**Speech Analysis Using NLP and Deep Learning to Detect**

**Scam Advertisements**

**Abstract**

In today’s modern world advertisements play a vital role in selling various goods and make people aware of different products. Advertisements provide an excellent marketing tool for convincing people to buy their products, and marketers are cunning to use the right words at the right time. Business companies are trying to make different kinds of attractive advertisements to attract people to their products. It seems that many people are easily attracted by watching different kinds of false and unrealistic advertisements and believing that it might be true. But most of the time it results in a bad investment for a product or good just by watching their advertisement. So our main goal is to detect lying and unrealistic advertisements using speech recognition. We try to detect the pattern of their speech and try to classify the advertisement which is a spam or which is more trusted. Using Natural Language Processing (NLP) we can analyze their speech patterns and recognize whether they are lying or telling the truth. By using their speech as input, we then break down the sentences and use semantic analysis to analyze the words, tones, and phonetics to detect any abnormalities in their speech.

**Introduction**

With the explosion of internet resources, users are depending on the internet to find products or any kind of information. They are now more likely to depend on the internet to find any kind of solution or for decision-making. In such cases, a business company plays a vital role to motivate people to make different kinds of decisions. So businesses and other companies are very much interested in competing among themselves and making different kinds of advertisements to attract people. They follow different kinds of strategies to attract people to their products. And also in that case many fraud companies take the chance to influence people and get them into a trap. It's not a very easy task to detect which advertisements and offers are really good and which ones are fraudulent or spam. There are also a lot of problems finding trustworthy companies in social advertising. In machine learning, natural language is a field where we can detect the tone of the speech to find different kinds of fraud or spam detection using different kinds of machine learning algorithms. It is very popular nowadays to distinguish between real and spam advertisements and it helps people a lot by saving their time and money.

**Related works**

There are several words already done to detect false news or mail over different kinds of topics. It seems that nowadays a lot of spam mail goes around and we receive it almost every day. Sometimes this advertisement is more attractive and they give different kinds of interesting offers to attract the customers. Phayung Meesad introduces [1] the composition of a different kind of machine learning model to detect Thai fake news classification and the author shows that Long Short-Term Memory was the best model to deploy an automatic online fake news detection in web applications. Nida Aslam, Irfan Ullah Khan, Farah Salem Alotaibi, Lama Abdulaziz Aldaej, Asma Khaled Aldubaikil, [2] proposed an ensemble-based deep learning model to classify news as fake or real using LIAR dataset.Gohil, Nayanaba & Meniya, Arvind. [3] Shows a survey that contemplates showcasing details related to click fraud and domains working on its detection. Shaari, Hala & Ahmed, Nuredin [4] introduce a deeper understanding of vulnerabilities of online/mobile advertising ecosystems, the ad fraud's well-known attacks, their effective detection methods, and prevention mechanisms. Tran, Hung & Hornbeck, Thomas & Ha-Thuc, Viet & Cremer, James & Srinivasan, Padmini[5] propose a novel spam detection approach that takes into account the particular characteristics of this domain. Specifically, they introduce a novel set of features that could strongly discriminate between spam and legitimate advertisement posts. Raad, Mostafa & Mohd Yasin, Norizan & Alam, Gazi Mahabubul & Bahaa, Bilal & Zaidan [6] describe the impact of the anti-spam on the e-mail marketing system by surveying the expert on the area of marketing. They analyze every single question with more than 70 specialists on marketing and their opinions have been recorded and analyzed in this paper.

**Methodology**

By comparing the analyzed data with regular human speaking patterns we can deduce whether the advertisement is true or false. The advised model will use a Support Vector Machine (SVM) to work out all the features of the speech being inputted. The data is then fed into an Artificial Neural Network (ANN) for analysis of all the features made by the SVM. The data goes through layers and is compared with regular human speech recognition to find out if the advertisement is fake or real. Here SVM is used as they require fewer data since most advertisements are short.

We need to input the advertisement recording. Then we have to decode the audio into its semantics through semantic analysis. Then we have to group the data from the semantics together by using the SVM algorithm. We have to input the grouped data into the neural network. Finally, after computation the neural network will show the results of whether it was true or not. The steps are given in details below.

First, we have to breakdown the advertisement to its semantics. To do this, semantic analysis which takes into account the words, phonetics and speed of speaking as some of the factors. For the semantic analysis we can use a Deep Belief Network (DBN) based on the K-SVD (K-Singular Value Decomposition) algorithm. According to Zhou, Zhao & Pan nonlinear characteristics are used to relate lie detection like auditory perception and psychological acoustics. These are difficult to extract and have a high computing cost. The DBN is comprised of a hierarchy. The internal structure of the characteristics in high dimension can be seen and extracted. The DBN is composed of multiple RBM (Restricted Boltzmann Machines) and there are Hj hidden layers. The DBN is used to extract the expressions of the audio. The DBN is trained in two steps- first, in an unsupervised manner then secondly in a supervised manner, tuning and changing the necessary factors. After training is over the activation probability vector of the hidden layer nodes is used by the RBM to move to the next layer. At the top of the network the SVM classifier can be used. The labeled data is used to train the model. To account for dynamic changes of basic acoustic parameters, the spectrogram signal is analyzed. The length of the lie is assumed as a variable. The dimension of the spectrogram is reduced using PCA. To reduce number of calculations the spectrogram needs to be divided and overlapped. Discreet Cosine Transform (DCT) complete dictionary is chosen with a set dictionary redundancy. This is decomposed to the sparse coefficients using the K-SVD. The average value is calculated and the coefficients are put into the DBN model to obtain deep features. The lies of the speech features are extracted and is based on the K-SVD sparse representation and DBN. Feature extraction is comprised of two phases. First, Hamming Window is added and the dimension is reduced then sparse decomposition for the lies using K-SVD. The calculations are made. Then, based on the sparse coefficient matrix which is obtained by K-SVD, the DBN characteristics learning is done. The deep lie features are found around here [7].

Second, after obtaining all the features necessary to detect the lies and semantics, we use the SVM for classification. The data is then classified. According to Lurz, two linearly separable categories exist. The SVM finds the most optimal hyper plane to separate the two categories. It is obtained by maximizing the margin. The margin is where the distance between the separation line and the closest data points coincide. Using support vectors- which are a fraction of the data that is used to recount the separation line [8]. At first we have to choose the type of SVM model, here we choose classification. Second we have to choose the classification type, here as we are going for the extracted features of our speech which are texts, we use semantic analysis. We then take input the extracted features obtained from the DBN. We then define the tags. We train the SVM model and then put it to the test. Note that before putting the extracted features the SVM was already trained before-hand. Again as per [8], up until CNNs the SVM outperformed the ANN on predicting accuracy, but that’s starting to shift. SVM is excellent for binary classifications while ANNs for multi-class. To tackle all of these issues we have decided to make use of an ANN after the SVM is done with the classification task.

Third, we are already done with the classification, but now one task remains, to find out and compare the classified data with the data of regular human speaking patterns. For this purpose an ANN is best suited for the ANN can handle multi-class categorization. We then load the classified data into the already trained ANN network for comparison and finding the result of the audio advertisement. According to Lurz, ANNs are based on the functioning of neurons in a human brain. The neurons in this case are the computational nodes. They receive information via a numerical weight given to the node. There are many dense hidden layers. The information only moves in a forward manner. The problem of ANN is taking too much time to train and an issue of over-fitting. From a mathematical perspective it is an approximation function, y(x: w), where y is the output, x is the input and w is the weight matrix, which is a parameter that is optimized on the labelled data di. The optimization process is gradient descent. This involves calculation of the gradients of an error function with respect to the weights using a method called back-propagation. The output of each output-node is calculated using its activation function [8]. Here we can use the Adam optimization algorithm for the gradient descent. For the activation function we can use ReLu (Rectified Linear Unit), it is preferred by most researchers, as it has the ability to prevent vanishing gradient problems even for large datasets according to [8].

In addition, the Adam optimization is used for the optimization process. The Adam optimization is an extension of the stochastic gradient descent, and has recently surfaced as an excellent optimization algorithm in natural language processing (NLP). Adam optimization combines the advantages of two extensions of stochastic gradient- Adaptive Gradient Algorithm (AdaGrad) which maintains a per-parameter learning rate that boosts performance on issues with sparse gradients (a pressing problem in NLP). Root Mean Square Propagation (RMSProp) which maintains a per-parameter learning rate which are modified based on the average of recent sizes of gradients for the weights. Adam optimization understands the advantages of both these extensions. Adam has 4 configuration parameters, learning rate or step size (α), exponential decay rate of first moment estimate (β1), exponential decay rate of second moment estimate (β2), a very small number to prevent division by zero (𝜀). Here α= 0.001, β1=0.9, β2=0.999 and 𝜀=10-8. The Adam optimization algorithm is very effective in achieving our task. It is also very easy to configure.

Furthermore, we have opted to use ReLu for it ensures that disappearing gradients don’t occur during the activation. It is a linear piecewise function that will output the given input if it’s positive else it will output zero. This avoids easy saturation and provides more sensitivity to the activation. We can set the rectification. ReLu has some advantages. It has computational simplicity – a simple max () function like max(0.0, x) where x is a variable, representational sparsity – can output a true zero value, linear manner- acts like a linear function and can train deep networks.

Finally, after feeding the ANN with our classified and labeled data we can get the results of our advertisement audio. The results are then displayed using various display formats- such as graphs, bar-charts and tables. The attributes found by comparisons are displayed in a table with all the necessary measurements and differences.

**Conclusion**

In this paper, we propose a method where we can easily detect spam or fraudulent advertisements by using natural language processing and machine learning. We hope that by using it we can reduce a lot of fraudulent advertisements over the internet. We can filter out what kind of advertisement is really trusted or what kind of information or advertisement is supposed to be false or spam. Also, there are many offers, of different kinds of products that come over the phone. Whenever anyone wants to know any information over the phone the companies describe their offers and other benefits. Our research is also helpful to detect speech tones and find the possibility of whether the information is trusted or not.

**References**

1. Meesad, P. Thai Fake News Detection Based on Information Retrieval, Natural Language Processing and Machine Learning. SN COMPUT. SCI. 2, 425 (2021). Retrieved from: <https://doi.org/10.1007/s42979-021-00775-6>
2. Nida Aslam, Irfan Ullah Khan, Farah Salem Alotaibi, Lama Abdulaziz Aldaej, Asma Khaled Aldubaikil, "Fake Detect: A Deep Learning Ensemble Model for Fake News Detection", Complexity, vol. 2021, Article ID 5557784, 8 pages, 2021. Retrieved from: <https://doi.org/10.1155/2021/5557784>
3. Gohil, Nayanaba & Meniya, Arvind. (2020). A Survey on Online Advertising and Click fraud detection. Retrieved from: <https://www.researchgate.net/publication/344105501_A_Survey_on_Online_Advertising_and_Click_fraud_detection>
4. Shaari, Hala & Ahmed, Nuredin. (2020). An Extensive Study on Online and Mobile Ad Fraud. Retrieved from: <https://www.researchgate.net/publication/346972554_An_Extensive_Study_on_Online_and_Mobile_Ad_Fraud>
5. Tran, Hung & Hornbeck, Thomas & Ha-Thuc, Viet & Cremer, James & Srinivasan, Padmini. (2011). Spam detection in online classified advertisements. ACM International Conference Proceeding Series. 11. 10.1145/1964114.1964122.
6. Raad, Mostafa & Mohd Yasin, Norizan & Alam, Gazi Mahabubul & Bahaa, Bilal & Zaidan, A.. (2010). Impact of spam advertisement through e-mail: A study to assess the influence of the anti-spam on the e-mail marketing. African Journal of Business Management. 4. 2362-2367.
7. Zhou, Y., Zhao, H., & Pan, X. (2015). Lie Detection from Speech Analysis Based On K–SVD Deep Belief Network Model. *Springer International Publishing Switzerland*, pp. 189-196. Retrieved from:  <https://sci-hub.mksa.top/10.1007/978-3-319-22180-9_19>
8. Lurz, KK. (2018). *Natural Language Processing in Artificial Neural Networks Sentence analysis in medical papers.* [Master’s Thesis, Lund University]. Retrieved from: <https://lup.lub.lu.se/luur/download?func=downloadFile&recordOId=8948103&fileOId=8948109>