

BRAND REVIEW ANALYSIS USING TWITTER DATASET

Teena Varma¹, Albert Keerimolel², Brandon Gowray³, Janeya Limbachia⁴ And Kshitij Karekar⁵

¹ *Computer Engineering Department, Xavier Institute of Engineering, India*

² *Computer Engineering Department, Xavier Institute of Engineering, India*

³ *Computer Engineering Department, Xavier Institute of Engineering, India*

⁴ *Computer Engineering Department, Xavier Institute of Engineering, India*

⁵ *Computer Engineering Department, Xavier Institute of Engineering, India*

Abstract— A Machine Learning based model to analyze and study brand reviews on different products. Reviews consist of various special characters different hash tags and various emoticons. Here using machine Learning we first clean the dataset obtained and then using it predict the review as positive, negative or neutral regarding the particular product of the particular brand. This in turn helps companies keep a track of how the different products launched are doing in the market and what upgrade and changes can be made in order to get full customer satisfaction. Using the various decision algorithms of Machine Learning we can achieve better results and compare it with other competitors.

Keywords— Review, brand review, multinomial naive bayes, countvectorise, confusion matrix.

1. INTRODUCTION

Brand review is the basic of all customer services offered in today's day and age. Taking feedback and reviews from customers helps the companies to monitor the different goods and facilities being offered are up to standard of the customer and what they want, making it easy for them to use. Be it a new software of a new product launched, always tested for reviews and basis that the changes are made in order to give 100% customer satisfaction and keep up with the cooperators. In today's day and age with social media so prevalent reviews come like second nature to help keep the brand name and stay in touch with what is in demand and what improvements can be made. This model will then in turn help companies determine sale charts and projections on the success or failure of the product launched.

Being an integral part in today's day and age of technology where everything is almost reliant on

computers, this model helps companies keep up with the demands of the people and keeping in mind what is the new thing needed to be added. An example of this is when a product is not doing as well, it is first seen the reviews are like and then the decision is made to improve or recall the product.

2. LITERATURE SURVEY

Review system has been used for almost all things we use. Making a computer based Machine Learning model helps us reduce our work by doing everything for us and giving the final output regarding the product. Here we use decision making algorithms. Here we have used the Multinomial Naïve Bayes to predict the model accuracy. We already have in place algorithms which are used by companies to take the reviews and classify them based on the product and check how a product is doing.

3. PROBLEM STATEMENT

Since reviews are a vital part for any product or business, this project **Brand Review Analysis** helps us not only understand the different aspects of Machine Learning as in different algorithms and features used to implement them, but also gives us a view on how company strategies regarding their products are affected by this. This helps give an insight to the various aspects that go with comments reviews and social media.

4. PROPOSED SYSTEM

Brand Review Analysis performs the main action of determining how good a company's product and in turn how well like the company is with the people at large. A robust detection system is still a task because of the various different words and special characters used to express the emotion. In this project we use data preprocessing techniques to clean the data. Then using

countvectorizer and the commands related to it we use the Multinomial Naive Bayes Algorithm. After that we predict the model and check the accuracy. Finally preparing the confusion matrix and check the data. Then the graph is plot using scatter plot and it can be used for comparisons.

Block diagram is given as follows:

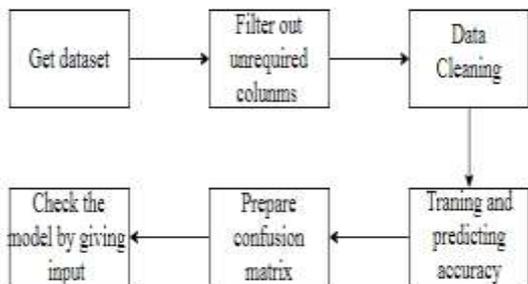


Fig.1. Block Diagram

Data flow diagram is given as following:

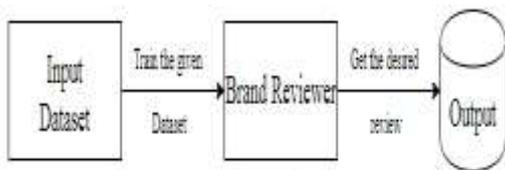


Fig.2. Data Flow Diagram

5. TEST CASES

Screenshot of the Application:

```

runfile('C:/Users/admin/.spyder-py3/multi_brand.py', wdir='C:/Users/admin/.spyder-py3')
Index(['_unit_id', '_golden', '_unit_state', '_trusted_judgments',
      '_last_judgment_at', 'sentiment', 'sentiment:confidence', 'date', 'id',
      'query', 'sentiment_gold', 'text'],
      dtype='object')

Number of words:4762

Most common words: [('need', 399), ('one', 384), ('aapf', 341), ('future', 327), ('computers', 316),
 ('batteries', 313), ('studio', 310), ('it', 310), ('outlet', 309), ('iphone', 293), ('store', 261), ('new', 176),
 ('phone', 164), ('agree', 159), ('stage', 157)]

neutral 2156
negative 1214
positive 422

**** Accuracy of naive bayesMultinomialNB is:****
0.7460456942003515
  
```

- Above results tabulated:

Accuracy of Model: 74.60%

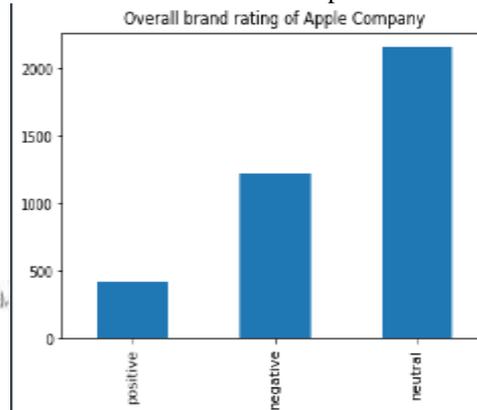
Confusion Matrix			
Predicted	Neg	Neut	Pos
Actual neg	242	132	0
Neut	57	605	0
Pos	18	82	2

Classification Report				
Sentiments	Precision	Recall	F1-score	Support
Negative	0.76	0.65	0.70	374
Neutral	0.74	0.91	0.82	662
Positive	1.00	0.02	0.04	102

- User Input Example Case:
After the model is trained, to check the model the user can give any input to get the desired output.

Classification Report	
User Input	Predicted Output
App comp is love	Positive

- Screenshot of Result Graph:



- Result of Test Case:
According to tests we performed, we were able to develop the model using the different python libraries and spyder software. We were able to detect the reviews and classify then for the company, in our case apple, using the twitter data set.

6. ADVANTAGES AND APPLICATION

A. Advantages:

Convenient to use: - This model is convenient as only training is required and then we can use it for any review.

Effective: - This is effective to use as we getting above 70% accuracy for the model.

B. Applications:

It can be used by companies to keep a check on their product's and how it is doing in the market.

7. CONCLUSION

Thus we have developed a brand review analysis system, which predicts how the brand is doing for the consumer.

- Software:

Using spyder and the different python libraries for the model and to clean data, apply the algorithm.

- Model:

We used the different models related to countvectorizer and the Multinomial Naive Bayes Algorithm.

Using all of the above we have implemented our Brand Review Analyzer.

ACKNOWLEDGMENTS

We would like to acknowledge our mentor and project guide Ms. Teena Varma for her support and guidance during the duration of this project.

REFERENCES

- [1] Apple Twitter Sentiment database by crowdflower.
- [2] Twitter Sentiment Analysis with Machine Learning
www.monkeylearn.com/blog/sentiment-analysis-of-twitter
- [3] Gurneet Kaur, Abhinash Singla Sentimental Analysis of Flipcart reviews using Naive Bayes and Decision Tree Algorithm. 2016.