OPTIMIZATION OF NEURAL NETWORK FORNEWS CLASSIFICATION

Under the guidance of Mr. T. Thirupathitrao, MTech,
P.Kiranmai, V.Navya, T.Ayyappa, P.Bhanu Pranay

LIET,JNTU-GV, AP
Tokenization, Activation Function,Convolutional Neural Network.

ABSTRACT: The evolution of the information and communication technologies hasdramatically increased the number of people with access of Internet, which lead to increase fake news. This amplified the old problem of fake news, which became amajor concern now-a-days due to the negative impact it brings to the communities. In order to tackle the rise and spreading of fake news. We created a machine learning model using verified dataset from Kaggle which classifies the data using term frequency and inverse document frequency to predict whether thegiven news is fake or not. For optimizing the output, now we are building a new deeplearning model and will compare theaccuracy between the previous Multinomial Naive Bayes and current Neural Network. For testing this model we will create an integration of the project into web to test and produce the output.

Key Words: Fake news detection,Deep learning, Neural Network

Introduction

Social media's explosive growth has made it possible for people to disseminate information quickly, cheaply, and with fewer checks and filters than ever before. This amplified the old problem of fake news, which became a major concern now-a-days due to the negative impact it brings to the communities. People all across the world should be grateful for the world wideweb significant contribution to sharing of information and transmission in making modern life so comfortable. Without a doubt, the internet has made life easier andaccess to a wealth of information possible. This is a development in human history, butit also blurs the distinction between legitimate media and media that has been purposefully falsified. Today, anyone may publish content for the internet, regardless of how reliable it is. Fake news, particularlyon social media, sadly attracts a lot of attention online. People fall for deception and don't hesitate to distribute such inaccurate information to the other side of the globe. Such information quickly fades away.

1. Literature Survey

There are numerous tools available for identifying fake news. To do that, we analyse data using various classifiers in various study papers. Random Forest,CNN, SVM, KNN, Logistic Regression,Naive Bayes, Long Short Term Memory, and SGD are the predictors. The accuracy achieved with Random Forest is 83%, the accuracy with CNN is 97%, the accuracy with SVM is 94%, the accuracy with KNNis 79%, the accuracy with Logistic Regression is 97%, the accuracy with NaiveBayes is 90%, the accuracy with Long Short Term Memory is 97%, the accuracy with SVM &NB combined is 78%, and theaccuracy with SGD is 77.2% ,compared to all CNN,LR and LSTM obtains highaccuracy.

2. Related Work

Earlier, to stop and eradicate it is Email spam, better known as unwanted email messages, is the practice of sending unsolicited electronic messages with different intentions, commonly commercialpurposes, or trying to commit criminal actions. Despite the numerous anti-spammeasures nowadays, spam still being a problem all over the internet due to the low-cost and high impact that represents elaborate a spam campaign. Many differentsolutions exist to categorize incomingmessages such as white list, grey list, blacklist, Machine Learning, Rule-based filtering, etc. However, no one definitively.A possible reason is since spammers are high resilient, once a spam filtering methodis compromised spammers adapt to it. The aim of the present work has the objective ofdetecting in a more effective way spam email with the Multinomial Naïve Bayes approach, in addition to text sanitation. Results given by the proposed model givesan accuracy improve than Multinomial Naïve Bayes by its own.

2.1 Multinomial Naïve Bayes Classifier

Multinomial (MB) NB is one of the three types of NB models. Multinomial (MB) is used for discrete counts. Applies when the data features have discrete frequency counts. (In our exercise we count the wordsin the body of the email), then is necessary to use word counts in the email's body.Considering
X = [x1,x2,x3…..Xn] as vector of n distinctfeature in the dataset, and V = [v1,v2,v3,v4, ……. Vn] express all possible classes belonging to NB Multinomial model asa division of both the vectors of X&V respectively.

2.2CNN and LSTM

To combat the automatic false news identification from text, several networkdesigns have been explored. The detection of fake news in texts with political statements was done using convolutional neural networks. Moreover, writers utilizedmetadata to describe themselves or to provide details on claims made by other authors. The authors employed a convolutional layer of a neural network to capture the dependency between the metadata vectors and randomly initialized an embedding vector matrix to encapsulatethe data and metadata. The LSTM layer of the recurrent neural network was then applied, followed by a maximum association operation in the latent space. Tocreate the final prediction, the author integrated the representations of the texts and the metadata representation from LSTM into a fully linked layer. The embedding was made using the Word2vec tool. The creators of also used the Capsule neural network to identify bogus news. By giving each source and destination node distinct features, they were able to enhancethe traditional CNN and RNN. The approach is used to recognise bogus news stories of various lengths. The authors employed two distinct designs based on thedimensions of the parts. Pre-trained vectorsare used by the model to start learning. Theauthors created a framework for the brief texts that has layers just like the first model,but only two parallel networks are taken into account. Static word embedding, whichrepresents pre-learned vectors kept static throughout training, is utilised in this model. Only other parameters are trained. Using a non-static word embedding, themodel with medium and long texts had the greatest accuracy (99.8%). While it was more challenging to identify fake communications, the model with brief sentences was nevertheless tested using metadata.

3. Existing System

Parts-of-Speech (POS) labelling and PCFG by Conroy, Rubin, and Chen(Probabilistic Context Free Grammars). Feng, Banerjee, and Choi: by Feng and Hirst. Rubin, Lukoianova, andGTatiana: Use a vector space model to analyse rhetorical construction. Previously,we used the Multinomial Naive

Bayes method to construct a model using machine learning. By creating a model, wewere able to categorise the data using multinomial naive bayes classification.Machine learning techniques for fraud detection have been the subject of much study, with the majority of it concentratingon categorising online reviews and publiclyaccessible social media posts. Particularly since late 2016 during the American Presidential election, the question of determining ‘fake news’ has also been the subject of particular attention within the literature.A number of strategies are outlined by Conroy, Rubin, and Chen with the purpose of accurately classifying the deceptive publications. They point out thatsuperficial parts-of-speech (POS) tagging and straightforward content-related n- grams have frequently failed to account forcrucial context information, rendering themineffective for the classification task. Thesetechniques have only been shown effectivewhen used in conjunction with more sophisticated analytical techniques. Combining n-gram techniques with deep syntax analysis utilising probabilisticcontext free grammars (PCFG) has proven to be especially beneficial. Using online review corpora, Feng, Banerjee, and Choi [2] are able to classify cases of fraud with 85%-91% accuracy. On top of Feng's first deep syntax model, Feng and Hirst added asemantic analysis that checks 'object : descriptor' pairings for inconsistencies withthe text for further advancement. With similar effectiveness, Rubin, Lukoianova, and Tatiana use a vector space approach to examine rhetorical structure.

4. Proposed System

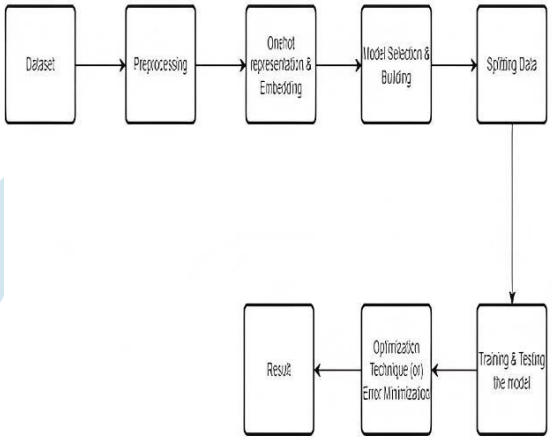
NLTK prepossessing of data. We are using a database with a set number of words.For translation, stemming is being used. Text is converted into a single hot representation before being converted tointegers according to the corpus' length creating padding and embedding features toregulate the flow of sequential data with varying or shifting lengths. Dividing the information into sections for training and testing. Selecting the model's layers. Both the model's training and testing are done. Dropout layer addition helps minimize errors. Assessing the programme. Data pre-processing using NLTK is done priorily. Weare using corpus of fixed Vocabulary size for the process of skimming. We are using stemming for conversion of data properly. Conversion of text into one hot representation is the most important thing her. Here, one hot representation is a process in whichthe conversion of text into numbers which is done based on the corpuses length. Later,Creating Embedding features and padding for the control of sequential data flow withdifferent or changing lengths. Splitting the
Dataset for training and testing purposes isdone in the next step itself. As a part of it, layers selection for the model is done. Then the next phase is training the model, later after the training testing of the model is done. This is the special layer which isadded after the hidden layers namelydropout layer. The error minimization is done at the dropout layer which is used foroptimization. Then evaluating the modelwill be done. These series of steps are required for our proposed system. Also,
Neural networks for news categorizationcan be optimised using a variety of methods. Here are a few suggested actions:

A. Preprocessing: the data is the first stagein optimising a neural network for news classification. The data must be cleaned, stop words must be eliminated, words mustbe stemmed or lemmatized, and the text must be transformed into number vectors.

B. ***Feature engineering:*** This procedure iscarried out after the data has been In conclusion, a number of steps, includingdata preprocessing, feature engineering, model selection, hyperparameter, regularization, evaluation, and iterative improvement, are required when optimisinga neural network for news classification. These methods can be used to develop a highly accurate news classification model which is also an optimized model. This is the model which is made with optimizers and used for classification.

5. Architecture

A suggested design for enhancing aneural network for news classification is provided below:



The preprocessed news text data is presented as number vectors in the input layer. Then,
Embedding Layer: In a low-dimensional environment, the embedding layer transforms the numerical vectors into distributed representations of words. Eitherstarting from fresh or using pre-trained word embeddings like Glove, this layer canbe trained.

Convolutional Layer: To extract features from the news text, the convolutional layeruses a collection of filters of varying sizes to the embeddings. Feature maps areproduced by the filters as they pass over theembeddings. A collection of feature maps for each filter is the layer's output.

Pooling Layer: The highest value for each feature map is chosen by the max pooling layer, which then down samples the feature maps. This aids in lowering the feature maps' dimensionality while retaining the most crucial characteristics.

Fully Connected Layer: Using a group of neurons and the output from the max pooling layer, the fully connected layer conducts classification. A chance distribution over the classes is the layer's output.

Output Layer: The output layer is made upof a group of neurons that correspond to theamount of predicted classes. The expected class label is the layer's output.

Loss Function: The difference between thepredicted class name and the actual label ismeasured by the loss function. Categorical cross-entropy is a well-liked loss function for multi-class categorization.

Optimization: The optimisation algorithm modifies the neural network's weights to minimise the loss function. Adam, RMSprop, and Stochastic Gradient Descent(SGD) are a few well-known optimisation techniques.

Regularization methods: Regularizationmethods like dropout, L1 or L2 regularisation can be used to avoid overfitting.

Hyperparameters: Grid search or random search can be used to adjust thehyperparameters of the neural network, including the learning rate, group size, number of filters, filter sizes, and number ofhidden layers.

Metrics for Evaluation: The neuralnetwork's success can be assessed usingmetrics like accuracy, precision, recall, andF1-score.

Iterative Improvement: Until the desired performance is attained, the architecture can be enhanced iteratively byexperimenting with varioushyperparameters, regularisation methods, and optimization for different optimizers.
The input layer, embedding layer, convolutional layer, max pooling layer, fully connected layer, output layer, loss function, optimisation algorithm, regularisation techniques, hyperparameters,evaluation metrics, and iterative improvement comprise the proposed architecture for optimising a neural networkfor news classification. A extremely accurate news classification model can be created by adhering to this architecture.

6. Results

Using the above algorithms of Naïve bayes,Random Forest and other algorithms we have come up with the accuracy of approximately 95-97, but here we are tryingto get more accuracy than previously used methods and also the main specification here is adding a dropout layer. This dropoutlayer is mainly used in error minimization. When going through this process, we couldn't find time decay as our major problem.
Also, another key feature here is the use ofoptimizers. These are added at the dropout layer , which is named as Adaptive Moment Estimation(ADAM). There are many optimizers namely Gradient Descent,Stochastic Gradient Descent, Mini Batch Stochastic Gradient Descent, SGD with momentum, Nesterov Accelerated Gradient, Adaptive Gradient, Ada Delta, RMSprop. But, currently we are using the Adaptive Moment Estimation(ADAM) in order to improve accuracy, increase performance of the model and error minimization.

7. Future Work

Here is the work based on multinomial based on classification and then accuracy Here are some possible directions for futureneural network news classification optimisation research:
We can also use audio, video and all kinds of languages data can be taken and can be classified as per the given context. It can also include images and its classification. Also, many number of optimizers can be used and can compare each one which givesan efficient and optimized output as per thedata given by the user and its choice. *Transfer Learning:* For the purpose of classifying news, transfer learning methodscan be investigated to make use of pre-trained models on sizable datasets. By initialising the weights with pre-trained models and fine-tuning them on the news dataset, this method can aid in enhancing the neural network's performance *Attention Mechanisms:* In order to concentrate on the most crucial portions ofthe news text, attention mechanisms can beintegrated into the neural network. This canmake the model easier to understand and give it the ability to manage longer news articles.
Multi-Modal Learning: To include extradata such as images, videos, and audio in the neural network for news classification, multi-modal learning techniques can be investigated. This can aid in enhancing the model's performance by adding supplementary data from various modalities. These methods can be investigated to revealinformation about how the neural network generates its predictions. This can increase the model's credibility and make it possiblefor stakeholders to comprehend how the model makes choices. *Adversarial assaults:* Adversarial assaults can be investigated to gauge how well- defended the neural network is. Researchers may be able to create more reliable models by using this to help spot potential weaknesses in the model.
Online Learning: To allow the neural network to adapt to shifting news trends in real-time, online learning techniques can be investigated. This can assist the model'saccuracy and give it the ability to handle breaking news stories.
In conclusion, future research in neural network news categorization optimization can concentrate on areas like transfer learning, attention mechanisms, multi-modal learning, adversarial attacks, and online learning. Building more accurate,reliable, and trustworthy models for news classification is feasible by examining theseareas.

References

Here we are providing the references:

1. Shivam B. Parikh and Pradeep K. Atrey,"Media-Rich Fake News Detecton: A Survey", IEEE Conference on Multimedia Information Processing and Retrieval, 2018.
2. Mykhailo Granik and Volodymyr Mesurya, "Fake News Detection Using Naïve Bayes Classifier", IEEE First Ukraine Conference on Electrical andComputer Engineering, 2017
3. Shlok Gilad, "Evaluating Machine Learning Algorithms for Fake News Detection" , IEEE 15th Student Conferenceon Research and Development, 2017.
4. Akshay Jain and Amey Kasbe, IEEE International Students' Conference on Electrical, Electronics and Computer Sciences, 2018.
5. QIN Yumerg, Dominik Wurzer and TANG Cunchen, "Predecting Future Rumours",Chinese Journal of Electronics, 2018.
6. Arushi Gupta and Rishabh Kaushal, "Improving Spam Detection in Online Social Networks" International Conferenceon Cognitive Computing and Information Processing , 2015.
7. Veronica Perez-Rosas, Bennett Kleinberg, Alexandra Lefevre and Rada Mihalceal, "Automatic Detection of Fake News", 2018.
8. Supanya Aphiwongsophon and PrabhasChongstitvatana, "Detecting Fake News with Machine Learning Method",CP journal, 2018.
9. Stefan Helmstetter and Heiko Paulheim,"Weakly Supervised Learning for Fake News Detection on Twitter",IEEE/ACM International Conference onAdvances in Social Networks Analysis andMining 2018.
10. Cody Buntain and Jennifer Golbeck, "Automatically Identifying Fake News in Popular Twitter Threads",IEEE International Conference on Smart Cloud 2017.