Improving Credit Card Services

1. **Problem Statement and Business Case**

Consumer complaint resolutions are important for business. This research aims at reviewing and analyzing records of credit card relevant complaints, and providing suggestions for financial institutions on how could they improve their services, minimize complaints, and improve consumer experiences. Through the research, the following questions are expected to be answered:

* What are the patterns of these complaints?
* Is there any improvement in some areas?
* Is there any emerging trend on complaints?
* How did the financial institutions respond to the complaints?
* What are the areas the financial institutions should focus on?
* Recommendations on how financial institutions could improve?
* Present consumer complaints by year with in a financial institution?
* What is the response time for each financial institution?
* Forecasting complaints trend within a financial institution?

1. **Data collection and resources**

The Consumer Complaint Database of Customer Financial Protection Bureau (CFPB) was introduced in July 2012, underscoring CFPB’s intent to fulfill two core objectives: enforcing federal consumer protection laws more vigorously and analyzing consumers, financial services providers and market activities[[1]](#footnote-1). The Consumer Complaint Database is a collection of complaints about consumer financial products and services that CFPB sent to companies for response. Complaints are published after the company responds, confirming a commercial relationship with the consumer, or after 15 days, whichever comes first. Complaints referred to other regulators, such as complaints about depository institutions with less than $10 billion in assets, are not published in the Consumer Complaint Database. The database generally updates daily and is open for public access.

Currently there are about 2 million records/entries in the database, ranging from December 1, 2011 to current date. The data could be retrieved online from: <https://catalog.data.gov/dataset/consumer-complaint-database>.

This project focuses on the sector of credit card. More than 8% of the complaints in the database are relevant to general purpose credit card issued by financial institutions. In addition, the Consumer Financial Protection Bureau (CFPB) received approximately 154,500 credit or consumer reporting complaints in 2019, a 23% increase from 2018. Complaints with credit card accounts 8% of the total, amount at 29,900, a 4% increase from 2018.

1. **Research procedures**

The research will start with retrieving data from online and extracting entries that are relevant to credit card. Cleaning the dataset requires identifying outliers and handling missing data. In this process we will also standardize the timestamp since it is a time series dataset. Other techniques include proper indexing. Outliers will be neglected and missing data will be handled accordingly.

Then exploratory data analysis will be deployed to derive patterns and trends from the dataset. This section focuses on not only the complaints itself but also how financial institutions resolve the complaints, how long it takes to handle a complaint, and whether the consumer disputed.

Lastly, machine learning models will be built to further analyze the complaints and advanced analytic skills will be applied to identify more subtle patterns presented by the dataset.

1. **Methodologies**

As the dataset is mostly text, Natural Language Process algorithm will be adopted throughout the process. In addition, the following methodologies will be applied:

* 1. Exploratory data analysis

EDA studies the distribution of complaints by region, company, issue, etc. as well as explores what the complaints looked like and the information stored in those complaints. Understanding the underlying themes in the data, we will use visualization tools such as bar plots, pie charts, maps and word clouds to illustrate the patterns.

* 1. Classification

While filing a complaint, customers may be unaware of business terms so they often choose the wrong issue. This misclassification may lead to long wait times and incomplete resolution of customer complaints. We will leverage topic algorithms such as TF-IDF and LDA to reclassify customer complaints based on the exact language used in the complaint.

Different models based on Logistic Regression, Decision Tree, Random Forest will be run to find the best model. Through classification, it is expected to find out the correct issue relevant to the complaints, automatically route them to the right team, and analyze them for immediately actionable insights.

* 1. Predictive models

After financial companies responded to a complaint, the consumer may feel not satisfied therefore they dispute. The aim of building predictive models are to investigate the dataset and make a prediction whether or not a consumer would likely dispute. If we are able to predict this, consumer who is more likely to dispute a conclusion can be given more attention as to how the complaints are handled as well as how persuasively the final conclusions are conveyed to them.

This work will be based on the result of classification and the prediction will reduces in time (and money) spent analyzing thousands of complaints filed on an ongoing basis with the CFPB.

1. Consumer Financial Protection Bureau. “About us/Core functions.” Retrieved from: <http://www.consumerfinance.gov/the-bureau/>. [↑](#footnote-ref-1)