

**Project Report**

**Analysis of Amazon cell phone reviews**

*Project Report submitted to fulfill the requisites for the award of Smart Bridge Internship Certificate*

**Work Submitted by :**

Dharshini M

**Table of Contents**

Project Details ---------------------------------- 3

acknowledgment ---------------------------------- 4

1. Introduction ---------------------------------- 5
   1. Overview
   2. Purpose
2. Literature Survey ---------------------------------- 6
   1. Existing Problem
   2. Proposed Solution
3. Theoretical Analysis ---------------------------------- 7
   1. Block Diagram
   2. Hardware/Software Designing
4. Experimental Investigations-------------------------------- 8
5. Flowchart -------------------------------- 10
6. Result -------------------------------- 11
7. Advantages and Disadvantages----------------------------- 12
8. Conclusion ------------------------------------------------- 14
9. Future scope -------------------------------- 15
10. bibilography- ------------------------------- 16

Appendix -------------------------------- 17

A. Source code --------------------------------

# Project Details

**Project ID** : sps\_pro\_180

**Project Title** : Analysis of Amazon Cell Phone Reviews

**Internship title** : RSIP Career basic AI 022

**Duration** : 21 days

**NAME** : Dharshini M

**Project Support** : The Smart Bridge Educational

Services

**Kickoff Date**  : June 1st, 2020

**Finish Date** : June 30th, 2020

# ACKNOWLEDGEMENT

This project has taken a considerable amount of time and resources. I would like to acknowledge the help of all of those who have made this project possible. In finical I would like to thank my supervisor Mr. Rammohan Bethi Gahlot for his time, patience and guidance, and also for allowing the idea to be pursued primitively. I would also like to thank Mr. Vinay kumar Nomula for his help. Further to these people I would like to thank the members of the Smartbridge career workshop for their technical help in setting up various codes and faults. Also, I would like to thank all of my co-interns who have worked on the Open Source projects without whose efforts this project would not have been possible.

# 1.Introduction

1. **a .Overview**

This project aims at building a model to predict the helpfulness of the review and the rating based on the review text. Corpus-based and knowledge-based methods can be used to determine the semantic similarity of review text. We will be using Natural language processing to analyse the sentiment ( positive or a negative) of the given review . A sample web application is integrated to the model built.

1. **b Purpose**

The purpose of this project is to predict the rating and review of particular cell phone in the

Amazon. So that the consumers read online reviews before they decide to purchase any Mobile phone from any e-commerce website. Online Mobile applications has revolutionized the way consumers purchase mobile phones online as these apps have all the information regarding any mobile phone at users finger tips. Amazon is one of the best mobile applications which is considered as a treasure trove of all mobile reviews, and their review system is accessible across all channels presenting reviews in an easy-to-use format. So,There should be a system which analyses thousands of reviews of unlocked mobile phones sold on Amazon.com to find insights with respect to reviews, ratings, price and their relationships.

# 2.Literature Survey

1. **a Existing Problem**

As we have considered Amazon as the source of data, the rating process which is used to rate the products in Amazon will lead to False Positive which in turn leads the customers to purchase the products by seeing the false rated products . For example, consider a sample product review which is given by a user in Amazon as shown below.

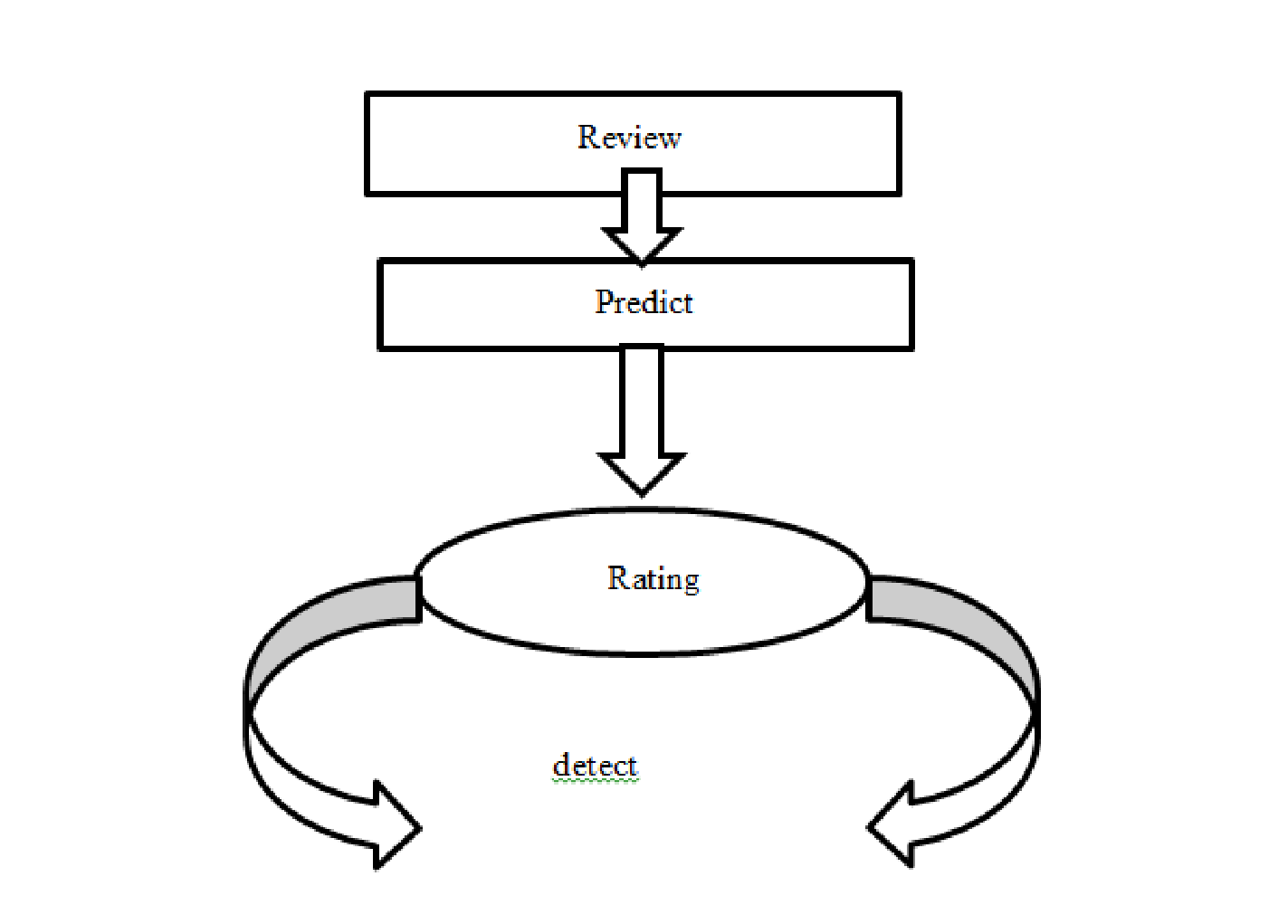
The below figure shows that the process of giving the ratings to products. As on when if the rating is in the form of stars we can understand that highest number of stars indicates that the product overall rating is good and the customer satisfies to the maximum extent. If it goes to the least number the customer completely disappointed with the product and it is not up to the range of customer's needs and the remaining rating process gives the satisfactory ranges of the moderate customers who ever partially satisfy with that product.

1. **b.Proposed Solution**

By analyzing all the existing review and rating processes we have observed that there is no clarity on the products rating and there is no feature level rating of the product which helps the customers to buy a product based on a specific feature. In our proposed approach we are trying to concentrate on the feature level rating process of the product by considering the reviews given by the users.

# 3.Theoritical Analysis:

1. **a Block Diagram**



1. **b Hardware/Software Designing** **Hardware:**

**Manufacturer:** Microsoft Corporation

**Processor:**I ntel@Core i5-8650u cpu@1.90GHz 2.11Ghz

Installed memory(RAM):8.0GB

**System Type:** 64-bit Operating System,x64-based Processor

**Software:**

**Operating System:**Windows 10

**Anaconda version:** 3.7 for windows

**Python version:** 3.7

**Tensorflow version:** 1.14.0 using anaconda

**Keras version:** 2.2.4 using anaconda

# Experimental Investigations:

Various experiments are applied on dataset which is based on Python, Python Web Frame Works, Natural Language Processing(NLP) concepts like removing Punctuations and numbers, converting each word into lower case, stemming ,Adding input layers, output layers, hidden layers,optimizing the model all these experiments are performed in model classify the data set accurately

**asin** **name** **rati date** **verifi helpfulV outp review**

**ng** **ed** **otes** **ut**

B0000SX October 11, Def not best, but not worstI had the

1. Janet 3 False 1.0 0 2UC 2005 Samsung A...

B0000SX Luke January 7, Text Messaging Doesn't WorkDue to

1. 1 False 17.0 0

2UC Wyatt 2004 a software i...

B0000SX December 30, Love This PhoneThis is a great,

1. Brooke 5 False 5.0 1 2UC 2003 reliable phone...

B0000SX amy m. March 18, Love the Phone, BUT...!I love the

1. 3 False 1.0 0

2UC teague 2004 phone and al...

B0000SX tristazbim August 28, Great phone service and options,

1. 4 False 1.0 1

2UC mer 2005 lousy case!Th...

Epoch 1/8

54388/54388 [==============================] - 760s 14ms/step - loss: 0.2664 - acc: 0.8964

Epoch 2/8

54388/54388 [==============================] - 751s 14ms/step - loss: 0.1564 - acc: 0.9405

Epoch 3/8

54388/54388 [==============================] - 753s 14ms/step - loss: 0.0722 - acc: 0.9757

Epoch 4/8

54388/54388 [==============================] - 742s 14ms/step - loss: 0.0329 - acc: 0.9897

Epoch 5/8

54388/54388 [==============================] - 743s 14ms/step - loss: 0.0284 - acc: 0.9921

Epoch 6/8

54388/54388 [==============================] - 749s 14ms/step - loss: 0.0258 - acc: 0.9933

Epoch 7/8

54388/54388 [==============================] - 740s 14ms/step - loss: 0.0254 - acc: 0.9935

Epoch 8/8

54388/54388 [==============================] - 739s 14ms/step - loss: 0.0185 - acc: 0.9947

**Pre Processing:**

It refers to transformation applied to our data before feeding it to the algorithm. The messages have to be pre-processed for the removal of unwanted punctuations, stemming, converting to lower case etc..

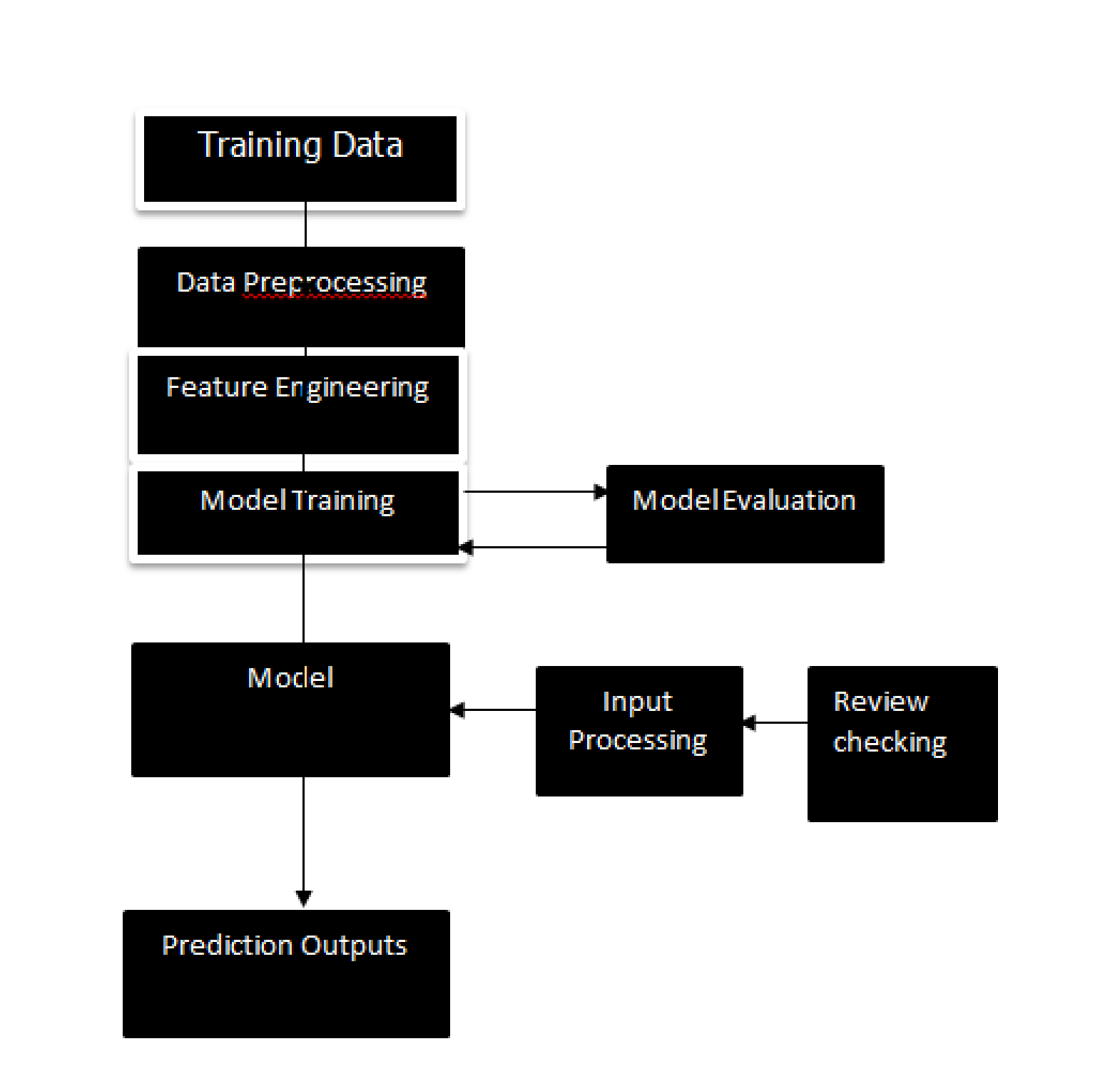
**Stemming:**

Stemming is another pre-processing step that normalize sentences. Stemming is a way to account for the variations of words and sentences which often have a same meaning; furthermore, it will help us shorten the sentences and shorten our lookup.

Stemming helps reduce a word to its stem form. It often makes sense to treat related words in the same way. It removes suffices, like “ing”, “ly”, “s”, etc. by a simple rule-based approach. It reduces the corpus of words but often the actual words get neglected.

# 5.FLOW CHART

Various experiments are performed to evaluate the performance of our proposed Analysis of Amazon Phone review. Initially we have selected features on the basis of behavior of analysing the rating and review of previous customers who buyed the phone earlier through Amazon and extracted these features from the dataset to get the feature vector.So that the consumer can easily recognize the usage of the phone.



# Result

Various experiments are performed to evaluate the performance of our proposed Analysis of Amazon Phone review. Initially we have selected features on the basis of behavior of analysing the rating and review of previous customers who bought the phone earlier through Amazon and extracted these features from the dataset to get the feature vector. So that the consumer can easily recognize the usage of the phone.



As a result, My trained model is analyzing the reviews with an accuracy of nearly 0.99

1. **ADVANTAGES and DISADVANTAGES:**

**ADVANTAGES**

1. **A Larger Market**

Amazon Cell Phone Reviews allows you to reach customers all over the country and around the world. Your customers can make a purchase anywhere and anytime, especially more people are getting used to shopping on their mobile devices.

1. **Customer Insights Through Tracking And Analytics**

Whether you're sending visitors to your Amazon website through SEO, PPC ads or a good old postcard, there is a way to track your traffic and customers' entire user journey to get insights into keywords, user experience, marketing message, pricing strategy, and more.

1. **Fast Response To Consumer Trends And Market Demand**

The streamlined logistics, especially for merchants who do "drop ship," allow businesses to respond to market and Amazon trends and consumer demands in a nimble manner. Merchants can also create promotions and deals on the fly to attract customers and generate more sales.

1. **Lower Cost**

With the advance in Amazon platform technologies it has become very easy and affordable to set up and maintain an Amazon store with a low overhead. Merchants no longer have to spend a large budget on TV ads or billboard, nor worry about the expense for personnel and real estate.

1. **More Opportunities To "Sell"**

Merchants can only provide a limited amount of information on a product in a physical store. On the other hand, Amazon websites allow the space to include more information such as demo videos, reviews, and customer testimonials to help increase conversion.

**DISADVANTAGES:**

1. **Lack Of Personal Touch**

Some consumers value the personal touch they get from visiting a physical store and interacting with sales associates. Such personal touch is

Particularly important for businesses selling high-end products as customers not only want to buy the merchandise but also have a great experience during the process.

1. **Lack Of Tactile Experience**

No matter how well a video is made, consumers still can't touch and feel a product. Not to mention, it's not an easy feat to deliver a brand experience, which could often include the sense of touch, smell, taste, and sound, through the two-dimensionality of a screen.

1. **Price And Product Comparison**

With online shopping, consumers can compare many products and find the lowest price. This forces many merchants to compete on price and reduce their profit margin.

1. **Need For Internet Access**

This is pretty obvious, but don't forget that your customers do need Internet access before they can purchase from you! Since many eCommerce platforms have features and functionalities that require high-speed Internet access for an optimal customer experience, there's a chance you're excluding visitors who have slow connections.

1. **Credit Card Fraud**

Credit card fraud is a real and growing problem for online businesses. It can lead to chargebacks that result in the loss of revenue, penalties, and bad reputation.

# 10.Future Scope

Amazon business is the best option available for the people to build a better business world for insuring success in future rather than doing a traditional mode of business. For any business person, to have an e-commerce business is added advantage for their business. Several factors for the importance of Amazon business:-

**Convenience**

Sometimes, the question arises that why do you use Amazon website for online shopping? The most prominent answer to this question is convenience. One person can shop, buy and sell products while sitting at home at any time.

**Round the clock services**

Amazon provides us round the clock services at all times even in midnight. So the customers do not require visiting a physical market if they need something during the night. It is the most convenient option for the people who are usually busy with their working schedules. So it helps you to be available for your customer 24×7.

**Wide Platform**

Amazon brings a wide range of customers across the nation or globe to your business. Therefore, it is a wise choice to choose an Amazon platform to cross the geographical barriers for your business.

**Business partner**

Amazon is directly link to your business promotions, as it is the age of digital media. Making your business available online is crucial to your business development such as, highly convenience, wide exposure, global customer, easy to run, etc. and it will help in creating a strong & global brand image for your business.

**Low cost**

If the inventory management of goods and services is an automated process then not only there will be a reduction in costs, but also in risk. Also having an Amazon business is much more cost effective than a physical store as it saves your extra expenses like rent, electricity, etc.,

# 11.Bibliography

**Dataset :**  <https://www.kaggle.com/datasets?search=Amazon+cell+phone+reviews>

**Html file:**

<https://www.w3schools.com/bootstrap/bootstrap_forms_inputs.asp>

**APPENDIX**

**Source code:**

**mymodel.py:**

I had run this code in Jupyter notebook. so it has been saved as mymodel.ipynb

import pandas as pd import numpy as np data = pd.read\_csv('reviews.csv') data.head() data.shape data['rating'].unique() data.insert(8,"output","") data.head() for i in range(0,len(data['rating'])): num=data['rating'][i] if num>3:

data['output'][i]=1 else:

data['output'][i]=0 data.head() data['review']=data['title']+data['body'] data.head() y=data.iloc[:,6:7].values y.shape import re import nltk#natural language tool kit nltk.download("stopwords") from nltk.corpus import stopwords from nltk.stem.porter import PorterStemmer data['review'].isnull().any() data['review'].fillna(data['review'].mode()[0],inplace=True) data['review'].isnull().any() list1=[] for i in range(len(data['review'])):

review=data["review"][i]

review=re.sub("[^a-zA-Z]"," ",review)

review=review.lower()

review=review.split()

review=[ps.stem(word) for word in review if not word in set(stopwords.words('english'))]

review=" ".join(review)

list1.append(review) #Vectorization from sklearn.feature\_extraction.text import CountVectorizer cv=CountVectorizer(max\_features=3000) x=cv.fit\_transform(list1).toarray() x.shape !mkdir -p count1 import pickle

Filename="count1/countvectorizer.pkl" with open(Filename,"wb") as file:

pickle.dump(cv,file) from sklearn.model\_selection import train\_test\_split x\_train,x\_test,y\_train,y\_test=train\_test\_split(x,y,test\_size=0.2,random\_state=0) x\_train.shape from keras.models import Sequential from keras.layers import Dense model = Sequential() model.add(Dense(units =3000 ,init = "random\_uniform",activation = "relu")) model.add(Dense(units =3000 ,init = "random\_uniform",activation = "relu"))

model.add(Dense(units = 1 ,init = "random\_uniform",activation = "sigmoid")) model.compile(optimizer = "adam",loss = "binary\_crossentropy",metrics = ["accuracy"]) model.fit(x\_train,y\_train,epochs = 8) y\_pred=model.predict(x\_test) y\_pred=(y\_pred>0.5) y\_pred y\_test y\_p=model.predict(cv.transform(['bad phone'])) y\_p y\_p>0.5

model.save('mymodel.h5')

**app.py:**

# -\*- coding: utf-8 -\*-

"""

Created on Sat Jun 27 07:30:23 2020

@author: DELL

"""

from keras.models import load\_model import tensorflow as tf import os

global graph graph = tf.compat.v1.get\_default\_graph() from flask import Flask , request, render\_template,url\_for import pickle import re import nltk#natural language tool kit nltk.download("stopwords") from nltk.corpus import stopwords from nltk.stem.porter import PorterStemmer ps=PorterStemmer() with open(r'countvectorizer.pkl','rb') as file:

cv=pickle.load(file) cla=load\_model('mymodel.h5',compile=False) cla.compile(optimizer='adam',loss='binary\_crossentropy') app = Flask(\_\_name\_\_)

@app.route('/') def home():

return render\_template('home.html')

@app.route('/select',methods=['POST']) def select():

return render\_template('base.html')

@app.route('/page') def page():

return render\_template('base.html')

@app.route('/predict',methods=['GET','POST']) def predict():

if request.method == 'GET':

img\_url = url\_for('static',filename = 'style/3.jpg') return render\_template('base.html',url=img\_url) if request.method == 'POST':

topic = request.form['ms']

review=re.sub("[^a-zA-Z]"," ",topic)

review=review.lower()

review=review.split()

review=[ps.stem(word) for word in review if not word in set(stopwords.words('english'))]

topic=" ".join(review)

print("I m" +topic)

topic=cv.transform([topic]) print("\n"+str(topic.shape)+"\n") with graph.as\_default():

y\_pred = cla.predict(topic) print("pred is "+str(y\_pred)) if(y\_pred>0.5):

img\_url = url\_for('static',filename = 'style/1.jpg') topic = "happy" elif(y\_pred<0.5):

img\_url = url\_for('static',filename = 'style/2.jpg')

topic = "sad"

return render\_template('base.html',ypred = topic)

if \_\_name\_\_ == '\_\_main\_\_':

app.run(debug = False, threaded = True)

**home.html:**

<!DOCTYPE html>

<html lang="en">

<head>

<meta charset="UTF-8">

<title>home.html</title>

<style> html, body { height: 100%; } body {

background:radial-gradient(ellipse at center, rgba(255,254,234,1) 0%, rgba(255,254,234,1) 35%,

#B7E8EB 100%);

overflow: hidden;

}

body { height: 100vh;

display: flex;

justify-content: center;

align-items: center;

overflow: hidden;

}

.background {

height: 5%;

width:100%;

position:absolute;

bottom:0;

left:0;

background: #015871;

}

p{

font-family: Impact, Charcoal, sans-serif;

font-size:100px;

}

.btn {

box-sizing: border-box;

-webkit-appearance: none;

-moz-appearance: none;

appearance: none;

background-color: transparent;

border: 2px solid #1459DE;

border-radius: 0.6em;

color: #1459DE;

cursor: pointer;

display: -webkit-box; display: -webkit-flex;

display: -ms-flexbox;

display: flex;

-webkit-align-self: center;

-ms-flex-item-align: center;

align-self: center;

font-size: 1rem;

font-weight: 400;

line-height: 1;

margin: 20px;

padding: 1.2em 2.8em;

text-decoration: none;

text-align: center;

text-transform: uppercase;

font-family: 'Montserrat', sans-serif;

font-weight: 700;

}

.btn:hover, .btn:focus { color: #fff;

outline: 0;

}

.first {

-webkit-transition: box-shadow 300ms ease-in-out, color 300ms ease-in-out; transition: box-shadow 300ms ease-in-out, color 300ms ease-in-out;

}

.first:hover {

box-shadow: 0 0 40px 40px #1459DE inset;

}

</style>

</head>

<body>

<div class="background">

</div>

<center>

<form name="passdata" action="/select" method="POST">

<link href="https://fonts.googleapis.com/css?family=Montserrat:900" rel="stylesheet">

<p class="rock">AMAZON CELL<br>PHONE REVIEWS<br>

<button class="btn first">CLICKHERE</button>

</p>

</form>

</center>

</body>

</html>

**base.html:**

<!DOCTYPE html>

<html lang="en">

<head>

<meta charset="UTF-8">

<title>base.html</title>

<style> html, body { height: 100%; } body {

background:radial-gradient(ellipse at center, rgba(255,254,234,1) 0%, rgba(255,254,234,1) 35%,

#B7E8EB 100%);

overflow: hidden;

}

body { height: 100vh;

display: flex;

justify-content: center;

align-items: center;

overflow: hidden;

}

.background {

height: 5%;

width:100%;

position:absolute;

bottom:0;

left:0;

background: #015871;

}

.box {

position: absolute;

top: 50%; left: 50%;

transform: translate(-50%, -50%);

width: 35rem;

padding: 2.5rem;

box-sizing: border-box;

background: #fff;

border-radius: 0.625rem;

}

.box h2 {

margin: 0 0 1.875rem;

padding: 0;

text-align: center;

}

.box .inputBox {

position: relative;

}

.box .inputBox input { width: 100%; padding: 1rem 0;

font-size: 1rem;

color: black ;

letter-spacing: 0.062rem;

margin-bottom: 3.875rem;

border: none;

border-bottom: 0.075rem solid black;

outline: none;

background: transparent;

}

.box .inputBox label {

position: absolute;

top: 0; left: 0;

padding-bottom: 1rem;

font-size:1.2rem;

pointer-events: none;

transition: 0.5s;

}

.box input[type="submit"] {

border: none;

outline: none; color: #black;

padding: 0.625rem 1.25rem;

cursor: pointer;

border-radius: 0.312rem;

font-size: 1rem;

}

.button {

border-radius: 9px;

background-color: #f4511e;

border: none;

color: #FFFFFF;

text-align: center;

font-size: 10px;

padding: 5px;

width: 50px;

transition: all 0.5s; cursor: pointer; margin: 5px;

}

.glow-on-hover {

width: 220px;

height: 50px;

border: none;

outline: none;

color: #fff;

background: blue;

cursor: pointer;

position: relative;

z-index: 0;

border-radius: 10px;

}

</style>

</head>

<body>

<div class="background">

</div>

<div class="box">

<h2>AMAZON'S REVIEW ANALYSIS</h2>

<br>

<br>

<center><form name="passdata" action="/predict" method="POST">

<div class="inputBox">

<label>Enter your text below:</label>

<br>

<input type="text" name="ms">

</div>

<br>

<button class="glow-on-hover" type="submit" value="submit" >SUBMIT</button>

</form>

</center>

<center> <p>

{% if ypred == "happy" %}

<p>Positive Feedback</p>

<img src="/static/1.jpg" height="60" width="70" alt="Positive" class="tab3">

{% else %}

{% if ypred == "sad" %}

<p>Negative Feedback</p>

<img src="/static/2.jpg" height="60" width="70" alt="Negative" class="tab3">

{% endif %}

{% endif %}

</p>

</center>

</div>

</body>

</html>