**Introduction:**  
  
Artificial intelligence (AI) is a field of computer science that involves the creation of intelligent machines that can perform tasks that typically require human intelligence. One of the most exciting applications of AI is in natural language processing (NLP), which involves teaching machines to understand and generate human language. ChatGPT is an NLP-based language model that uses AI to generate human-like responses to text-based inputs.  
  
Overview of ChatGPT:  
  
ChatGPT is a member of the family of Generative Pre-trained Transformer (GPT) models developed by OpenAI. GPT models are based on a deep learning architecture called a transformer, which is designed to process sequential data such as natural language. ChatGPT is trained on a massive dataset of text, which allows it to generate responses to a wide range of inputs.  
  
ChatGPT works by predicting the most likely word or phrase to follow a given input, based on the patterns it has learned from the training data. It uses a technique called "sampling" to generate responses that are not simply copies of the training data. This means that ChatGPT can generate new and creative responses to inputs, rather than simply repeating what it has learned from the training data.  
  
Applications of ChatGPT:  
  
ChatGPT has many potential applications, including:  
  
Chatbots: Chatbots are computer programs that are designed to simulate human conversation. ChatGPT can be used to create chatbots that are more human-like and can engage in more complex conversations than traditional rule-based chatbots.  
  
Content creation: ChatGPT can be used to generate content such as news articles, product descriptions, and social media posts. This can save time and effort for content creators and allow them to focus on more creative tasks.  
  
Language translation: ChatGPT can be used to translate text from one language to another, potentially improving the accuracy and fluency of machine translation systems.  
  
Challenges and limitations:  
  
While ChatGPT and AI have many exciting applications, there are also challenges and limitations to consider. One major challenge is the potential for bias in AI systems, which can lead to unfair or discriminatory outcomes. This is particularly important to consider in applications such as hiring, lending, and criminal justice, where biased AI systems could have serious consequences.  
  
Another limitation of ChatGPT and AI more broadly is the "black box" problem. AI systems can be difficult to interpret, meaning that it can be hard to understand why they make the decisions they do. This can make it difficult to identify and correct errors or biases in the system.  
  
  
ChatGPT is an exciting example of the potential of AI and natural language processing. While there are challenges and limitations to consider, the applications of ChatGPT and similar AI systems are vast and could transform many areas of our lives. As AI technology continues to develop, it is important to ensure that it is used ethically and responsibly, to maximize its benefits for society.

**Project and goals:**

A gloomy future when machines run the world was once feared, while others believed AI would solve all our problems. Based on the information that the researcher independently gathered from a variety of sources, including news stories, reports, blogs, social media, and academic papers, this study intends to assess the evolution of public sentiment toward AI. We intend to use the analysis of this data to understand how people's attitudes regarding AI have evolved and extrapolate what those sentiments might be in the future. This is an NLP task that calls for the use of natural language processing methods to glean valuable information from the gathered information. This project becomes much more complex due to the self-collection of data and content. The script and report will include the source data, as well as a detailed analysis of the trends in public opinion towards AI over time.

Introduction to a dataset:

So for analysis, we use four sets of data first

we use a dataset that is

**dataset4 df4** **is the file** we use is saved with the name of the file here the dataset has three columns where the first column is the serial number then followed by the tweets and the link of the tweets and the final column is labeled with the tag which says whether that tag is good, bad or neutral.

**Dataset 3 df3** **is chatgpt (1)**

This dataset contains a collection of tweets with the hashtag #chatgpt. The tweets were scraped from Twitter and cover a range of topics related to the ChatGPT language model. The dataset includes the following information for each tweet:

Tweet text

User information (username, user ID, location, etc.)

Tweet timestamp

Retweet and favourite count

Hashtags used in the tweet

URLs

The dataset provides a glimpse into the online conversation surrounding the ChatGPT language model and can be used for various natural language processing and machine learning tasks, such as sentiment analysis, topic modeling, and more. It allows understanding the community, the level of interest, and the use of chatGPT.

**Data2** **df2** **is chatgpt** This file contains all the tweets about ChatGPT from 30/11/2022 to 31/12/2022.This data is obtained by Obtained by Twitter API v2 for Academic Research.This file contains 219295 data . This file provides the tweet id from which the tweet has been made, created at is the date at which the tweet has been made, the like count is the number of like the tweet has got, quote count is the number of the tweet has been counted and the reply count is number of counts the customer is received, and retweet the number of retweets it has received, tweet is the tweet that has been made, and the photo URL is if any photos have been uploaded. For some tweets even the city and the country where it has been uploaded from is given.

**Steps followed for analysis:**

The code begins by importing necessary libraries and modules such as numpy, pandas, seaborn, sklearn, matplotlib, and wordcloud.

Next, three csv files are loaded into three different dataframes using the pd.read\_csv() function. The encoding format is specified as 'ISO-8859-1' and the engine used to parse the csv file is specified as 'python'.

The third dataframe df4 contains both tweets and their corresponding labels, which are printed and plotted using value\_counts(), plot(kind='barh'), and plot(kind='pie') functions.

A histogram is plotted to show the distribution of the labels using the plt.hist() function.

Sentiment analysis is performed on the df2 dataframe using the TextBlob library. The get\_sentiment() function is defined to apply the TextBlob sentiment analysis function to each row of the 'tweet' column. A new column 'sentiment\_score' is created to store the sentiment score for each tweet. The 'sentiment\_score' column is used to assign a 'sentiment\_label' to each tweet using a lambda function. The updated dataframe is printed using the print() function and the count and distribution of the sentiment labels are plotted using the value\_counts(), plot(kind='barh'), plot(kind='pie'), and sns.pairplot() functions.

Sentiment analysis is also performed on the df3 dataframe using the same approach as for df2. The updated dataframe is printed and the count and distribution of the sentiment labels are plotted using the same functions as before.

The code execution ends.

**Output:**

First I am presenting the output from df4 the dataset named file which is an opinion of people on chatgpt month end of December.This data shows that the number of neutral comments is 107796, the number of positive comments are 56011 and the number of bad comments are 55487]

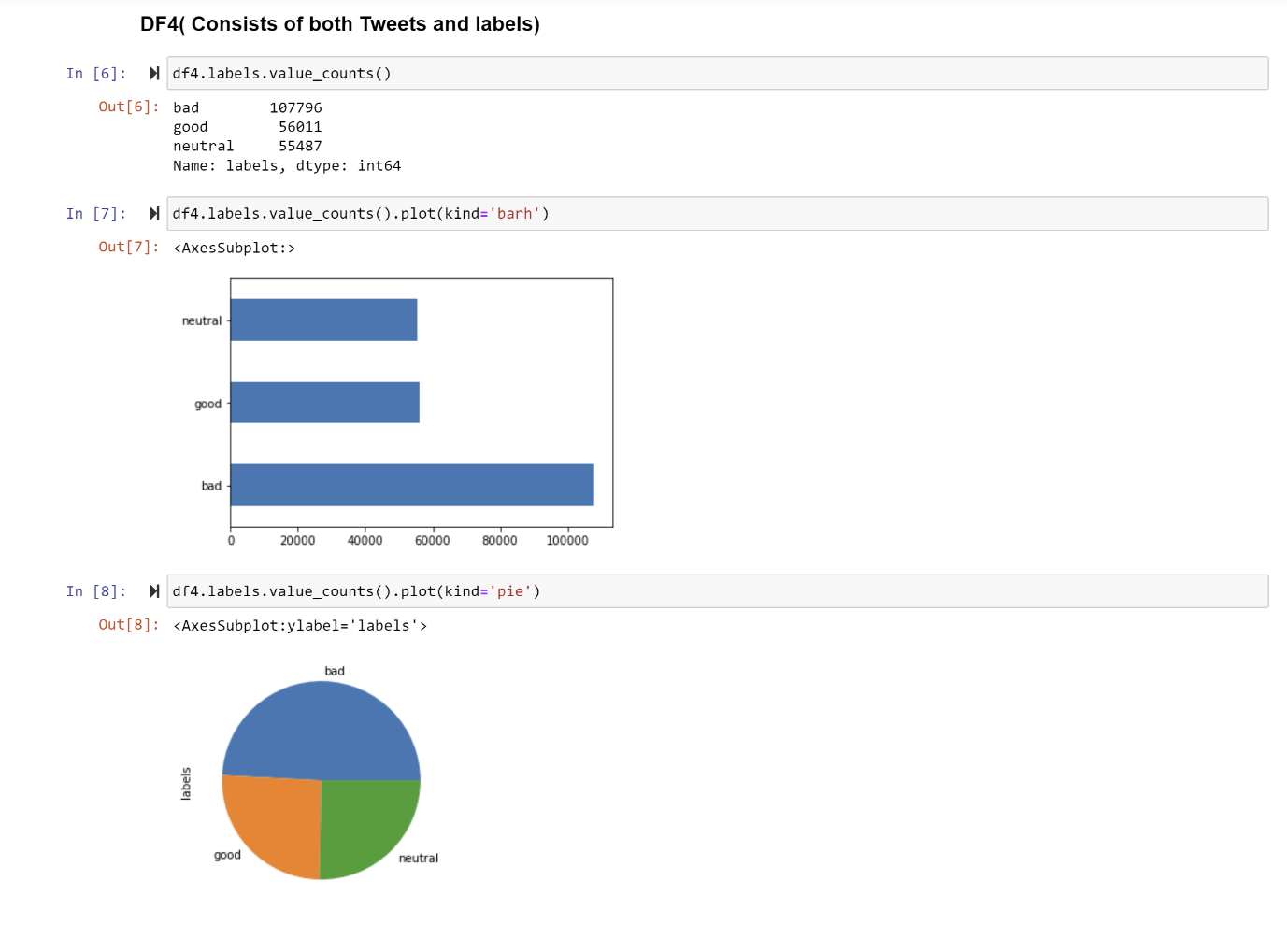


Figure 1 Data plot of positive , negative and neutral comments

The second dataset consisting of df2 is chatgpt This file contains all the tweets about ChatGPT from 30/11/2022 to 31/12/2022 has 105987 positive comments , 79698 neutral comments and 33609 negative comments

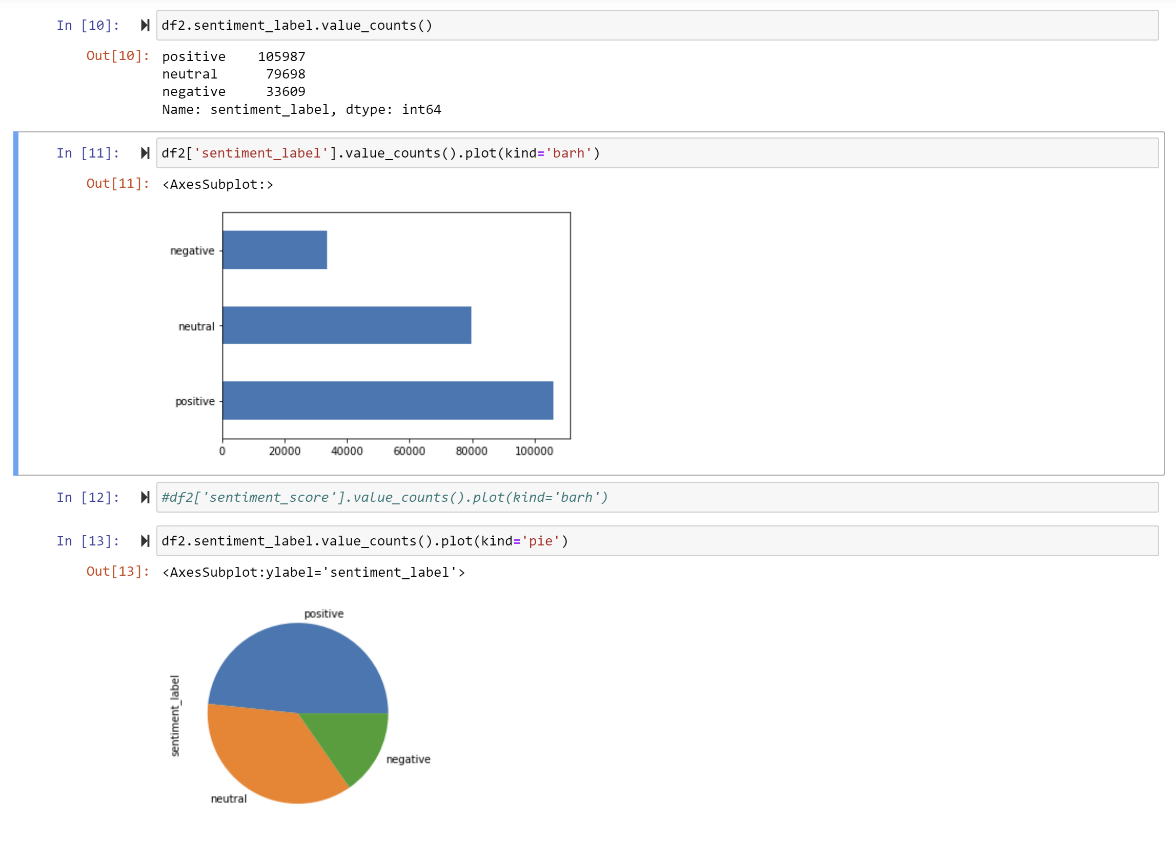


Figure 1 Data plot of positive , negative and neutral comments

.Chart

Description automatically generated with low confidence

Fig 3. Pear plot of Tweetid, Reply counts, Retweet counts, like counts , Quote counts , Conversation id, Hastag counts , sentiment counts on df3

Finally, the data set df3 shows 28849 neutral comments, 15396 positive comments, and 5756 negative comments.

Graphical user interface, chart, application

Description automatically generated

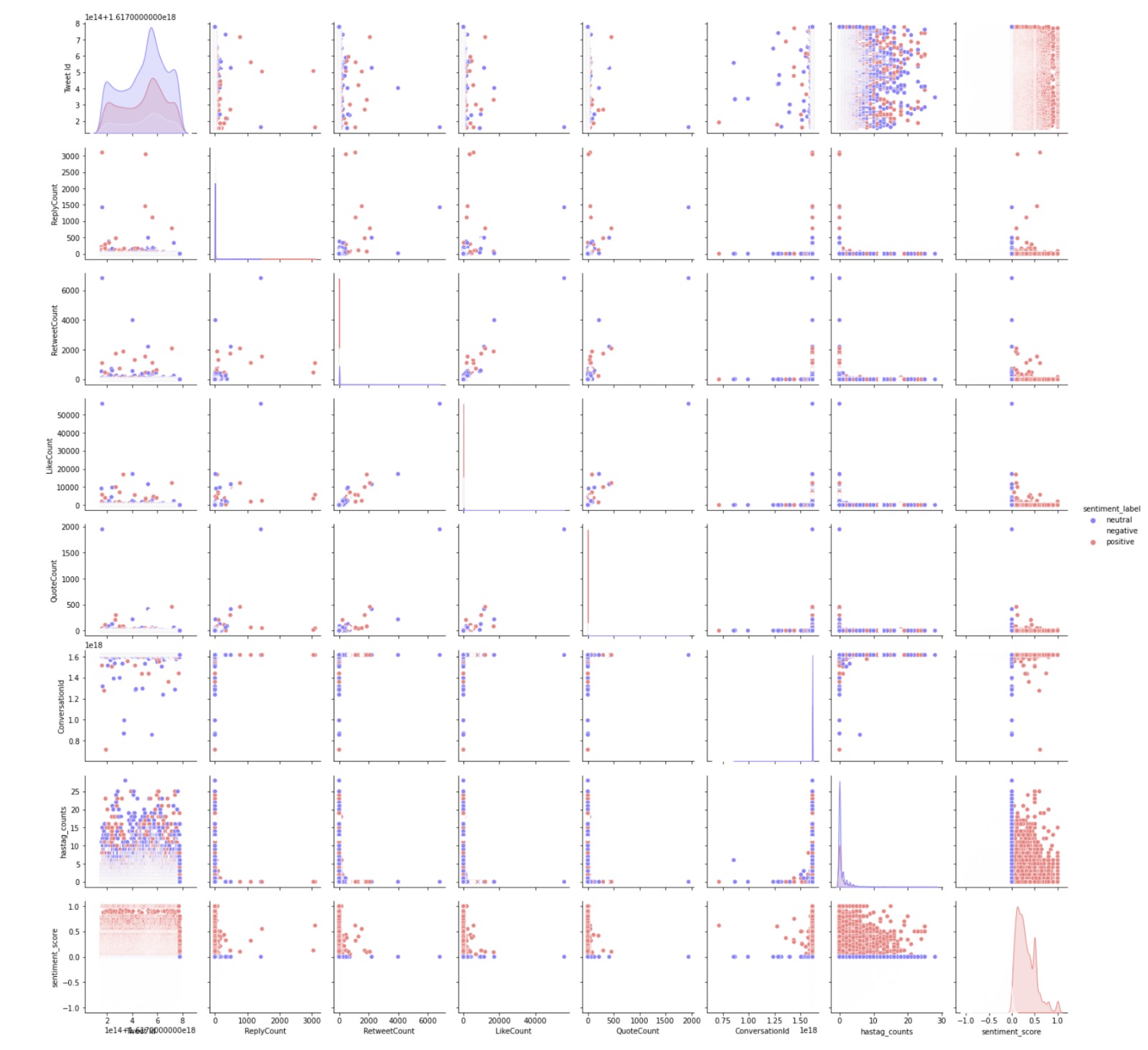


Fig 5. Pear plot of Tweetid, Reply counts, Retweet counts, like counts , Quote counts , Conversation id, Hastag counts , sentiment counts on df3

**Conclusion:**

So form the sensitivity analysis produces on dataset 2 or df2 where data set 2 represents the tweets on chatgpt and the opinion of people on chatgpt on the first month of its launch and in dataset df 4 which represents end of the month of December 2022 there has been more of negative neutral comments compared to the positive ones But dataset 3 or df 3 represents people’s opinion on 22 January in the year 2023 shows a significant rise in number of positive comments and the positive a significant reduction in the neutral comments. There has been also a slight rise in the negative comments on chatbot. But nothing compared to the positive rise in people’s opinion about AI.Hence this shows there is a rise in the positive opinion of people perception on AI

**References for datasets :**

[1] https://www.kaggle.com/datasets/charunisa/chatgpt-sentiment-analysis

[2] <https://www.kaggle.com/datasets/tariqsays/chatgpt-twitter-dataset>

[3] <https://www.kaggle.com/datasets/pcminh0505/chatgpt-twitter>