sciendo

Comparison of Supervised Learning Algorithms for Fake News Detection

Andreea – Oana RADU
Bucharest University of Economic Studies, Romania
raduandreea22@stud.ase.ro

Sergiu – Alexandru IONESCU
Bucharest University of Economic Studies, Romania
ionescusergiu14@stud.ase.ro

Abstract. *Today, national and international information and news have greatly increased in volume. Almost all people have access to everything they want through devices and the internet. Information is just a click away. But besides this very good aspect, there are also negative sides. A negative aspect, but also the most important, is the fact that news can be distorted by those who process it and post it online. Thus, you can read information, but it may not be the real one, thus becoming misinformed. Various researches have been carried out that have tried to discover methods and algorithms to detect this fake news, but not all algorithms used are effective in detecting whether online content is false or true. There are machine learning algorithms that can detect such news and tell whether the information is true or false. This paper aims to discover which supervised learning algorithm is the most efficient and best at detecting fake news from real ones. Thus, several supervised learning algorithms were trained on a dataset taken from Kaggle and finally the most efficient algorithm that can be used for fake news detection will be presented.*

Keywords: Supervised Learning, Big Data, Fake News, Detection, Algorithm.

Introduction

Nowadays, the more information increases in volume, the greater the danger that it will be modified and thus no longer be the real thing. News is increasingly posted on the internet, so a typical social media user cannot know how true or false it is.

To find out if a news item is true or false, a user can obtain information from several sources, but the question arises: are all the sources accessed by the user fake? Thus, artificial intelligence plays an important role in solving this dilemma and coming up with concrete solutions. Artificial intelligence has an important component, that of learning. This component is divided into three broad categories: supervised learning, unsupervised learning, and reinforcement learning (Iqbal, Shahzad, Khan, & Chaudhry, 2023).

In this paper, we aim to analyze in detail four supervised learning algorithms and apply them to a dataset of fake and true news taken from Kaggle in order to be able to detect fake news from real news so that the end user can be correctly informed. This includes an evaluation of the efficiency of the algorithms on the dataset with a focus on precision, recall, f1-score and accuracy.

The four algorithms that will be presented and tested in this research are: Random Forest Classifier, Logistic Regression, SVM and DeepText.

In this study we will answer the following research questions:

RQ1: What impact could the implementation of supervised learning algorithms have in the field of news on social platforms?

RQ2: Which algorithm among those studied is the most effective for detecting fake news?

DOI: 10.2478/picbe-2025-0166

Given the current trend of misinformation among social media users, this paper aims to add value to solving this problem by presenting the supervised learning algorithm that can most quickly help users know whether the information is real, truthful or false. It also makes a considerable contribution to understanding the importance of artificial intelligence and especially supervised learning in the context of the large volume of information in current social media platforms.

Literature review

Fake news detection is a topic of great interest these days in the field of machine learning. There are studies that analyze various algorithms for detecting fake or true news. These algorithms are specifically from supervised learning because they are based on classification algorithms that are easier to run and with fewer resources, all to improve the accuracy of fake news detection (Gongane, Munot, & Anuse, 2024), (Agarwala, Sultanaa, Malhotra, & Sarkar, 2019).

Because information is found on any search engine at any time, this means that we have a large volume of data that we need to analyze to clarify whether we are correctly informed or the information is distorted. In the context of big data, something automated is needed to analyze and process the data. Thus, the notion of artificial intelligence is essential to help in this context.

Artificial intelligence is a field that focuses on developing systems that are capable of imitating human intelligence. Thus, these systems can perform tasks that human intelligence can normally do, and here we can think of learning, analysis, decision-making, and information processing. Thus, an essential component in artificial intelligence is machine learning, which is divided into three branches: supervised learning, unsupervised learning, and reinforcement learning (Saravanan & Sujatha, 2018), (Obaido, și alții, 2024). In this paper we will focus on supervised learning.

We will introduce what supervised learning is. This is a type of machine learning that trains an algorithm on a set of labeled or classified data. It is very important that the data set is larger in order to be as accurate as possible in the predictions it makes. There are several classification algorithms, but this study focuses on the following four algorithms: Random Forest Classifier, Logistic Regression, SVM și DeepText.

Random Forest Classifier is an algorithm that is based on a large number of decision trees, each trained on a different subset of data. It is used for fraud detection or credit risk assessment but is also widely used in fake news detection (Schonlau & Zou, 2020), (Ren, Cheng, & Han, 2017).

Logistic Regression is used in binary classification problems. And this algorithm is also a classification algorithm and is used for detecting spam emails, in economics and can also be applied for detecting fake news (de Menezes, Liska, Cirillo, & Vivanco, 2016).

SVM stands for Support Vector Machine and is an algorithm used for classification and regression and is most commonly used in the context of large volumes of data. It is often found in problems of facial recognition and image or news classification (Bai & Zhang, 2009).

DeepText is an algorithm that is developed by Facebook and is based on deep neural networks used for natural language processing (NLP for short) (Zhong, Jin, & Huang, 2017).

To conduct the current study, we consulted various international databases such as Web of Science, Scopus, IEEE Xplorer and Google Scholar.

Methodology

For this research, the observation methodology was used to identify the data set necessary to conduct the study. Thus, two .csv files containing fake news and real news were chosen as the data

set. The goal of the research is to identify which learning algorithms are best to be trained and used in detecting real news from social media or fake news.

To compare the performance of supervised learning algorithms for fake news detection, various prototypes for the following algorithms were created for the datasets taken from Kaggle: Random Forest Clasifier, Logistic Regression, SVM și DeepText. Mathematical analysis methodology was also used to identify algorithms and their calculation methods.

The datasets were concatenated into a single file using the Python programming language in the Google Colab test environment. The data was processed and labeled as false or true, respectively 0 and 1. The text was cleaned by removing punctuation marks and transformed into numerical representations using TF-IDF (Term Frequency – Inverse Document Frequency). And the dataset after being cleaned was divided into two: 80% used for training the model and 20% for testing the model.

The implementation of the algorithms was carried out using the scikit-learn libraries for Logistic Regression, Random Forest and SVM and TensorFlow/Keras for the DeepText algorithm.

To evaluate the performance of the algorithms, we used metrics such as model accuracy to see the percentage of accuracy of the predictions, Precision, Recall and F1-score to assess the balance in identifying fake and true news, and the Confusion Matrix to analyze the classification errors of the model.

Finally, after running the experiments, we highlighted which is the most suitable algorithm for detecting fake news by comparing the performance metrics of the models.

The source code was inspired by a notebook available on kaggle under the Apache 2.0 license. We modified the code and added our own code for the SVM and DeepText algorithms. The code for this research can be found by accessing the following link: Source Code and data sources can be found by accessing the following link: Data Sources.

Algorithm presentation

In order to carry out the research, the emphasis was placed on knowing the four algorithms presented previously, both from a mathematical and an application point of view, or where they can be commonly used.

Random Forest Classifier Algorithm

It is a supervised learning algorithm that is used in both classification and regression problems. The prediction calculation formula is as follows:

$$\mathrm{y} = \arg max \sum_{t-1}^{T} 1(h_t(x) = c)$$

The terms used in the formula are: c which represents a possible class, 1(ht (x)=c) represents an indicator function which will take the value 1 if the tree t predicts class c and 0 otherwise, finally the number of times each class is predicted is counted and the class with the most votes will be chosen.

Logistic Regression Algorithm

It is an algorithm used for binary classification. It estimates the probability that an observation belongs to a particular class using the sigmoid function. The function is as follows:

$$P(y = 1|X) = \frac{1}{1 + e^{-z}}$$

The terms used are: $z = w_0+w_1x_1+ w_2x_2+...+ w_nx_n$ this is a linear combination of the predictors x_i, w_i are the model coefficients and e is the base of the natural logarithm.

SVM algorithm

SVM stands for Support Vector Machine and is a classification algorithm used in facial recognition or image classification problems. To mathematically calculate this algorithm, the equation of the hyperplane and the constraint or imposition of the class separation condition are needed.

Hyperplane equation:

$$w * x + b = 0$$

Class separation condition:

$$Y_i(w*x_i+b)\geq 1$$

The terms used are: w represents the weight vector, x represents the input data, b is the bias term, $w * x$ is the dot product between the vectors, and $y_i \varepsilon \{-1, 1\}$ represents the class label for point x_i.

DeepText algorithm

DeepText is an algorithm developed by Facebook that can be used in binary classification problems. It is designed to be used in fake news detection of content uploaded to the social media platform. The calculation formula is as follows:

$$L = -\frac{1}{N}\sum_{i=1}^{N}[y_i \log y_i + (1 - y_i)\log(1 - y_i)]$$

The terms used are: y_i represents the actual label 0 și 1, y_i is the probability predicted by the model and also uses Gradient Descent which is an optimizer for updating the weights.

Algorithm implementation

In order to develop and test algorithms, datasets were needed. These were downloaded from the Kaggle platform. The Kaggle platform is an online platform that provides a space for Data Science and provides essential resources for the development and application of machine learning algorithms. So we chose to test the four algorithms on datasets downloaded from Kaggel containing various news stories. There are two .csv files containing false information/fake news and also a .csv file containing true information. Through concatenation methods, we merged the two files into one using a DataFrame.

The tools used for data analysis, modelling, and visualization are Python libraries and frameworks. We used NumPy for numerical calculations within the learning algorithms, Pandas for data analysis, Matplotlib for diagrams, histograms and displaying the Confusion Matrix for each algorithm, Scikit-learn which is very useful for machine learning algorithms in Python, it contains classification, regression, clustering and many other aspects for analysis.

Tools such as supervised machine learning algorithms were used. They helped in analyzing the data and making predictions. The algorithms used are: Logistic Regression, Random Forest, SVM and DeepText.

With the help of learning algorithms, we wanted to classify or divide news into fake news and real news.

The data retrieved from the Kaggle platform is divided into two files in .csv format, one file contains fake news and one file contains true news. To test the algorithms, we loaded both files into two DataFrames. We then visualized the data in the DataFrame as shown in Figure 1.

	title	text	subject	date
0	Donald Trump Sends Out Embarrassing New Year'...	Donald Trump just couldn t wish all Americans ...	News	December 31, 2017
1	Drunk Bragging Trump Staffer Started Russian ...	House Intelligence Committee Chairman Devin Nu...	News	December 31, 2017
2	Sheriff David Clarke Becomes An Internet Joke...	On Friday, it was revealed that former Milwauk...	News	December 30, 2017
3	Trump Is So Obsessed He Even Has Obama's Name...	On Christmas day, Donald Trump announced that ...	News	December 29, 2017
4	Pope Francis Just Called Out Donald Trump Dur...	Pope Francis used his annual Christmas Day mes...	News	December 25, 2017
5	Racist Alabama Cops Brutalize Black Boy While...	The number of cases of cops brutalizing and ki...	News	December 25, 2017
6	Fresh Off The Golf Course, Trump Lashes Out A...	Donald Trump spent a good portion of his day a...	News	December 23, 2017
7	Trump Said Some INSANELY Racist Stuff Inside ...	In the wake of yet another court decision that...	News	December 23, 2017
8	Former CIA Director Slams Trump Over UN Bully...	Many people have raised the alarm regarding th...	News	December 22, 2017
9	WATCH: Brand-New Pro-Trump Ad Features So Muc...	Just when you might have thought we d get a br...	News	December 21, 2017

Figure 1. News in the dataset

Source: Authors' own research.

I noticed the organization of the news which is done on four columns, title, text, subject and date. Later I concatenated the two DataFrames into one so that the news is mixed, the fake ones with the real ones. The news indexes were reorganized to be done in chronological order. And in order to know which news is fake and which is true, we added a new column called "class" which took the value 1 for real news and the value 0 for fake news. All these details can be seen in Figure 2. We later exported the data to a .csv file so we could apply the algorithms to the new data.

	text	class
0	ROME (Reuters) - The Syrian conflict needs to ...	1
1	James Kallstrom is the former Assistant Direct...	0
2	PRINCETON, N.J. (Reuters) - A citizen's compla...	1
3	Conservative donors are shaking in their boots...	0
4	WASHINGTON (Reuters) - U.S. House of Represent...	1

Figure 2. Viewing data in DataFrame after concatenation

Source: Authors' own research.

After concatenating the two DataFrames, I used the process_text(text) function to process the text with the true and false news, thus converting all the text to lowercase, removing the text between the square brackets, we removed everything that means alphanumeric character, we removed URLs and HTML tags, we removed punctuation marks, we removed blank lines from the news, we removed words containing numbers and later we returned the processed and cleaned text. In Figure 3, you can see the news grouped by topics.

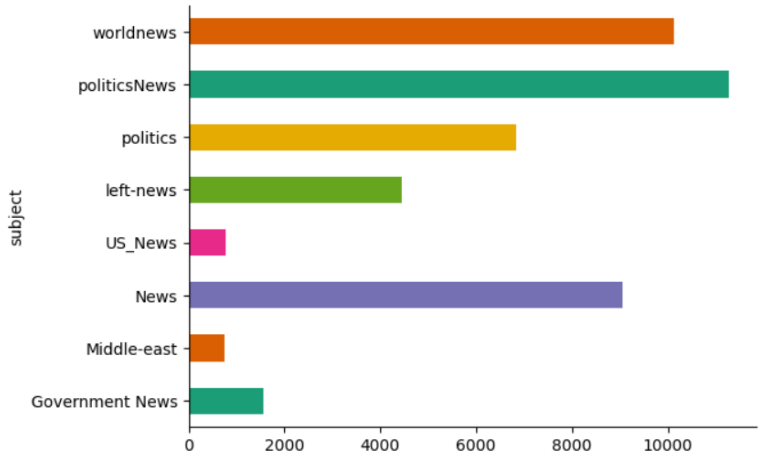

Figure 3. Grouping news by topics

Source: Authors' own research.

Before training the algorithms, we used the TfidfVectorizer class from the feature_extraction.text module of the scikit-learn library to transform the text into a TF-IDF term matrix. Basically, we transformed the text into vectors so we could apply machine learning algorithms.

To implement the logistic regression algorithm, we divided the data set into two, one part for training and one part for testing, we imported the LogisticRegression class from the sklearn.linear_model library, we initialized the regression model, trained the logistic regression model on the training data, and later tested the model on the test data and calculated the accuracy.

We trained the model with the logistic regression algorithm and the confusion matrix can be seen in Figure 4. Based on which it can be evaluated whether this algorithm is good or whether another classification algorithm should be chosen. It can be seen that for precision we have 0.99 for fake news and 0.98 for true news, which means that 99% of fake news predictions are correct and 98% of true news predictions are correct. The accuracy is 0.9844% which means that 98.44% of the news are classified correctly. Thus it can be seen that the model is very good, precise.

```
print(classification_report(y_test, prediction_lr))
```

```
              precision    recall  f1-score   support

           0       0.99      0.98      0.99     11731
           1       0.98      0.99      0.98     10698

    accuracy                           0.98     22429
   macro avg       0.98      0.98      0.98     22429
weighted avg       0.98      0.98      0.98     22429
```

Figure 4. Confusion Matrix – Logistic Regression

Source: Authors' own research.

To implement the random forest algorithm, we divided the data set into two, one part for training and one part for testing, we imported the RandomForestClassifier class from the

sklearn.ensemble library, we initialized the random forest model, we trained the random forest model on the training data and later we tested the model on the test data and calculated the accuracy.

We subsequently trained a model with the Random Forest algorithm, the results can be seen in Figure 5. It can be seen that for precision we have 0.99 for fake news and 0.99 for true news, which means that 99% of fake news predictions are correct and 99% of true news predictions are correct. The accuracy is 0.9880% which means that 98.80% of the news are classified correctly. So it can be seen that the model is very good, accurate. And it is better in this situation than the previous model based on the results.


```
print(classification_report(y_test, prediction_rfc))
```

	precision	recall	f1-score	support
0	0.99	0.99	0.99	11731
1	0.99	0.99	0.99	10698
accuracy			0.99	22429
macro avg	0.99	0.99	0.99	22429
weighted avg	0.99	0.99	0.99	22429

Figure 5. Confusion Matrix – Random Forest

Source: Authors' own research.

To implement the SVM algorithm, we divided the data set into two, one part for training and one part for testing. We imported the SVC class from the sklearn library, we initialized the SVC model, trained the SVC model on the training data, and then tested the model on the test data and calculated the accuracy.

Later, we tested the SVM algorithm on the same data set, which we have seen from research to be the most famous for classification and regression problems. In figure 6, we can see the result of running this algorithm. It can be seen that for precision we have 0.99 for fake news and 0.99 for true news, which means that 99% of fake news predictions are correct and 99% of true news predictions are correct. The accuracy is 0.9907% which means that 99.07% of the news are classified correctly. So it can be seen that the model is very good, accurate. And it is better in this situation than the previous model based on the results.

```
print(classification_report(y_test, prediction_svc))
```

	precision	recall	f1-score	support
0	0.99	0.99	0.99	11731
1	0.99	0.99	0.99	10698
accuracy			0.99	22429
macro avg	0.99	0.99	0.99	22429
weighted avg	0.99	0.99	0.99	22429

Figure 6. Confusion Matrix – SVM

Source: Authors' own research.

To create the DeepText algorithm, we imported TensorFlow, which is a widely used framework for machine learning but also for neural networks. We created several layers, an input layer, we used the Relu (Rectified Linear Unit) function to speed up the learning process of the network, we also created hidden layers for example Droppou(0.3) which means that it deactivates 30% of the neurons to prevent overfitting, this happens after each layer. There is also an output layer with a single neuron that will tell whether the news is false or true. The model was trained for 5 epochs with 100 examples (batches) in a single iteration.

The model was tested on the same dataset used in the previous algorithms. The results can be seen in Figure 7. It can be seen that for precision we have 0.99 for fake news and 0.99 for true news, which means that 99% of fake news predictions are correct and 99% of true news predictions are correct. The accuracy is 0.9891% which means that 98.91% of the news are classified correctly. Thus it can be seen that the model is very good, accurate. And it is better in this situation than the first two models but it is a little less efficient than the SVM algorithm.

```
print(classification_report(y_test, predictions))
```

	precision	recall	f1-score	support
0	0.99	0.99	0.99	11731
1	0.99	0.99	0.99	10698
accuracy			0.99	22429
macro avg	0.99	0.99	0.99	22429
weighted avg	0.99	0.99	0.99	22429

Figure 7. Confusion Matrix – DeepText

Source: Authors' own research.

Machine learning algorithms are very useful for detecting fake news. Search engines and social media platforms can use these algorithms to help stop misinformation or misinforming users. In response to RQ2, the impact would be very good because the population would be correctly informed and thus all kinds of false ideas related to certain political aspects or even everyday life would no longer spread.

Results and discussions

In this research, we trained four supervised learning algorithms. The algorithms were trained to detect fake news from real news on a dataset taken from Kaggle. We summarized the results in Table 1 for easy viewing.

Table 1. Confusion Matrix Results

Algorithms	*Precision*		*Recall*		F1-score		Accuracy
	Fake	*True*	*Fake*	*True*	*Fake*	*True*	
Logistic Regression	0.99	0.98	0.98	0.99	0.99	0.98	0.9844
Random Forest	0.99	0.99	0.99	0.99	0.99	0.99	0.9880
SVM	0.99	0.99	0.99	0.99	0.99	0.99	0.9907
DeepText	0.99	0.99	0.99	0.99	0.99	0.99	0.9891

Source: Authors' own research.

The logistic regression model, following the analysis of the performance indicators, can be seen from the table that 0.9844% of the news were classified correctly using this model, thus the algorithm proves to be very good for detecting fake vs. true news.

The second trained algorithm, Random Forest, after testing and viewing the model's performance metrics, can be seen from the table that it managed to present an accuracy of 0.9980%, so it can be said that 99.80% of the news was classified correctly. Thus, this model can be seen to be more accurate in classifying both types of news and can be said to have superior performance to the previously presented algorithm.

The third trained algorithm, the SVM, presented the best performance among all four trained and tested models, presenting a precision of 0.99 for both news categories and the accuracy was 0.9907%, which means that 99.07% of the news were classified correctly.

The fourth trained algorithm DeepTest showed very good performance for fake news detection, it showed a precision of 0.99 for both news categories, just like the SVM, but the accuracy is slightly weaker at 0.9891%, which means that 98.91% of the news was categorized correctly.

Thus, it can be said after researching these algorithms that the SVM is the most effective for detecting fake news from true news using the hardware technical characteristics that we set in the collaboration and training the Facebook model with only 5 training epochs.

The results obtained from training the four models show that they all performed very well, with the SVM managing to prove itself to be better than the others through its accuracy score which was slightly higher than the accuracy of the other models. In response to RQ3, the best and most efficient algorithm among those analyzed proved to be the SVM, which achieved the highest accuracy score, thus for verifying fake and true news, the best model is the SVM. Given these results, it can be concluded that machine learning approaches are effective in correctly identifying and classifying news in the online environment.

Future Directions

In the future, we want to implement algorithms used by various search engines or social media platforms, such as Google's BERT, to test the accuracy of fake news detection. This algorithm is widely used by Google for detecting fake news.

We also noticed that running the algorithms on the Colab platform is very slow because hardware resources are limited. Thus, in the future, we want to implement containers that help run various algorithms in order to be able to do as many tests as possible and thus most accurately identify which algorithm has the best detection for news.

Conclusion

This research aimed to present four supervised learning algorithms that can be used in classification problems. These algorithms were implemented in python and tested on two datasets taken from Kaggle containing both fake and true news. Thus, algorithms were taught how to classify news from various social media platforms or search engines in order to detect fake news and thus be able to delete it from social media platforms.

Given that social networks are widely used and information spreads extremely quickly through them, this study comes up with the solution of implementing one of the presented algorithms or even a combination of several algorithms so that users are correctly informed and there is no more false information in the online environment.

Thus, on a data set of the size of the one in the present research, we can say from our study that the SVM algorithm followed by DeepText are the most effective in detecting fake news and can therefore be implemented in social networks or search engines.

In conclusion, the research shows that machine learning approaches, especially SVM and DeepText algorithms, are effective in detecting and correctly classifying news, thus helping to combat disinformation in the online environment, such as social networks or search and internet browsing engines.

References

Agarwala, V. (., Sultanaa, H. (., Malhotra, S. (., & Sarkar, A. (. (2019). Analysis of Classifiers for Fake News Detection. *2ND INTERNATIONAL CONFERENCE ON RECENT TRENDS IN ADVANCED COMPUTING ICRTAC -DISRUP - TIV INNOVATION , 2019* (pp. 377-383). Vellore Inst Technol, Chennai Campus, Chennai, INDIA: Procedia Computer Science.

Bai, X. (., & Zhang, Y. (. (2009). The Research on an Improved Fast SVM Classification Algorithm. *2nd International Symposium on Computational Intelligence and Design* (pp. 300-303). Changsha, PEOPLES R CHINA: SECOND INTERNATIONAL SYMPOSIUM ON COMPUTATIONAL INTELLIGENCE AND DESIGN, VOL 2, PROCEEDINGS.

de Menezes, F. (., Liska, G. (., Cirillo, M. (., & Vivanco, M. (. (2016). Data classification with binary response through the Boosting algorithm and logistic regression. *EXPERT SYSTEMS WITH APPLICATIONS*, 62-73.

Gongane, V. (., Munot, M. (., & Anuse, A. (. (2024). A survey of explainable AI techniques for detection of fake news and hate speech on social media platforms. *JOURNAL OF COMPUTATIONAL SOCIAL SCIENCE*.

Iqbal, A. (., Shahzad, K. (., Khan, S. (., & Chaudhry, M. (. (2023). The relationship of artificial intelligence (AI) with fake news detection (FND): a systematic literature review. *Global Knowledge Memory and Communication*.

Obaido, G. (., Mienye, I. (., Egbelowo, O. (., Emmanuel, I. (., Ogunleye, A. (., Ogbuokiri, B. (., . . . Aruleba, K. (. (2024). Supervised machine learning in drug discovery and development: Algorithms, applications, challenges, and prospects. *MACHINE LEARNING WITH APPLICATIONS Volume 17*.

PRISMA. (2020, March 27). *PRISMA 2020 flow diagram*. Retrieved from PRISMA statement: https://www.prisma-statement.org/prisma-2020-flow-diagram

Ren, Q. (., Cheng, H. (., & Han, H. (. (2017). Research on Machine Learning Framework Based on Random Forest Algorithm. *International Conference on Advances in Materials, Machinery, Electronics (AMME)*. Wuhan, PEOPLES R CHINA: AIP Conference Proceedings.

Saravanan, R. (., & Sujatha, P. (. (2018). A State of Art Techniques on Machine Learning Algorithms: A Perspective of Supervised Learning Approaches in Data Classification. *Proceedings of the 2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS)* (pp. 945-949). Vaigai Coll Engn, Madurai, INDIA: IEEE.

Schonlau, M. (., & Zou, R. (. (2020). The random forest algorithm for statistical learning. *Stata Journal*, 3-29.

Zhong, Z., Jin, L., & Huang, S. (2017). DeepText: A new approach for text proposal generation and text detection in natural images. *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. New Orleans, LA, USA: IEEE Xplore.

This paper was co-financed by The Bucharest University of Economic Studies during the PhD program.