E-mail Fraud Detection

PRESENTED BY

- Anmol Jain
- Puneet Kumar Pal
- Abhyoday
- Sourabh Bhardwaj

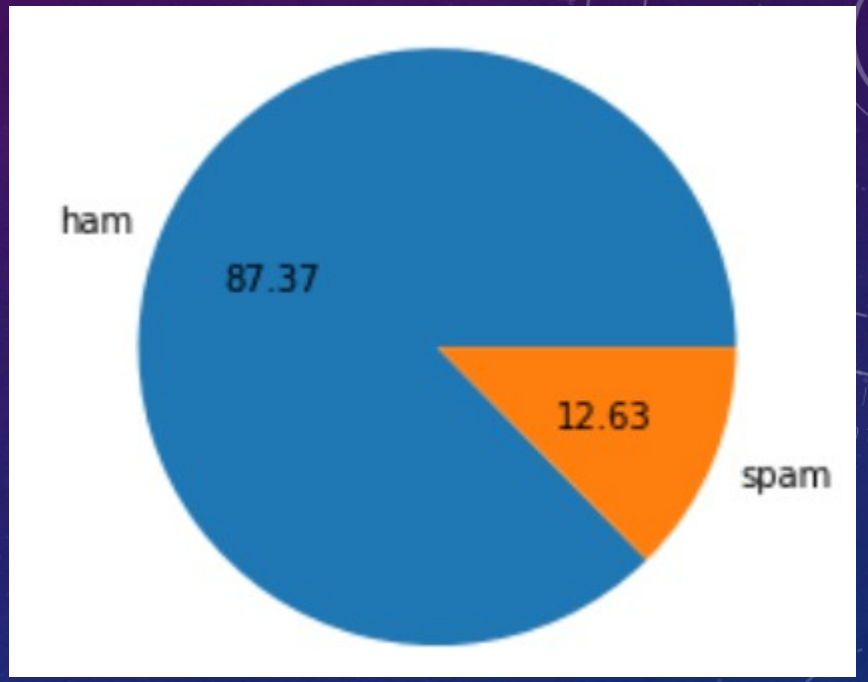
➢ Mentor : Mr Mohammed Ameer

Problem Statement

- Unwanted email irritating internet connection.
- Critical e-mail message are missed or delay.
- Identity theft.
- Resulted to untold financial loss to many users who have fallen victim of internet scams

Project Overview

- To give knowledge to the user about the false e-mails and relevant e-mails.
- It’s also work to identify Mobile SMS.
- To classify that mail spam or not.

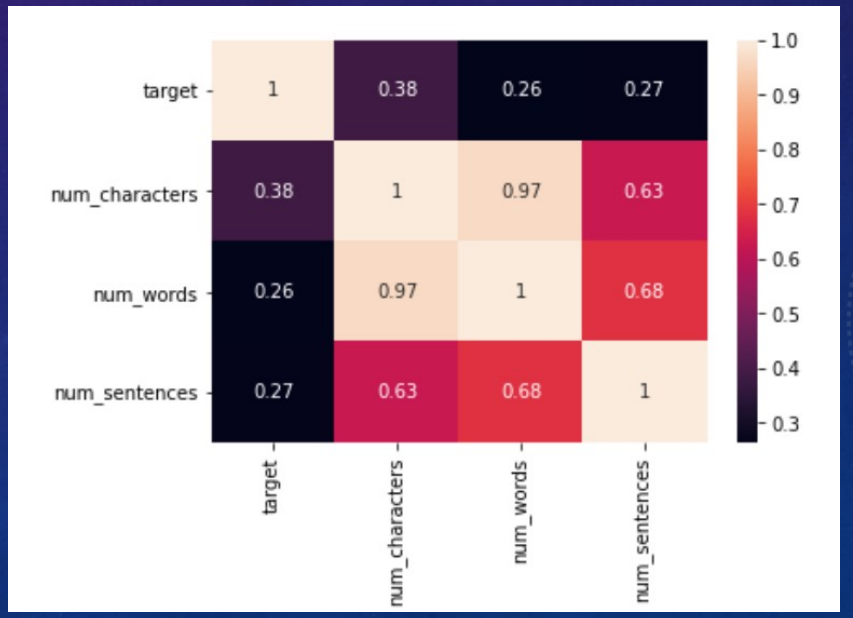


Who is experiencing the Problem and how they are impacted

- Play an important role in our lives because of many reasons such as personal communications, business-focused activities, marketing, advertising etc.
- Make life difficult when they are used outside of their purposes.
- Spam emails can be not only annoying receivers, but also dangerous for receiver's information security.

Solution

- NLP concerned with the processing and understanding of human language.
- Transform textual data into numerical representations Using nltk.
- Predict Spam/Ham Using Naive Bayes Classifier.



The WOW Solution

- Use NLTK library that work on human Dataset.

- Naive Bayes Algorithm Gives Better Accuracy Result

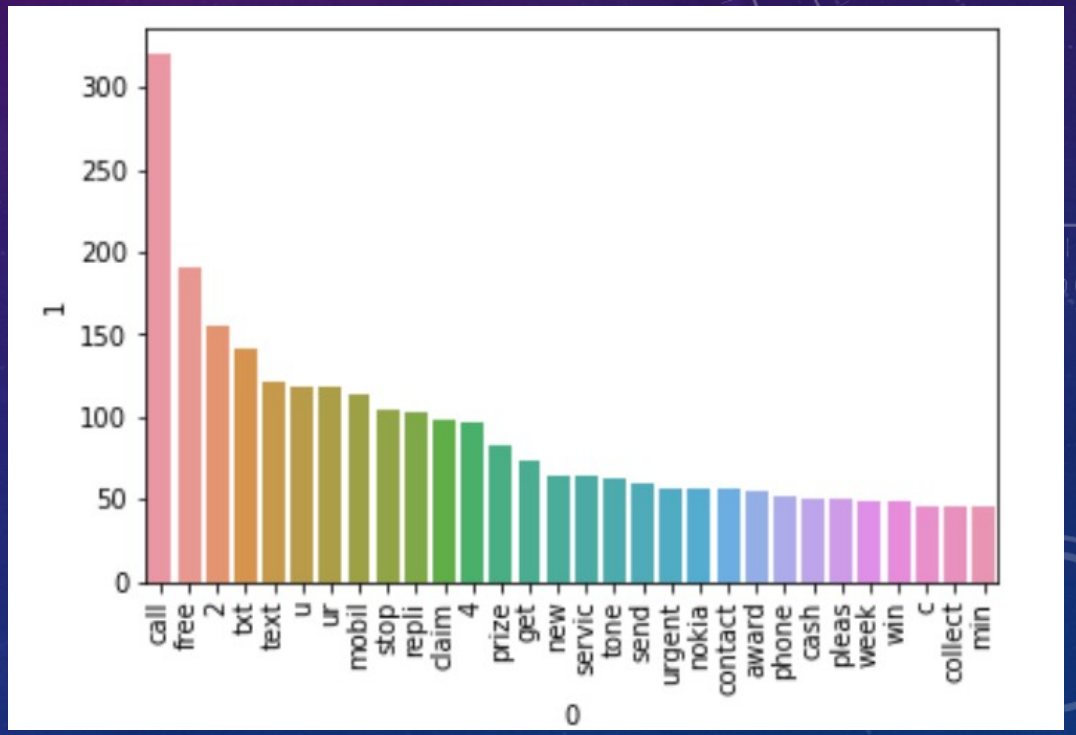
2	Multinomial NB	0.959381	1.000000

Naive Bayes Classifier

- Simple probability classifier that calculates a set of probability by counting the frequency and combination of values in given dataset.
- Represent as a vector of feature values.
- It is very useful to classify the e-mails properly.
- The precision and recall of this method is known to be very effective.

Results

	Algorithm	Accuracy	Precision
0	Gaussian NB	0.876209	0.523148
1	Bernoulli NB	0.970019	0.973451
2	Multinomial NB	0.959381	1.000000
3	Logistic Regression	0.946809	0.988235
4	SVM	0.966151	0.981308
5	Decision Tree	0.957447	0.835714
6	Random Forest	0.975822	1.000000



Live Analysis

Credits:

- Mrs khyati Nagpal
- Mr Mohammed Ameer