Review Sentilys

Yedla Jaswanth Naidu,Vamsi Jangyala
Computer Science Engineering
Bennett University
Greater Noida, India

I. Introduction

Our project is about building a machine learning model which analyzes review based on performance thereby generating a score and use sentimental analysis to check the polarity of situation wherein it can be a deciding factor. We also aim at building a model which generates a graphical representation of the data generated after cleaning the dataset, we tend to remove unpleasant or derogatory terms which affects the mental well being of users or people who reads the reviews, last but not least we plan on automating this manual labor thus reducing the time which the user of the product spends on it and creating a safe and healthy platform for all

The introductory part is further divided into the following section 1. Problem statement. 2. Motivation for solving the problem. 3. Solution. 4. Background knowledge

Sentiment Analysis is a popular in machine learning to analyze the text and find the polarity like positive and negative. We use the sentiment analysis in machine learning model to predict the emotion of a people based on the text written.

While working on this project we mainly relied on the sentimental analysis factor which these days has been a go to approach for almost each and every platform not only its one of the most used common approach but also it has a wide range of usefulness like • Emotion detection • Aspect based analysis. • Multilingual sentiment analysis. And comes with many benefits like • Scaling and sorting of data • Real time analysis • Time saving factor

• NOTE: we used open-source tools to create and train and build our own model

II. Motivation

For example, lets take a scenario wherein a visiting faculty gives a session on some topic and at the end of the day they want a review from the class now going through traditional methods of collection and reading each review the following things are likely to happen • Time consuming and tiresome(labor). • They may stumble upon some unpleasant words. • A chance for missing out some genuine to the point review in the masses. • A quick conclusion. .

III. Background Knowledge

DATA COLLECTION

• The real-time students' feedback is collected through a csv file.

DATA PREPROCESSING

• Wherein the given data which in raw or containing unwanted items get cleaned to form a new data set

• Tokenization: Data then gets separated as tokens for further processing

REMOVAL OF UNWANTED WORDS

• A web search tool or other natural language processing system may contain collection of stop records, or it may contain a solitary stop-list. • Most of the more frequently used stop words in English are "an", "a", "of", "the", "you", "and" these are some words which do not carry any meaning. • Words which contain no weight are then removed.

PART OF SPEECH TAGGING:

• For part-of-speech tagging we used the Stanford parts - of-speech tagger.

• This tagger works through splitting text data into sentences

• Consider following example "Staffs are amazing".

FEATURE EXTRACTION

• In the process of feature extraction, movie features are extracted from every sentence. For analyzing the overall polarity factor of the text or data received it is therefore necessary to form a relationship with the data from the words which we aim to train thereby getting a score which being generated through the relationship we built the model through the use of words

• Positive Sentiment Words: These are the words which are having a positive sentiment score according to Senti WordNet.

• For example: • Excellent • valuable, awesome etc.

• Negative Sentiment Words: word which carry weight that tend to polarize the factor by large margin thus reducing the accuracy factor, unpleasant words.

FEATURE REDUCTION

• One of the largest issues of sentimental analysis is addressing text information that square measure on the market in terribly high dimensions which can have an effect on the performance of classifier. So, there's a desire for such technique which can eliminate those options that aren't relevant and keeping solely those options that square measure abundant necessary and therefore the techniques which can facilitate to differentiate the sentences into category labels like positive and negative. • The Information Gain and Gain Ratio measure the foremost common techniques among variety of feature reduction techniques.

MODEL TRAINING

• Model Training The hybrid model for sentiment analysis was trained using unigrams, bigrams, TF-IDF and lexicon-based features

• Random Forest: Random Forest Algorithm was proposed by Leo Bierman. In this study, scikit-learn implementation of Random Forest algorithm was used. The hyper parameters were adjusted by means of threefold cross authentication.

Support Vector Machines (SVM)

Hence, we have developed an Image Caption Generator in Telugu. This design can be useful for projects on aiding the blind. For instance a camera appended White cane with an earpiece support that could effectively guide with audible instructions or a mobile traffic sign surveillance system for vehicles that alerts drivers for any new traffic signs detected. Further developments can be made for realtime captioning systems.

IV. Solution

Our main aim is to create a platform wherein we as a team are building a model which

• Receive data (csv format) • Clean data (remove derogatory words). • Analyze data (ranking based on certain words contained in review section) • Create a graphical representation of the data. • Generate a score (based on ranking of words). • Present a valid reliable score (formal review section).

V. Related Work

Sentimental analysis is used in a broad range of works relating from social media monitoring apps like twitter etc. to brand based review collection like Myntra/ajio therefore we attached examples of each under division

Social Media Monitoring Sentiment Analysis is widely used in the so many platforms to know the behavior of person based on their text. Sentiment Analysis is used in the social media monitoring to know the user feeling while using their product. Let's us take about the twitter is a big social media network around the world. Twitter use the sentiment analysis and they allow companies to understand their users felling about their brand and its helps to know their competitors and they can focus on keep in demand on their product and introduce the new trends and keep on update to present trends needs. To perform the sentiment analysis on twitter has to follow steps. Initially twitter has to gather twitter data and they have clear idea about type of tweets they want. After that prepare the data and clean it includes the removal of irrelevant information and delete the duplicated data. So, we need to create the sentiment analysis machine learning model in this step we have to train the model to predict the emotions based on the data. We train the model in this phase to predict the positive and negative of statements from the following the data. After the sentiment analysis model training, we will analyze the twitter data for sentiment. They final step is visualize the results. Twitter sentiment analysis allows us to keep track about your product on social media it will detects the customer feelings before they escalate.

Brand monitoring: Sentiment Analysis is used in brand monitoring. Brand monitoring is a basically know the what people think about their brand. Every company wants to know about their product in market it will beneficial to their brand. To get the information company tries to get it from the different number of channel and media. Company also searches the shopping sites like amazon reviews. With the help of this they can able to improve their service to their product.

First Companies has to find the data comes from different ways like twitter and Facebook and shopping apps reviews etc. After that we need to filter the statements like positive or negative with the brand sentiment analysis algorithm. The major thing is analyzing the information it is important task is not only finding the negative and positive but need to know who, where, what the particular terms we need to keep in mind. We need to find which type of targeted people are giving the positive feedback and which type of targeted people are given negative feedback we need to analysis. So, with the help of this information the company can alerts and focus to improve the product.

The major step is Prevent the crisis. Let's us take an example your product has gave the negative feedback in shopping app we get notified by the sentiment analysis model. So, we to apologize the customer for the product and give the compensation voucher to prevent loss or not a mistake from company the company has to defend them self.

Customer feedback:

Sentiment Analysis is used in customer feedback. Basically, the customer feedback is the process of detecting the emotions of a customer interact with the products and services. In the customer feedback we mainly use the Natural language Processing and use the algorithm to detect the emotions of a customers. The algorithm main focus on the two parameters they are polarity and magnitude. Polarity is used for detect the positive and negative whereas magnitude is used for emotions exhibited by customer.

For performance of customer feedback, we need the data sets from different resources. Collect the data from the live chat. We need to collect the customer feedback every live chat. With the help of it we need to classify the feedback into respective terms. We need to collect the customer sentiment analysis from the social media. Social media platform such as twitter and Facebook we need to collect the data like "good"," bad"," hate", "awesome" such kind of positive or negative word. With the help of it we can collect data. To know our product, conduct the online surveys to get better in view of the customer side. Monitor the reviews and ratings in shopping apps like amazon, google play, Ajio. It is best way to predict the customer feelings and sentiment analysis.

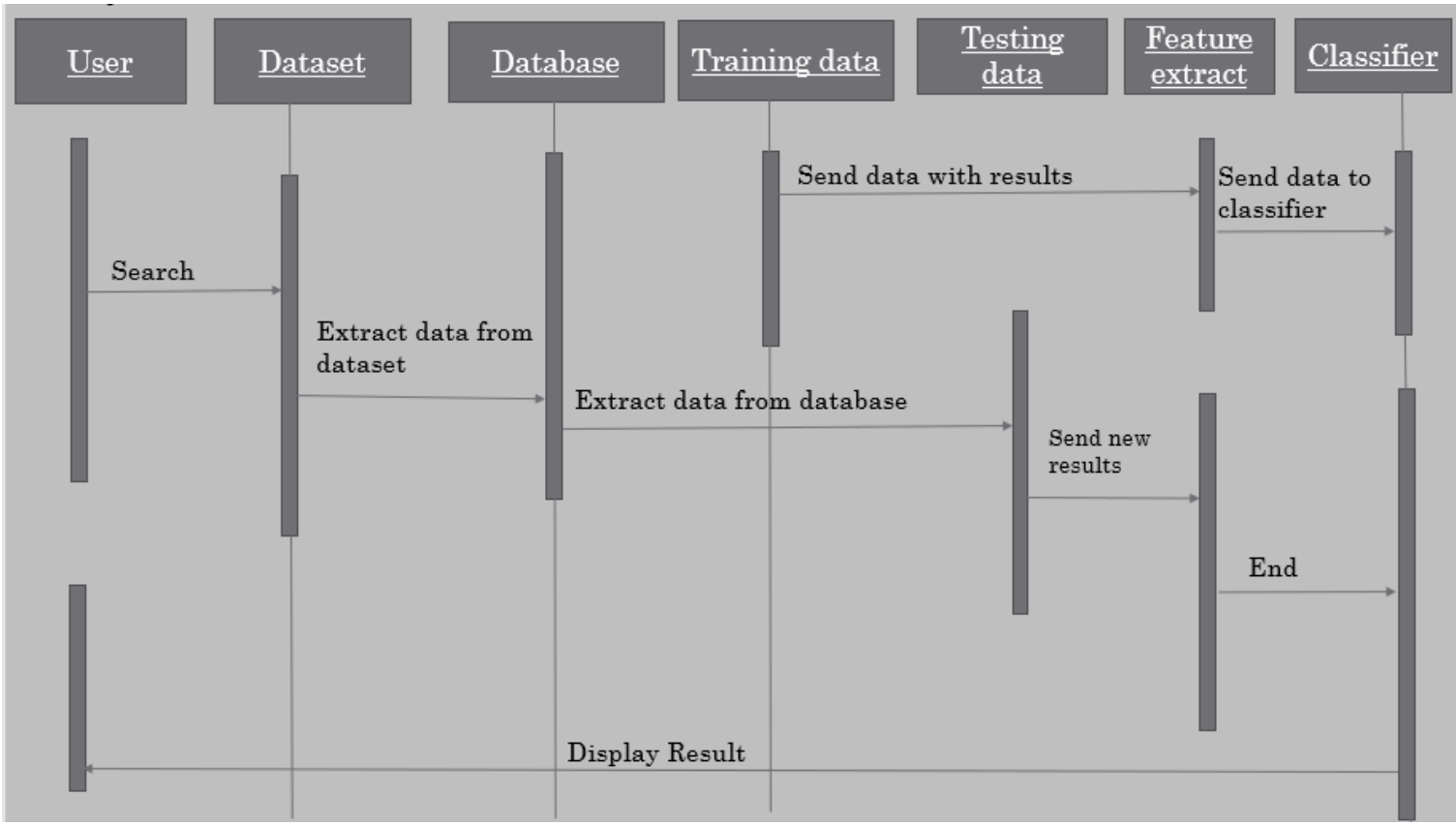

Fig. 1. Flow Diagram 1

With the help of the data, we can conduct the customer sentiment analysis. It will make you to know the problems and weakness, thereby adapting to policies which suits the customers needs and helps in improving the brand image or value thereby not only benefiting the customer experience but also sales

REFERENCES

[1]S. Rani and P. Kumar, "A Sentiment Analysis System to Improve Teaching and Learning," in Computer, vol. 50, no. 5, pp. 36-43, May 2017, doi: 10.1109/MC.2017.133.

[2]K. L. S. Kumar, J. Desai and J. Majumdar, "Opinion mining and sentiment analysis on online customer review," 2016 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), 2016, pp. 1-4, doi: 10.1109/ICCIC.2016.7919584.

[3]D. Zimbra, M. Ghiassi and S. Lee, "Brand-Related Twitter Sentiment Analysis Using Feature Engineering and the Dynamic Architecture for Artificial Neural Networks," 2016 49th Hawaii International Conference on System Sciences (HICSS), 2016, pp. 1930-1938, doi: 10.1109/HICSS.2016.244.

[4] M. H. Abd El-Jawad, R. Hodhod and Y. M. K. Omar, "Sentiment Analysis of Social Media Networks Using Machine Learning," 2018 14th International Computer Engineering Conference (ICENCO), 2018, pp. 174-176, doi: 10.1109/ICENCO.2018.8636124.

((Learning from Students' Perception on Professors Through Opinion Mining) [online] Available from:https://link.springer.com/chapter/10.1007/978-3-030-61702-8$_2 3 https://towardsdatascience.com/sentiment-analysis-concept-analysis-and-applications-6c94d6f58c17$

•https://www.ciodive.com/news/companies-using-sentiment-analysis-software-to-understand-employee-concerns/407357/

•https://www.wsj.com/articles/how-do-employees-really-feel-about-their-companies-1444788408

•https://www.researchgate.net/publication/51969319$_Scikit-learn_Machine_Learning_in_Python$

•https://scikit-learn.org/stable/modules/svm.html.

•https://towardsdatascience.com/naive-bayes-in-machine-learning-f49cc8f831b4

•https://www.tokenex.com/resource-center/what-is-tokenization: :text=Tokenization

•https://monkeylearn.com/ •https://scikit-learn.org/stable/