Decoding Google Play Store Apps User Reviews

Sujan Shrestha
University ID: 226673
6th Semester, BSc IT
Herald College Kathmandu
np03cs4s220179@heraldcollege.edu.np

Nayan Raj Khanal
University ID: 2227486
6th Semester, BSc IT
Herald College Kathmandu
np03cs4s220389@heraldcollege.edu.np

Abstract—Attributes such as App, Translated_Review, Sentiment, Sentiment_Polarity and Sentiment_Subjectivity are available in the Google Play Store User Reviews dataset which gives us insights into users feelings. Google Play store is massive with lots of and lots of app that are both successful and failures. Veteran and novice developers both struggle when pushing apps to the store and most often it feels like a hit or miss. The mvxcain aim of our task is to eliminate this feeling of hit or miss by providing developers clear insights into which app users prefer and which they don't. Utilizing text classification techniques designed for sentiment analysis we have provided detailed methodology of the task undertaken. This methodology allows informed decision making for developers enhancing user satisfaction. The report findings will provide significant insights for the research community helping optimize user experience. [1] [2] [3] [4]

Keywords: **Sentiment Analysis, Logistic Regression, Google Play Store-User Reviews, Text Classification, Machine Learning, Data Preprocessing, Model Evaluation, Visualization, Spark ML Pipeline**

I. INTRODUCTION

Google Play store is a massive platform. It contains many apps, some successful and popular while others are failures and forgotten. We aim to identify what influences the success of an app by looking into the user reviews for different apps. By doing so our goal is to help app developers get the proper recipe for success and users the benefits of enjoying user-centric apps. [2]

To explore the Google Play Store User Reviews Dataset, which contains attributes App, Translated_Review, Sentiment, Sentiment_Polarity and Sentiment_Subjectivity. From these attributes we will perform data analysis, data visualization and pre-processing. By analyzing, cleaning and visualizing we can look into user trends which will help app developers to design user-centric apps. [1]

The main focus is on the mobile app development industry, specifically for Android app development in the Google Play Store ecosystem. All app developers, marketers, and users will benefit by understanding users' feelings towards various apps. By analyzing the Google Play Store User Reviews dataset, we aim to provide insights to stakeholders in the app industry to improve app quality, discoverability, and user satisfaction. [3]

II. BACKGROUND OF THE STUDY

A. Generic Information:

The Google Play Store is a massive hosting platform which hosts millions of mobile applications which caters to the various interests and needs of the users but not every application accomplishes the similar level of success. For the developers and the users of the mobile application, they need to understand the sentiment users harbors towards the app. This study's objective is to discover the insights related to the dynamics of the market of the application through the analysis of user reviews and visualization of the reviews. [2]

B. Problem Statement:

Every user is different, we can never build something that 100% of the masses agree or like.

Hence, user experience influence whether a mobile app becomes a hit or gets lost in the digital sea. With the Google Play Store dataset that is exactly what we aim to do, we will break down these reviews into simple, understandable terms. We analyze the data and create a model that classifies reviews into Positive or Negative and display it in easy-to-understand graphs, by doing so we can uncover the secrets behind app success and help both stakeholders and developers navigate and develop the app world with ease. [4]

C. Aim/objective of the work:

The main objective of this study is to examine the dataset of Google Play Store User Reviews in order to recognize the users reviews which influences the success of the application. The main aim is to provide valuable information for the developers and the users of the mobile application by analyzing the various reviews. This contributes to the success of the application as it provides valuable information to the developers to optimize their application and to the users to easily make decisions during the selection of the applications. [4] [5]

D. Contributions of the work:

The cleaning and the preprocessing the dataset help to ensure the quality of the data and consistency. The techniques such as visualization of the reviews and text classification techniques designed for sentiment analysis helps to discover users' attitude in the dataset. Recognizing and studying the reviews helps to recognize the valuable factors which influences and affects the success of the application. [5]

E. Organization of the Report:

This study report of Google Play Store User Reviews dataset includes various topics such as introduction, data cleaning and preprocessing, data analysis, model building and training, findings and insights, and conclusion. Each topic consists of detailed information regarding the analysis and understanding of the dataset which aims to provide valuable information to the developers and the users of the application.

A. RELATED WORK

A handful of studies have been carried out on the mobile app development field. While the main focus is mostly on the App Store for the Apple ecosystem, few studies delve into the Google Play Store for the Android ecosystem.

In the paper "Android Apps Success Prediction Before Uploading on Google Play Store" aims to predict the success of the app via the prediction of user rating and installation number. They used a scrapped dataset containing 267000 unique app data which was cleaned and converted into appropriate forms. After that, they created bar charts to analyze different attributes and finally used machine learning algorithms such as Random Forest, K-Nearest Neighbor, and Support Vector Machine to assume the success of the new apps. They concluded that the KNN and SVM algorithms predicted the success rate more accurately than the Random Forest. [6]

Similarly, the paper "Android App Success Prediction Based on Reviews" aims to determine the success of an app by applying sentiment analysis to the apps and related attributes. They scrapped websites with user reviews to create a dataset. They then performed data preprocessing using "re library" in Python and sentiment analysis following that. For regression, they used Stochastic gradient descent (SDG) and Support Vector Machine Regression (SVM) and trained their model. The conclusion drawn was that SVR is superior to SDG as it predicted success more accurately. [7]

Also, the paper "Success and Failure Rate Prediction of Android Application Using Machine Learning" also aimed to predict the likelihood of success of Android applications using machine learning models. They gathered a dataset consisting of 30000 apps and 184 features from third-party means. Then they divided the apps into two groups and dropped redundant information. Data cleaning, dimension reduction, and mathematical analysis were carried out respectively. For machine learning algorithms Logistic Regression (LR), Decision Tree (DT), Support Vector Machine (SVM), and Extreme Gradient Boosting (XG-Boost) are used. The

conclusion being that the XG-Boost classifier obtained the highest accuracy of 84.44%. [8]

Finally, the paper 'Play Store App Analysis & Rating Prediction Using Classical ML Models & Artificial Neural Network' used Play Store Analysis Data to analyze various attributes present and develop models accordingly and predict the rating of the app. Using simple common accuracy metrics, the models were judged and real-time predictions of new apps in the play store was carried out. ANN model performed the best by giving an accuracy of 81.61%. The paper concludes by saying LSTM architecture can be utilized to improve accuracy and that will be addressed in their future works. [9]

While existing works have made valuable contributions to understanding various aspects of app analysis, our work is different. We consider just the core attribute 'Reviews' which allows us to solely focus on developing a model that classifies reviews allowing us for a more in-depth analysis of an app's success by figuring out users mood towards the app. Secondly, most of the research utilized traditional data analysis techniques which may have limited their ability to uncover complex trends and patterns, whereas we'll be employing advanced data analysis techniques and text classification model to gain valuable insights that might have been missed.

III. METHODOLOGY

The logistic regression which is further categorized into various phases is used as the methodology for the sentiment analysis project on Google Play Store user reviews. For the comprehensive and systematic approach, each phase deals with a particular aspect of the process. Each phase is mentioned and explained below along with a block diagram. [10]

Phases:

- Data Collection
- Data Pre-processing and Cleaning
- Feature Extraction
- Model Selection and Training
- Model Evaluation and Prediction

Block Diagram:

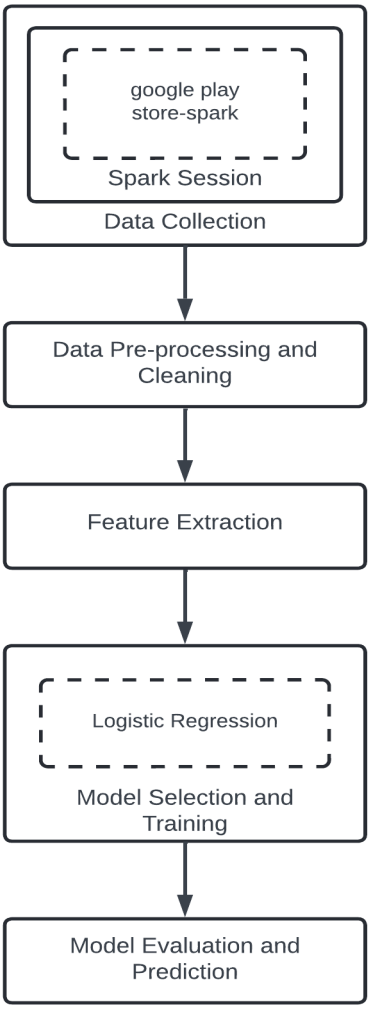

Fig. 1 Block Diagram

A. Phase 1: Data Collection

A collection of user reviews from the google play store is used as the dataset for this project. The attributes such as Translated_Review, Sentiment, App, Sentiment_Polarity, and Sentiment_Subjectivity are involved in this dataset. The mentioned attributes provide important information regarding the experiences and opinions of the users which also allow for prediction and sentiment analysis.

- **Translated_Review:** The Insights regarding users' experiences and opinions are provided by the text of user reviews.

- **Sentiment:** For each review, the sentiment labels including Positive, Negative and Neutral are allocated to indicate the emotional tone.
- **App:** The reviews of the specific application name allow for app-specific sentiment analysis.
- **Sentiment_Polarity and Sentiment_Subjectivity:** The auxiliary parameters, such as Sentiment_Polarity and Sentiment_Subjectivity gives additional information regarding the sentiments which are expressed in the reviews.

App	Transl	Sentin	Sentiment_Po	Sentiment_Sub
10 Best Fo	I like e	Positiv	1	0.533333333
10 Best Fo	This he	Positiv	0.25	0.288461538
10 Best Fo	nan	nan	nan	nan
10 Best Fo	Works	Positiv	0.4	0.875
10 Best Fo	Best ic	Positiv	1	0.3
10 Best Fo	Best w	Positiv	1	0.3
10 Best Fo	Amazi	Positiv	0.6	0.9
10 Best Fo	nan	nan	nan	nan
10 Best Fo	Lookin	Neutra	0	0
10 Best Fo	It help	Neutra	0	0
10 Best Fo	good y	Positiv	0.7	0.6

Fig. 2 Column Details

B. Phase 2: Data Pre-processing and Cleaning

Data cleaning ensures the reliability and quality of the model, as it is one of the important steps. Data cleaning includes the removal of noise and unnecessary information from the dataset. The performance of the model is directly affected by this essential step because the quality of input data directly influences the model's performance. The steps of data cleaning involve importing the data into Spark DataFrame, removal of unnecessary columns, managing missing values, removal of duplicates, cleaning the text data which is done by using a pipeline which involves the steps of converting text to lowercase, removal of URLs, unnecessary characters, tokenizing text, removal of stopwords and dividing tokenized text into array of strings. [11]

C. Phase 3: Feature Extraction

The process of converting the cleaned text data into a format which is suitable for machine learning algorithms is known as feature extraction. The StringIndexer is used to convert the sentiment column into numerical labels and to convert the cleaned reviews into a matrix of token counts, CountVectorizer is used. [12]

D. Phase 4: Model Selection and Training

For the sentiment analysis, logistic regression is used as the model because it is efficient and simple for binary classification tasks. In binary classification problems like sentiment analysis, it demonstrates effectiveness. It provides probabilities for the prediction of class because it is a straightforward and interpretable model. This also helps to understand the model's confidence in its predictions. In order to facilitate seamless data processing and model training, the model was incorporated into a Spark ML pipeline. [13]

- **Pipeline Creation:** A pipeline is created to simplify and speed up the process of training the model and feature transformation. The steps such as the string indexing, count vectorization and logistic regression are involved in the pipeline.
- **Model Training:** The transformed features are used in the training of the logistic regression model.

E. Phase 5: Model Evaluation and Prediction

After the training is completed, the evaluation of model is done and then used for prediction.

- **Model Evaluation:** Metrics like precision, recall, accuracy, and the F1-score are used to evaluate the performance of the model. The reports of confusion matrix and classification are created because it provides detailed information or insights.
- **Prediction:** The sentiment of new user reviews is predicted by using the trained model.

F. Phase 6: Conclusion

In conclusion, a structured methodology including collection of data, pre-processing, feature extraction, selection of the model, training, evaluation, and prediction is demonstrated using

logistic regression for the sentiment analysis project on Google Play Store user reviews. The project provides a reliable framework which is achieved through the collection of rich datasets, cleaning and preparing the data, extraction of the important features and utilization of logistic regression within a Spark ML pipeline to understand the sentiments of the users. The model's efficiency is ensured by the evaluation metrics and its implementation for sentiment prediction helps to enable current analysis of trends in the app market. The important insights into the experiences and preferences of the user are achieved by implementing this approach. This helps in decision-making for the developers of the app and stakeholders. [10]

IV. RESULT AND DISCUSSION

A. Experimental Setup

The sentiment analysis of Google Play Store user reviews is included in the experimental setup. The dataset which are used includes the reviews of the user for different apps which also involves the related sentiments like Positive, Neutral, Negative; and sentiment scores such as Sentiment_Polarity and Sentiment_Subjectivity. The steps of the data preprocessing and analysis procedures are mentioned below:

- Data Import and Initial Inspection:
- Data Cleaning:
- Text Preprocessing:
- Filtering for Sentiment Analysis: [10]

B. Discussion and Analysis of the Findings

The evaluation of the sentiment distribution, visualization of the common words and assessment of the performance of the sentiment analysis model is involved in the analysis phase by implementing different metrics. The findings are mentioned and explained below: [14]

C. Sentiment Distribution:

To demonstrate the distribution of positive and negative reviews, a bar chat was created. A quick overview of the sentiment landscape is provided by this chart. This chart aids in determining if the users leave more positive or negative feedback.

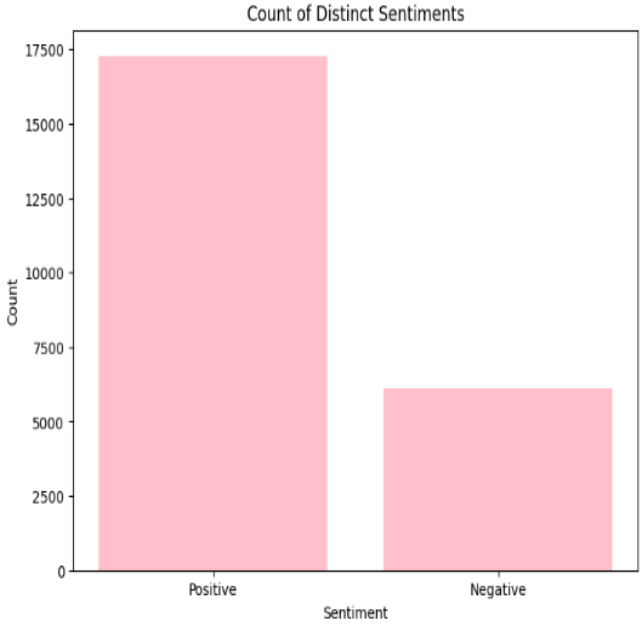

Fig. 3 Sentiment Distribution in User Reviews

D. Word Cloud Visualization:

In order to visualize the most often occurring words in the positive reviews as well as in the negative reviews, a word cloud was created. The key terms and themes which are respectively commonly found in the positive and negative sentiments are highlighted by these word clouds. The words such as "easy", "great", "love", and many more are usually mentioned in the positive reviews whereas the words such as "poor", "crash", "bug" and many more are mentioned in the negative reviews. [15]

Fig. 4 Word Cloud for Positive Sentiment Reviews

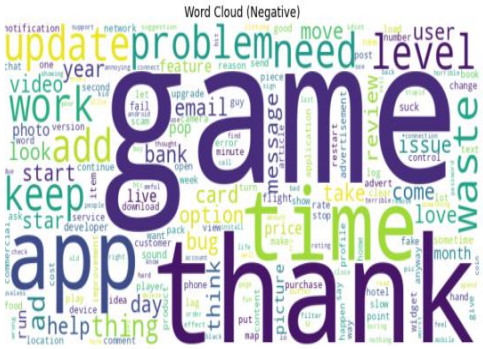

Fig. 5 Word Cloud for Negative Sentiment Reviews

E. Model Evaluation:

For the evaluation of the performance of the sentiment analysis model, classification metrics like precision, recall, and F1-score were used. In a classification report (fig 5), these metrics were calculated and displayed.

	precision	recall	f1-score	support
0	0.89	0.84	0.87	3365
1	0.61	0.71	0.66	1187
accuracy			0.81	4552
macro avg	0.75	0.78	0.76	4552
weighted avg	0.82	0.81	0.81	4552

Fig. 6 Classification Report for Sentiment Analysis Model

The model's performance is good and displays good performance with balanced precision and recall scores. This indicates that it is effective at recognizing both positive and negative sentiments accurately. [16]

F. Confusion Matrix**:**

The model's performance is demonstrated by the confusion matrix (fig 6). This provides a detailed insight regarding the performance of the model by displaying the number of true positives, true negatives, false positives, and false negatives. The confusion matrix demonstrates that there is slightly higher rate of accurately detecting positive reviews compared to negative reviews. [17]

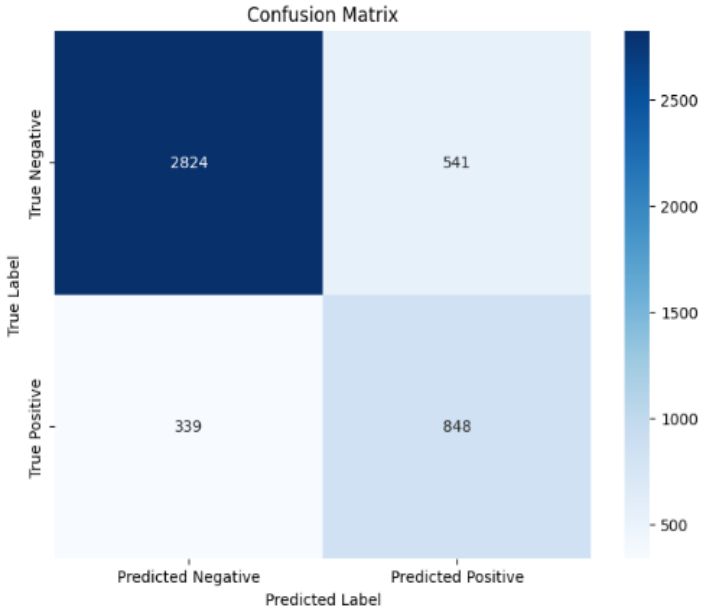

Fig. 7 Confusion Matrix of Sentiment Analysis Model

V. CONCLUSION

The sentiment analysis of Google Play Store user reviews has helped to gain important insights or information regarding the opinions and experiences of the users. This is very helpful for the developers, marketers and the stakeholders of the app. In sentiment classification, the effectiveness of text preprocessing and machine learning techniques were emphasized by the experimental setup and analysis. The preprocessing steps involves text cleaning, tokenization, and removal of stopwords which were very important as it helped in the process of converting raw review text into a suitable format for analysis. This also ensured that the input data were clean and consistent. The analysis of sentiment distribution shows that the positive reviews were commonly found. This indicates general user satisfaction. The common themes along with positive reviews usually mentioning the terms such as "easy", "great," and "love" as well as the negative reviews which mentioned the terms such as "poor," "crash," and "bug" were showed by the analysis of the sentiment distribution. The performance of the logistic regression model was well along with balanced precision, recall and F1-scores which showed efficiency in recognizing the both

positive and negative sentiments, however it was slightly better at identifying positive reviews. The actional insights or information regarding these findings are very important as it helps the app developers to enhance the quality of the app and satisfaction of the user whereas the marketers can create effective strategies. These findings are important and useful for the stakeholders as well because it helps to make decisions regarding the investment and developments of the app. Future research could emphasize on the model's accuracy for detecting negative sentiment and discovering advanced machine learning techniques in order to receive more insights. This study offers a reliable framework is provided for understanding the experiences and preferences of the user by demonstrating a structured methodology for the analysis of the sentiment. In the competitive market, this analysis provides valuable insights which are important for making proper business decisions, optimization of app development and enhancement of the satisfaction of user.

VI. REFERENCES

[1] LAVANYA, "Google Play Store Apps," 2019. [Online]. Available: https://www.kaggle.com/datasets/lava18/google-play-store-apps. [Accessed 1 May 2020].

[2] R. Sheldon, "Google Play," 2021. [Online]. Available: https://www.techtarget.com/searchmobilecomputing/definition/Google-Play-Android-Market. [Accessed 1 May 2024].

[3] A. Gupta and D. Kamthania, "Study of Sentiment on Google Play Store Applications," 2021. [Online]. Available: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3833926. [Accessed 1 2014 May].

[4] Y. N, "Binary Sentiment Classification," 2023. [Online]. Available: https://medium.com/@yeshsurya/binary-sentiment-classification-157bce6688f. [Accessed 1 May 2024].

[5] N. Bhole, "Sentiment Analysis of Google Play Store reviews using Python," 2021. [Online]. Available: https://medium.com/@nikita.bhole05/sentiment-analysis-of-google-play-store-reviews-using-python-a7df195b42fc. [Accessed 1 May 2024].

[6] M. S. H. D. K. M. J. K. Golam Md. Muradul Bashir, "Android Apps Success Prediction Before Uploading on Google Play Store," 2020. [Online]. Available: https://ieeexplore.ieee.org/document/9068071. [Accessed 03 04 2024].

[7] S. U. Keerthana Pramudi Suresh, "Android App Success Prediction based on Reviews," 2020. [Online]. Available: https://ieeexplore.ieee.org/document/9051529. [Accessed 03 04 2024].

[8] U. S. I. A. M. N. M. K. G. Swagatam Jay Sankar, "Success and failure rate prediction of Android Application using Machine Learning," 2023. [Online]. Available: https://ieeexplore.ieee.org/document/10113988. [Accessed 03 04 2024].

[9] B. Moharana, B. B. Biswal, S. Dey, M. K. Rath and S. Banerjee, "Play Store App Analysis & Rating Prediction Using Classical ML Models & Artificial Neural Network," Pune, 2023.

[10] A. BIYANI, "A Complete Guide to Sentiment Analysis," 2023. [Online]. Available: https://careerfoundry.com/en/blog/data-analytics/sentiment-analysis/. [Accessed 17 05 2024].

[11] D. E. Lee, "Text Preprocessing and Classification with Logistic Regression," 2024. [Online]. Available: https://drlee.io/text-preprocessing-and-classification-with-logistic-regression-ea4fe3cfcaac. [Accessed 10 May 2024].

[12] D. Agarwal, "What is Feature Extraction and Feature Extraction Techniques," 2024. [Online]. Available: https://www.analyticsvidhya.com/blog/2021/04/guide-for-feature-extraction-techniques/. [Accessed 10 May 2024].

[13] V. Haswani, "Create a Pipeline to Perform Sentiment Analysis using NLP," 2020. [Online]. Available: https://www.analyticsvidhya.com/blog/2020/11/cr

eate-a-pipeline-to-perform-sentiment-analysis-using-nlp/. [Accessed 10 May 2024].

[14] D. L. Lei Lei, "Conducting Sentiment Analysis," 2021. [Online]. Available: https://www.cambridge.org/core/elements/abs/conducting-sentiment-analysis/B00BACADE638BF1AD5F61972FEE4183D. [Accessed 17 05 2024].

[15] Amanatulamriyah, "Sentiment Analysis — WordCloud," 2023. [Online]. Available: https://medium.com/@amanatulamriyah66/sentiment-analysis-wordcloud-c09787c27803. [Accessed 10 May 2024].

[16] S. Kohli, "Understanding a Classification Report For Your Machine Learning Model," 2019. [Online]. Available: https://medium.com/@kohlishivam5522/understanding-a-classification-report-for-your-machine-learning-model-88815e2ce397. [Accessed 10 May 2024].

[17] Deepanshi, "In-depth understanding of Confusion Matrix," 2023. [Online]. Available: https://www.analyticsvidhya.com/blog/2021/05/in-depth-understanding-of-confusion-matrix/. [Accessed 10 May 2024].

VII. APPENDIX

A. Peer Review

This report was a collaborative task between Sujan Shrestha and Nayan Raj Khanal, below is the peer review which states what each member's contribution to the task:

- Nayan Raj Khanal for the report section was tasked with choosing the appropriate dataset and its title, then research on related works revolved around the dataset of choice and finally explaining the code in layman's term in the methodology section. For the code section, Nayan was tasked with pre-processing of the data that included cleaning the dataset by removing unwanted headers, characters, stopwords and tokenizing the texts, encoding, feature extraction etc alongside model selection and training.
- Sujan Shrestha for the report section was tasked with providing the abstract of the report alongside background of the study which answered the problem statement, aims/objectives, research etc. and interpreting the results and providing discussion interpreting the results and providing discussion. For the code section, Sujan was tasked with data gathering and visualization which included plotting of wordclouds, bar graphs, classification report, confusion matrix and drawing predictions table.

The completion of the task is due to the equal contribution from both individual during the data collection, pre-processing and cleaning, feature extraction, model selection and training and evaluation and predictions tasks and designated report sections. At the end of the task both members are satisfied with the roles their partner played and have no complaints.

B. CSV File and Code Link:

- https://drive.google.com/drive/folders/1pgUaoUQW3IUYcmZqrAu5y1mtgpH4imnR?usp=sharing