PROJECT REPORT

ON

TWITTER SENTIMENTAL ANALYSIS

Certificate

THIS IS TO CERTIFY THAT AVI PRUTHI OF BTECH CSE IV SEM HAS SUCCESSFUL COMPLETED HIS SENTIMENTAL ANALYSIS ON TWITTER PROJECT ON THE PRESCRIBED DURING THE ACADEMIC YEAR 2021 AS PER THE GUIDELINES ISSUED BY GRAPHIC ERA UNIVERSITY.

ACKNOWLEDGEMENT

I would like to express our sincere gratitude to Mr. Deepak Uniyal whose role as project guide was invaluable for the project. I are extremely thankful for the keen interest he took in advising us, for the books and reference materials provided for the moral support extended to us. Last but not the least we convey our gratitude to all the teachers for providing us the technical skill that will always remain as our asset and to all non-teaching staff for the gracious hospitality they offered us.

CONTENTS

1.INTRODUCTION

2.PREREQIUISITES

3.PREPARING THE TEST SET

4.CLASSIFIED TWEETS

5.CONCLUSIONINTRODUCTION

Sentiment Analysis is a term that you must have heard if you have been in the Tech field long enough. It is the process of predicting whether a piece of information (i.e. text, most commonly) indicates a positive, negative or neutral sentiment on the topic. We will go through making a Python program that analyzes the sentiment of tweets on a particular topic. Also known as “Opinion Mining”, Sentiment Analysis refers to the use of Natural Language Processing to determine the attitude, opinions and emotions of a speaker, writer, or other subject within an online mention.

*Essentially, it is the process of determining whether a piece of writing is positive or negative. This is also called the Polarity of the content.*

As humans, we are able to classify text into positive/negative subconsciously. For example, the sentence “The kid had a gorgeous smile on his face”, will most likely give us a positive sentiment. In layman’s terms, we kind of arrive to such conclusion by examining the words and averaging out the positives and the negatives. For instance, the words “gorgeous” and “smile” are more likely to be positive, while words like “the”, “kid” and “face” are really neutral. Therefore, the overall sentiment of the sentence is likely to be positive.

A common use for this technology comes from its deployment in the social media space to discover how people feel about certain topics, particularly through users’ word-of-mouth in textual posts, or in the context of Twitter, their tweets*.*

PREREQUISITES

* Basic programming knowledge

Although Python is highly involved in this mini-project, it is not required to have a deep knowledge in the language, as long as you have basic programming knowledge.

* Installed tools

For this program, we will need Python to be installed on the computer. We will be using the libraires tweepy, TextBlob, Pandas, re, matplotlib.

The required libraries tweepy, pandas, textblob, matplotlib can be installed using pip.

The rest already come with the python interpreter.

* Data set splitting concept

This is critical to fully understand the process pipeline. This is not crucial, but it could help. We will be using the Twitter API here and there in the code, making normal calls to the API and dealing with the JSON objects it returns.

* Machine Learning basic concepts

As an additional value to the project, We will be using sklearn library to train the dataset obtained..

PREPARING THE TEST SET

* Step A.1: Getting the authentication credentials
* Step A.2: Authenticating our Python script
* Step A.3: Creating the function to build the Test set

**Classified Tweets:**

We labelled the tweets in three classes according to sentiments expressed/observed in the tweets: positive, negative and neutral We gave the following guidelines to our labellers to help them in the labelling process.

**Positive:**

If the entire tweet has a positive/happy/excited/joyful attitude or if something is mentioned with positive connotations. Also if more than one sentiment is expressed in the tweet but the positive sentiment is more dominant.

**Negative:**

If the entire tweet has a negative/sad/displeased attitude or if something is mentioned with negative connotations. Also if more than one sentiment is expressed in the tweet but the negative sentiment is more dominant.

**Neutral:**

If the creator of tweet expresses no personal sentiment/opinion in the tweet and merely transmits information. Advertisements of different products would be labelled under this category.

CONCLUSION

Sentiment Analysis is an interesting way to think about the applicability of Natural Language Processing in making automated conclusions about text. It is being utilized in social media trend analysis and, sometimes, for marketing purposes. Making a Sentiment Analysis program in Python is not a difficult task, thanks to modern-day, ready-for-use libraries. This program is a simple explanation to how this kind of application works. This is only for academic purposes, as the program described here is by no means production-level.

BIBLIOGRAPHY

**Books-**

* Python-The complete reference by Martin C.Brown

**Internet-**

For Machine learning

* <https://scikitlearn.org/stable/modules/generated/sklearn.linear_model.LogisticRegression.html>

For getting credentials

* <https://developer.twitter.com/en>

For sentimental analysis

* <https://towardsdatascience.com/my-absolute-go-to-for-sentiment-analysis-textblob-3ac3a11d524>