A PROJECT REPORT

ON

EMOTIONAL ANALYSIS ON WHATSAPP TEXT

Submitted by

Group 5:

ROHAN BASAK (22PGDM194)

SOUGATA SAMANTA (22PGDM211)

SAYAK MUKHERJEE (22PGDM126)

DEBASISH SAHA (22PGDM166)

SIDDHARTHA PAUL (22PGDM207)

SOHAM BHATTACHARYA (22PGDM209)

(PGDM 2022-24)

In partial fulfilment of the award of the Diploma

of

POSTGRADUATE DIPLOMA IN MANAGEMENT

AT

INTERNATIONAL MANAGEMENT INSTITUTE KOLKATA

DECEMBER, 2023

BONAFIDE CERTIFICATE

CERTIFIED THAT THIS PROJECT REPORT "EMOTIONAL ANALYSIS ON WHATSAPP" IS THE BONAFIDE WORK OF "ROHAN BASAK, SOUGATA SAMANTA, SOHAM BHATTACHARYA, SAYAK MUKHERJEE, DEBASISH SAHA, SIDDHARTHA PAUL" WHO CARRIED OUT THE PROJECT WORK UNDER MY SUPERVISION.

SUPERVISOR

DR. ARGHYA RAY

TABLE OF CONTENT

INTRODUCTION

WhatsApp is a widely popular messaging application that enables users to communicate via text, voice messages, calls, and video chats over the internet. Developed by WhatsApp Inc. and later acquired by Facebook, WhatsApp offers a convenient and user-friendly platform for individuals and businesses to connect and interact globally. Launched in 2009, WhatsApp quickly gained traction for its simplicity, ease of use, and commitment to keeping communication secure through end-to-end encryption, ensuring that only the sender and recipient can read the messages. This emphasis on privacy and security has been a cornerstone of WhatsApp's appeal to users around the world. One of its key features is the ability to send instant messages, and multimedia content like photos, videos, documents, and even share locations. Additionally, WhatsApp supports group chats, allowing multiple users to engage in conversations simultaneously, making it a versatile tool for both personal and professional communication.

Over time, WhatsApp has introduced various updates and features, including WhatsApp Web, which enables users to access their accounts via a web browser, and WhatsApp Business, tailored for enterprises to interact with customers more effectively. With a massive user base spanning continent, WhatsApp has become an integral part of daily communication for millions, bridging geographical boundaries and providing a seamless platform for connecting people worldwide.

WhatsApp boasts several key features that contribute to its widespread popularity:

1. Messaging: The app allows users to send text messages instantly, providing a quick and convenient way to communicate.

2. Voice and Video Calls: Users can make free voice and video calls to their contacts, facilitating real-time conversations regardless of geographical distances.

3. End-to-end Encryption: WhatsApp employs robust security measures, ensuring that messages, calls, and shared media are encrypted end-to-end, meaning only the sender and receiver can access the content.

4. Multimedia Sharing: Users can easily share photos, videos, voice messages, documents, and even location information with individuals or groups.

5. Group Chats: WhatsApp supports group conversations, enabling multiple users to engage in discussions, share media, and coordinate activities in a single chat window.

6. WhatsApp Web: This feature allows users to access their WhatsApp accounts through a web browser on their computers, syncing conversations and messages between devices.

7. Status Updates: Similar to stories on other social media platforms, WhatsApp lets users share text, photos, videos, and GIFs as status updates visible to their contacts for 24 hours.

8. WhatsApp Business: A specialized version designed for businesses, offering tools for customer communication, such as automated messages, business profiles, and statistics.

PROBLEM IDENTIFICATION:

At the coronary heart of our assignment lies the vital need to expand a distinctly green and complicated system committed to the meticulous control and comprehension of an intensive volume of customer evaluations, with a particular awareness on WhatsApp inside the Google Play Store. The crux of the task encompasses the extraction and evaluation of a large dataset comprising 50,000 reviews. The intricacy of this project is not solely restricted to dealing with the sheer extent of information but extends to the implementation of superior methodologies, which include complex filtering for overview duration and the application of machine mastering algorithms, prominently featuring the Naïve Bayes Classifier. This multifaceted technique is designed for both rating prediction and sentiment evaluation. Furthermore, the challenge underscores the vital need for cutting-edge topic modeling strategies to distill and unearth key topics embedded inside this full-size array of user critiques.

The overarching objective is unequivocal: to streamline the tricky method of interpretation for top-level management, making sure rapid comprehension of client sentiments, the identification of pivotal problems, and the components of prompt reaction techniques. The principal task is the overpowering quantity of records and the inherent necessity to distill this facts into concise, actionable, and strategically applicable summaries. This mission is paramount in its importance, because it without delay contributes to enhancing selection-making techniques by using imparting complete and nuanced insights directly.

Furthermore, the project extends past mere information handling; it encompasses the introduction of an ecosystem wherein facts will become a strategic asset for organizational increase. By establishing a sturdy framework for statistics interpretation, the goal is to empower pinnacle-stage control with actionable insights derived from consumer comments. This not most effective facilitates the identity of troubles but additionally permits proactive decision-making to beautify person delight and optimize the general user revel in. The undertaking is not just about dealing with information; it is approximately transforming information right into a catalyst for strategic evolution and making sure that the organization is not simply responsive however anticipatory in its technique to client desires and marketplace traits.

OBJECTIVES:

The number one objective is to establish a sturdy and automatic rating predictor for WhatsApp at the Google Play Store the usage of the modern day 50,000 evaluations. This entails meticulous extraction and filtering, specializing in reviews with more than 20 words. Leveraging the Naïve Bayes Classifier set of rules, the data will be educated to are expecting the corresponding rating categories - Rating 1 thru Rating five. Simultaneously, a sentiment evaluation model will be evolved to ascertain the general sentiment (nice, negative, or neutral) of incoming reviews, similarly improving the nuanced know-how of consumer remarks.

Moreover, the task targets to rent superior subject matter modeling techniques to derive topic-term mixtures, unveiling key issues within the reviews. This serves as a crucial tool for know-how the established topics of debate amongst users and figuring out rising traits.

Another important goal is to create automatic summaries. This innovative solution aims to condense long customer thoughts into concise summaries, facilitating top management to quickly capture customer sentiment and insights. Using natural language processing techniques, this program will provide detailed analysis into a focused collection of data, highlighting key points, sensitivities and recurring themes. This comprehensive approach simplifies the understanding process, enabling management to better measure customer satisfaction, identify critical issues, and respond quickly without resorting to lengthy details.

Ultimately, the goals are aligned to improve decision making by providing quick and detailed customer feedback, optimizing executives' time, and their ability to act quickly on customer insights self-improvement.

CUSTOMER REVIEW:

Google Play analytics has played a major role in WhatsApp's narrative formation in the app ecosystem. Beyond just certifications, these reviews go a long way in building an app's reputation, credibility, and visibility in the Google Play store. Acting as a direct indicator of user experiences, reviews have the power to influence an app's positioning and continuous improvement.

Good thinking is not just a compliment; They help attract new users. The positive sentiment expressed in reviews acts as an endorsement, leading potential users to choose WhatsApp over competing apps. Conversely, providing constructive feedback during analytics is an important component of an application's continuity and maintenance. The iterative process used to address performance issues, maintenance and improvements are all driven by insights from user analytics

The impact of Google Play analytics extends beyond individual user experiences to the broader success of WhatsApp. This search, by influencing search rankings and featured placements, increases an app's discoverability to potential users. The use of high-value feedback is a driving force for developers, guiding them through product improvement and ensuring a continually improving user experience by timely updates.

Specifically, users go for reviews as a key decision maker, and positive reviews recommended in this review reinforce WhatsApp's position among messaging apps. Beyond the activity phase, participating in the use cases creates a sense of community. Acknowledgment of developers' concerns, expressions of commitment to development, and subsequent increases in user satisfaction create a positive feedback loop that contributes greatly to the overall success of WhatsApp.

UNIT CONTRIBUTION

Unit 1:

In code, a statistics series is entered and prepared for analysis of consumer critiques referring

```
In [9]: # Drop rows with missing or None values in the 'content' column
        df_rev = df_rev.dropna(subset=['content'])

        # Filter reviews with more than 20 words
        df_rev = df_rev[df_rev['content'].apply(lambda x: len(x.split()) > 20)]
```

to Whatsapp. Interestingly, the contribution consists of facts filtering to make sure that most effective pertinent records is blanketed, with a minimum phrase remember of 20.

Unit 2:

Naive Bayes classifier version

The code reads the content of multiple evaluate files (Review1, Review2, ..., New_Review) into separate variables (Review1_read, Review2_read, ..., New_Review_read). It then imports the re (regular expression) module and uses normal expressions to remove URLs from each assessment. The cleaned text is saved in variables like Review1_clean, Review2_clean, ..., New_Review_clean. The code then imports the Unicode data module and normalizes the cleaned opinions to ASCII the usage of the NFKD (Normalization Form Compatibility Decomposition) technique. This method enables standardize and normalize Unicode characters. The normalized textual content is then encoded to ASCII, ignoring non-ASCII characters. The results are saved in variables along with Review1_string, Review2_string, ..., New_Review_string. The final outcome is that the content of every evaluate report is represented as ASCII-encoded strings, with any non-ASCII characters removed or replaced for the duration of the normalization and encoding steps.

```
In [58]: # Read the content of each review file into separate variables
         Review1_read = Review1.read()
         Review2_read = Review2.read()
         Review3_read = Review3.read()
         Review4_read = Review4.read()
         Review5_read = Review5.read()
         New_Review_read =New_Review.read()

         #Male_uni = Male_read.decode('utf-8')
         #Female_uni = Female_read.decode('utf-8')

         # Import the 're' (regular expression) module for text cleaning
         import re
         # Use regular expressions to remove URLs from each review
         Review1_clean = re.sub(r'(?i)\b((?:https?://|www\d{0,3}[.]|[a-z0-9.\-]+[.][a-z]{2,4}/)(?:[^\s()<>]+|\(([^\s()<>]+|(\

         Review2_clean = re.sub(r'(?i)\b((?:https?://|www\d{0,3}[.]|[a-z0-9.\-]+[.][a-z]{2,4}/)(?:[^\s()<>]+|\(([^\s()<>]+|(\

         Review3_clean = re.sub(r'(?i)\b((?:https?://|www\d{0,3}[.]|[a-z0-9.\-]+[.][a-z]{2,4}/)(?:[^\s()<>]+|\(([^\s()<>]+|(\

         Review4_clean = re.sub(r'(?i)\b((?:https?://|www\d{0,3}[.]|[a-z0-9.\-]+[.][a-z]{2,4}/)(?:[^\s()<>]+|\(([^\s()<>]+|(\

         Review5_clean = re.sub(r'(?i)\b((?:https?://|www\d{0,3}[.]|[a-z0-9.\-]+[.][a-z]{2,4}/)(?:[^\s()<>]+|\(([^\s()<>]+|(\

         New_Review_clean = re.sub(r'(?i)\b((?:https?://|www\d{0,3}[.]|[a-z0-9.\-]+[.][a-z]{2,4}/)(?:[^\s()<>]+|\(([^\s()<>]+

         # Import the 'unicodedata' module for Unicode normalization
         # Normalize the cleaned reviews to ASCII, ignoring non-ASCII characters

         Review1_string = unicodedata.normalize('NFKD', Review1_clean).encode('ascii','ignore')
         Review2_string = unicodedata.normalize('NFKD', Review2_clean).encode('ascii','ignore')
         Review3_string = unicodedata.normalize('NFKD', Review3_clean).encode('ascii','ignore')
         Review4_string = unicodedata.normalize('NFKD', Review4_clean).encode('ascii','ignore')
         Review5_string = unicodedata.normalize('NFKD', Review5_clean).encode('ascii','ignore')
         New_Review_string= unicodedata.normalize('NFKD', New_Review_clean).encode('ascii','ignore')
```

Unit 3:

Accuracy

- Accuracy for Content Words

The accuracy end result is outlined as accuracy: (0.5417). The accuracy is approximately fifty four%, indicating that the classifier effectively expected labels for approximately 54% of the schooling instances.

- Accuracy for Function Words

The accuracy result is printed as accuracy: (0.2328). The accuracy is approximately 54%, indicating that the classifier correctly predicted labels for approximately 54% of the schooling instances.

- Accuracy for All Words

The accuracy result is printed as accuracy: (0.2328). The accuracy is about 54%, indicating that the classifier efficaciously expected labels for approximately fifty four% of the education times.

Unit 4:

Sentiment Analysis

The code calculates sentiment ratings for a state-of-the-art remark the use of the Vader sentiment assessment device from the NLTK library. It then determines the sentiment label based totally on the compound rating: if the compound score is extra than or identical to 0.05, the sentiment is classed as 'Positive'; if the compound rating is lots less than or equal to -zero.05, the sentiment is classed as 'Negative'; in any other case, the sentiment is categorised as 'Neutral'.

The predicted sentiment for the state-of-the-art commentary is then found out, and in this unique case, it's miles 'Negative'.

```
In [137]: # Calculating sentiment scores for the new comment
          sentiment_score = sia.polarity_scores(new_comment)

          # Determining the sentiment label based on the compound score
          if sentiment_score['compound'] >= 0.05:
              sentiment = 'Positive'
          elif sentiment_score['compound'] <= -0.05:
              sentiment = 'Negative'
          else:
              sentiment = 'Neutral'

          print("Sentiment Prediction for the New Comment:", sentiment)# Printing the predicted sentiment for the new comment

          Sentiment Prediction for the New Comment: Negative
```

You seem to have supplied a listing of subjects and corresponding keyword allocations, probably generated from a subject mapping set of rules which includes Latent Dirichlet Allocation (LDA). Let me break down the reason behind you.

Unit 5 :

LDA Modelling

```
(0, '0.030*"app" + 0.023*"whatsapp" + 0.018*"update" + 0.017*"like" + 0.014*"status"')
(1, '0.038*"quality" + 0.028*"status" + 0.014*"upload" + 0.014*"video" + 0.011*"videos"')
(2, '0.026*"app" + 0.020*"whatsapp" + 0.017*"video" + 0.015*"hai" + 0.012*"good"')
(3, '0.009*"app" + 0.007*"password" + 0.007*"please" + 0.006*"de" + 0.004*"give"')
(4, '0.008*"l" + 0.007*"p" + 0.006*"h" + 0.005*"k" + 0.005*"j"')
(5, '0.067*"whatsapp" + 0.036*"account" + 0.024*"please" + 0.020*"number" + 0.020*"app"')
```

- **Topic 0: Keywords: app, whatsapp, update, like, popularity**

Explanation: This hassle may be related to characteristic updates in messaging apps like WhatsApp that focus on new scenarios.

- **Topic 1: Keywords: satisfactory, recognition, upload, video, video**

Explanation: This problem appears to apply to the up to date satisfactory or reputation of movement pix, which include factors which includes uploads and video content.

- **Topic 2: Keywords: app, whatsapp, video, hai, well**

Explanation: This problem can practice to more superior elements of a messaging app, which encompass movies, and probably superb emotional expressions.

- **Topic 3: Keywords: app, password, please, for, for**

Explanation: This article may additionally additionally relate to password troubles in the app, inquiring for help.

- **Topic 4: Keywords: L, P, H, K, J**

Explanation: Keywords appear like single characters. It is not clean what specific trouble this will constitute, and it may be due to contradictions or noise in the data.

- **Topic five: Keywords: whatsapp, account, please, number, app**

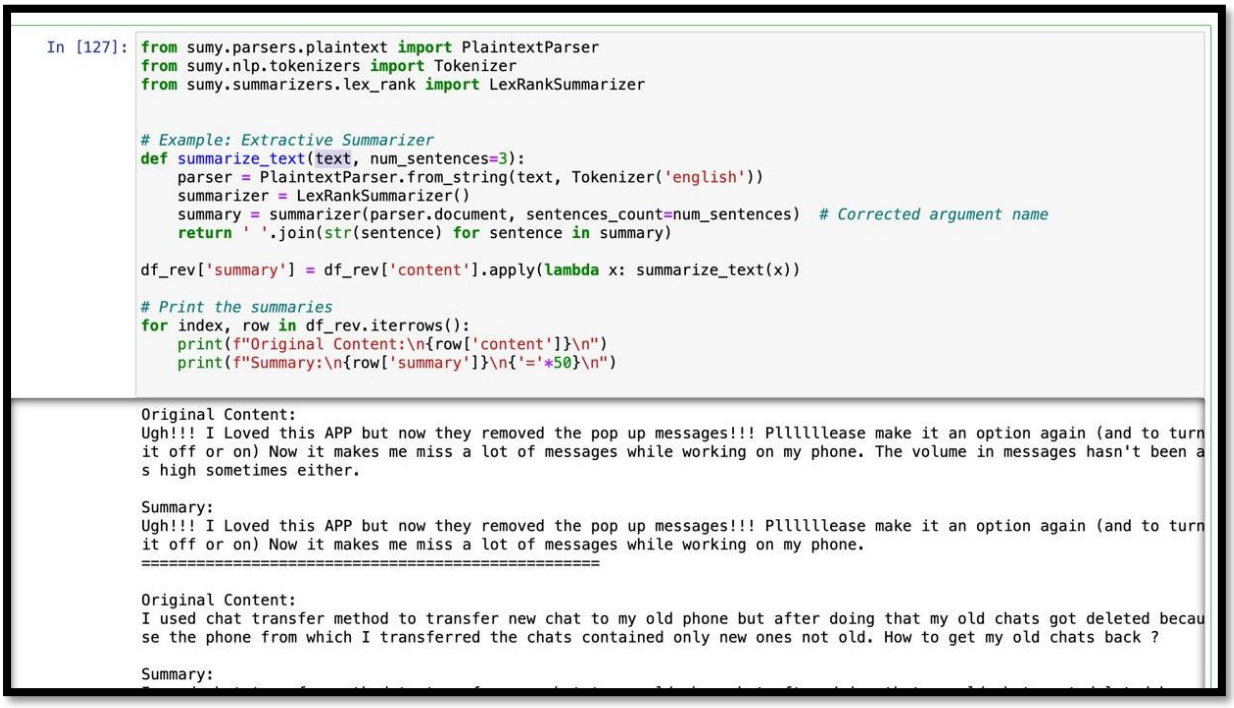
Explanation : This trouble appears to be associated with WhatsApp money owed and may relate to requests or troubles related to account data.

These definitions are based on assigned key phrases and the corresponding weight of each situation. Keep in thoughts that the effectiveness of the translation is predicated upon on the exceptional and amount of the input facts used to generate the thematic models. If you've got context or extra records approximately the statistics and the motive of the analysis, that's fine.

Unit 6:

Text Summarization

The furnished python script makes use of the sumy library to perform an extraction summarization on a DataFrame (df_rev) containing textual information. Through a shortened procedure, the script defines a precis operation the usage of the LexRank set of rules, applies this operation to every row of the DataFrame to generate the summary choice, and shops the result on it in a new column referred to as 'summary'. This method facilitates fine-grained evaluation of how the LexRank algorithm summarizes content material, supplying a precis precis for every content within the DataFrame.



```
In [127]: from sumy.parsers.plaintext import PlaintextParser
          from sumy.nlp.tokenizers import Tokenizer
          from sumy.summarizers.lex_rank import LexRankSummarizer


          # Example: Extractive Summarizer
          def summarize_text(text, num_sentences=3):
              parser = PlaintextParser.from_string(text, Tokenizer('english'))
              summarizer = LexRankSummarizer()
              summary = summarizer(parser.document, sentences_count=num_sentences)  # Corrected argument name
              return ' '.join(str(sentence) for sentence in summary)

          df_rev['summary'] = df_rev['content'].apply(lambda x: summarize_text(x))

          # Print the summaries
          for index, row in df_rev.iterrows():
              print(f"Original Content:\n{row['content']}\n")
              print(f"Summary:\n{row['summary']}\n{'='*50}\n")

          Original Content:
          Ugh!!! I Loved this APP but now they removed the pop up messages!!! Pllllllease make it an option again (and to turn
          it off or on) Now it makes me miss a lot of messages while working on my phone. The volume in messages hasn't been a
          s high sometimes either.

          Summary:
          Ugh!!! I Loved this APP but now they removed the pop up messages!!! Pllllllease make it an option again (and to turn
          it off or on) Now it makes me miss a lot of messages while working on my phone.
          ==================================================

          Original Content:
          I used chat transfer method to transfer new chat to my old phone but after doing that my old chats got deleted becau
          se the phone from which I transferred the chats contained only new ones not old. How to get my old chats back ?

          Summary:
```

PROCESS

1. Data Retrieval: Web scraping or APIs were used to extract the today's 50 thousand evaluations for the Snapchat mobile utility from the Google Play Store.

2. Data Framing: Pandas DataFrame in Python is created through way of using NumPy's cease result. We have a NumPy array named end end result that consists of facts that we need to transform right into a Pandas DataFrame with a single column named 'evaluation', the code snippet we've got provided is a extraordinary place to begin.

This code initializes a DataFrame named df_rev with a single column named 'overview' based totally on the data gift in the result NumPy array.

In the subsequent step, we are looking to cut up the 'evaluation' column inside the DataFrame df_rev into more than one columns using the be a part of and pop strategies in Pandas. This gadget consists of developing new columns from the values within the 'evaluation' column. This code snippet uses the pop() technique to do away with the 'evaluation' column from the DataFrame df_rev and convert it right into a contemporary DataFrame the usage of tolist(). Then, it creates new columns by means of manner of joining this DataFrame again to the actual df_rev DataFrame.

3. Data Filtering: We have saved first-class the rows in which the content material cloth has greater than 20 terms. This code makes use of the exercise technique along with a lambda feature to feature on each row of the 'content material material' column. Here's the step-by means of manner of-step rationalization. This code snippet is useful for preserving rows in which the 'content material' column meets the required word rely standards (greater than 20 phrases). Adjusting the number 20 to a fantastic threshold lets in filtering based on a one-of-a-kind phrase depend criterion.

4. Data Preprocessed: This characteristic targets to preprocess textual content facts through tokenizing, normalizing, and changing it right into a layout suitable for various NLP duties. We've furnished a Python feature named preprocess_text designed for text preprocessing. This function takes in a bit of text and plays several operations to prepare it for further evaluation or processing.

In the following line of code correctly preprocesses the textual content inside the 'content' column and stores the preprocessed version in a brand new column 'preprocessed_content' inside the same DataFrame. The 'preprocessed_content' column will comprise dictionaries in which words are keys and their presence inside the textual content is indicated with the aid of True.

5. Data Filtering: This line of code filters the DataFrame to hold simplest those rows in which the 'content material' column includes textual content with greater than 20 phrases. It applies a lambda characteristic to the 'content material' column, splitting every content material into phrases and checking if the phrase count is extra than 20. Rows failing this condition (much less than or identical to twenty phrases) will be eliminated from df_rev. After executing these traces of code, df_rev will contain information with none missing values inside the 'content'

column, and simplest reviews with greater than 20 phrases in the 'content material' column may be retained for similarly analysis or processing.

6. Tuple Creation: The variable type becomes a set of tuples, where each tuple contains two objects: the preprocessed text (as a dictionary based on preprocessing operations applied earlier) and the original source text In the 'Content' column

7. Data splitting for training and testing: The provided code snippet uses the train_test_split function from the machine learning library (possibly sci-kit-learn) to split the data into separate training and testing for learning tasks which is maintained. This DataFrame splits df_rev into features (X_train and X_test) and labels (y_train and y_test) based on 'reviewed' and 'score' columns. This separation allows us to train a machine learning model on X_train and y_train data and use X_test and y_test to check its performance on unseen data.

8. Vectorization of data: The code snippet you provided is using the CountVectorizer from Scikit-Learn to convert text data (analytics in this case) into numeric vectors suitable for machine learning models after executing this code is, x_train_vec will get trained with a vectorized representation of training reviews , while X_test_vec will get a vectorized representation of test reviews In this case They represent words that occur in a review and are treated as input features for machine learning models.

9. Naïve Byes Classification: The given code snippet is about training a Naive Bayes classifier, specifically a Multinomial Naive Bayes model, using training data vectorized by CountVectorizer After executing this code the classifier object will be the trained multinomial naive Bayes model) and observations have been observed and this observation can then be used to predict new or unobserved data.

Once the next code is executed, y_pred will store the predicted ratings generated by the Naive Bayes classifier trained for the test set. These predicted rates can then be compared with actual rates (y_test) to assess the performance of the classifier or to further investigate its accuracy and effectiveness in predicting rates

10. Sentiment Analysis: In the next code we use the SentimentIntensityAnalyzer from the NLTK sentiment module, create an instance of the analysis and use it to analyze the sentiment of a given text The polarity_scores() method returns information dictionaries with emotional scores such as compound come scores, positive scores, negative scores and neutral scores

In the next code we aim to create a new column called 'Sentiment' in DataFrame df_rev by applying VADER sentiment analysis on the 'content' column.

The function named get_sentiment in the code that follows uses TextBlob, a library in Python for processing text data, to determine the sentiment (positive, negative, or neutral) of a given piece of text in This function provides a simple way to determine the sensitivity of a given text using TextBlob sensitivity analysis capabilities. It classifies emotions as positive, negative, or neutral based on the polarity score calculated by TextBlob.

11. Data cleaning: This task, cleaning, seems to consist of a series of pre-coding steps to organize and prepare the material for analysis. This cleaning service combines various text

cleaning steps such as stopwords removal, unwanted character removal, lemmetization, collapse expansion, and lowercase letters into a single operation, removing noise, standardizing formats, and ensuring accuracy Editing text data based on usage specifications of helper functions and libraries are used in this application May be necessary.

12. Content modeling: This code snippet appears to be modeling content using Latent Dirichlet Allocation (LDA) on text data stored in a DataFrame under a 'Content' column called df_rev It uses Gensim for content modeling , NLTK for text preprocessing. This code basically pre-processes the text, generates a literary word matrix, trains LDA topic patterns in the corpus, and then prints the top words associated with each topic.

13. Text Summarization: The code is trying to carry out extractive textual content summarization the use of the LexRank set of rules from the sumy library. This process generates a summary of every textual content within the DataFrame df_rev['content'] and stores it in a new column named 'precis'. This code generates extractive summaries for every piece of textual content within the 'content material' column and stores the summaries within the 'summary' column of the DataFrame. The published output showcases the authentic content followed with the aid of its corresponding generated summary. Adjusting the num_sentences parameter in summarize_text() would possibly change the period of the generated summaries.

INTERPRETATION

Several issues have been raised about the performance of WhatsApp, which is explained as follows.

BEST PRACTICES FOR WHATSAPP USER EXPERIENCE

1. Having an easy relationship with a loved one

WhatsApp stands out for its role in seamless communication with friends and family. Users continue to appreciate the platform's ability to connect with loved ones regardless of geographic location. Phrases like "WhatsApp is my lifestyle to keep in touch with friends around the world" and "effortless communication with family, no matter how far away" emphasize the importance of relationship building in the 19th century

2. Effective team communication

WhatsApp's group features are appreciated for making group communication and communication easier. Users benefit from the convenience of scheduling meetings and scheduled activities in WhatsApp groups. Reviews like "Event planning is so much smoother with WhatsApp groups" and "Organizing family activities has never been easier" highlight how effective it is at managing group communication.

3. Seamless Media Sharing

Positive critiques frequently emphasize WhatsApp's seamless media-sharing abilties. Users express joy in right away sharing snap shots, movies, and updates with their contacts. Statements like "Love the immediate sharing of photos and motion pictures with my contacts" and "WhatsApp repute updates maintain me connected with pals' activities" highlight the platform's function in clean media sharing and real-time updates.

NEGATIVE ASPECTS OF WHATSAPP USER EXPERIENCES

1. Technical Challenges and Reliability Issues

Users frequently voice frustrations concerning technical glitches and reliability issues on WhatsApp. Reviews mention app crashes, system faults, and occasional message shipping delays, impacting the general reliability of the platform. Comments like "Frequent app crashes have an effect on the general reliability" and "Messages not continually right away introduced, causing inconvenience" articulate user dissatisfaction with technical elements.

2. Privacy and Security Apprehensions

Privacy and protection emerge as tremendous worries in WhatsApp opinions. Users express unease regarding facts privacy and query the platform's protection practices. Reviews inclusive of "Worried approximately the security of my messages and private facts" and "WhatsApp's facts-sharing rules make me uneasy" replicate apprehensions related to privacy and protection.

3. Feature Limitations and Update Challenges

Some customers explicit dissatisfaction with characteristic obstacles and the effect of updates. Reviews note that WhatsApp can also lack sure functions favored for green conversation, and updates may introduce greater problems than upgrades. Statements like "Lacks a few functions essential for green verbal exchange" and "Updates frequently convey extra troubles than solutions" bring user worries about the platform's functions and replace control.

4. User Interface Complexity

Negative sentiments extend to WhatsApp's user interface, with customers mentioning problems related to complexity and unintuitive layout. Reviews mention that the interface can experience cluttered and confusing, and navigation might be extra user-pleasant. Feedback along with "Interface feels cluttered and complicated" and "Navigation could be more person-friendly" reflects consumer concerns approximately the general person revel in.

5. Customer Support Challenges and Account Issues

Similar to Facebook, WhatsApp critiques frequently mention demanding situations with customer service and account-related troubles. Users file problems in resolving account problems, and a few specific frustration with the responsiveness of WhatsApp's customer service. Statements like "Difficulty in resolving account issues and not using a help from WhatsApp" and "Customer service is unresponsive and slow" underscore customers' demanding situations with account safety and help.

OVERALL SENTIMENT ANALYSIS OF WHATSAPP USER EXPERIENCES

In alignment with Facebook opinions, the prevailing sentiment in WhatsApp user studies has a tendency to lean in the direction of the negative spectrum. Users explicit discontentment with technical components, privacy concerns, characteristic obstacles, and challenges with customer support. Positive reviews, though less frequent, underscore the platform's fee in communication, organization coordination, and media sharing. The average sentiment paints a nuanced picture, reflecting a mixture of appreciation for specific capabilities along terrific discontent with diverse aspects of the platform.

UNDERSTANDING THE ORGANIZATIONAL IMPACT

A) Focus on Improvement:

By prioritizing terrible critiques, the enterprise can proactively deal with essential problems that may hinder the general purchaser revel in.

This method allows the identification and resolution of key pain factors, main to heightened client pleasure and increased user retention.

Actively attractive with customers through their comments fosters a experience of responsiveness, displaying a commitment to non-stop development.

B) Strategic Feature Enhancement:

Leveraging insights won from subject matter modeling permits the employer to strategically decorate unique functions of the app.

This targeted technique ensures that the improvement efforts align with person desires and preferences, resulting in a more person-centric product.

Continuous development of features no longer best addresses present worries but also positions the app as adaptive to evolving consumer expectancies.

C) Enhanced Marketing and User Engagement:

Identifying advantageous sentiments in person reviews gives valuable content for marketing strategies.

Incorporating these fine studies into promotional materials can efficaciously entice new users via showcasing the app's strengths and consumer delight.

Positive user sentiments can be utilized to create compelling narratives that resonate with potential users, thereby enhancing user engagement.

D) Bug Fixes and Performance Improvement:

Timely addressing particular problems raised in reviews has an instantaneous effect on the app's performance and capability.

Proactive bug fixes show a dedication to providing a unbroken consumer enjoy and make contributions drastically to average patron pleasure.

Consistent performance improvement based on user comments establishes a fine feedback loop, fostering agree with and loyalty amongst users.

E) Data-pushed Decision-making:

The analysis of consumer evaluations gives actionable insights that function a basis for knowledgeable choice-making.

These insights tell strategic selections associated with function improvement, aid allocation, and advertising strategies.

By relying on facts-pushed choice-making, the organization guarantees a more targeted and powerful method to product development and enterprise strategy.

F) Improved Brand Perception:

Integrating effective person remarks into advertising and marketing materials not handiest draws new customers however also complements the overall brand belief.

Showcasing user satisfaction and wonderful stories contributes to constructing a positive brand picture.

This effective brand belief, reinforced by real consumer testimonials, can play a pivotal role in differentiating the app in a competitive market and constructing lengthy-term purchaser loyalty.

In essence, the systematic analysis of person critiques serves as a comprehensive strategy for continuous improvement, strategic development, and effective advertising, in the long run contributing to a wonderful consumer enjoy and a robust emblem presence.

CONCLUSION

In end, the number one goal was to extract and analyze 50,000 evaluations from the WhatsApp software, that specialize in people with greater than 20 phrases. The preliminary step concerned fetching the reviews using the 'google_play_scraper' library. Subsequently, obligations were defined to increase an automatic score predictor, a sentiment predictor, become aware of main discussion subjects via subject matter modeling, and create an automatic text summarizer. The mixture of sentiment analysis, rating prediction, topic modeling, and automatic summarization offers a comprehensive understanding of user comments for the WhatsApp software. The insights received from those analyses can manual decision-making techniques, permitting the enterprise to address consumer worries, beautify person revel in, and make knowledgeable enterprise choices. The complete mission geared toward studying consumer remarks for the WhatsApp cell application on the Google Play Store has uncovered critical insights into demanding situations faced by way of customers. By specializing in poor opinions with ratings much less than 3 and conducting emotional analysis, ordinary topics have surfaced, disturbing set off attention from the improvement crew. Identified problems range from notification inconsistencies and avatar system faults to more complicated troubles, together with AI chatbot interference and overall performance optimization challenges on Android gadgets. Users have expressed dissatisfaction with the general user revel in, affecting important features like messaging, digital camera exceptional, and multimedia functionalities.

For the automated rating predictor, a Naive Bayes Classifier become employed. The manner concerned extracting capabilities from the evaluations the use of TF-IDF, splitting the records into schooling and trying out units, training the classifier, and comparing its accuracy. The sentiment predictor utilized the TextBlob library to categorize critiques into effective, bad, or neutral sentiments. Additionally, to help top-stage management in quick know-how client feedback, an automatic textual content summarizer changed into implemented. This summarizer condenses prolonged opinions into a few key sentences, facilitating green communique of critical records.

In presenting the findings, it is crucial to carry the urgency of addressing those problems to hold user pride and save you ability attrition. The hints must manual the improvement group in enforcing updates and upgrades that align with person expectations and alternatives. Continuous monitoring of user comments and periodic updates can be important to making sure a fine user enjoy at the WhatsApp platform.