html>. Accessed 19 Apr 2021
5. Chowdhury, M.F.M., Lavelli, A.: Assessing the practical usability of an automatically annotated corpus. In: Proceedings of the 5th Linguistic Annotation Workshop, pp. 101–109. Association for Computational Linguistics, Portland, Oregon, USA, Jun 2011. <https://www.aclweb.org/anthology/W11-0412>
6. Cruz, A., Rocha, G., Sousa-Silva, R., Lopes Cardoso, H.: Team Fernando-Pessa at SemEval-2019 task 4: Back to basics in hyperpartisan news detection. In: Proceedings of the 13th International Workshop on Semantic Evaluation, pp. 999–1003. Association for Computational Linguistics, Minneapolis, Minnesota, USA, Jun 2019. <https://doi.org/10.18653/v1/S19-2173>
7. Devlin, J., Chang, M.W., Lee, K., Toutanova, K.: BERT: pre-training of deep bidirectional transformers for language understanding (2019)
8. Álvaro Figueira, Oliveira, L.: The current state of fake news: challenges and opportunities. *Procedia Computer Science* (2017). <https://doi.org/10.1016/j.procs.2017.11.106>
9. Flesch, R.: A new readability yardstick. *J. Appl. Psychol.* **32**(3), 221–233 (1948)
10. Gunning, R.: *The Technique of Clear Writing*. McGraw-Hill, New York (1952)
11. Hahn, U., Tomanek, K., Beisswanger, E., Faessler, E.: A proposal for a configurable silver standard. In: Proceedings of the Fourth Linguistic Annotation Workshop, pp. 235–242. Association for Computational Linguistics, Uppsala, Sweden, July 2010. <https://www.aclweb.org/anthology/W10-1838>

12. Hancock, J.T., Curry, L.E., Goorha, S., Woodworth, M.: On lying and being lied to: A linguistic analysis of deception in computer-mediated communication. *Discourse Process*. **45** (2007). <https://doi.org/10.1080/01638530701739181>
13. Harrower, T.: *Inside Reporting: A Practical Guide to the Craft of Journalism*. McGraw-Hill Companies, Incorporated (2007)
14. Horne, B.D., Adali, S.: This just. In: Fake news packs a lot in title, uses simpler, repetitive content in text body, more similar to satire than real news (2017)
15. Kincaid, J.P., Aagard, J.A., O'Hara, J.W.: Development and test of a computer readability editing system (CRES). Technical report, TRAINING ANALYSIS AND EVALUATION GROUP (NAVY) ORLANDO FL (1980)
16. Laboreiro, G., Oliveira, E.: What we can learn from looking at profanity, pp. 108–113 (2014). [https://doi.org/10.1007/978-3-319-09761-9\\_11](https://doi.org/10.1007/978-3-319-09761-9_11)
17. Laughlin, G.H.M.: Smog grading-a new readability formula. *J. Reading* **12**(8), 639–646 (1969). <http://www.jstor.org/stable/40011226>
18. Litvinova, O., Seredin, P., Litvinova, T., Lyell, J.: Deception detection in Russian texts. In: Proceedings of the Student Research Workshop at the 15th Conference of the European Chapter of the Association for Computational Linguistics (2017)
19. Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., Galstyan, A.: A survey on bias and fairness in machine learning. arXiv (2019). [arXiv:1908.09635](https://arxiv.org/abs/1908.09635)
20. Mourão, R.R., Robertson, C.T.: Fake news as discursive integration: an analysis of sites that publish false, misleading, hyperpartisan and sensational information. *Journalism Stud.* **20**(14), 2077–2095 (2019). <https://doi.org/10.1080/1461670X.2019.1566871>
21. Pérez-Rosas, V., Kleinberg, B., Lefevre, A., Mihalcea, R.: Automatic detection of fake news. arXiv preprint [arXiv:1708.07104](https://arxiv.org/abs/1708.07104) (2017)
22. Sousa-Silva, R.: Computational forensic linguistics: an overview of computational applications in forensic contexts. *Language and Law/Linguagem e Direito* **5**(2), 118–143 (2019)
23. Sullivan, M.: What it really means when trump calls a story 'fake news'. [https://www.washingtonpost.com/lifestyle/media/what-it-really-means-when-trump-calls-a-story-fake-news/2020/04/13/56fbc2c0-7d8c-11ea-9040-68981f488eed\\_story.html](https://www.washingtonpost.com/lifestyle/media/what-it-really-means-when-trump-calls-a-story-fake-news/2020/04/13/56fbc2c0-7d8c-11ea-9040-68981f488eed_story.html) (2020). Accessed 20 Apr 2021
24. Tandoc, E., Lim, Z., Ling, R.: Defining “fake news”: a typology of scholarly definitions. *Digital Journalism* **6** (2017). <https://doi.org/10.1080/21670811.2017.1360143>
25. Vorhaus, M.: People increasingly turn to social media for news. <https://www.forbes.com/sites/mikevorhaus/2020/06/24/people-increasingly-turn-to-social-media-for-news/> (2020). Accessed 5 Apr 2021
26. Weber, G.: Top languages. *The World's 10* (2008)