# VIGNANA BHARATHI INSTITUTE OF TECHNOLOGY

**A PROJECT REPORT ON**

**“ Applying suitable ML technique on any dataset and if possible deploying it using heroku and streamlit i.e.,Applying Clssification(Supervised learning) on the dataset 'amazon.csv' and deploying it using streamlit“**

## SUBMITTED BY:


**ADAPA HARSHITHA**

**ROLL NO: 20P61A0502**

**BRANCH: CSE**

**BATCH: 2020-2024**


**UNDER THE GUIDANCE OF**


**MR. AMEEN MANNA**

**DEPARTMENT OF COMPUTER SCIENCE ENGINEERING**

**VIGNANA BHARATHI INSTITUTE OF TECHNOLOGY**

**GHATKESAR - 501301**

## Required python code

```
#predicting whether an amazon review is positive or not(1-POSITIVE REVIEW,0-NEGATIVE REVIEW)

Import pandas as pd

df=pd.read_csv('/content/amazon.csv')

df


df.info()


df['Positive'].value_counts()


df['Positive'].value_counts().plot(kind='bar')


#dividing data inyo i/p and o/p

x=df.iloc[:,0].values

y=df.iloc[:,1].values


#train_test_split

from sklearn.model_selection import train_test_split

x_train,x_test,y_train,y_test=train_test_split(x,y,random_state=0)


#Apply TFIDF Vectorizer

from sklearn.feature_extraction.text import TfidfVectorizer

vect = TfidfVectorizer()
```

```
x_train_v = vect.fit_transform(x_train)#for training data it is vect.fit_transform

x_test_v = vect.transform(x_test)#for testing data ,it is vect.transform


#Apply CLASSIFIER(SVC)

from sklearn.svm import SVC

model = SVC()


#fitting the model

model.fit(x_train_v,y_train)


#Predictor variable

y_pred = model.predict(x_test_v)

y_pred # predicted values


y_test # actual values


#Accuracy

from sklearn.metrics import accuracy_score

accuracy_score(y_pred,y_test)*100


a = df['reviewText'][19995]

a
```

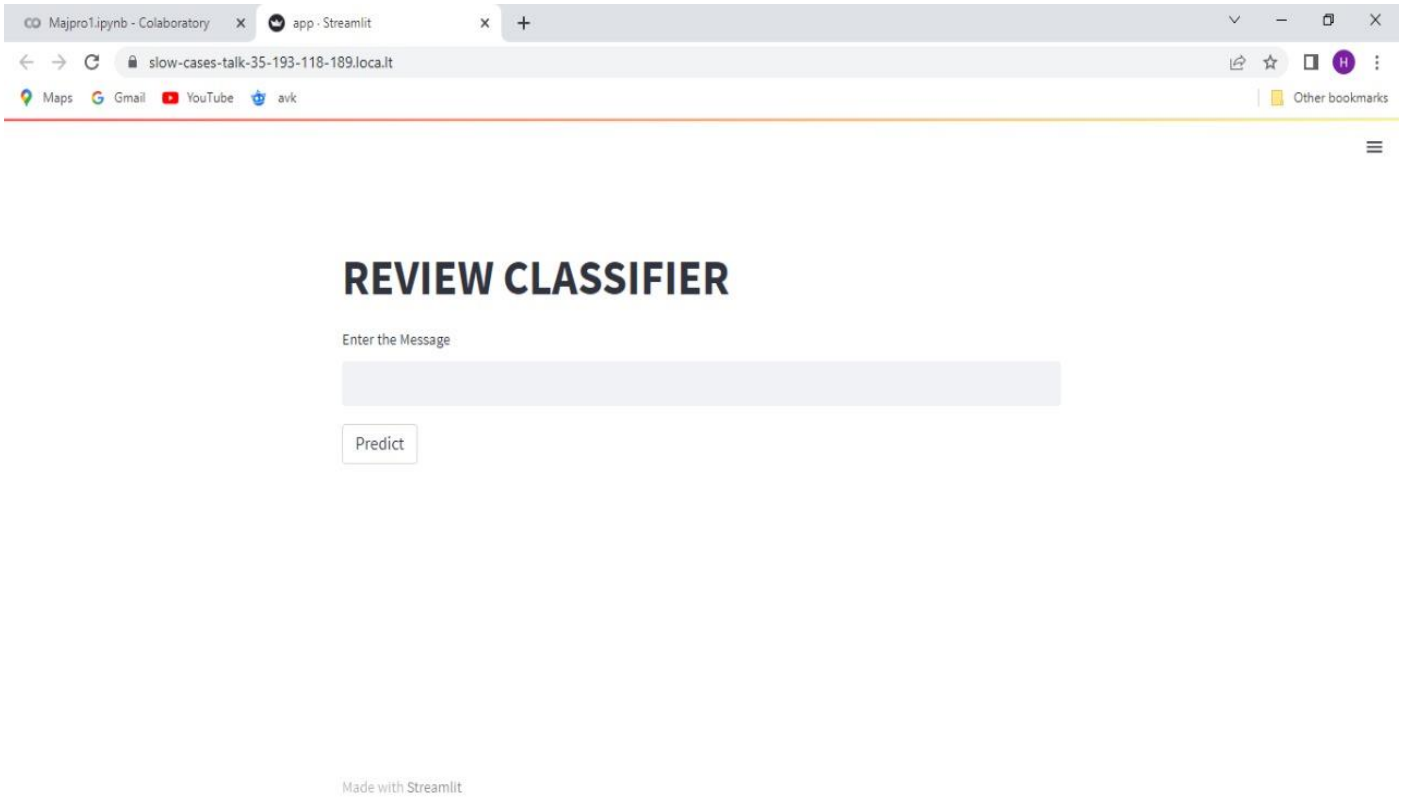
```
a = vect.transform([a])
model.predict(a)

#Pipelining
from sklearn.pipeline import make_pipeline
text_model = make_pipeline(TfidfVectorizer(),SVC())
text_model.fit(x_train,y_train)

#predictor variable
y_pred1 = text_model.predict(x_test)
y_pred1

y_test

accuracy_score(y_pred1,y_test)*100



a1 = df['reviewText'][2]
a1

text_model.predict([a1])

text_model.predict(['ver bad'])
```

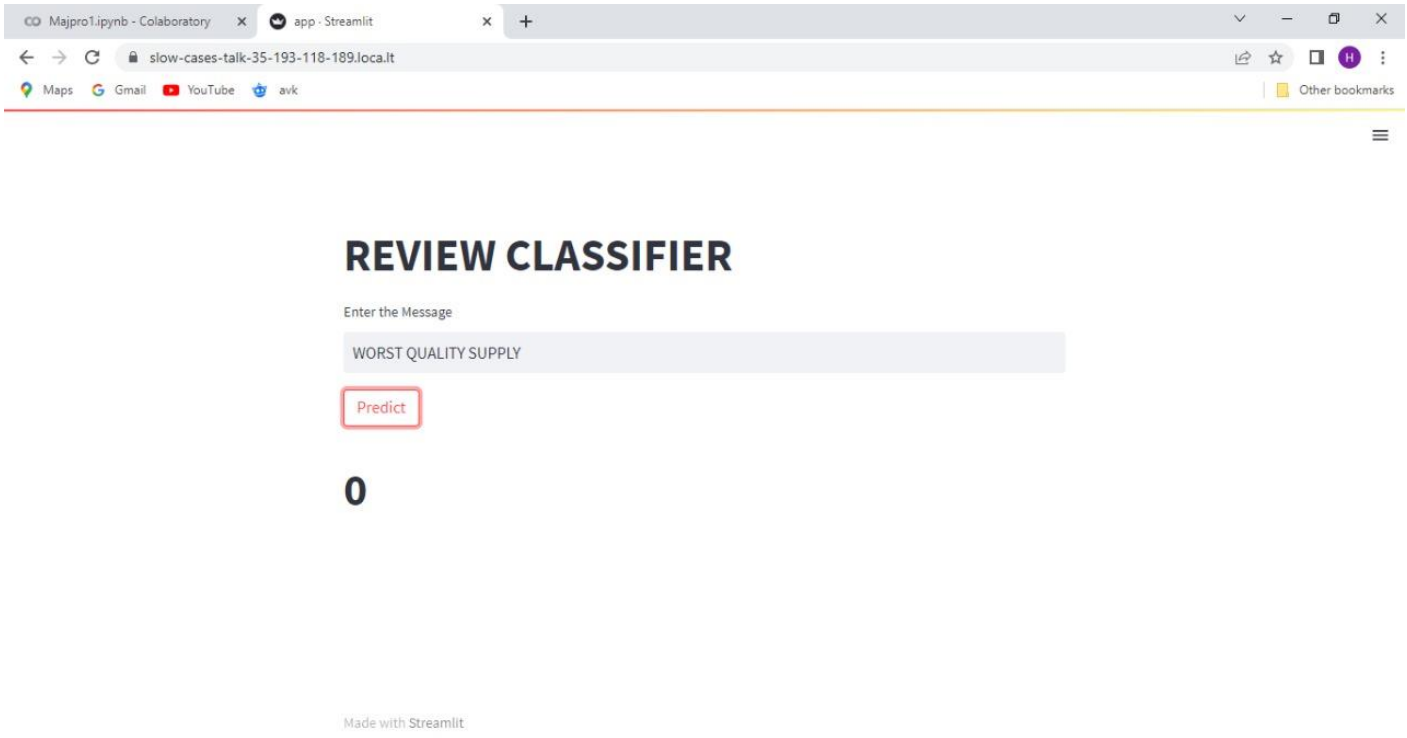
```
#Joblib has 2 types -  1.Dump and 2.Load

import joblib

joblib.dump(text_model,'amazon.csv')

#We are creating a new file called spam-ham,and we are dumping the pipelined model inside it


#CREATE A WEB APP USING SREAMLIT


!pip install streamlit --quiet # Installing the streamlit library


%%writefile app.py

#%%writefile app.py is a magic command which creates a file named app.py

import streamlit as st

import joblib

model = joblib.load('amazon.csv')

st.title('REVIEW CLASSIFIER')# creates a title with name review classifier

ip = st.text_input('Enter the Message')#creates a text box

op = model.predict([ip])

if st.button('Predict'):

  st.title(op[0])#st.button will create a button with name predict

  #st.title(op[0]) - the output will be displayed as a title
```

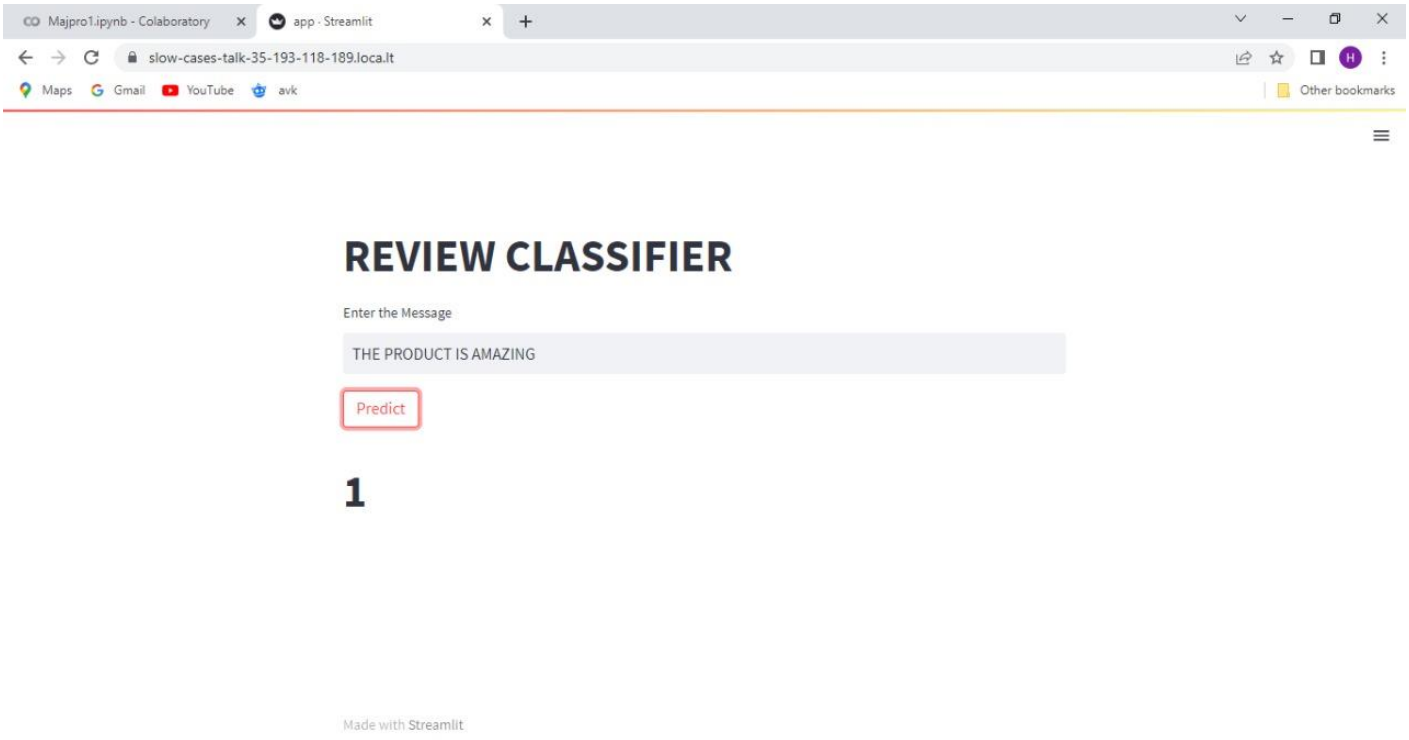
**#TEMPORARY DEPLOYMENT**

**!streamlit run app.py & npx localtunnel --port 8501**

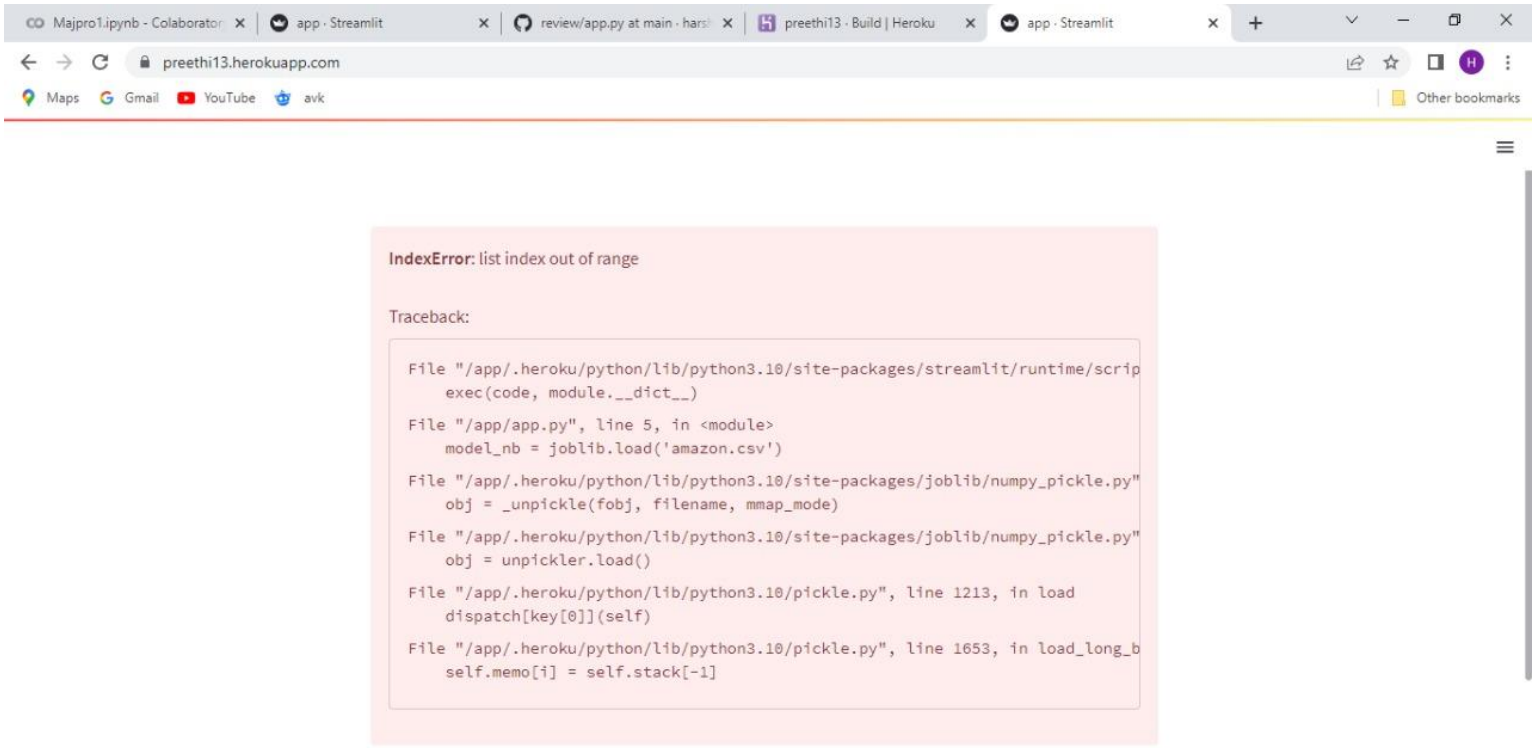
**#8501 is the default port number for local tunnel**


Majpro1.ipynb - Colaboratory
app · Streamlit
slow-cases-talk-35-193-118-189.loca.lt
Maps
Gmail
YouTube
avk
Other bookmarks
REVIEW CLASSIFIER
Enter the Message
Predict
Made with Streamlit


Majpro1.ipynb - Colaboratory
app · Streamlit
slow-cases-talk-35-193-118-189.loca.lt
Maps
Gmail
YouTube
avk
Other bookmarks
REVIEW CLASSIFIER
Enter the Message
WORST QUALITY SUPPLY
Predict
0
Made with Streamlit




**When tried to deploy it through heroku got this.**


Majpro1.ipynb - Colaborator
app · Streamlit
review/app.py at main · hars
preethi13 · Build | Heroku
app · Streamlit
preethi13.herokuapp.com
Maps
Gmail
YouTube
avk
Other bookmarks
IndexError: list index out of range
Traceback:
File "/app/.heroku/python/lib/python3.10/site-packages/streamlit/runtime/scrip
exec(code, module.__dict__)
File "/app/app.py", line 5, in <module>
model_nb = joblib.load('amazon.csv')
File "/app/.heroku/python/lib/python3.10/site-packages/joblib/numpy_pickle.py"
obj = _unpickle(fobj, filename, mmap_mode)
File "/app/.heroku/python/lib/python3.10/site-packages/joblib/numpy_pickle.py"
obj = unpickler.load()
File "/app/.heroku/python/lib/python3.10/pickle.py", line 1213, in load
dispatch[key[0]](self)
File "/app/.heroku/python/lib/python3.10/pickle.py", line 1653, in load_long_b
self.memo[i] = self.stack[-1]