SOCIAL MEDIA SPAM DETECTION USING DEEP LEARNING

Rumi Juwairiyyah*, Nanditha Sriram, Jyotshna Bhushan Sharma, Babeetha
SRM Institute of Science and Technology, Chennai, Tamil Nadu, India
* rumi1241@gmail.com

ABSTRACT: *The use of social media in the 21st century has generated a humongous amount of data. This technology is a very cheap, efficient and popular means of communication and can be used for sharing information and resources. Data can be of any type, such as facts, statistics, graphs, charts, trends that can be accessed by a computer for interpretation. People these days count on the data they get from social media to make their decisions. As the data available is huge and free of cost, it has enabled social spammers to exploit the world of internet by spreading different kinds of spam messages in order to promote blogs, advertisements and so on. The liberty given to a user to write a review of their will has given spammers a golden opportunity to exploit the trust of the customers by spamming the site with spam reviews about products based on different interests.*

ECLIPSE: Eclipse is an IDE (Integrated Development Environment) that is used in computer programming. It provides the user an environment to compile and execute his code. This platform was primarily created for the JAVA language, however other programming languages could also be coded used various plug-in. This IDE provides the users a workbench where they can arrive at the end result by integrating various tools.

PYTHON: Python is regarded as a general-purpose high-level interpreted language. It supports dynamic typing and the concept of garbage collection. The significant property of this language is the use of notable whitespaces. It follows the concept of object-oriented programming which helps the user to distinctly understand the logic, the function and the execution of the code.

MODULES:

Module 1:
Exploratory Data Evaluation

The first step to the data analysis part of the project is the EDA, elaborately called as the Explanatory Data Analysis. It is first important to understand the kind of data we are dealing with and connect all the dots in order to arrive at conclusions so that it becomes easy to figure out the questions we want to ask, the possibility to manipulate the interpreted data to get the answers we need. This can be done by studying the various patterns, trends, forecasts, outliners and test case results in the existing data base, using various tools that use quantitative, qualitative and visual methods to comprehend the data that we deal with. In data science, EDA holds immense value as it helps us to reach to the point of certainty in a short span of time in which the future results are correctly interpreted and are applied to the various ends of businesses. It is only possible to reach this level of accuracy when the data collected is validated and checked for anomalies, to ensure that there are no errors in the data set. Using EDA it is possible to find the hidden insights of the stakeholders and the data scientists that were not captured during the investigation, these details could prove to be very useful to understand specific businesses. The primary goal of EDA is to enable better feature selection that can be used in technologies like deep learning, machine learning, and deep neural networks and so on. Once the scientists understand the type of features that are being selected they tend to come back to the initial parts of the concept, as it might not have resulted in the purpose that they were serving. Once this stage is complete, the scientists have a distinct set of data that they use for deep learning and machine learning algorithms.

Module 2: Pre-handling

There is a possibility that we might at times come across situations in which certain data might be missing from the dataset. It is important that we are well equipped when such issues arise. If we decide to eliminate a section of data there is a tendency to eliminate something crucial. In such cases it is better to take the mea n value of the data column and replace it in the missing spots of it. We make use of the Scikit learn preprocessing as the library. It comprises of a class named imputer, which takes care of the missing data. At times there comes a situation in which the data happens to be in text form and this makes it harder for the system to interpret such types of data. We need to encode the categorical data so that the system interprets it accurately. The data set will now split itself into two sets, a training set and a testing set. Now all the models are first tested on the training set to understand how they form their correlation and once it is completed the models are now tested on the test set to check for its validation and accuracy based on which how well it predicts. There is a general thumb rule that a major of 80% of the data set is allocated to training set and the rest of 20% to the test set.

Module 3: Feature Engineering

The step of selecting the feature is considered to be the most important step. The selection feature functions irrespective to the machine learning algorithms. The features are mostly chosen on the basis of their scores such that there is a favorable outcome.

Pearson's Correlation: It is used to measure the linear dependence between two continuous variables X and Y. The values vary from -1 to +1.

LDA: Linear discriminant analysis finds the linear combination of features that separates two or more classes of a categorical level.

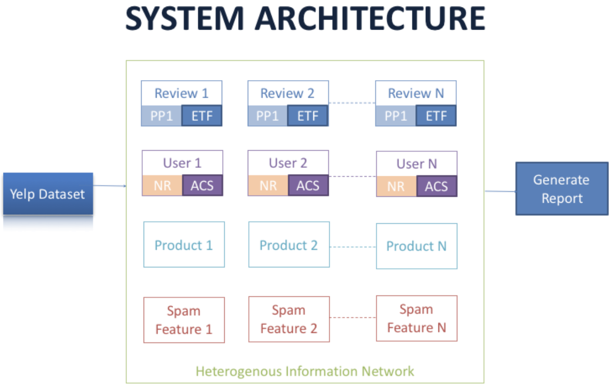
ANOVA: It stands for Analysis of Variance. It is used to operate using one or more independent features and one dependent feature. It provides a statistical test that proves if the means of groups are equal or not.

Module 4: Prediction

The final step of the module is to check if all the processes have been executed correctly using evaluation techniques. In this step we test our model against the dataset that we have acquired but not used for training. It gives us an insight of how the model would appear in the real world. The training or evaluation splits on the basis of thumb rule, which might either be 80-20 or 70-30. This entirely depends on the size of our dataset. After the evaluation is complete, it is possible to check if any further improvements can be made. Tuning the existing parameters of the module can do this.

Architecture:

The architecture for the concept elaborately explains how the spam is visible to the scientists that detect it. At first the reviews are posted based on the product. The spammers do not follow creative strategies as their aim is to fill more content rather than useful content. The text could be of a particular type of font, color, style and pattern. The estimated time frame is an entity that they keep track off as they follow a schedule in which they post at a particular fixed time. The user now could be of many types; they follow a particular trend unique to them. They have a fixed allocation of the type and size of content and they remain fixed to it. From this a product can be formed as a combination of review and user. By following this trend we get our first spam feature. There are many possibilities. By using deep learning it is possible to integrate various algorithms into a single unit to eliminate categorization when it comes to detecting spam or even errors.



Methodology:

In order to efficiently detect spam on social media there are multiple algorithms that can be used. We use the concept of deep learning where various algorithms are combined into one unit to eliminate the loophole of having slightly less accurate data. We use the system features of: Review-behavioral model, User-behavioral model, Review-linguistic model, User-linguistic model. The review behavioral model primarily studies the types of reviews that get posted based on the time frame, the day of post, the time of post, the category of post, the method of post and so on. By this we get an understanding of the whereabouts of the reviews that are frequently posted by spammers. The user behavior model gives us an explanation on the type of user, depending on the trends the user follows and the type the user sticks to. As a spammer no user would follow various trends but would prefer to stick to a narrow range. The review linguistic model focuses on the minor details of the review such as the font, texture, color, type, domain, size and so on. It makes sure that these characteristics are categorized efficiently for the user to stand review to stand out from the rest making it more unique and easier to spot. The user linguistic model tells about the minor details of the user, if the user appears anonymously or the user uses multiple fake accounts or if the user pretends to be someone that he is not. Considering these methods of categorization, it becomes easier to detect the spam by using deep learning concepts in order to integrate various options and make it a single model of detection.

Conclusion:

We belong to an era of infinite possibilities. This can be used for the best as well as for the worst. We are a technologically driven generation and hence we should make use of the resources around us to eliminate the misuse of technology and find better ways to make new inventions that could result in the progress and betterment of mankind. The deep learning concepts have made it easier by giving us a platform to integrate various dimensions of a problem and to solve it without any exceptions or categorization.

References:

[1]. G. Stringing, C. Kruegel, and G. Vigna, “Detecting spammers on social networks,” in Proceedings of the 26th Annual Computer Security Applications Conference. ACM, 2010, pp. 1–9.

[2]. A. Sala, L. Cao, C. Wilson, R. Zablit, H. Zheng, and B. Zhao, “Measurement-calibrated graph models for social network experiments,” in Proceedings of the 19th International Conference on World Wide Web (WWW’10). ACM, 2010, pp. 861–870.

[3]. F. Ahmed, A. Muhammad, An MCL-Based Approach for Spam Profile Detection in Online Social Networks 2012 IEEE 11th International Conference on Trust, Security and Privacy in Computing and Communications, pp 1-7

[4]. Z. Halim, M. Gul, N. ul Hassan, R. Baig, S. Rehman, and F. Naz, “Malicious users’ circle detection in social network based on spatio-temporal co-occurrence,” in Computer Networks and Information Technology (ICCNIT), 2011 International Conference on, July, pp. 35–39.

[5]. https://www.researchgate.net/publication/220846842_Social_spam_detection

[6]. https://link.springer.com/referenceworkentry/10.1007%2F978-1-4939-7131-2_110199

[7]. https://www.sciencedirect.com/science/article/pii/S0925231215002106

[8]. https://ieeexplore.ieee.org/document/7582661