**Project Name:** Omicron Tweets Sentiment Analysis

**Github Link:** https://github.com/utkrisht2000/Omicron-Sentiment-Analysis.git

**Why was this project created?**

Predict the sentiment from continually creating Big Data on Twitter in relation to the exchange of knowledge, comments, thoughts, perspectives, and experiences concerning the COVID-19 pandemic, with a concentration on the Omicron version, which is now the worldwide dominating variation of SARS-CoV-2. The Omicron variation was the subject of 12028 tweets in total, and the particular features of those tweets that were examined were sentiment, language, source, type, and embedded URLs.

**What problem is it solving?**

On the recommendation of WHO's Technical Advisory Group on Virus Evolution, WHO identified the variety B.1.1.529 as a variant of concern and gave it the name Omicron on November 26, 2021. (TAG-VE). This choice was made in light of the information provided to the TAG-VE that Omicron had a number of mutations that might affect how it acts. We are analyzing the sentiment of the tweets on Twitter using the hashtag #Omicron.

**Entire explanation of project**

* **PROPOSED APPROACH**

Omicron sentiment analysis employs a dataset that was originally gathered from Twitter when people were discussing the Omicron version and is now loaded from Kaggle. In the dataset, null values will be handled. Remove all the rows containing the null values as there are null values in three of the dataset's textual data fields. People's tweets expressing their thoughts regarding the Omicron variant are included in the dataset's text column. To get ready for sentiment analysis, we must clean and prepare this column. Then, to see which terms were used the most frequently by users in their tweets, we create a word cloud from the text column.

We will next determine the sentiment ratings for the tweets mentioning the Omicron version. By computing the sentiment ratings of the text column, add three extra columns to this dataset as Positive, Negative, and Neutral. The majority of opinions were then revealed to be neutral, indicating that neither favourable nor negative views of the Omicron variety were being expressed.

Algorithm for creating next word prediction model :

**Step 1:** Dataset is imported

**Step 2:** The data is preprocessed

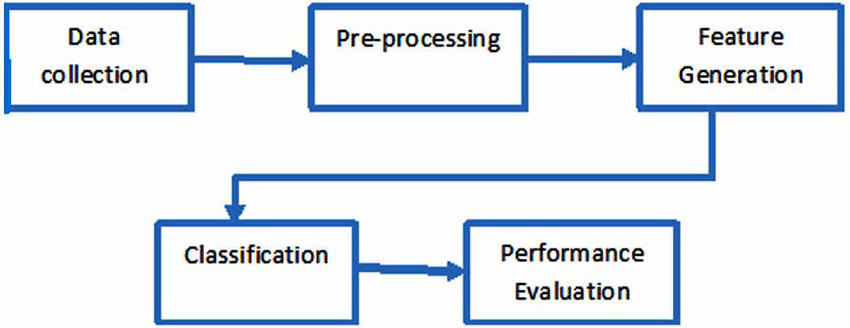
**Step 3:** Word Cloud generated

**Step 4:** Sentiment Score Calculated

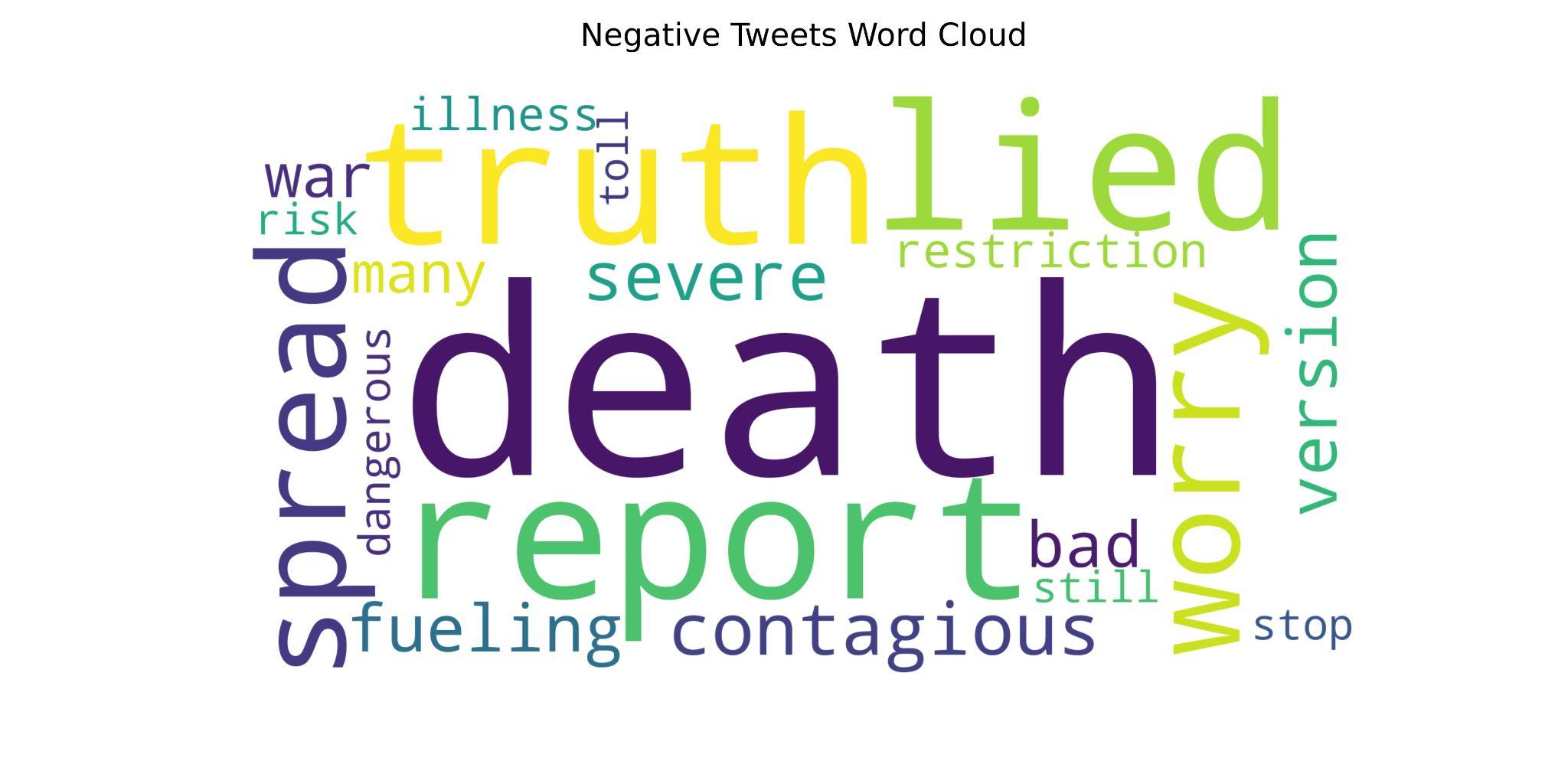
**Step 5:** Add columns positive, negative and neutral

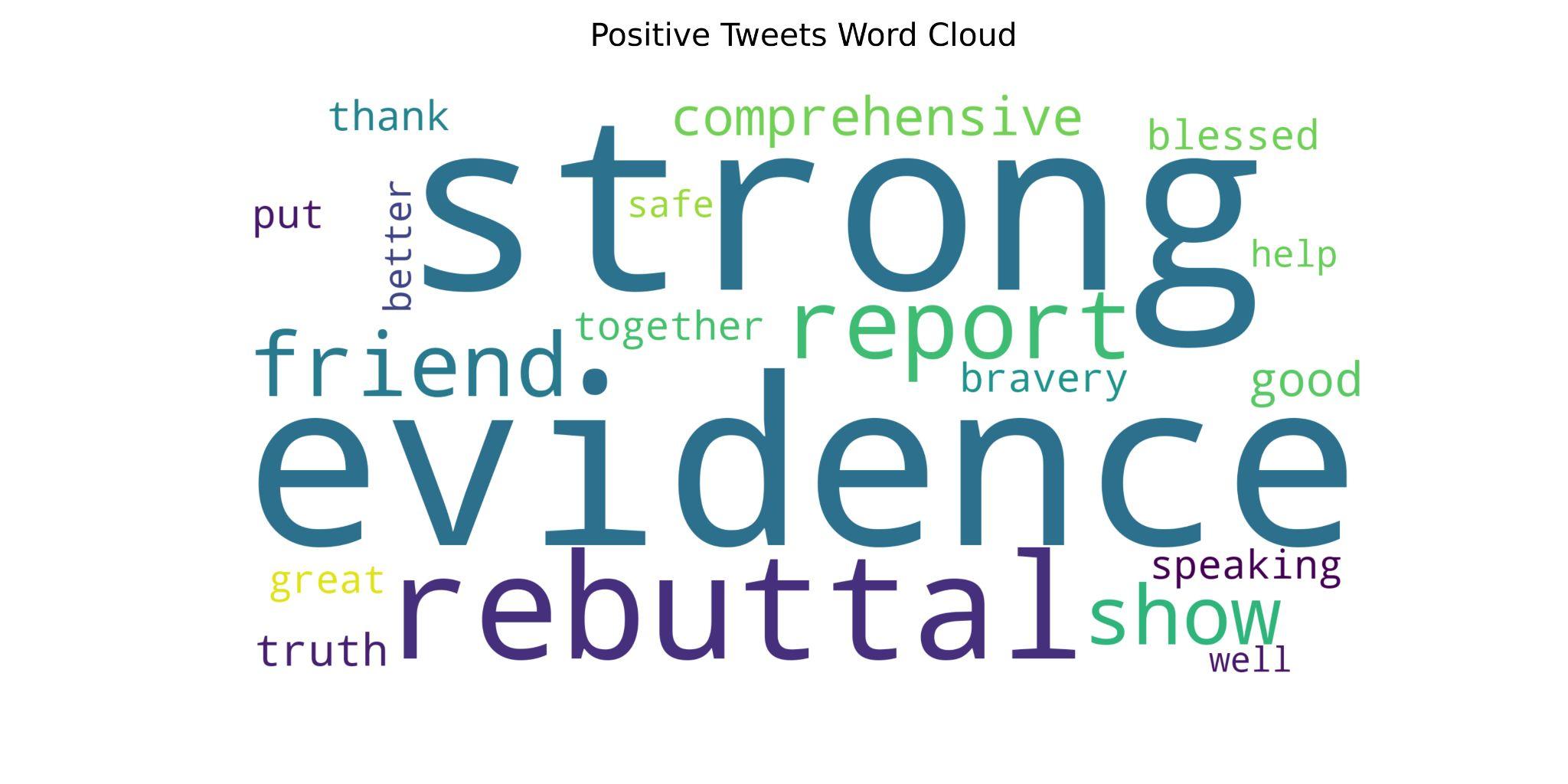
**Step 6:** Generate word cloud for positive, negative and neutral

* **DATA FLOW DIAGRAM**



* **RESULT**





* **CONCLUSION**

People frequently utilize social media platforms for communication, information, news, thoughts, opinions, ideas, knowledge, feedback, and experiences connected to the pandemic they are now confronting, according to research done during pandemics in the past. One such social media site is Twitter, which is well-liked by people of all ages. As a result, this research used a thorough strategy to find, examine, and analyze tweets about the SARS-CoV-2 Omicron variant in order to comprehend, classify, and explain the associated dynamics and distinctive aspects of social media activity. The sentiment (excellent, good, neutral, poor, and dreadful), kind (original tweet, retweet, or reply), source (such as "Twitter for Android," "Twitter Web App," "Twitter for iPhone," etc.), and number of tweets (12028) concerning the SARS-CoV-2 Omicron strain were all extracted from Twitter.We looked at language (like English, Spanish, Bengali, etc.) and embedded URLs (URLs that were contained inside the tweet content). This exploratory investigation yielded a variety of results. First, the majority of tweets (50.5% of all tweets) had a "neutral" emotion, which was followed by "negative" (15.6% of all tweets) and "positive" (14.0% of all tweets).