CREDIT CARD COMPLAINTS ANALYSIS USING NLP BY

LEVERAGING TRANSFORMERS

BANSARI SHAH

KOMAL PRAKASH LONDHE

JEEVAN SAI SREELA

VAISHNAVI VEJELLA

University of Maryland, Baltimore County (UMBC)

DATA 690: SPECIAL TOPICS IN DATA SCIENCE: INTRODUCTION TO NLP

DR. TONY DIANA

6th MAY, 2022

ABSTRACT

Over the epidemic, the number of complaints about being overcharged, having problems with the payment process, and with customer support has quadrupled. The analysis of credit card complaints is important these days since it determines the overall satisfaction of customers who use various companies' credit cards. This study examines numerous factors to comprehend the issues that lead to complaints being filed. Transformers have been employed to deal with sequential input data as well as text summarization. It also aids in the classification of complaints along with interpreting the intent for the complaints filed. The most prevalent classification types observed were Functionality concern, Security threat, and Overcharged. Customers' emotions have been revealed to be recognized by Zero-Shot Classification when it comes to resolving problems. Furthermore, machine learning is incorporated with transformers to build a model that can be used in the future to assess the relevance of credit card services that can be addressed to improve customer happiness by reducing the overall number of complaints.

INTRODUCTION

Over the last few years and through 2019, the credit card market, the largest U.S. consumer lending market measured by the number of users, continued to rise in practically all measures during the last few years and into 2019, until abruptly reversing course in March 2020 (BUREAU OF CONSUMER FINANCIAL PROTECTION, 2021). The Consumer Financial Protection Bureau, a US federal body, collects data on financial institutions to ensure that customers are treated fairly by all businesses (Consumer Financial Protection Bureau, n.d.). It provides a platform for the consumers to report complaints and support related information regarding their concerns about the financial services or products they utilize (2021 Consumer Response Annual Report, 2022).

Consumer Complaint Database is a collection of complaints received by the Consumer Financial Protection Bureau on a wide range of consumer financial products and services provided by financial organizations in the United States (Consumer Complaint Database, 2018). The complaints reported to Bureau increased dramatically in the second quarter of the year 2020 and remained elevated throughout the year (BUREAU OF CONSUMER FINANCIAL PROTECTION, 2021). Analyzing consumer complaints can assist credit-card companies to enhance or update the functional and security features offered based on the feedback or emotions provided by the consumers.

Sentiments play a significant role in gaining an overview of the target customers' expectations regarding individual aspects and issues without perusing a large number of customer remarks at once (Sudhir & Suresh, 2021). Sentiment analysis enables identifying emotional states in human speech or prose (Robinson, 2021). Credit card companies can utilize sentiment analysis to evaluate customer expressions and reviews through emotional states for the products they offer, and by implementing appropriate measures, companies can enhance the market's profitability. Credit card businesses can deploy a machine-learning-based sentiment analysis model to evaluate positive and negative sentiments regarding challenges and concerns about the products they deliver.

Recent advancements have aided data scientists in the development of transformer NLP models capable of understanding text, performing sentimental analysis, answering queries, and summarizing data (Meadows, 2021). The zero-shot model had been implemented to determine and classify the data that has not been previously used to build the model. The Hugging Face transformers package leverages the pre-trained zero-shot model to perform classification (Hugging Face, n.d.). The research conducted in this paper examines the consumer complaint narratives by implementing a sentiment analyzer through the transformer to observe the nature of the complaints registered. The stacked machine learning model by combining classifiers is incorporated with the results obtained from the transformers to predict and enhance the credit-card services provided by various financial firms.

LITERATURE RESEARCH

Few data analysts, scientists, and academicians have conducted thorough research to explore and analyze the publicly available consumer complaint database, resulting in invaluable insights that can help financial institutions reframe their policies and regulations to better serve consumers. Several research studies have documented the study of the CFPB database. The following is a summary of these studies.

Ayres et al. (2013) undertake an early examination of the CPFB consumer complaint database at the institution level and the demographic level using zip codes. Ayres et al. examined the statistics of the data in three distinct ways. Firstly, they observed the total complaints considering each company. Secondly, they looked at the percentage of complaints resolved by each company. Finally, the percentage of responses by each company, that are unable to settle complaints promptly, particularly it was noticed, in the case of mortgages. A regression analysis of demographic data observed a substantial increase in mortgage complaints when the population of specific groups, such as seniors and high school and college students, expanded Ayres et al. (2013). Ayres et al. (2013) did extensive research and developed beneficial consumer complaint insights for the CPFB, financial institutions, and their consumers. The authors, however, considered consumer complaint narratives to be a significant source of data that could be leveraged to provide useful insights to financial institutions and the regulatory organizations that set the rules and regulations that regulate them. Hence, this research inspired us to utilize the consumer complaint narratives to determine the major reasons behind the complaints filed. Instead, we used a zero-shot classifier to analyze the emotions and sentiments of the consumers.

Tom Sabo (2017) uses SAS contextual analytics software to conduct research and explore sentiment analysis on the CFPB database. They determined to extract the actionable content from customer complaint narratives and use it to predict if the complaints resulted in monetary gains. Tom Sabo entails examining the complaints' sentiments and incorporating to simulate the natural language available in each free-form complaint against a disposition code for the complaint, primarily focusing on whether a company paid out money (Tom Sabo, 2017). The aim is to determine whether the complaint resulted in a financial payout from the financial institution. This method could be employed in the retail industry to see if the financial institution that backs them up has a detrimental effect on their brand. The above-mentioned research supports our assumption to assess the narratives by various trends such as location, time of the complaint registered including the greatest concerns for consumers. Also, the machine learning models such as Random Forest, Linear SVM, and Logistic Regression enable to predict of the future concerns of the consumers.

Ryan Bluteau et al. (2021) recommended employing additional embeddings to represent emotion inputs to improve sentiment categorization for transformers (based on BERT and DistilBERT). Using HuggingFace's zero-shot prediction pipeline, the probabilities were generated for whether the emotions apply to a specific sample. After, a zero-shot probability for 1.6 million samples using a sentiment classification dataset was obtained along with the smaller sentiment airline dataset with 63 emotions. Later, new tokens and tokenizers to BERT's embeddings and tokenizers for each predicted emotion represent varying levels of emotion. Finally, depending on the probability of each emotion, the custom token representing that level was prepended to the text input of the model to analyze and train for classification (Ryan Bluteau et al. ,2021). In our study, we will implement a zero-shot prediction pipeline using HuggingFace to classify the emotions and sentiments of the consumers about the complaints registered.

Akhtar et al. in a research paper suggested a stacked ensemble technique for sentiment intensity prediction in the financial domain (Akhtar et al., 2020). Boosting, bagging, voting (weighted, the majority), and other traditional ensemble-building techniques are used. The authors presented an ensemble technique based on Particle Swarm Optimization to solving the challenge of aspect-based sentiment analysis (PSO). They developed three deep learning models using LSTM, CNN, and GRU, as well as a feature-driven classical supervised model using SVR (Akhtar et al., 2020). To combine them, a Multilayer Perceptron (MLP) classifier is used. We used a similar method to aggregate the outputs of three models (KNN, Random Forest, and Gaussian) using a meta-classifier (In our case Logistic regression for classification).

PROBLEM STATEMENT

Based on the literature review we will address the following questions.

- Which states and credit-card companies are receiving the highest number of complaints?
- What are the common sub-issues related to credit-card products that are customers dissatisfied with?
- How does the consumer feel about the credit-card services offered considering the emotions and intent?
- What are the major concerns related to credit-card services that affect the emotions and sentiments of the consumers?

DATASET DESCRIPTION

Dataset Source: Consumer Financial Protection Bureau (Consumer Complaint Database)

Dataset Files: ‘Complaints.csv’

The dataset file ‘Complaint.csv’ has been taken from Consumer Complaints Database. The dataset used for analysis is from July 2017 to January 2022. Consumer Complaints Database provides the data that has been gathered by Consumer Financial Protection Bureau from the complaints filed against the financial services such as credit cards, mortgage loans, student loans, credit reporting, and debt collection. In the dataset, there are in total 23 different products mentioned which are offered by various financial firms about which the complaints are registered. Here, we will be focusing on the issues related to Credit-card to perform complaint analysis.

There are multiple sub-products provided under the credit-card category that are general-purpose credit cards or charge cards, gift cards, payroll cards, government benefit cards, and Student prepaid cards (Figure 1). There are a total of 65981 unique complaints filed by consumers for different sub-products under credit-card. ‘Complaint.csv’ provides in-depth information related to the complaints such as date received, state, company, zip code, time, location, and purpose of complaint registered across the United States (Table 1). It also consists of columns that state the information about the platform used to register the issue.

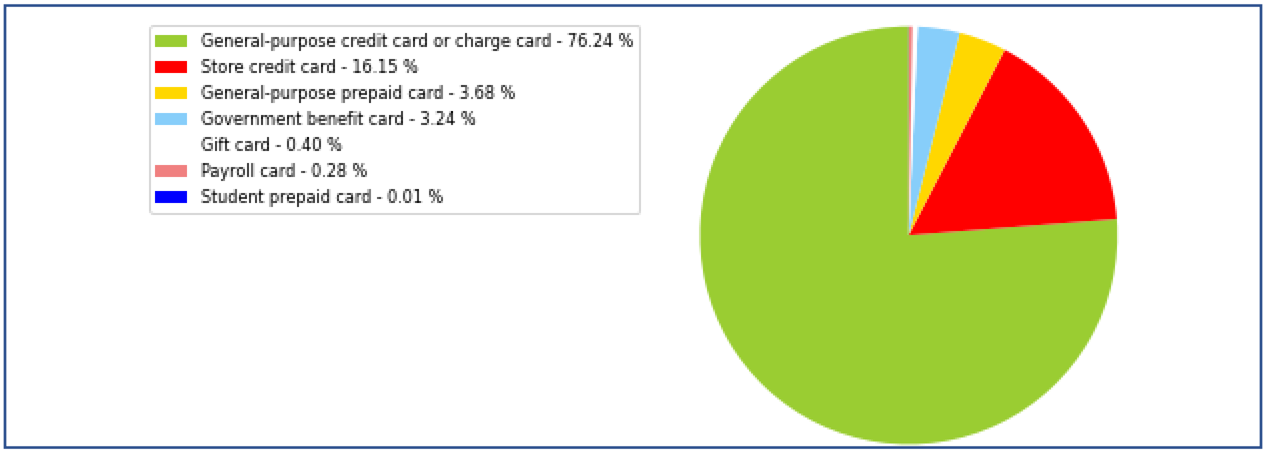

Figure 1. Pie-chart depicting the percentage of complaints associated with each sub-product.

Column Names	Description
Date received	The date at which the complaint was registered
Product	Product on which consumer registered a complaint
Sub-product	Sub-product of the product mentioned
Issue	the reason behind complaint registration
Sub-Issue	Briefing the issue registered
Consumer complaint narrative	A detailed description of the complaint
Company	Company for whom the complaint was registered
State	State in which the company is present
Date sent to the company	The date at which the complaint was received by the company
Company response to consumer	The date at which the company responded to the complaint
Timely response?	Whether the response from the company was on time or not
Complaint ID	Unique identification is given for each complaint

TABLE 1: COLUMN DESCRIPTION FOR DATASET

METHODOLOGY

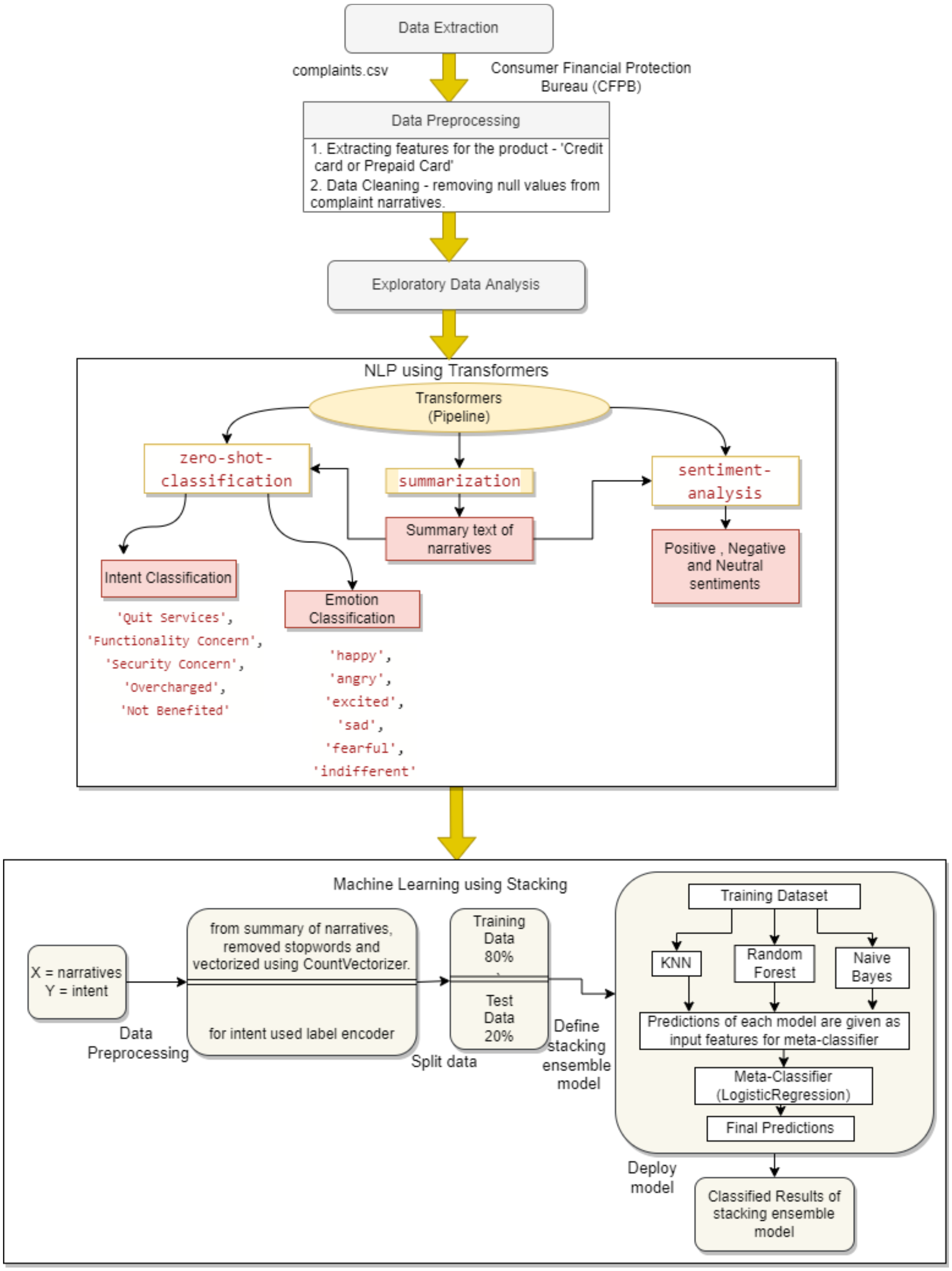

Figure 2. Workflow for the analysis of the credit-card complaints.

In this section, we will get across the process for performing data analysis on consumer complaints. Data Extraction, Data pre-processing, and dimensionality reduction are the primary steps to be demonstrated on the dataset before proceeding further with EDA (Figure 2). The goal is to analyze the complaints regarding the issues among the services offered by credit-card companies. The credit card is considered the product to demonstrate the analysis of consumer complaints. Hence, we further identified the key columns that seem to be significant while performing a detailed analysis of the consumer complaints after determining the correlation. Among all the phases in the framework, dealing with missing values was the most fundamental step (Figure 2). Dropping the null values is the only viable approach to handle the missing values, as the data consists of narratives that cannot be replaced by any other values.

Later, we attempted to comprehend the data through various visualizations that are addressed in the results and findings section. The plots such as the total number of complaints received by each credit card company and state (Figure 4), the most common issues that our customers are unhappy about, timely response to the complaints received (Figure 6), and the most prominent sub-product based on the number of issues were analyzed in detail (Figure 7).

Advancing to the next stage in the workflow, implementing transformers plays a critical role in recognizing and categorizing consumer satisfaction, concern, and feedback regarding various financial institution services (Figure 2). Before the implementation of the transformer, Vader SentimentIntensityAnalyzer was used for sentiment analysis, however, it presented us with skewed results by displaying more positive narratives than negative, which is not the case with complaints. The goal was therefore to improve sentiment predictions for which transformers outperformed significantly in categorizing the positive and negative sentiments.

Considering the length of the narratives (description of the complaint), a text summarization transformer is employed on the narratives to summarize them accordingly. Further, summarized narratives were provided as input to the sentiment analyzer transformer to categorize them into positive and negative sentiments. Most of the narratives were recognized under the negative sentiment category. Later, the Zero-shot classification model was utilized to determine the intent and emotions for narratives rather than a general sentiment. Functional issues, overcharged, security worries, not benefited, and quit services are considered as intent labels to evaluate narratives. The most frequent cause of complaints noticed is due to the lack of functionality and security services offered to the consumers (Figure 10). Emotions such as happiness, anger, joy, pleased, fear, sad, and indifference are used to assess the narratives under the emotion labels.

Moving to the final stage of the workflow, the machine learning model is implemented to predict the consumer's intent to improve the functionality issues and security concerns, respectively. In the Data cleaning stage, Stop words and punctuation were removed from the narratives using the NLTK (Natural Language Toolkit). TF-IDF vectorizer is used to vectorize the narratives so that they can be provided as an input to the machine learning model. The categories of Intent classification are considered as a target variable to classify the intents for which label encoding is used to encode the intent labels that are functional issues, overcharged, security worries, not benefited, and quit services as 0, 1, 2, 3, and 4, respectively. The data is split into 75% and 25% for training and testing purposes. The stacking model is built using KNN, RandomForestClassifier, and Naive Bayes also known as base models (Level-0). Further, the predictions of all the models are passed to the meta-model (Level-1) for which Logistic Regression is trained on the predictions provided by the base models to classify the intent of the summarized narratives.

RESULTS AND FINDINGS

The correlation plot below depicts the positive and negative correlation of the columns (Table 2) mentioned in the plot concerning consumer complaints narratives (Figure 3). Complaint-Id, Issues, sub-products, and company are positively correlated which means that as the number of businesses, issues, and sub-products grow, similarly the frequency of consumer complaints will also increase. Timely response on the other hand is inversely proportional which implies that if the complaints will be resolved on time, the number of complaints will start decreasing.

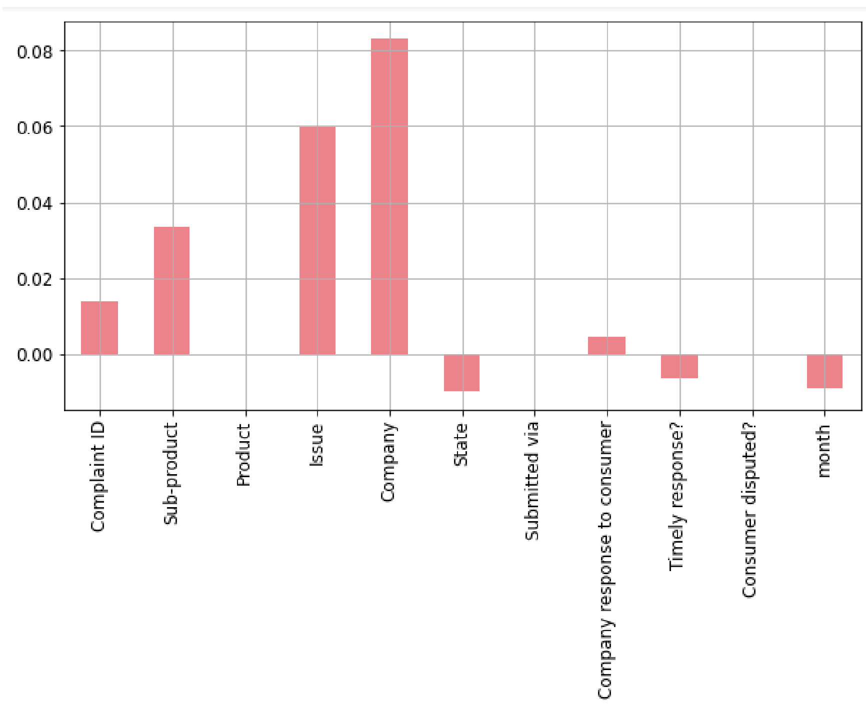

Figure 3. Correlation Plot

Column Name	Description
Complaint ID	Unique ID associated with each complaint
Sub-product	Sub-product of the product 'Credit card or prepaid card'
Product	Product - Credit card or Prepaid card
Issue	The issue behind the consumer complaints
Company	Name of the company for which the complaint was received
State	State in which the company is established

Submitted via	Whether the complaints are submitted via web, phone, referral, etc
Company response to consumer	Description of the response given by the company to the complaints received
Timely response?	Yes/No, whether the response from the company was within the time frame
Consumer disputed?	Yes/No, whether the consumer was disputed
Month	Month in which the complaint was received

TABLE 2: Column description for the correlation plot

We evaluated the data in three distinct ways during the exploratory data analysis. We began by looking for the state and firm that received the most complaints. California had the highest number of complaints, according to the state-by-state complaint breakdown (Figure 5). We discovered that Citibank received the most complaints, followed by Capital One Financial Corporation and JPMorgan Chase & Co. after visualizing the number of complaints against the top-ranked corporations (Figure 4).

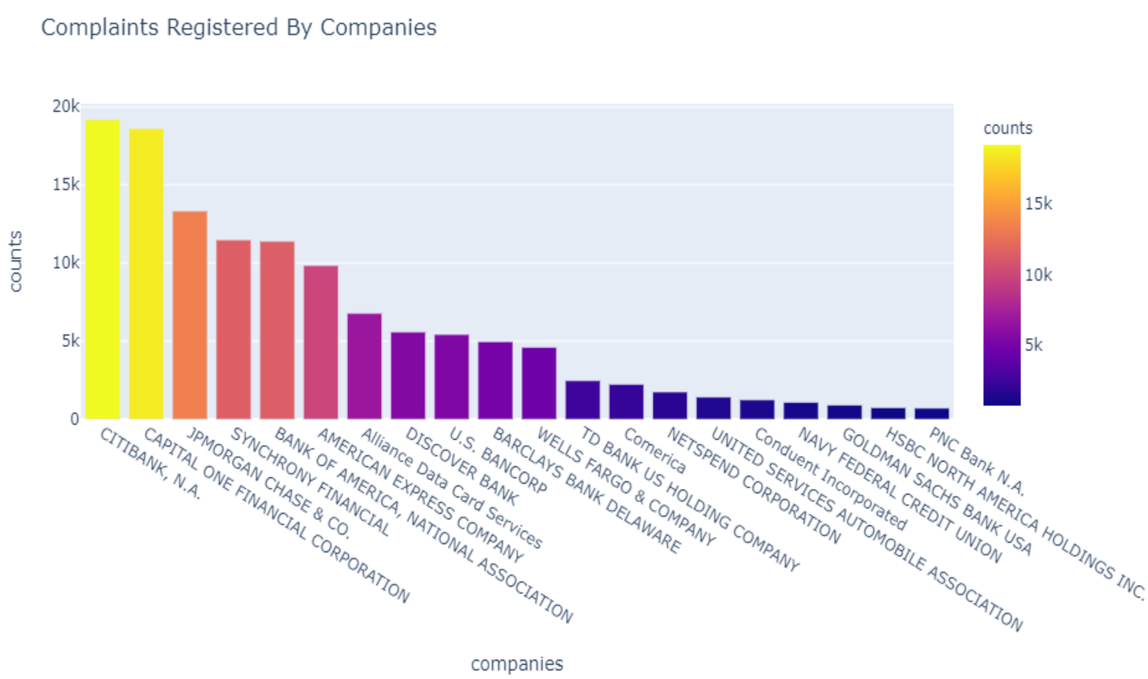

Figure 4. Distribution of Complaints by Companies

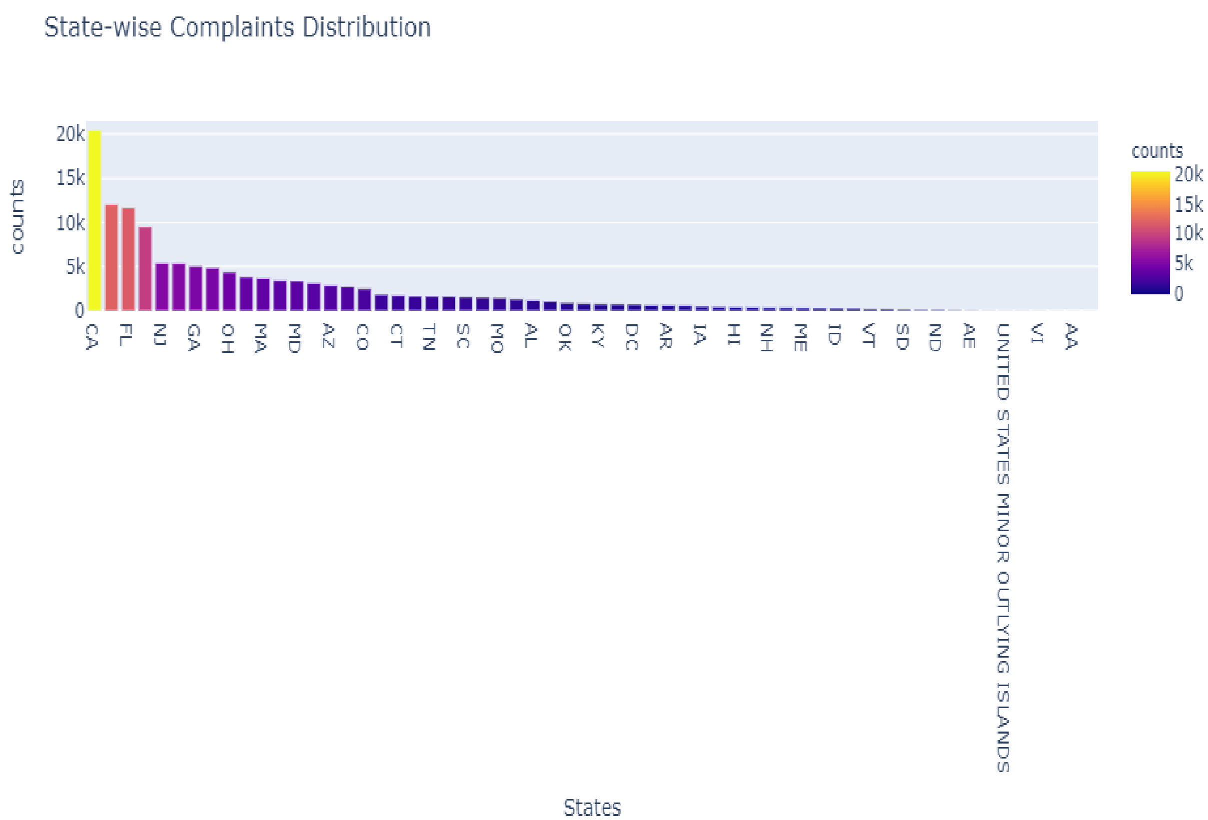

Figure 5. Distribution of Complaints by State

Analyzing the frequent issues that the customers are dissatisfied with credit-card services is critical since it reflects customer satisfaction with credit-card products such as student prepaid cards, store credit cards, gift cards, and payroll cards (Figure 6). The overall number of complaints filed against the issues is a significant source of information for credit-card issuers to enhance the functionality and security services offered to the consumers. Topmost Issues such as a problem with a purchase shown on your statement, fees or interests, problems while making payments, getting a credit card, and closing an account were observed (Figure 6).

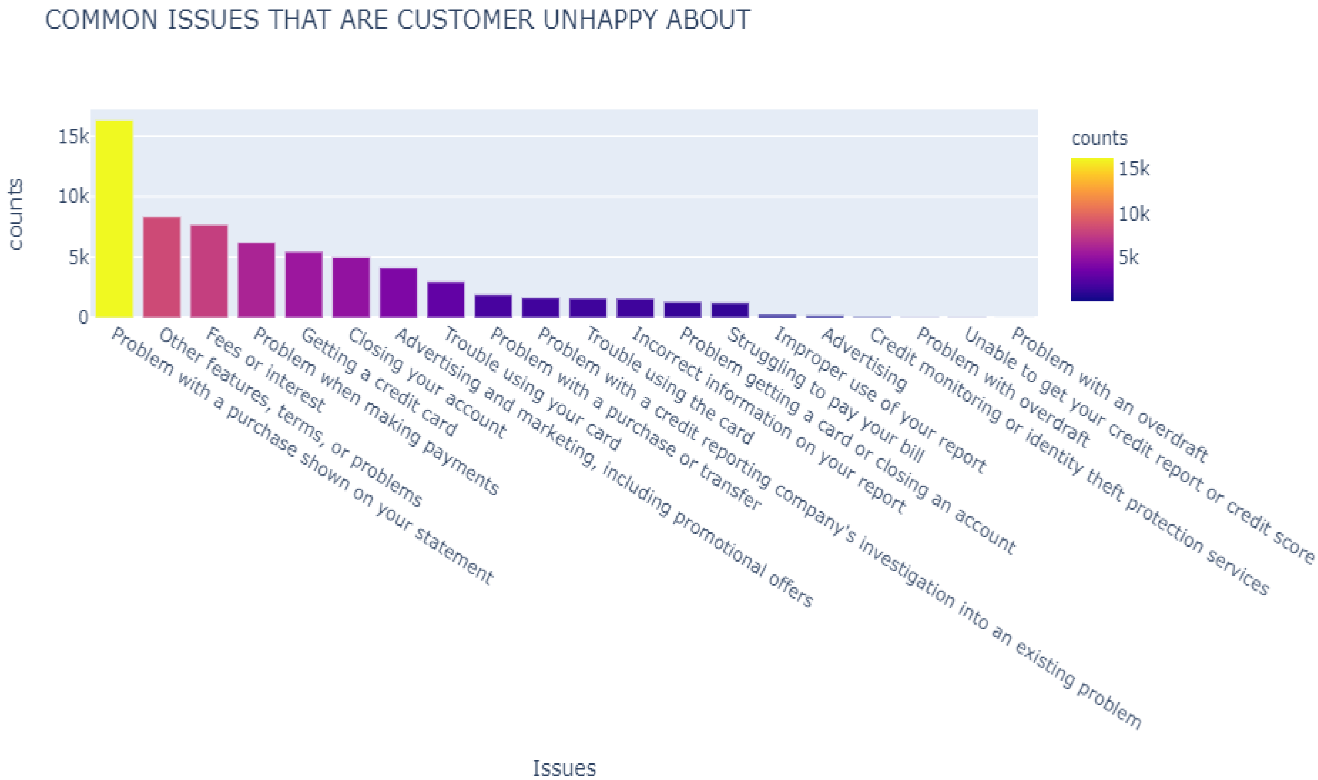

Figure 6. Common Issues that are consumers unhappy about

The time series plot of the complaints was the second key element of our analysis (Figure 7). It was determined that after January 2020, the number of complaints skyrocketed. The cause of such a big surge in complaints could be the pandemic situation in which people were having multiple difficulties managing their finances (BUREAU OF CONSUMER FINANCIAL PROTECTION, 2021) ,(Figure 7). While dealing with the financial difficulties, consumers were overcharged and faced difficulties while doing a transaction.

Figure 7. Time Series Analysis by month of Consumer Complaints

The below visualizations represented in the box plots demonstrates the sentiments of consumer complaints narratives categorized into positive, negative, and neutral regarding the sub-products using the Vader Sentiment Intensity Analyzer (Figure 8). The sentiments for the sub-products were distorted as a result of the analysis. When it comes to general-purpose credit cards, retail credit cards, and government benefit cards, positive sentiments outnumber negative sentiments, which is not the case when focusing on the complaint narratives (Figure 9).

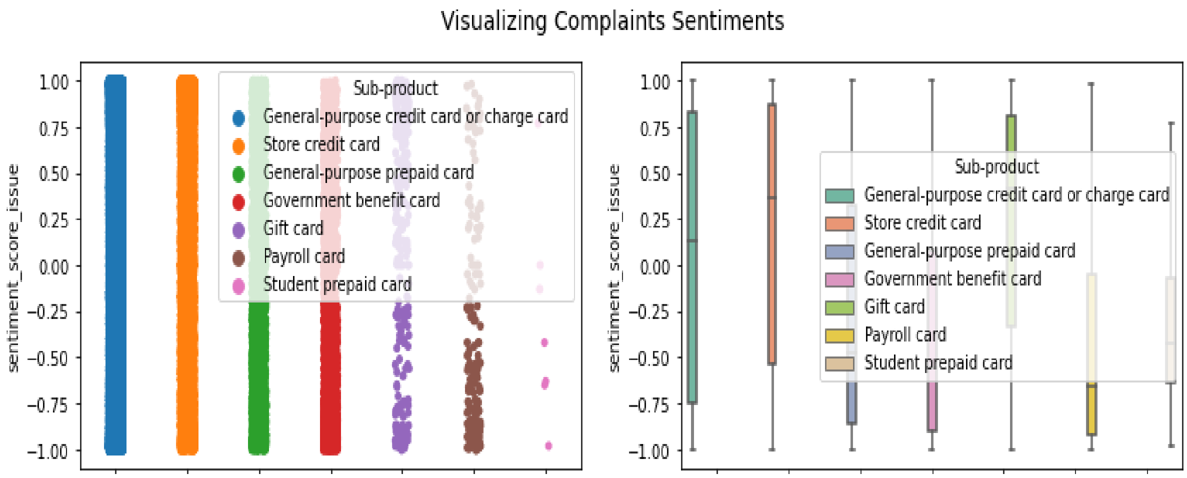

Figure 8. Box plot of sentiments for sub-products

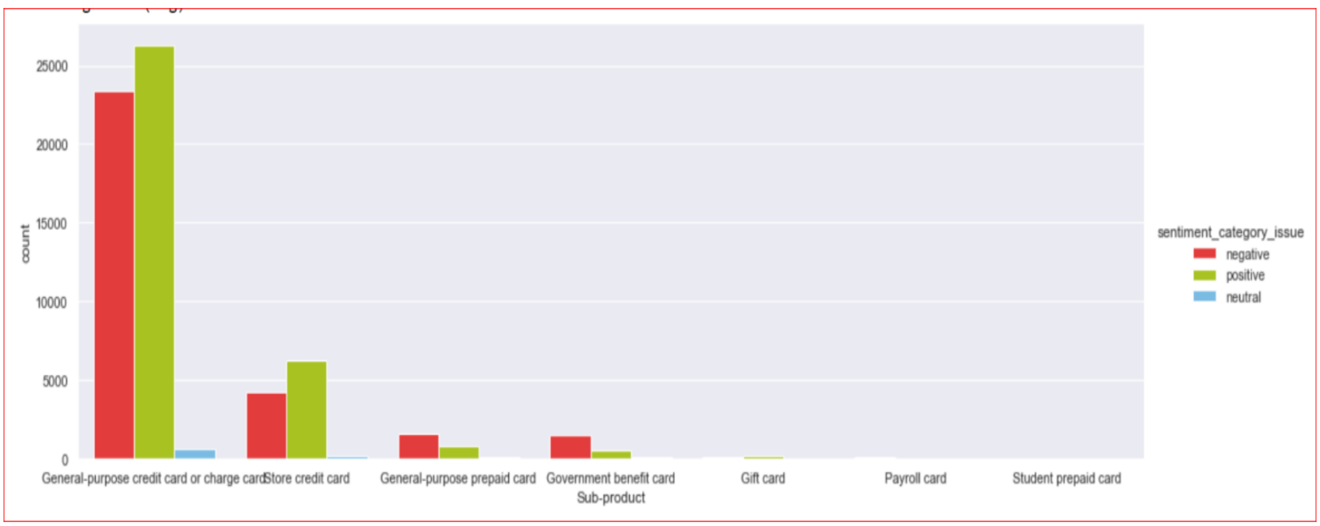

Figure 9. Comparison among the sentiments for sub-product

Later, we used an alternative approach of implementing transformers to classify and analyze the sentiments, intents, and emotions of the consumers about the complaints registered. Sentiment Analyzer transformer by Hugging Face outperformed the results of Vader Sentiment Analyzer by recognizing the overall complaints as negative with 99.78%. Further, we analyzed the intent classification into five distinct classes that are functionality concern, overcharged, security concern, Not benefited, and Quit Services. The pie chart represents the overall percentage concerning the complaint narratives (Figure 10). The percentage for the Functional Concerns, Overcharged and security concerns were 31%, 37.3%, and 22% respectively. As they were the top-ranked intent among all other intent categories.

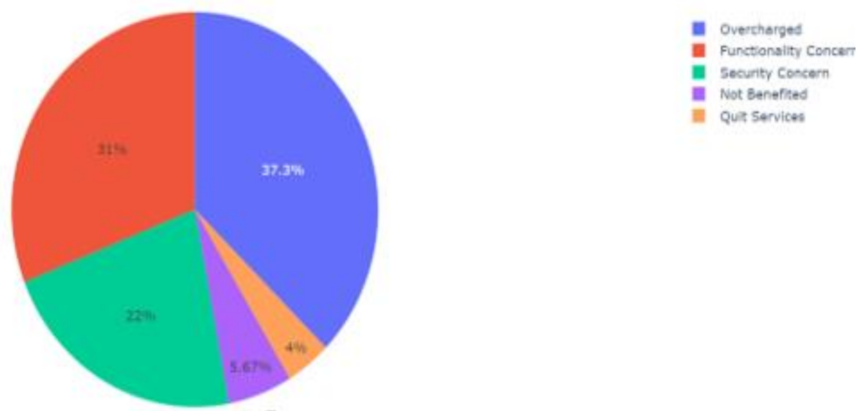

Figure 10. Pie-chart [Distribution by categories of intent]

After observing the emotions of the consumers regarding the sub-product, the majority of the consumers that is 74.6% were fearful and angry about the sub-products provided by the various credit-card institutions (Figure 11). 11.3% of the consumers were sad whereas 10.3% of the consumers were excited which is very low compared to the percentage of fearful and angry (Figure 11).

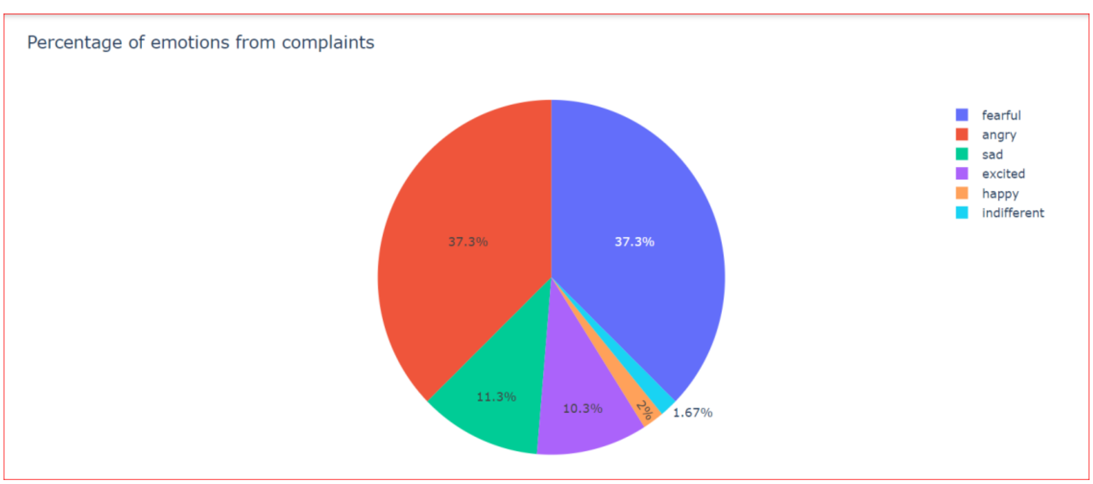

Figure 11. Pie-chart [Distribution by categories of emotions]

The most significant or frequent words in consumer complaints have been projected in a word cloud (Figure 12). Most of the words, such as security, internet, credit, misleading, review, and cancel, have negative connotations.

Figure 12. Word Cloud of Summarized Consumer Complaint Narratives

The emotion categories like afraid, angry, sad, thrilled, happy, and neutral were also analyzed among individual intent categories using a stacked bar plot (Figure 13). The underlying analysis is reflected in the below plot (Figure 13): 1) Considering the overcharged intent, consumers were majorly fearful and angry. 2) For the not-benefited category none of the consumers were excited implying that no consumers were also happy because they were not benefited. 3) Security concerns are the most essential intent to examine as it associated with the angry, fearful, and sad emotions of the consumers.

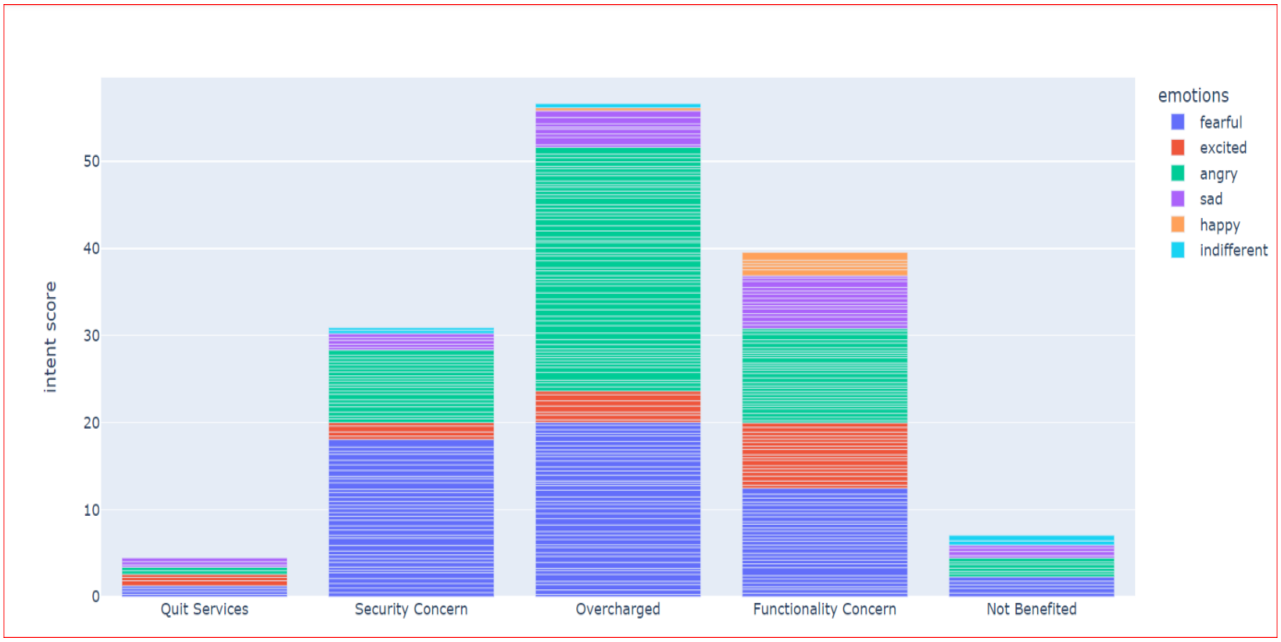

Figure 13. Stacked Bar Plot of Emotion Categories for individual intent

Predicting the categories of intents related to consumers is the last part of Exploratory Data Analysis. Using a stacked machine learning model built by KNN, Random Forest, and Naïve Bayes, below are the accuracy results obtained (Figure 14). KNN represents 69.85% of accuracy while Random Forest and Naïve Bayes outperformed with 75.42% and 98.85% of accuracy, respectively (Figure 14). Meta Classifier after obtaining the predicted results from the base models provides highly accurate results while classifying the intent categories by 94.28 %.

```
Accuracy from KNN (base model)           : 69.85714285714286
Accuracy from Random Forest (base model) : 75.42857142857143
Accuracy from Naive Bayes (base model)   : 98.85714285714286

Accuracy from Stacking (Meta Classifier) : 94.28571428571429
```

Figure 14. Classification Accuracy of Stacking models

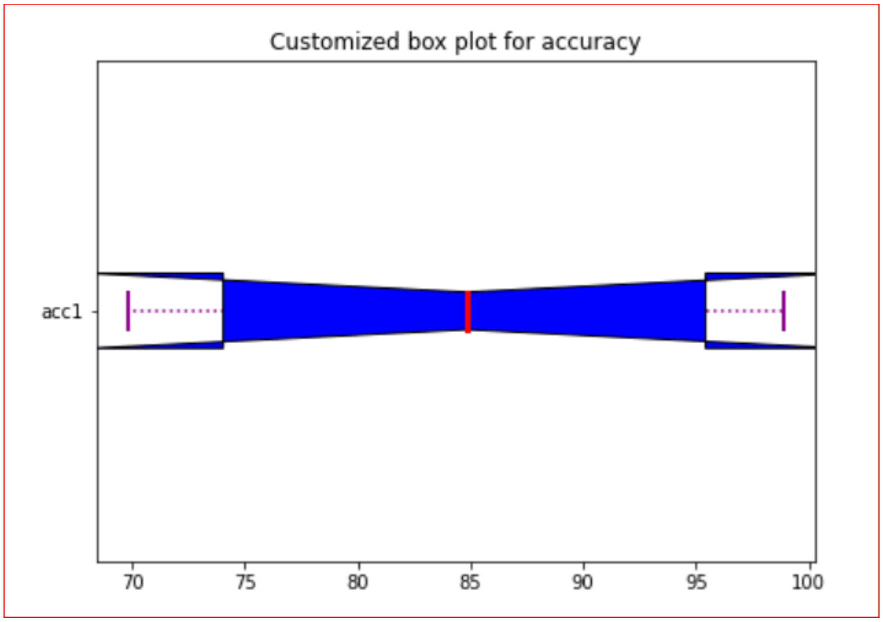

Figure 15. Box Plot for Classification Report of Logistic Regression Model

CONCLUSION

The goal of the research is to decrease the overall number of complaints registered by enhancing the features and services of various credit cards offered to the consumers. Consumer complaints narratives are a valuable source of information to extract insightful analysis of emotions about the consumers related to the various products under credit-card. Transformers are demonstrated to study the sentiments, intents, and emotions about the consumer complaints analysis that can be useful to classify and predict the nature of the consumer complaints. Results obtained after implementing the transformers incorporated with the stacked model(machine learning model with base and meta-models) display that the consumers are more fearful considering the security concerns. When it comes to functionality concerns and overcharged, consumers feel angrier about the credit-card services which impact the total counts of complaints being registered.

ACKNOWLEDGEMENT

We would like to thank our professor Dr. Tony Diana for providing us with the dataset available as a consumer complaints dataset. We appreciate the methodology suggested for implementing transformers that helped us in leveraging the results of overall sentiments about the consumer complaints narratives. We are extremely grateful to Dr. Tony Diana for providing us with his relentless encouragement, assistance, and guidance, which was invaluable in the successful completion of this project and paper.

REFERENCES:

BUREAU OF CONSUMER FINANCIAL PROTECTION. (2021, September). *The Consumer Credit Card Market*. https://files.consumerfinance.gov/f/documents/cfpb_consumer-credit-card-market-report_2021.pdf

Consumer Financial Protection Bureau. (n.d.). *Credit cards*. Retrieved April 30, 2022, from https://www.consumerfinance.gov/consumer-tools/credit-cards/

2021 Consumer Response annual report. (2022, March 31). Consumer Financial Protection Bureau. https://www.consumerfinance.gov/data-research/research-reports/2021-consumer-response-annual-report/

Consumer Complaint Database. (2018). Consumer Financial Protection Bureau. https://www.consumerfinance.gov/data-research/consumer-complaints/

Sudhir, P., & Suresh, V. D. (2021). Comparative study of various approaches, applications, and classifiers for sentiment analysis. *Global Transitions Proceedings, 2*(2), 205–211. https://doi.org/10.1016/j.gltp.2021.08.004

Robinson, S. (2021, April 12). *Sentiment analysis: Why it's necessary and how it improves CX.* SearchCustomerExperience. https://www.techtarget.com/searchcustomerexperience/tip/Sentiment-analysis-Why-its-necessary-and-how-it-improves-CX

Meadows, Y. (2021, November 3). *AI Customer Service: Employing Transformer NLP Models*. Employing Transformer NLP Models. https://ymeadows.com/en-articles/ai-in-business-employing-transformer-nlp-models-in-customer-service

Hugging Face. (n.d.). *Hugging Face – The AI community building the future.* The AI Community Building the Future. https://huggingface.co/

Ayres, I., Lingwall, J., & Steinway, S. (2013). Skeletons in the Database: An Early Analysis of the CFPB's Consumer Complaints. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2295157

Sabo, T. (2017). *Applying Text Analytics and Machine Learning to Assess Consumer Financial Complaints*. SAS Institute Inc. https://support.sas.com/resources/papers/proceedings17/SAS0282-2017.pdf

IMPROVING SENTIMENT CLASSIFICATION USING 0- SHOT GENERATED LABELS FOR CUSTOM TRANSFORMER EMBEDDINGS. (2016). *IMPROVING SENTIMENT CLASSIFICATION USING 0- SHOT GENERATED LABELS FOR CUSTOM TRANSFORMER EMBEDDINGS*. ICLR 2022.

Akhtar, M. S., Ekbal, A., & Cambria, E. (2020). How Intense Are You? Predicting Intensities of Emotions and Sentiments using Stacked Ensemble [Application Notes]. *IEEE Computational Intelligence Magazine*, *15*(1), 64–75. https://doi.org/10.1109/mci.2019.2954667