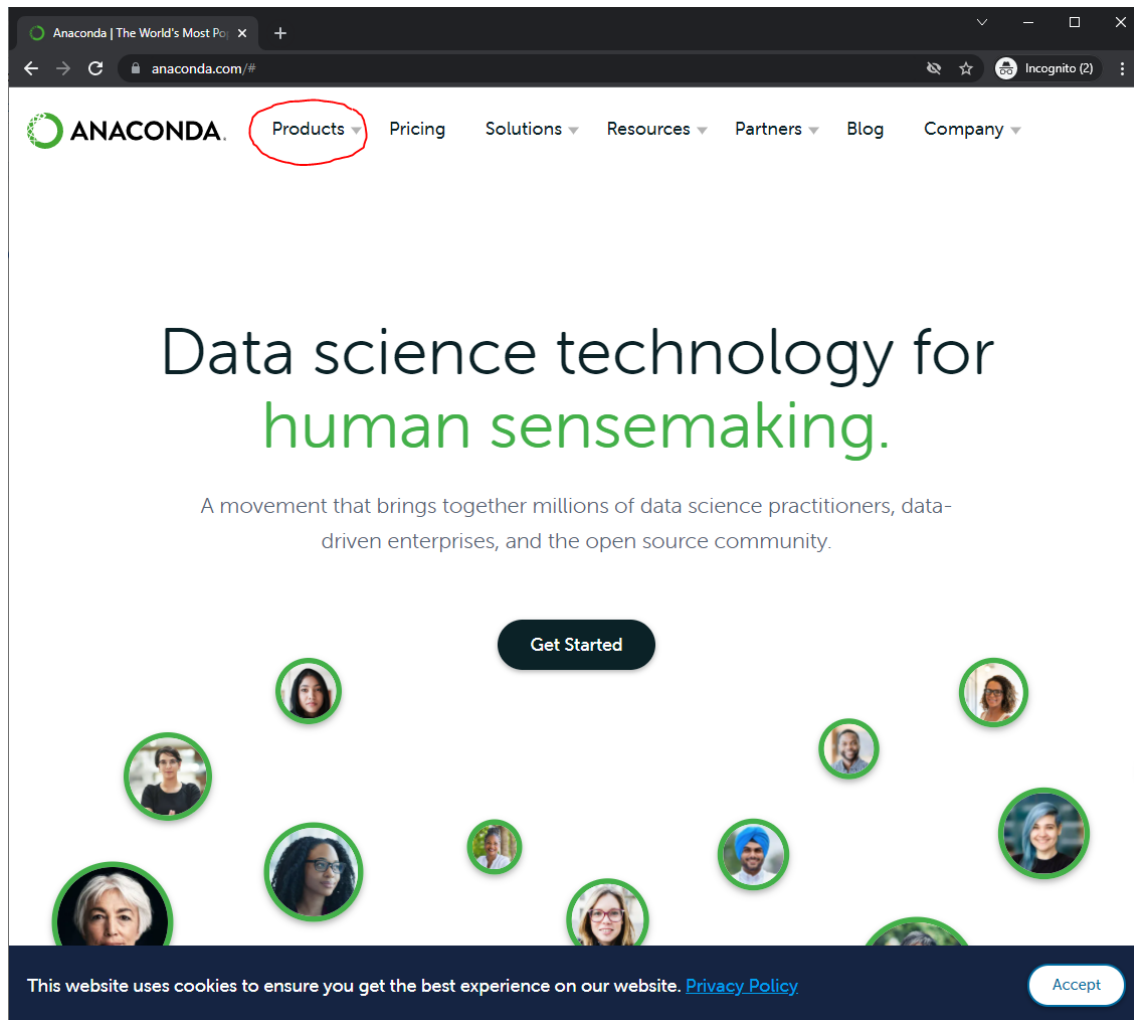


Text Classification

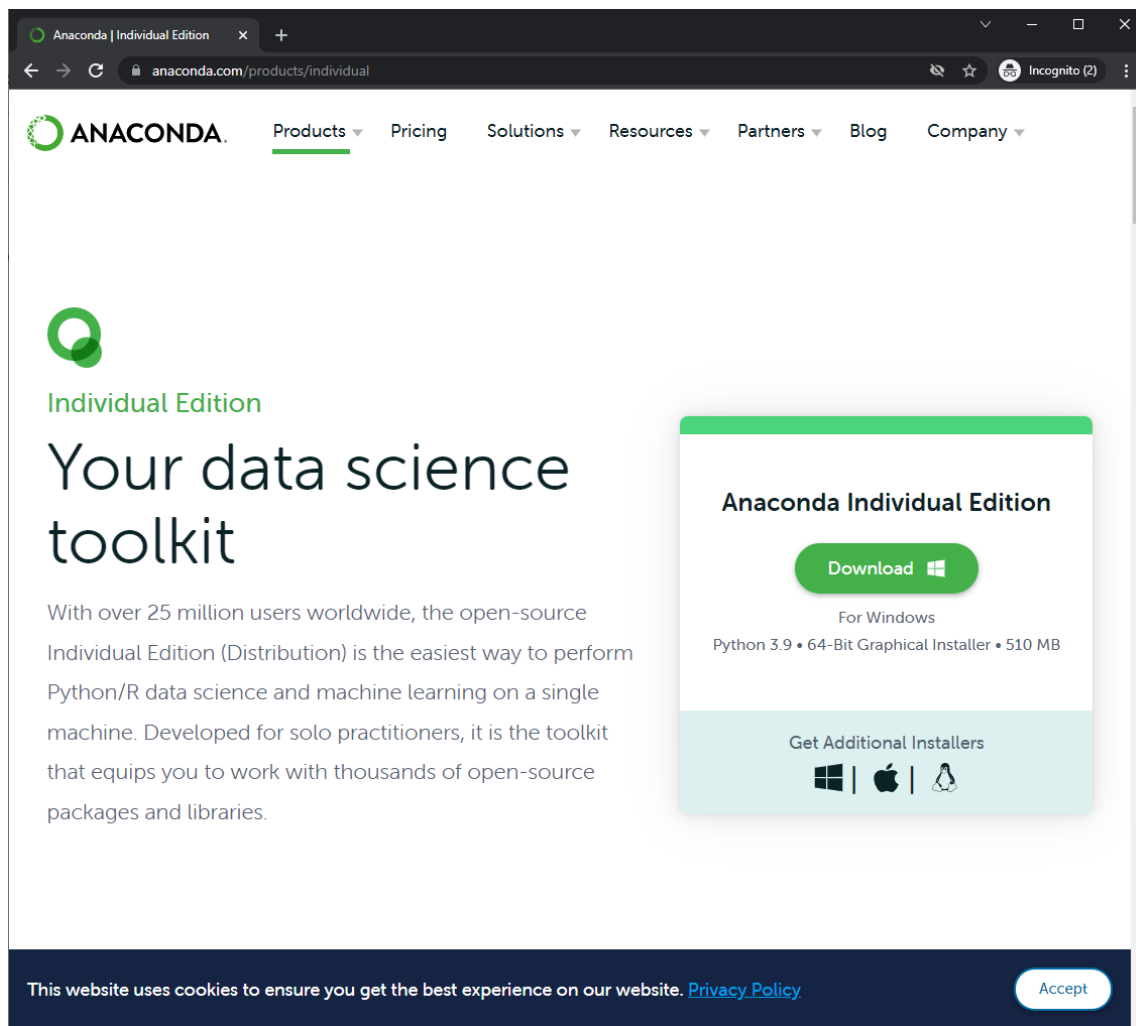
Download the latest version of Anaconda Individual edition.

<https://www.anaconda.com/>



Click on Products dropdown and select Individual edition.

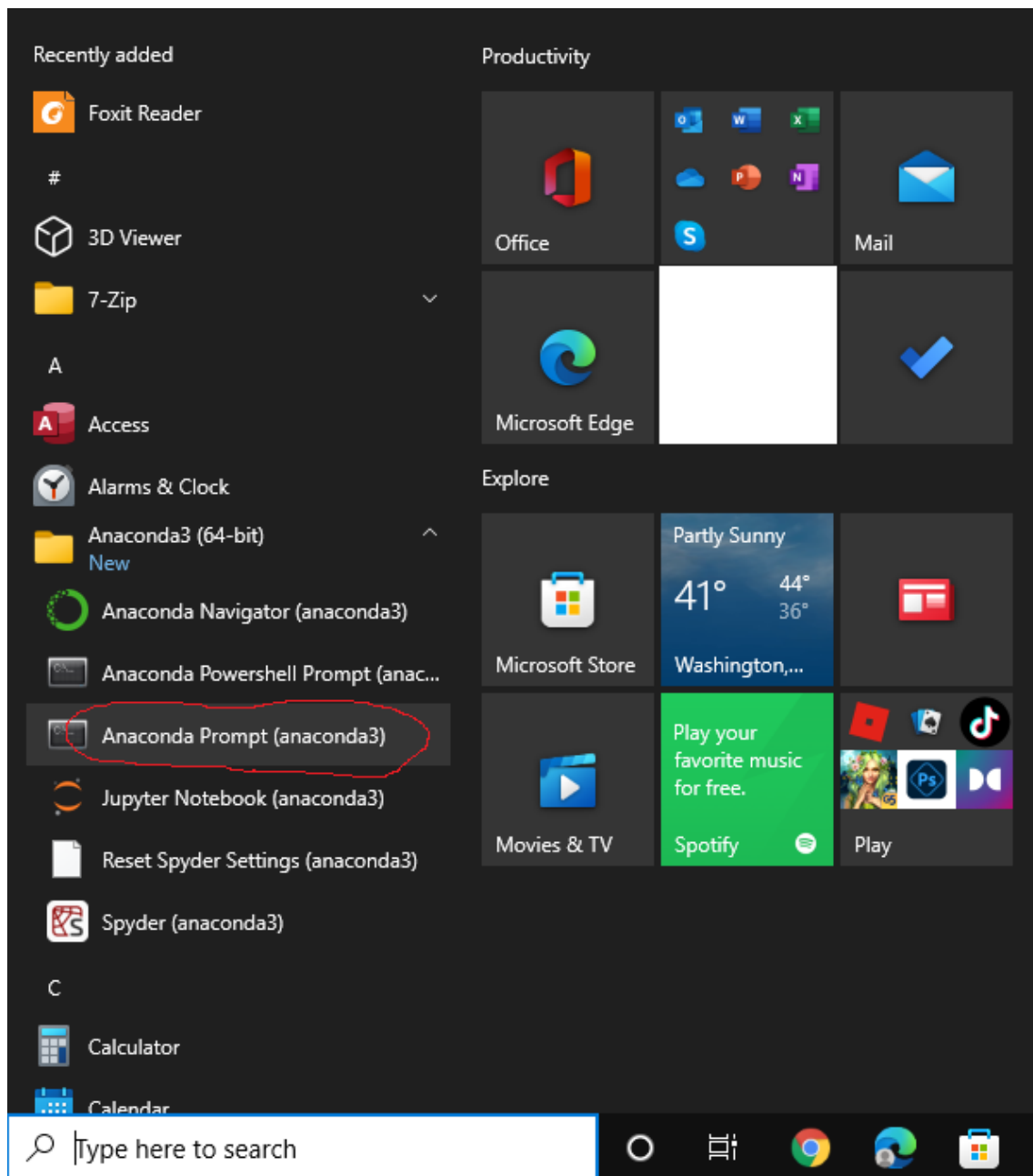
If anaconda3 is already installed then go to the section (page #3) for installing steamlit 0.84.2 version.



Click on Download.

After the software is downloaded, click on the .exe file to install.

Once the software is installed, click on the start button and look for “Anaconda Prompt (anaconda3)” within Anaconda3 (64-bit)



Click on the Anaconda Prompt (anaconda3). It opens up a command line window.

Now do a pip install to install streamlit version 0.84.2

pip install streamlit==0.84.2

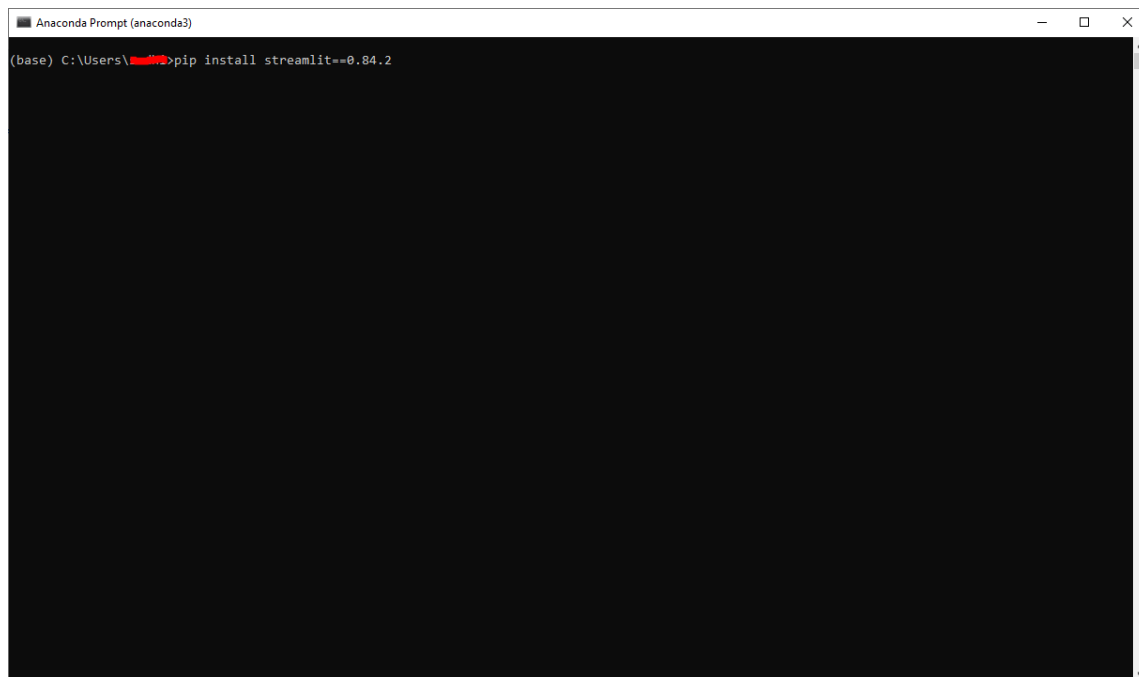
If streamlit is already installed ensure it is version 0.84.2 else install 0.84.2 as code requires this version. Installation removes any prior version and then installs 0.84.2

Please ensure the below packages are installed. (The latest version of Anaconda comes with these packages installed)

sklearn

matplotlib

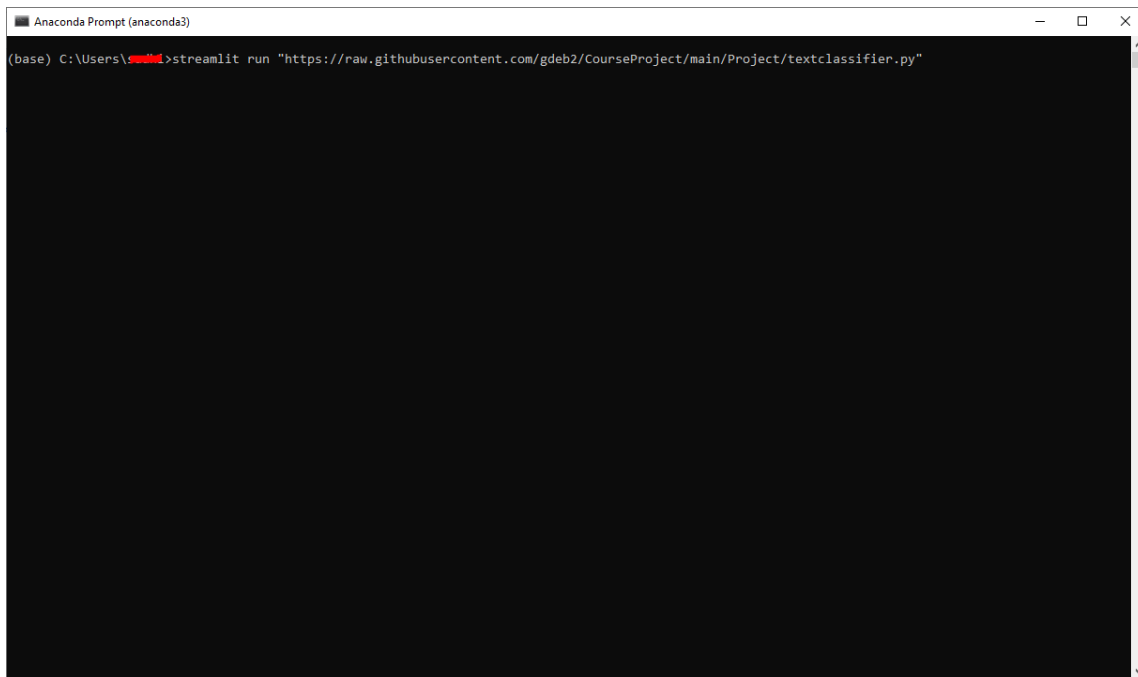
pandas

A screenshot of an Anaconda Prompt window. The title bar reads "Anaconda Prompt (anaconda3)". The command prompt shows the text "(base) C:\Users\redacted>pip install streamlit==0.84.2". The rest of the window is black, indicating the command is still running or the output is not visible.

After installing streamlit, execute the below command to run the “Text Classifier” project.

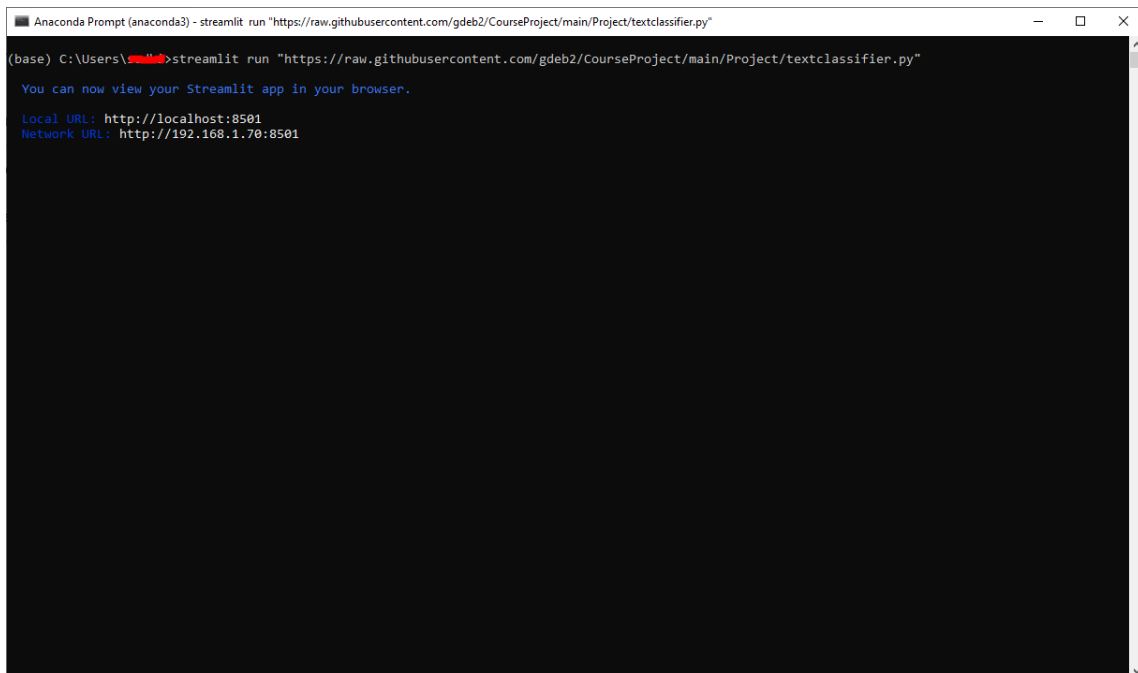
streamlit run

“<https://raw.githubusercontent.com/gdeb2/CourseProject/main/Project/textclassifier.py>”



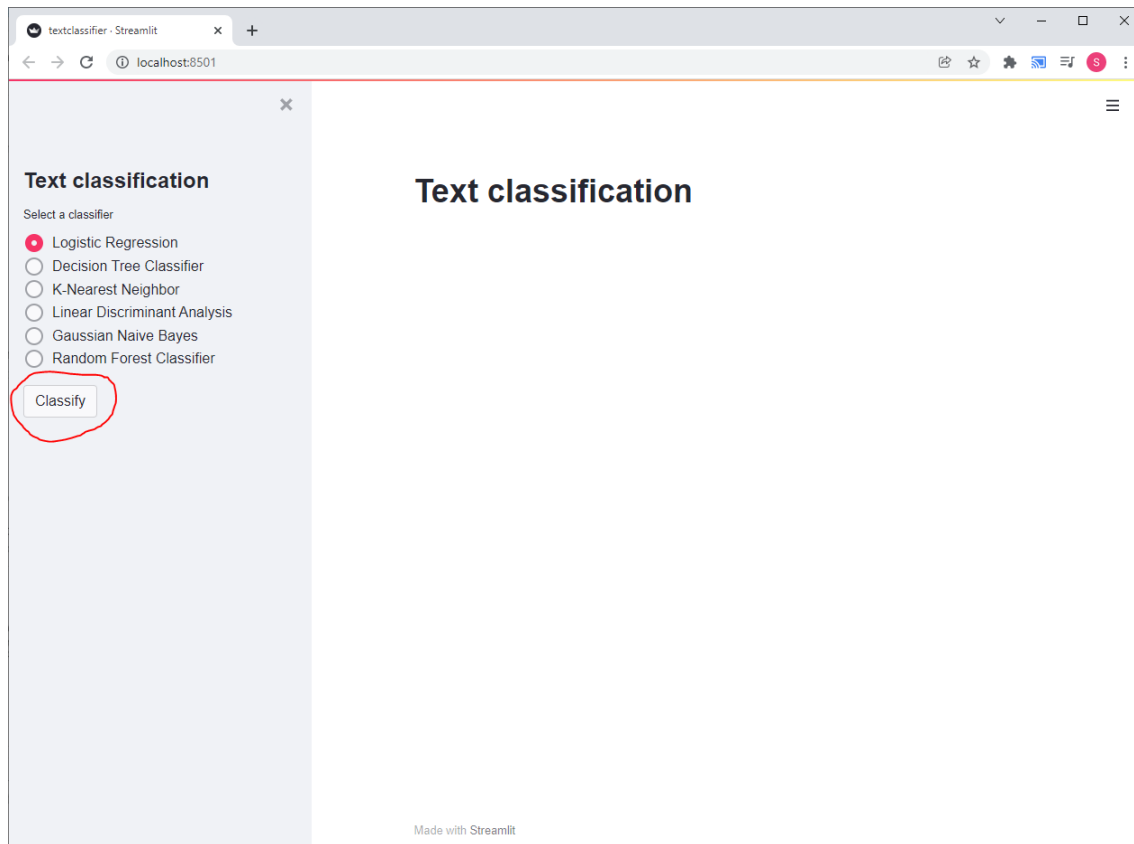
```
Anaconda Prompt (anaconda3)
(base) C:\Users\redmi>streamlit run "https://raw.githubusercontent.com/gdeb2/CourseProject/main/Project/textclassifier.py"
```

Once the streamlit run command is executed, Local and Network URL will be displayed as shown below.



```
Anaconda Prompt (anaconda3) - streamlit run "https://raw.githubusercontent.com/gdeb2/CourseProject/main/Project/textclassifier.py"
(base) C:\Users\redmi>streamlit run "https://raw.githubusercontent.com/gdeb2/CourseProject/main/Project/textclassifier.py"
You can now view your Streamlit app in your browser.
Local URL: http://localhost:8501
Network URL: http://192.168.1.70:8501
```

A new window opens up Text Classification.



This is how the main screen looks like. The left panel lists all the classification algorithms implemented.

Check the radio button to select an algorithm and click the “**Classify**” button.

textclassifier - Streamlit

localhost:8501

Text classification

Select a classifier

☒ Logistic Regression
 ☐ Decision Tree Classifier
 ☐ K-Nearest Neighbor
 ☐ Linear Discriminant Analysis
 ☐ Gaussian Naive Bayes
 ☐ Random Forest Classifier

Classify

Text classification

Exploratory Data Analysis

First 5 rows of the dataset

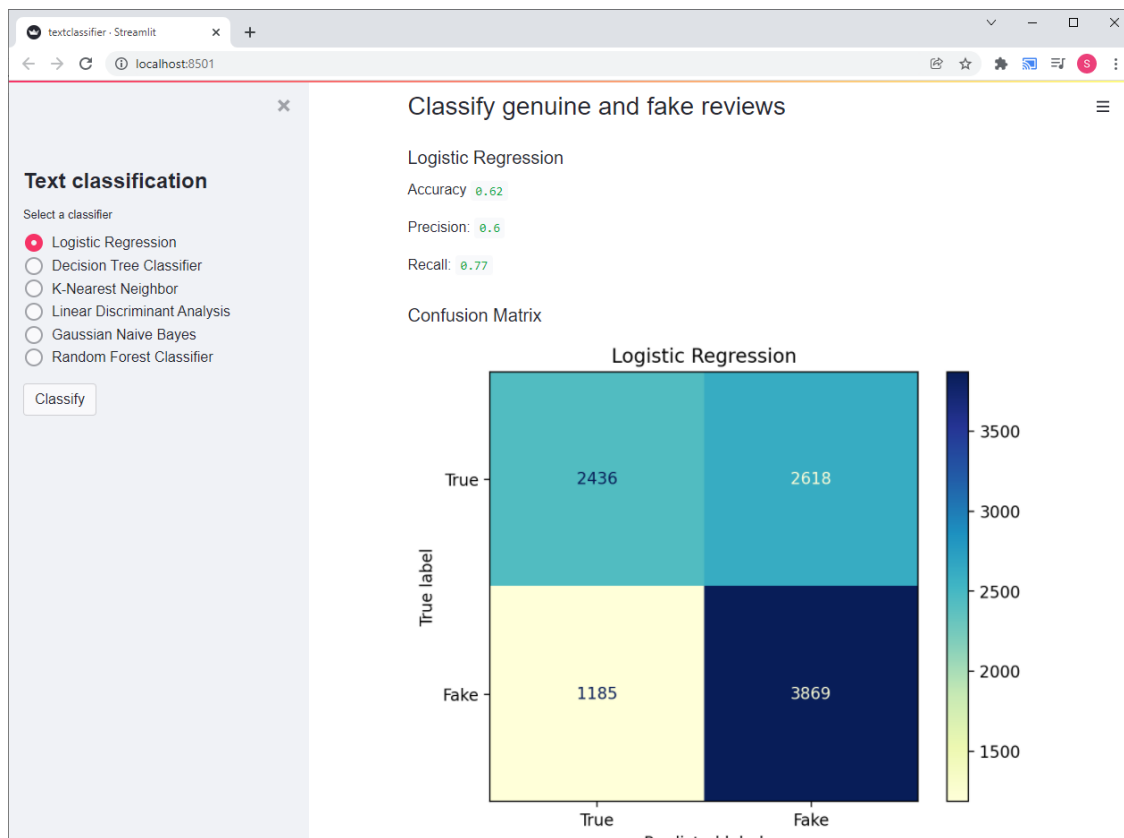
	category	rating	label	original_text	clear
0	Home_and_Kitchen_5	5	CG	Love this! Well made, stu...	love well made sturdy
1	Home_and_Kitchen_5	5	CG	love it, a great upgrade -	love great upgrade or
2	Home_and_Kitchen_5	5	CG	This pillow saved my back...	pillow saved back lov
3	Home_and_Kitchen_5	1	CG	Missing information on ho...	missing information u
4	Home_and_Kitchen_5	5	CG	Very nice set. Good quali...	nice set good quality

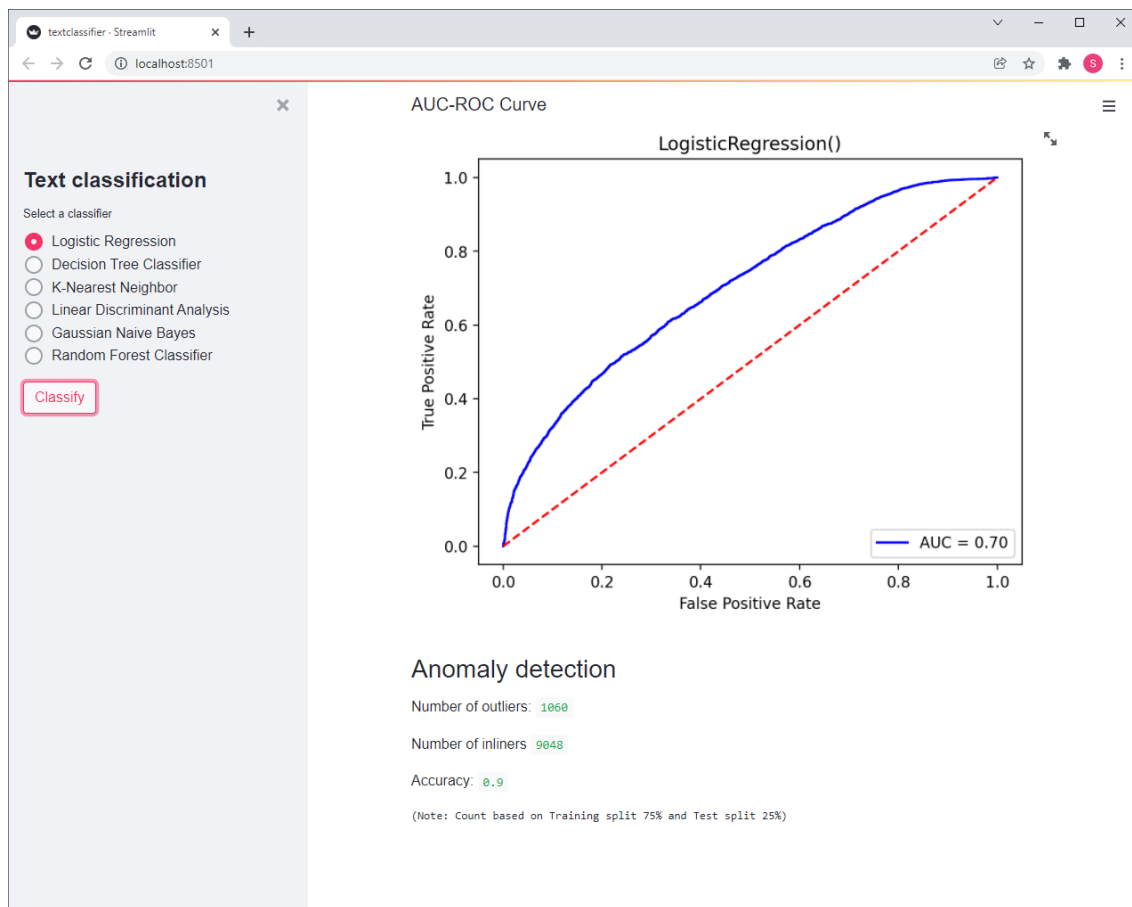
Descriptive statistics

	category	rating	label	original_text	clea
count	40432	40432	40432	40432	
unique	10	NaN	2	40412	
top	Kindle_Store_5	NaN	CG	Easy to put together and -	dog love kept occup
freq	4730	NaN	20216	2	
mean	nan	4.2566	nan	nan	
std	nan	1.1444	nan	nan	
min	nan	1	nan	nan	
25%	nan	4	nan	nan	
50%	nan	5	nan	nan	
75%	nan	5	nan	nan	
max	nan	5	nan	nan	

Exploratory Data Analysis (EDA), Model metrics (Accuracy, Precision and Recall), Confusion Matrix and ROC curve for the selected algorithm will be displayed on the right side as shown in the above screenshot. The selected algorithm will train the model prior to displaying the results.

Scroll down the screen to view all the information.





After reviewing the EDA, model metrics, confusion matrix and ROC curve, repeat the above step for all other algorithms.

To exit, click the “x” button on the top right corner.