PROJECT REPORT OF RECOMMENDATION SYSTEM BASED ON CUSTOMER REVIEWS

by

Team United

Members:

Mani Shankar Reddy Nayakanti Harshith Beeneedi Uday Sai Krishna Yarlagadda Sai Krishna

A Report submitted as per
of the requirements for the Python and Deep Learning course of
Master of Science
in the

School of Computing and Engineering
University of Missouri Kansas city

INTRODUCTION

The project report will introduce Recommendation system based on customer reviews using python NLTK and Django Framework. This is web application that will take product name as input and gives the customer the best product based on the ratings. In this project, the user first login to his/her account. If user does not exist, so he/she can create new account through new user options. After login into the system, the user can type the product name and our system will recommend them a best product based on the ratings. In order to log out of the system, user need to click on quit button. In this project a simple database is used to store customer login credentials. With this application customer will easily get to know the best product instead of thinking about which project to buy.

PROBLEM STATEMENT

Customers now a days are spending hours to find the best product available in the segment they are willing to buy. This is due to the poor browsing experience provided by the ecommerce websites. By searching through these sites they are finding it more difficult and are prone to buy the product which is not durable and worthy to the penny they are spending. Customers require a system that would help them to recommend the best product based on the rating and review of the product given by customers who have already bought it.

SOLUTION

To solve this problem Recommendation system based on customer review is developed. It uses unsupervised algorithm to find out the best review of the product and display it on the screen. By this way customer can benefit without spending hours for searching the product. Our goal for this project is to create one that is using the reviews of products, and ratings given and provide a best review to the customer.

DATASET USED

Amazon reviews for sentiment analysis is used . This data set contains two
columns i.e amazon reviews of the customers and their ratings.In the
ratings coloumn we have the classeslabel1 andlabel2, and there
is only one class per rowlabel1 corresponds to 1- and 2-star reviews,
andlabel2 corresponds to 4- and 5-star reviews. (3-star reviews i.e.
reviews with neutral sentiment were not included in the original), The
review titles, followed by ':' and a space, are prepended to the text. Most of
the reviews are in English, but there are a few in other languages, like
Spanish.

<u>PREPROCESSING</u>

 First step is to change the raw data to columns format. For example we have data in this format.

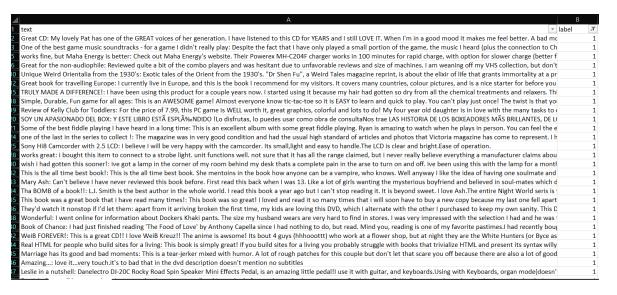
```
This book made me want to read again. Very enjoyable.

_label__1 The Worst!: A complete waste of time. Typographical errors, poor grammar, and a totally pathetic plot add up to absolutely nothing. I'm embarrassed for this author and very disappointed I actually paid for this book.

_label__2 Great book: This was a great book, I just could not put it down, and could not read it fast enough. Boy what a book the twist and turns in this just keeps you guessing and wanting to know what is going to happen next. This book makes you fall in love and can heat you up, it can also make you so angery. this book can make you go throu several of your emotions. This is a quick read romance. It is something that you will want to end your day off with if you read at night.

_label__2 Great Read: I thought this book was brilliant, but yet realistic. It showed me that to error is human. I loved the fact that this writer showed the loving side of God and not the revengeful side of him. I loved how it twisted and turned and I could not put it down. I also loved The glass castle.
```

 We have converted this to two columns. One column having reviews and another column having ratings. We have coded Label 1 as 0 and Label 2 as 1 for our convenience and have placed two columns.
 Named text and label.



We have ignored languages other than English by using stop words.

stop words = set(stopwords.words('english'))

Converting all upper-case letters to lower- case letters.

MODEL

For processing the data and finding the insights of the data we have used Sentiment intensity Analyzer model. NLTK already has a built-in, pretrained sentiment analyzer called VADER. As our model is sentiment based analyzer we have to find positive, negative scores based on the sentiment using polarity scores. We have taken the score greater than 0.05 as positive and less than that score as negative.

```
stop_words = set(stopwords.words('english'))
sid = SentimentIntensityAnalyzer()

def getReview(review):
    review_result = "none"
    review = review.lower()
    review = re.sub('[^A-Za-z]+', ' ', review)
    sentiment_dict = sid.polarity_scores(review.strip())
    compound = sentiment_dict['compound']
    if compound >= 0.05 :
        review_result = 'Positive'
    return review_result
```

POST PROCESSING

If the product is not found in the dataset it will give the below Output.



If the product is found in the data set it will give the below output.



Product Name	Recommended Best Review	Recommended Best Rating	Suggestion
movie	"There Was Me": Great movie! Stanley Kubrick is personaly my favorite director after watching this movie. After I read the book I went out and bought the movie without seeing it. I think that the movie is as good as the book. The only downside to the whole movie is it did not add the additional chapter like in the book. But even the ending to the movie is great! Kubrick really scored big in this movie. Back then to make a movie as hard core as this one was crazy but he did it. This movie is so ahead of it's time. But it's my favorite movie of all time!	1	you have chosen best product

APPLICATION

 The application we have created is a web application using Django framework and have used my sql database to store the login credentials given by user during registration process. These tables are created initially when the course management system is deployed.

create table register(username varchar(30) primary key, password varchar(30), contact varchar(12), email varchar(100), address varchar(40));

• The user data is captured and stored in the table as below.



•	Once the user searches for the product name and clicks on submit button we will see the output with the output consisting of product review and rating.
•	Our application uses sentiment analysis to learn the hidden emotions in the product and based on polarity scores it will recommend us the best product with the good reviews.

REFERENCES

https://www.kaggle.com/bittlingmayer/amazonreviews

https://realpython.com/python-nltk-sentiment-analysis/

https://www.fullstackpython.com/django.html

https://www.digitalocean.com/community/tutorials/build-your-first-python-and-django-application

https://medium.com/coders-camp/60-python-projects-with-source-code-919cd8a6e512