

DNN based fake news identification and analysis

Dr.S.Suresh ¹, Rachakonda Chandini Devi ², Chiluvuri Reshma ³, Grandhi Harsha Nandini ⁴, Punyamanthula Veera Sai Akshitha ⁵, Kola Venkatesh ⁶

ASSOCIATE PROFESSOR

DEPT OF COMPUTER SCIENCE AND ENGINEERING

PRAGATI ENGINEERING COLLEGE(A), SURAMPALEM(EAST GODAVARI) A.P, INDIA

To Cite this Article

Dr.S.Suresh , Rachakonda Chandini Devi , Chiluvuri Reshma , Grandhi Harsha Nandini , Punyamanthula Veera Sai Akshitha , Kola Venkatesh, **DNN based fake news identification and analysis** ” *Journal of Science and Technology*, Vol. 08, Issue 04, -April 2023, pp32-39

Article Info

Received: 22-02-2023

Revised: 15-03-2023

Accepted: 22-03-2023

Published: 14-04-2023

ABSTRACT:

In this project we show an approach for detecting fake statements made by public figures by means of artificial intelligence. Several approaches were implemented as a software system and tested against a data set of statements. The best achieved result in binary classification problem (true or false statement) is 86%. The results may be improved in several ways that are described in the article as well. The progress in modern informational technologies brings us to the era where information is as accessible as ever. It is possible to find the answers to the questions we are interested in a matter of seconds. Availability of mobile devices makes it even more convenient for the users. This factor changed the way of how people get the news information a lot. Every mainstream mass media has its own online portal, Facebook account, Twitter account etc., so people can access news information really quickly.

I. INTRODUCTION

The progress in modern informational technologies brings us to the era where information is as accessible as ever. It is possible to find the answers to the questions we are interested in in a matter of seconds. Availability of mobile devices makes it even more convenient for the users. This factor changed the way of how people get the news information a lot. Every mainstream mass media has its own online portal, Facebook account, Twitter account etc., so people can access news information really quickly. Unfortunately, the news information that we get is not always true. Paradoxically, the Internet makes it harder to fact check the available information, because there are too many sources that often even contradict each other. All of this caused the emergence of fake news.

Mass media and social media have a great influence on a public. There are sides that are interested in using this to achieve their political goals with the help of fake news. They provide false information in form of news to manipulate people in different ways. There exist lots of websites with a single purpose of spreading of false information. They publish fake news, propaganda materials, hoaxes, conspiracy theories in disguise of a real news information. The main purpose of fake news websites is to affect the public opinion on certain matters. Examples of this may be found in Ukraine, United States of America, Great Britain, Russia and many other countries. Thus, fake news is a global issue and an important challenge to tackle. There is a belief that fake news problem may be solved automatically, without human interference, by means of artificial intelligence. This cause by the rise of deep learning and other artificial intelligence techniques showed us that they can be very effective in solving complex, sometimes even non-formal classification tasks. This article describes a way for classification of short political statements by means of artificial intelligence. Several approaches were implemented and tested on a data set of a statement made by real-life politicians The data set that was used for training and testing was collected by a RAMP studio team. It contains of short statements made by famous public figures. Two possible labels were available for the statement. They are:

1)'False'

2)'True'

Each entry in the data set, besides the statement itself, also contains a lot of metadata. It contains the date when the statement was made, the job of the public figure who made that statement, the source where the statement was taken from, some keywords that characterize the content of the statement and many more other features. The data set consists of 10460 entries in total (7569 of them were provided for training and 2891 for testing). There are more than 2000 different sources of the statements. The RAMP studio team collected the data set using PolitiFact website. The PolitiFact is a project operated by Tampa Bay Times in which reporters from the Times and affiliated media factcheck statements by members of the United States Congress, the White House, lobbyists and interests' groups. They publish original statements and their evaluations on the PolitiFact.com website, and assign exact h a "Truth-O-Meter" rating . PolitiFact.com was awarded the Pulitzer Prize for National Reporting in 2009 for "its fact-checking initiative during the 2008 presidential campaign that used probing reporters and the power of the World Wide Web to examine more than 750 political claims, separating rhetoric from truth to enlighten voters". At some points PolitiFact was criticized by both liberal and conservative wings of American politics, but nevertheless it is a viable source of fact-checked information. This makes a data set useful for creating a system which will classify statements as true or false. Before actually applying the artificial intelligence algorithms to the data, it should be pre-processed. First of all, it was decided to use only the statements themselves for classification purposes. This means that none of the metadata provided is used for classification. The classification algorithm might actually be improved in the future by taking into account this metadata. The steps that were used for the pre-processing are the following:

1) Splitting the statements into separate tokens (words).

2) Removing all numbers.

3) Removing all punctuation marks.

4) Remove all other non-alpha characters

5) Applying the stemming procedure to the rest of the tokens.

In linguistic morphology and information retrieval, stemming (or lemmatization) is the process of reducing inflected or derived words to their word stem, base or root form – generally a written word form. This helps to treat similar words (like “write” and “writing”) as the same words and might be extremely helpful for classification purposes. Removing stop words. Stop words are the words occur in basically all types of texts. These words are common and they do not really affect the meaning of the textual information, so it might be useful to get rid of them. Substitution of words with their tf-idf scores. In information retrieval, tf-idf, which is a short for term frequency–inverse document frequency”, is a numerical statistic measure reflects the importance of a certain word to a document in a collection or corpus. The tf-idf value increases proportionally to the number of times a word appears in the document and decreases proportionally to the frequency of the word in the corpus, which helps to adjust for the fact that some words appear more frequently in general. According to tf-idf, the weight of a term that occurs in a document is proportional to its frequency, and the specificity of a term can be calculated as an inverse function of the number of documents that contain the specified term. Several artificial intelligence algorithms were used for statement classification. All of them are implemented by scikit-learn (a library for Python programming language). For all of the algorithms two different metrics were measured:

- Classification accuracy based on six categories available
- Binary classification accuracy. This metric counts the

accuracy as if there were only 2 possible categories for the statement – true (based on the last three categories described above) and false (based on the first three categories described above) For all of the methods the provided data set with known labels was split into training and validation data sets. The training data set was used for the actual process of training of the machine learning models. The validation data set was used or some very basic model tuning. The idea is that having a validation data set we can iteratively tune the machine learning model by repeating the following process: 1. Change a subset of machine learning model meta parameters.

2. Train it on the training data set.

3. Measure its performance on the validation data set.

In the end, usually the model, which performed the best on the validation data set, is chosen as a final model. Its performance on the testing data set is considered as an unbalanced estimate of how well the model performs on previously unseen data.

A. Classification with logistic regression Logistic regression is a statistical method for analyzing a data set in which there are one or more independent variables that determine an outcome. The outcome is measured with a dichotomous variable (in which there are only two possible outcomes). For the cases when there are more than two labels, the strategy, which is called “One versus all”, is used. In this strategy every category is binary classified against its inverse a fictional category that states that the example does not belong to the current category). The category with the highest score is picked as a result of a classification. Logistic regression is one of the simplest machine learning techniques. It is usually a good idea to implement logistic regression classifier before proceeding with a more complex approach because it gives you an estimate of how well machine learning algorithms will perform on this specific task. The results that were achieved for logistic regression classifier are the following:

- classification accuracy - 72%
- binary classification accuracy - 75%

B. Classification with naive Bayes classifier in artificial intelligence, naive Bayes classifiers are a family of simple probabilistic classifiers based on applying Bayes theorem with strong independence assumptions between the features. Naive Bayes is a simple technique for constructing classifiers: models that assign class labels to problem instances, represented as vectors of feature values, where the class labels are drawn from some finite set. It is not a single algorithm for training such classifiers, but a family of algorithms based on a common principle: all naive Bayes classifiers assume that the value of a particular feature is independent of the value of any other feature, given the class variable. They were invented in the middle of the 90s and they were widely used for classification of e-mails as spam or not spam. Naive Bayes typically use bag of words features to classify texts. Naive Bayes classifiers usually correlate the use of tokens, with the classes that are used for classification, and then apply Bayes theorem to calculate a probability that a text belonging to a certain class. Using naive Bayes classifier is easy to use for both binary and multi-label classification. For the task, described in the paper it is possible to calculate probabilities of the fact that each given statement belongs to the specific group.

II. LITERATURE SURVEY

"A Fake News Detection Using Naive Bayes Classifier "

AUTHORS: M. Granik, V. Mesyura, Year: 2017

This paper shows a simple approach for fake news detection using naive Bayes classifier. This approach was implemented as a software system and tested against a data set of Facebook news posts. We achieved classification accuracy of approximately 74% on the test set which is a decent result considering the relative simplicity of the model. These results may be improved in several ways that are described in the article as well. Received results suggest, that fake news detection problem can be addressed with artificial intelligence methods. The news information can be easily accessed through Internet and social media. It is convenient for user to follow their interest events available in online mode. Mass-media playing a huge role in influencing the society and as it is common, some people try to take advantage of it. Sometimes mass media modulate the information in their own way to reach their goal. There are many websites which provide false information. They consciously try to bring out propaganda, hoaxes and misinformation under the guise of being authentic news.

"The Principles of The Truth-O-Meter: PolitiFact’s Methodology For Independent Fact-Checking "

AUTHORS: Angie Drobnic Holan, Year: 2022

Fact-checking journalism is the heart of PolitiFact. Our core principles are independence, transparency, fairness, thorough reporting and clear writing. The reason we publish is to give citizens the information they need to govern themselves in a democracy. Since our launch in 2007, we've received many questions about how we choose facts to check, how we stay nonpartisan, how we go about fact-checking and other topics. This document attempts to answer those questions and many more. PolitiFact started in 2007 as an election-year project of the Tampa Bay Times (then named the St. Petersburg Times), Florida's largest daily newspaper. From the beginning, PolitiFact focused on looking at specific statements made by politicians and rating them for accuracy. PolitiFact is run by the editors and journalists who make up the PolitiFact team. No one tells us what to write about or how to rate statements. We do so independently, using our news judgment. PolitiFact is owned by the nonprofit Poynter Institute for Media Studies. PolitiFact had been owned by the Tampa Bay Times, but in 2018 direct ownership of PolitiFact was transferred from the Times to Poynter, which is the newspaper's parent company. The move allows PolitiFact to function fully as not-for-profit national news organization.

"A Statistical Interpretation Of Term Specificity And Its Application In Retrieval "

AUTHORS: Sparck Jones, K, Year: 1972

The exhaustivity of document descriptions and the specificity of index terms are usually regarded as independent. It is suggested that specificity should be interpreted statistically, as a function of term use rather than of term meaning. The effects on retrieval of variations in term specificity are examined, experiments with three test collections showing, in particular, that frequently-occurring terms are required for good overall performance. It is argued that terms should be weighted according to collection frequency, so that matches on less frequent, more specific, terms are of greater value than matches on frequent terms. Results for the test collections show that considerable improvements in performance are obtained with this very simple procedure. We are familiar with the notions of exhaustivity and specificity: exhaustivity is a property of index descriptions, and specificity one of index terms. They are most clearly illustrated by a simple keyword or descriptor system. In this case the exhaustivity of a document description is the coverage of its various topics given by the terms assigned to it; and the specificity of an individual term is the level of detail at which a given concept is represented. These features of a document retrieval system have been discussed by Cleverdon et al. (1966) and Lancaster (1968), for example, and the effects of variation in either have been noted. For instance, if the exhaustivity of a document description is increased by the assignment of more terms, when the number of terms in the indexing vocabulary is constant, the chance of the document matching a request is increased. The idea of an optimum level of indexing exhaustivity for a given document collection then follows: the average number of descriptors per document should be adjusted so that, hopefully, the chances of requests matching relevant documents are maximized, while too many false drops are avoided. Exhaustivity obviously applies to requests too, and one function of a search strategy is to vary request exhaustivity. I will be mainly concerned here, however, with document descriptions.

"Artificial Neural Networks As Models Of Neural Information Processing"

Authors: Marcel van Gerven, and Sander Bohte, Year: 2017

In artificial intelligence (AI), new advances make it possible that artificial neural networks (ANNs) learn to solve complex problems in a reasonable amount of time (LeCun et al., 2015). To the computational neuroscientist, ANNs are theoretical vehicles that aid in the understanding of neural information processing (Van Gerven). These networks can take the form of the rate-based models that are used in AI or more biologically plausible models that make use of spiking neurons (Brette, 2015). The objective of this special issue is to explore the use of ANNs in the context of computational neuroscience from various perspectives. Biological plausibility is an important topic in neural networks research. That is, are ANNs simply convenient computational models or do they also inform about the computations that take place in our own brains? Marble stone et al. carefully lay out the rapid advances in deep learning and contrast these developments with current practice and views in neuroscience. Their main insight is that biological learning may be driven by the optimization of cost functions using successive neural network layers. The

algorithm computes the gradient of an objective function without relying on separate circuits for error propagation that integrate non-local signals. While acetylcholine (ACh) and dopamine (DA) are neuromodulators that are known to have profound and lasting effects on the neural responses to stimuli, it is unknown what their respective functional roles are. Holca-lamarre et al. develop a neural network model that is combined with the physiological release schedules of ACh and DA.

III.SYSTEM ANALYSIS

1.1. EXISTING SYSTEM

The news information that we get is not always true. Paradoxically, the Internet makes it harder to fact-check the available information, because there are too many sources that often even contradict each other. All of this caused the emergence of fake news. Mass media and social media have a great influence on a public. There are sides that are interested in using this to achieve their political goals with the help of fake news. They provide false information in form of news to manipulate people in different ways. There exist lots of websites with a single purpose of spreading of false information. They publish fake news, propaganda materials, hoaxes, conspiracy theories in disguise of a real news information.

DRAWBACKS OF EXISTING SYSTEM

- The main purpose of fake news websites is to affect the public opinion on certain matters (mostly political). Examples of this may be found in Ukraine, United States of America, Great Britain, Russia and many other countries. Thus, fake news is a global issue and an important challenge to tackle.
- This cause by the rise of deep learning and other artificial intelligence techniques showed us that they can be very effective in solving complex, sometimes even non-formal classification tasks.

1.2. PROPOSED SYSTEM

First of all, it was decided to use only the statements themselves for classification purposes. This means that none of the metadata provided is used for classification. The classification algorithm might actually be improved in the future by taking into account this metadata. Splitting the statements into separate tokens (words). Removing all numbers. Removing all punctuation marks. Remove all other non-alpha characters Applying the stemming procedure to the rest of the tokens. In linguistic morphology and information retrieval, stemming (or lemmatization) is the process of reducing inflected or derived words to their word stem, base or root form – generally a written word form. This helps to treat similar words (like “write” and “writing”) as the same words.

ADVANTAGES OF PROPOSED SYSTEM:

- Stop words are the words occur in basically all types of texts. These words are common and they do not really affect the meaning of the textual information.
- Term frequency–inverse document frequency is a numerical statistic measure reflects the importance of a certain word to a document in a collection or corpus

IV.SYSTEM DESIGN

4.1 SYSTEM ARCHITECTURE

Below diagram depicts the whole system architecture of fake news detection using artificial intelligence. A system architecture is the conceptual model that defines the structure, behavior, and more views of a system. An architecture description is a formal description and representation of a system, organized in a way that supports reasoning about the structures and behaviors of the system. Three types of system architectures are identified, integrated, distributed and mixed, (partly integrated and partly distributed). It is shown that the type of interfaces defines the type of

architecture. Integrated systems have more interfaces, which furthermore are vaguely defined. The architecture of a system describes its major components, their relationships (structures), and how they interact with each other. Software architecture and design includes several contributory factors such as Business strategy, quality attributes, human dynamics, design, and IT environment.

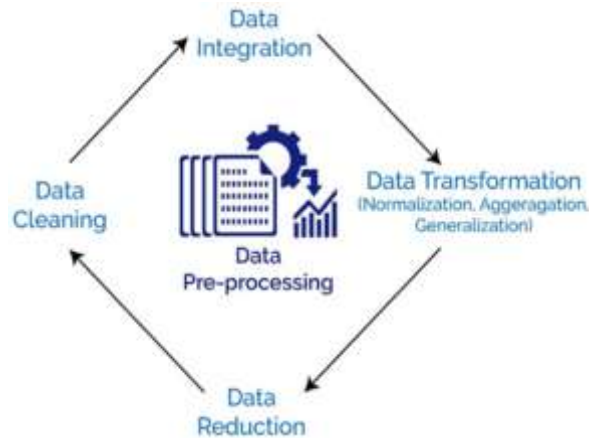


Fig. 4.1.1 System Architecture
V.SYSTEM IMPLEMENTATION

5.1. MODULES

- Data Pre-Processing
- Classification With Logistic Regression
- Classification With Naive Bayes Classifier
- Classification With Support Vector Machine

Module description:

5.1.1 Data Pre-Processing

Before actually applying the artificial intelligence algorithms to the data, it should be pre-processed. First of all, it was decided to use only the statements themselves for classification purposes. This means that none of the metadata provided is used for classification. The classification algorithm might actually be improved in the future by taking into account this metadata

5.1.2 Classification with Logistic Regression

Logistic regression is a statistical method for analyzing a data set in which there are one or more independent variables that determine an outcome. The outcome is measured with a dichotomous variable (in which there are only two possible outcomes). For the cases when there are more than two labels, the strategy, which is called “One versus all”, is used. In this strategy every category is binary classified against its inverse (a fictional category that states that the example does not belong to the current category). The category with the highest score is picked as a result of a classification.

5.1.3 Classification with Naive Bayes Classifier

In artificial intelligence, naive Bayes classifiers are a family of simple probabilistic classifiers based on applying Bayes theorem with strong (naive) independence assumptions between the features. Naive Bayes is a simple technique for constructing classifiers: models that assign class labels to problem instances, represented as vectors of feature values, where the class labels are drawn from some finite set. It is not a single algorithm for training such classifiers, but a family of algorithms based on a common principle: all naive Bayes classifiers assume that the value of a particular feature is independent of the value of any other feature, given the class variable.

5.1.4 Classification with Support Vector Machine

In machine learning, support vector machines are supervised learning models with associated learning algorithms that analyze data used for classification and regression analysis. A support vector machine model is a

representation of the examples as points in space mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible. New examples are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall. In the classification tasks for the cases when there are more than two labels, “One versus all” strategy is used

VI.RESULTS

S.No	Name	Email	Label	Probability Score	Statement	Statement Date
1	shaan	arunkata@gmail.com	Parts on Fire	0.43 %	says a u.s supreme court justice suggested that some u.s cases will be decided based on south african law	Sept. 13, 2019
2	shaan	arunkata@gmail.com	Mostly True	0.63 %	the national aeronautics and space administration of the united states will be joining efforts to locate and communicate with vikram the illated chandrayaan2 lander that went silent during its descent on to the moon last weekend	Sept. 13, 2019
3	shaan	arunkata@gmail.com	Half-True	0.63 %	the national aeronautics and space administration of the united states will be joining efforts to locate and communicate with vikram the illated chandrayaan2 lander that went silent during its descent on to the moon last weekend	Sept. 13, 2019
4	meghana	meghanfor@gmail.com	Mostly True	0.62 %	the seat-sharing talks between the two parties are likely to conclude in a day or two a big source said.	Sept. 13, 2019

Fig. 6.1 Output Screen for fake statement analysis

VI.CONCLUSION AND FUTURE WORK

In this paper, several algorithms for classifying statements made by public figures were implemented unsurprisingly, deep neural networks showed the best results both in classification accuracy based on six categories and binary classification. This encourages future research with extensive usage of deep neural networks. Achieved results might be significantly improved. It is possible to both improve the data which is used for training as well as the machine learning models themselves. This might be a subject for future research. Together with the text summarization (the problem that also can be solved by means of artificial intelligence), this approach might be used for classification of news articles as fake or true. This might also be a subject for future research.

REFERENCES :

- [1] M. Granik, V. Mesyura, "Fake news detection using naive Bayes classifier," 2017 IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON), Kiev, 2017, pp. 900- 903.
- [2] Cade Metz. (2016, Dec. 16). "The bittersweet sweepstakes to build an AI that destroys fake news". Available: <https://www.wired.com/2016/12/bittersweet-sweepstakes-build-aidestroys-fake-news/>
- [3] Fake news RAMP: classify statements of public figures. (n.d.) [Online]. Available: https://www.ramp.studio/problems/fake_news
- [4] The Principles of the Truth-O-Meter: PolitiFact’s methodology for independent fact-checking. (2018, Feb. 12) <http://www.politifact.com/truth-o-meter/article/2018/feb/12/principlestruth-o-meter-politifact-methodology-i/>. Accessed Mar. 24, 2018.
- [5] Rajaraman, A.; Ullman, J. D. "Data Mining". (2011) Available: <http://i.stanford.edu/~ullman/mmds/ch1.pdf>. Accessed Mar. 24, 2018.
- [6] Stemming and lemmatization. (n.d.) Available: <https://nlp.stanford.edu/IR-book/html/htmledition/stemming-andlemmatization-1.html>. Accessed Mar. 24, 2018
- [7] Sparck Jones, K. (1972). "A Statistical Interpretation of Term Specificity and Its Application in Retrieval". Journal of Documentation, vol 28, pp 11–21.
- [8] David A. Freedman (2009). “Statistical Models: Theory and Practice”. Cambridge University Press. p. 128.

-
- [9] Ho, Tin Kam (1995). "Random Decision Forests". Proceedings of the 3rd International Conference on Document Analysis and Recognition, Montreal, QC, 14–16 August 1995. pp. 278–282.
- [10] Cortes, Corinna; Vapnik, Vladimir N. (1995). "Support-vector networks". *Machine Learning*, vol 20, pp 273–297.
- [11] Artificial Neural Networks as Models of Neural Information Processing. *Frontiers Research Topic* (n.d.). <https://www.frontiersin.org/researchtopics/4817/artificial-neural-networks-as-models-of-neural-informationprocessing>. Accessed Mar. 24, 2018.
- [12] Opitz, D.; Maclin, R. (1999). "Popular ensemble methods: An empirical study". *Journal of Artificial Intelligence Research*. 11: 169.