Sentiment Analysis On Reddit Data And Analyzing Attacks on Social media

Sahithi NallaniChakravartula (01996604)

Abstract

Sentiment Analysis helps us to understand the subject in the present digital age.Using the sentimental analysis on reddit data the reviews were classified into positive ,negative and neutral using machine learning models like Naive Bayes. Any brand's presence on social networks has a significant impact on emotional reactions of its users to different types of posts on social media. If a company understands the preferred types of posts (photo or video) of its customers, based on their reactions, it could make use of these preferences in designing its future communication strategy.Using the kaggle dataset and the API used is the reddit's PRAW.The Total number of comments/tweets in the reddit dataset are : 36,799 Negative : 8277 Neutral: 13142 Positive: 15830.This Analysis was done on Amazon Alexa and compared it to Apple's Siri to see how the reviews on reddit would influence people from buying the product they liked.

Keywords: Sentiment analysis; Reddit dataset; Naive-Bayes, literature review

1 Introduction

Social media posts regarding how people are feeling can be utilized to gain a better knowledge of public health, everyday decision-making, and people's assessments of their quality of life. Evidence reveals that people are more prone to disclose their feelings online, particularly on social media platforms, during times of crisis. Sentiment analysis helps in real-time monitoring of products across public platforms. It helps in understanding customer feelings. Hackers are constantly developing new attacking tools and hacking strategies to gain malicious access to systems and attack social media networks thereby making it difficult for security administrators and organizations to develop and implement the proper policies and procedures necessary to prevent the hackers' attacks. The increase in cyber-attacks on social media platforms calls for urgent and more intelligent security measures to enhance the effectiveness of social media platforms.The product owners can streamline their promotions / advertisements concentrating majorly on a specific social media platform The limitation is filtering out spam tweets from twitter and reddit because of its ambiguity

2 Related Work

Social media adversarial attack refers to a wide range of hostile operations carried out through human interactions on social media. It manipulates users' minds to make them make security mistakes or reveal important information. Attacks on social media can take several forms. To carry out the assault, a perpetrator first examines the intended victim to obtain background information such as possible avenues of entry and weak security mechanisms. The attacker then attempts to acquire the victim's trust and give stimuli for later acts that violate security protocols, such as disclosing sensitive information or granting access to key resources.

3 Method And Evaluation

Sentiment Analysis on real-time Reddit data using Naive-bayes algorithm is the simplest and fastest technique for a large amount of data. Some of the applications for naive-bayes include sentiment analysis, text classification etc. The Naive Bayes Classifier uses the Bayes theorem to predict membership probabilities for each class, such as the probability that a particular record or data point belongs to that class. The most likely class is the one with the greatest chance of occurring.

Requirements: pandas – Python Data Analysis Library. pandas are open-source that provide high-performance, simple-to-use data structures, and data analysis tools.

Numpy – NumPy is a scientific computing fundamental package in Python. It contains among other things:

• a powerful N-dimensional array object
• sophisticated (broadcasting) functions
• tools for integrating C/C++ and Fortran code
• capabilities in linear algebra, Fourier transform, and random numbers NumPy can be used as a multi-dimensional container of generic data in addition to its apparent scientific applications.

sci-kit learn – Data mining and data analysis tools that are easy to use.

SciPy – SciPy is a Python-based ecosystem of open-source math, science, and engineering tools.

Reddit Api Keys: To fetch Reddit data

Data Preprocessing: The open available Reddit dataset on kaggle is used to train the model. For converting words into features the data preprocessing is essential. The countvectorizer here tokenizes and lower cases the text

Splitting Data: The columns are divided into dependent and independent variables first (or features and labels). The variables are then divided into train and test sets.

Vectorization: To transform each review/comment/tweet into a numerical representation

Model Generation: The initial idea was to use only Naive-Bayes algorithm for model generation. However, the results for Naive-Bayes in the first run were only 60%. In order to see if the model being used is a better performing model the randomforestclassifier, Logistic regression and Adaboostclassifier models were tested. These models seemed to slightly overfit the model with 99% accuracy except Adaboostclassifier which was at 68%. As the data was being tested multiple times, the initial Naive-Bayes algorithm model performance improved to 70% and this was finalized. Prediction Analysis and Reddit Instance: The data from Reddit was scrapped with Reddit API key generation and getting data using PRAW and the model was able to predict 1, 0, -1 , Positive, negative and neutral consecutively

4 Data and Evaluation

Statistics: The Total number of comments/tweets in the reddit dataset are : 36,799 Negative : 8277 Neutral: 13142 Positive: 15830 The dataset used here is the open source available dataset in Kaggle to train the model. The dataset is attached as part of the code. The results are evaluated based on model accuracy scores. One interesting observation from the Real-time reddit dataset was there were many words like “suicide” , “covid” that was not retreiving the data to analyze whereas twitter was able to do that. Another observation was that different products had varying numbers for positive, neutral and negative across both platforms. The accuracy score for the model achieved was 70% for the naive bayes classification.

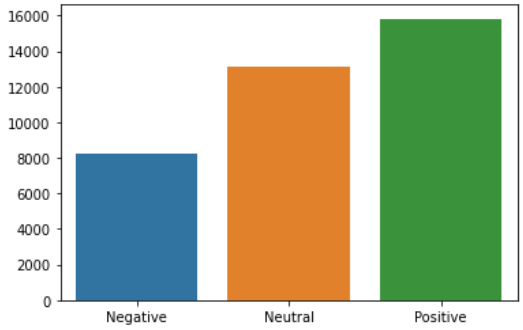

Figure 1 - Training Reddit Dataset

9 Results and Insights

Naive Bayes model has an accuracy score of 70% and performs well on reddit data. From the ML algorithm the data retrieval contained 500 real-time reddit posts for alexa and siri subreddits and classified them as positive, neutral and negative.

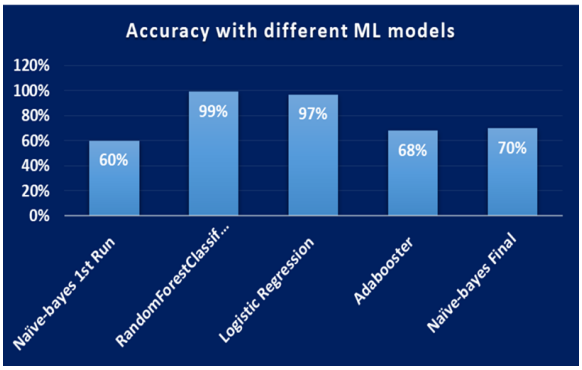

Figure 5: ML model accuracy

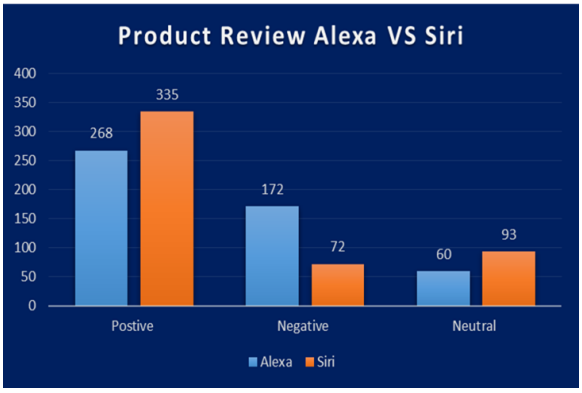

Figure 6: Alexa Vs Siri sentiment analysis

10 Future Improvements

The model can be created to work on multiple social media platforms such as twitter, facebook and instagram. The introduction of differential privacy in the form of a training bot ,what this essentially means is that during the introduction of differential privacy there is an introduction to noise , this noise can be generated using bots on the reddit dataset since it is based on which product is better to advertise on which platform and need user's personal details. Observations contained issues with the security and These bots can create fake credit card information or phone numbers and introduce the attacks which can later be used for the defense and then can be applied into the real life situation. THere is also a need to create a more robust defense mechanism for these social media platforms with large datasets.

12 References

[1] An Analysis of Security in Social Networks Weimin Luo1 , Jingbo Liu1 , Jing Liu2

[2] Adversarial Attacks and Defenses for Social Network Text Processing Applications: Techniques, Challenges and Future Research Directions Izzat Alsmadia , Kashif Ahmadb , Mahmoud Nazzalc , Firoj Alamd , Ala Al-Fuqahab , Abdallah Khreishahc , Abdulelah Algosaibie

[3] A Guide to Differential Privacy Theory in Social Network Analysis by Christine Task, Chris Clifton

[4] Evasion attacks against machine learning at test time Battista Biggio1 , Igino Corona1 , Davide Maiorca1 , Blaine Nelson2 , Nedim Srndi´c ˇ 3 , Pavel Laskov3 ,Giorgio Giacinto1 , and Fabio Roli1

[5] Adversarial Attack on Sentiment Classification Alicia Yi-Ting Tsai, Tobey Yang

[6] Privacy and security in online social networks: A survey Imrul Kayes a,* , Adriana Iamnitchi b