Working of the Machine Learning Model :

Natural Language Processing (NLP) is a branch of artificial intelligence that helps computers understand, interpret, and manipulate human language. NLP draws from many disciplines, such as computer science, linguistics, and statistics, to enable computers to perform various tasks with natural language data, such as text and speech.

A bag-of-words model, or BoW for short, is a simple way to represent text data for natural language processing (NLP) tasks. It involves two steps:

Creating a vocabulary of words from a collection of documents.
Counting the occurrence of each word in each document.

The result is a matrix of word counts, where each row represents a document and each column represents a word. The order and structure of the words in the documents are ignored, only the frequency of each word is considered. This is why it is called a "bag" of words.

A bag-of-words model can be used for various NLP applications, such as:

Text classification: assigning a label to a document based on its content, such as spam detection or sentiment analysis.

Text clustering: grouping similar documents together based on their content, such as topic modeling or document retrieval.

Text summarization: generating a short summary of a document based on its main points, such as news headlines or abstracts.

To evaluate our product, we sought original complainant statements for testing purposes. However, obtaining such statements posed challenges due to safety concerns surrounding potential sensitive information. In light of this, we opted to generate synthetic complainant statements to assess our AI's performance. Below, you'll find the outcomes of these tests and our observations.

How to Run a Custom Test Statement ?

1) In the given github link open the natural_language_processing.ipynb file in a Jupyter Notebook Environment
2) Download the SectionsCopy.csv file and paste it in the same directory as of the NoteBook
3) Paste your test FIR statement as a "t" variable in the 7th cell and run all the Cells
4) You can see the predicted outputs at the end of the Code

Test Statements :-

One –

I, Ram Prasad, son of Hari Prasad, resident of 10/5, Rajiv Nagar, Kurichedu, Andhra Pradesh, hereby lodge a complaint of murder against Shyam Lal, son of Ram Lal, resident of 8/3, Indira Colony, Kurichedu, Andhra Pradesh.

I state that Shyam Lal and I are neighbors and have a long-standing dispute over a piece of land adjacent to our houses. On 15/12/2023, at around 9 pm, I heard some loud noises and screams coming from Shyam Lal's house. I rushed to his house and saw that he had stabbed his wife Geeta with a kitchen knife and she was lying in a pool of blood on the floor. Shyam Lal also attacked me with the knife when I tried to stop him and injured me on my left arm. I managed to escape from his house and called the police from my mobile phone.

Predicted Output :-

328 - House Trespassing and House Breaking
99 - Murder

Two –

I, Ravi Kumar, son of Shyam Kumar, resident of 16/9, Rama Colony, Kurichedu, Andhra Pradesh, hereby lodge a complaint of public nuisance against Ram Prasad, son of Hari Prasad, resident of 10/5, Rajiv Nagar, Kurichedu, Andhra Pradesh.

I state that Ram Prasad and I are neighbors and have a long-standing dispute over a piece of land adjacent to our houses. On 15/12/2023, at around 10 pm, Ram Prasad started playing loud music and bursting firecrackers in his house. The noise was so loud that it disturbed the peace and tranquility of the entire neighborhood. Many people, including children, senior citizens, and patients, were unable to sleep or rest due to the noise. The firecrackers also created a risk of fire and air pollution.

I also state that this was not the first time that Ram Prasad had caused such a nuisance. He had done the same thing on several occasions in the past, despite being warned and requested by me and other neighbors to stop. He had also ignored the notices issued by the local authorities to refrain from creating such a nuisance.

Predicted Output :-
268 - Public Nuisance

Three –

I state that on 15/12/2023, at around 11 pm, I was returning home from my shop on my motorcycle. When I reached near the railway crossing, two men on a black bike came from behind and stopped me. They pointed a knife at me and snatched my gold chain, my wallet, and my mobile phone. They also took the keys of my motorcycle and fled away.

I also state that the robbers were wearing helmets and masks, so I could not see their faces. They were speaking in Telugu and had a local accent. The number plate of their bike was covered with a cloth, so I could not note it down.

I further state that I ran to the nearest shop and asked for help. The shopkeeper called the police and informed them about the incident. The police arrived at the spot and took me to the police station.

I request you to kindly register a FIR against the unknown robbers for the offense of robbery under Section 392 of the Indian Penal Code, 1860[12] and take necessary action against them. I have attached the copies of my identity card, driving license, and the

Predicted Output :-

302 - Snatching
307 - Robbery
141 - Trafficking of a person

Four –

I state that Suresh Kumar is a government officer working in the Revenue Department of Kurichedu. On 15/12/2023, at around 4 pm, I went to his office to get my land records updated. He asked me to pay him Rs. 50,000 as a bribe to do the work.
He said that if I did not pay him, he would delay the work or create problems for me. He also threatened me to implicate me in a false case of tax evasion. I also state that I refused to pay him the bribe and told him that I would file a complaint against him. He then became angry and abused me. He also snatched my land documents and tore them. He then threw me out of his office.

Predicted Output :-
168 – Bribery
302 – Snatching