CONSUMER COMPLAINTS CLASSIFICATION

CS 410/510: Natural Language Processing (NLP), Fall 2023

Team: Rutuja Padgilwar, Jithendra Bojedla, Dileep Kumar Boyapati

Project Overview

Motivation

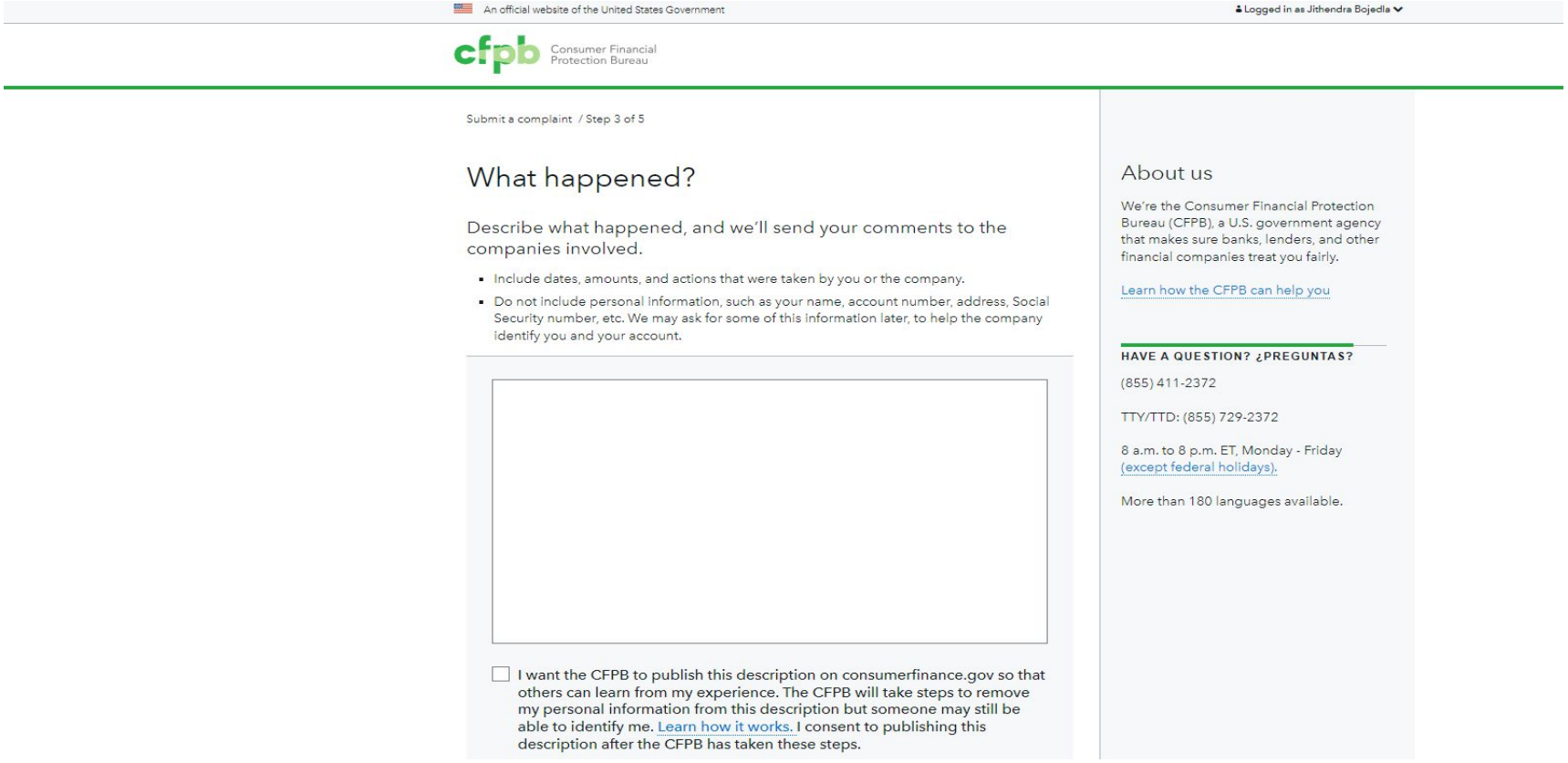
- Consumer complaints classification directly addresses real-world issues that individuals face when dealing with financial products and services.

- Demonstrating a commitment to understand and address customer concerns reinforces a customer-centric approach.

About the Data

- The data is submitted by consumers to the **Consumer Financial Protection Bureau**, a "U.S. government agency that makes sure banks, lenders, and other financial companies treat you fairly."
- Consumers can submit a **narrative** of their complaint through CFPB, an official website of the United States Government.
- Consumers are prompted to classify their complaint in four categories:
    - product
    - sub-product
    - issue
    - Sub-issue
- https://www.consumerfinance.gov/data-research/consumer-complaints

The portal of CFPB where the consumers file their complaints as a narrative



Sample Narrative

Hello my name is XXXX XXXX, I am being scammed by Comenity bank a credit card provider for companies The Children's place, New York & Co. , Forever 21 and Victoria Secret. My original credit from XXXX was {$500.00} Comenity bank then lowers my limit to {$300.00} and began to charge overage fees along with late fees. I then began to pay close attention to my other cards to find that my limits were also changed on them as well incurring overages and late fees.

I reached out to the company Comenity bank they stated that they would change my credit limit to its original limits but did not. I reached out to them again and told them I will not summit any payment until my accounts are corrected. Comenity bank credit cards has impacted my credit scores plummeted to a negative status.

I'm currently paying the price due to the corruption in which has affected me in detrimental way. I am now in debt over {$2000.00} due to the company charging overage fees as well as late fees even through COVID-19. The initial credit limits are fluctuating tremendously and the company charges major fees on each of those 4 accounts. They are not willing to correct my account and was nervous when I said I had an attorney, that is the reason I'm reaching out to you.

Purpose of Project

- Each week the Consumer Financial Protection Bureau sends thousands of consumer complaints about financial products and services to companies for a response.
- It's very tedious task to classify complaints manually.

Our Solution:

- Create a NLP model to classify all the complaints more effectively and in less time.

Business Case

- An NLP model will make the classification of complaints and their addressing to the appropriate teams more efficiently than manually tagged complaints.

EXPLORATORY DATA ANALYSIS

Data Overview

01

Decade of Consumer Complaints: 4.2M Submissions (2011-2023)

02

Narrative-Filtered Complaints (Removed Empty): 1.53M Remaining

03

Removed Duplicate Narratives: 1.29M Complaints

04

Consolidated into Five Product Categories: Complaints Initially Tagged in 21 Categories

Class Consolidation

Consolidated the products into five classes:

1. **Credit Reporting**
2. **Debt Collection**
3. **Mortgages and loans**
4. **Credit Cards**
5. **Retail Banking**

Workflow

Complaints Dataset

Pre-processing

1. Class Consolidation
2. Class Imbalances
3.Tokenization
4. Remove stop words
5. Vectorization

Train Data

Test Data

Build Baseline models

1. Naïve Bayes
2.Random Forest
3.Decision Tree
4.Gradient Boost
5.XGBoost

PLM

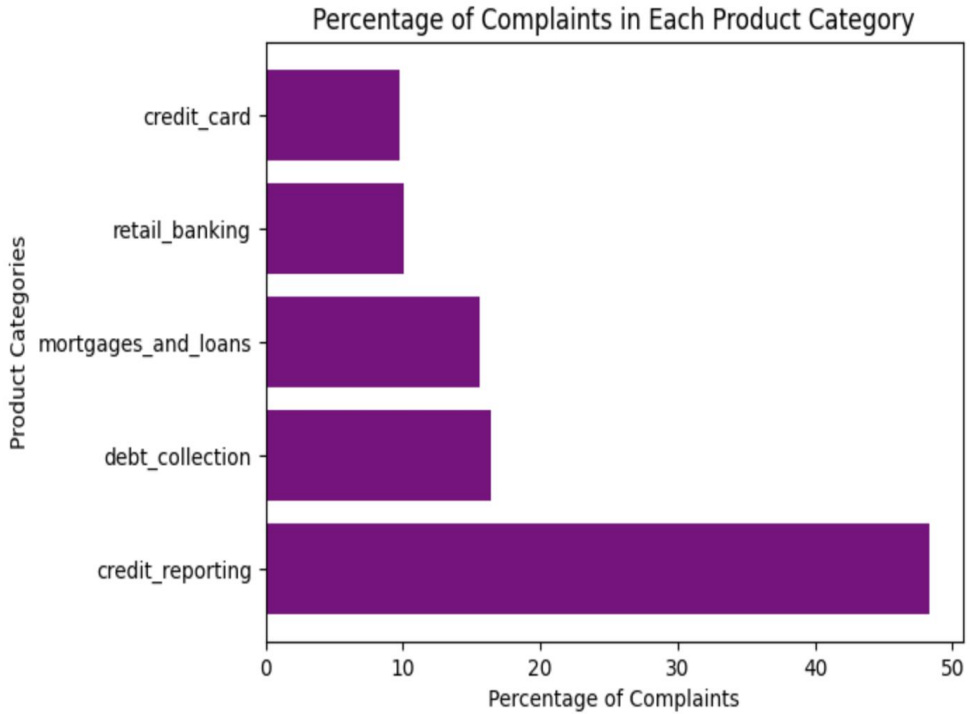
BERT

Classification of product category

1. Credit report
2.Mortgages and loans
3.Debit Collection
4.Retail Banking
5.Credit card

Performance Evaluation Metrics
Accuracy, Precision, Recall and F1-Score

Class Imbalances




```
df['product_category'].value_counts()
```

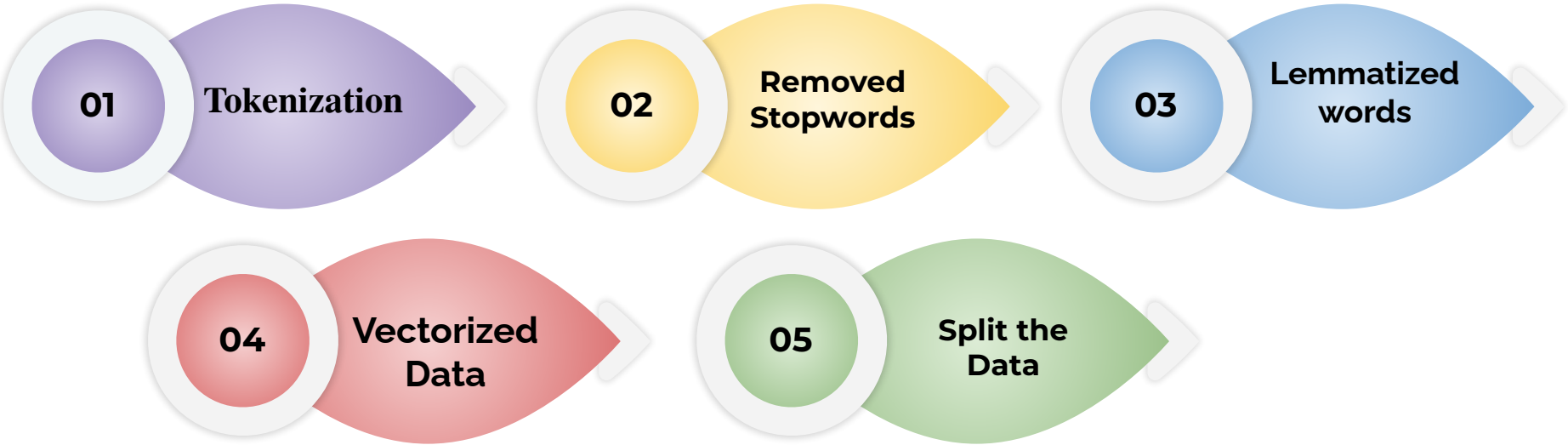
```
credit_reporting       2468167
mortgages_and_loans     581480
debt_collection         512306
retail_banking          337929
credit_card             307866
other_services            1058
Name: product_category, dtype: int64
```

After Balancing: (600k data)

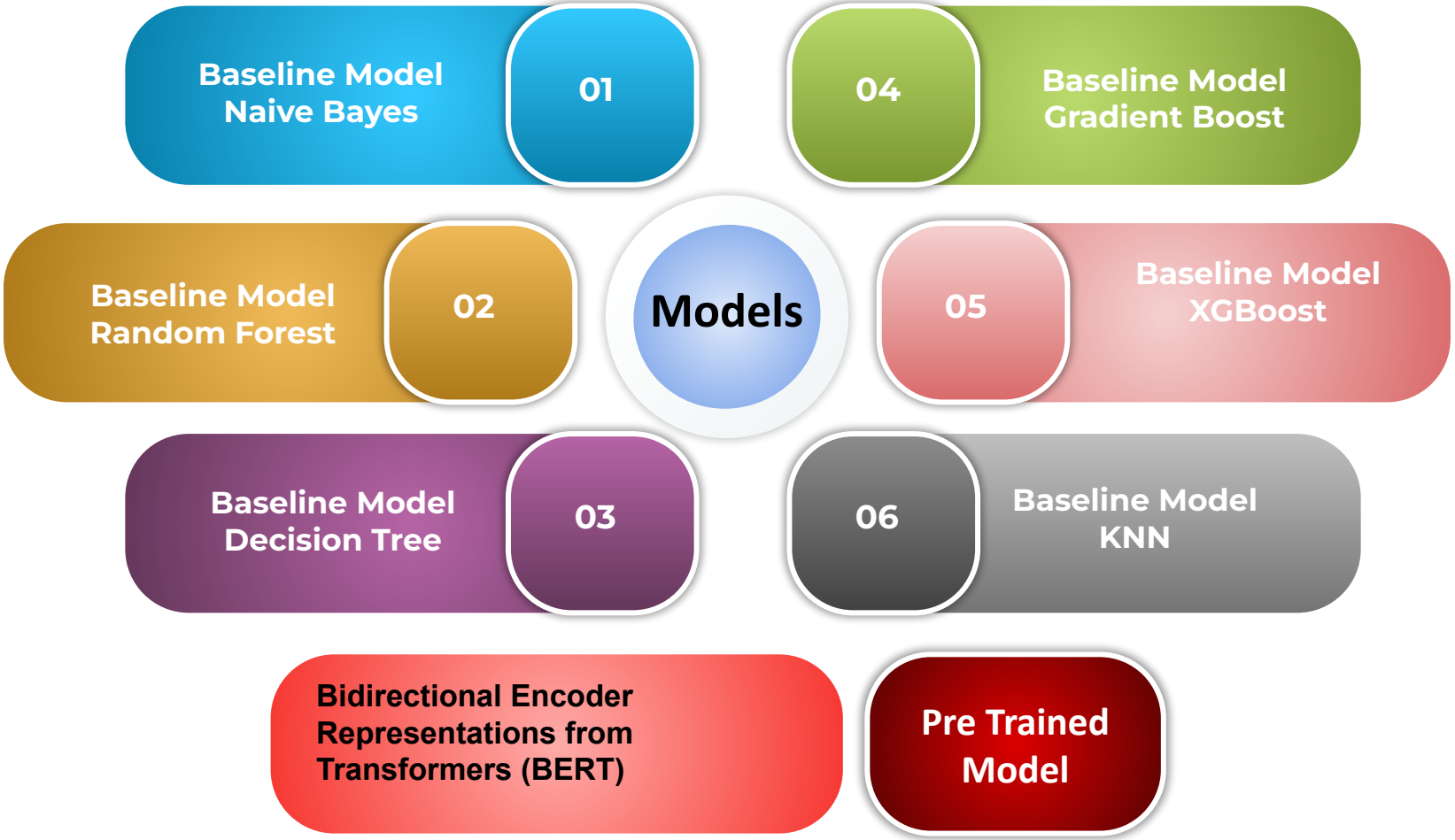
```
credit_card            125109
credit_reporting       125109
debt_collection        125109
mortgages_and_loans    125109
retail_banking         125109
Name: product_category, dtype: int64
```

To address this, we implemented a data balancing technique to ensure that each product category has an equal representation in the dataset.

Data Preprocessing



Model Training:



Modeling Process

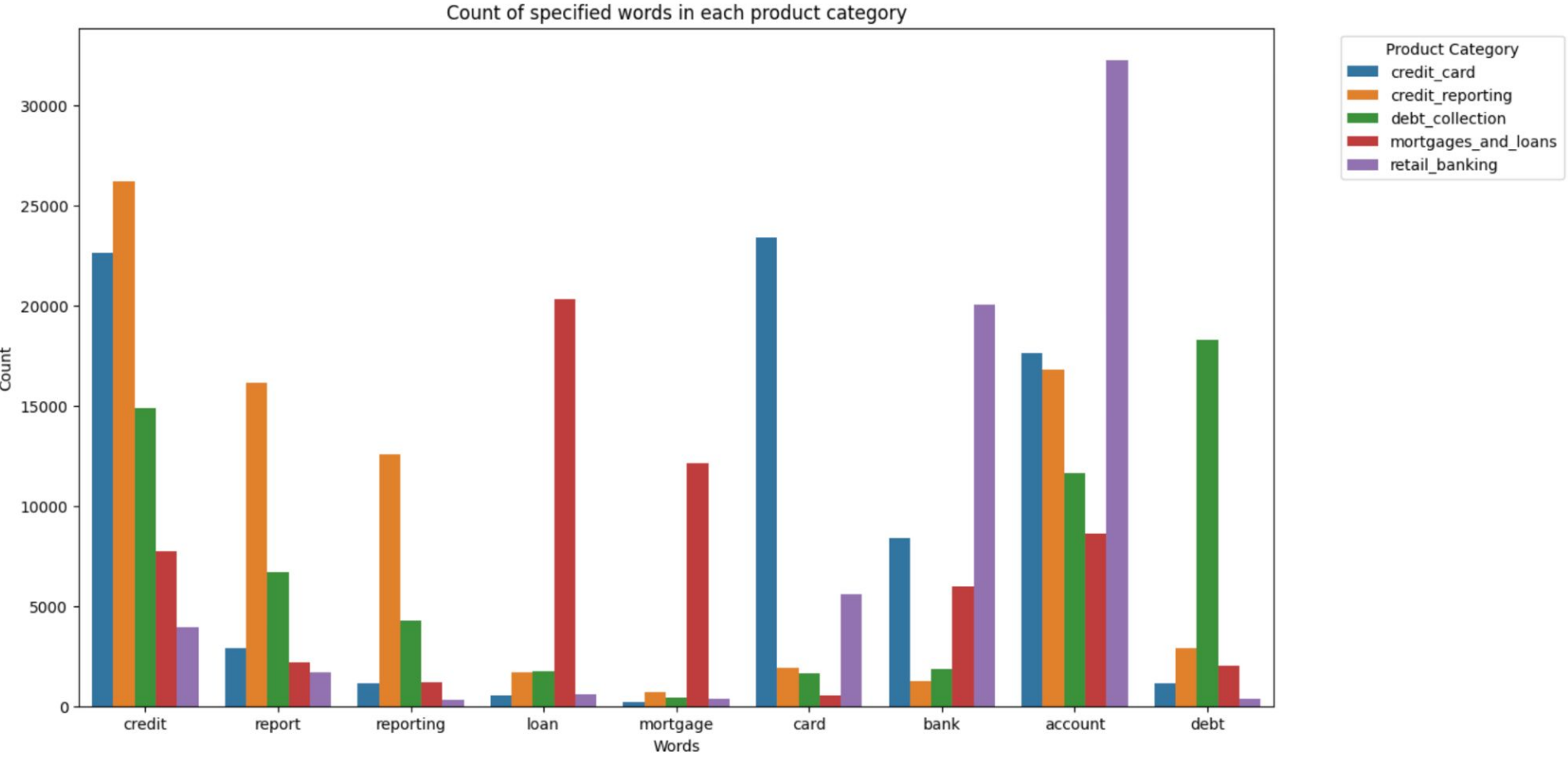
Scoring

- Used accuracy as the primary metric since the five classes are being balanced.
- Watched closely how our models performed on both the training and test datasets to eliminate the overfitting.

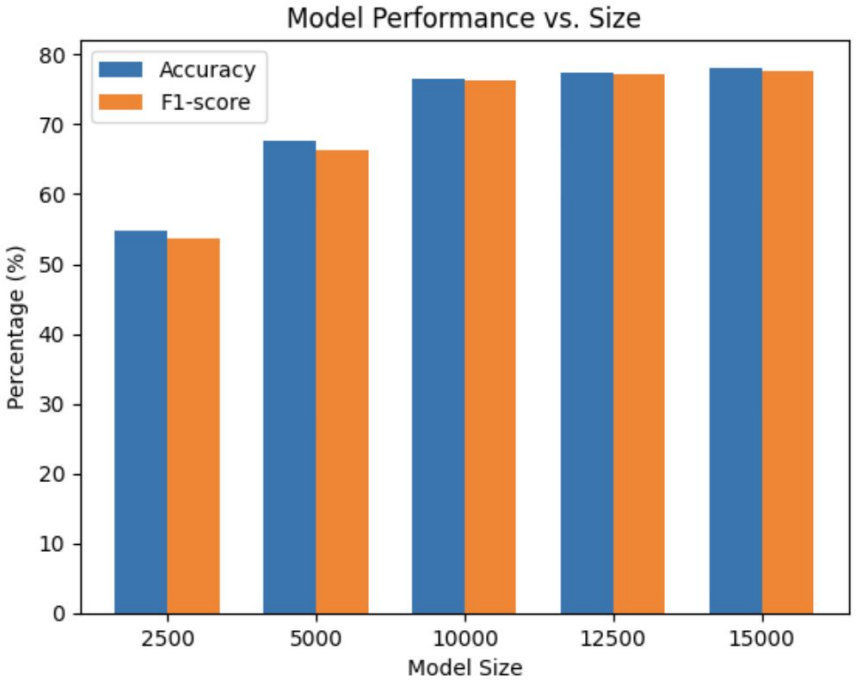
Refinement

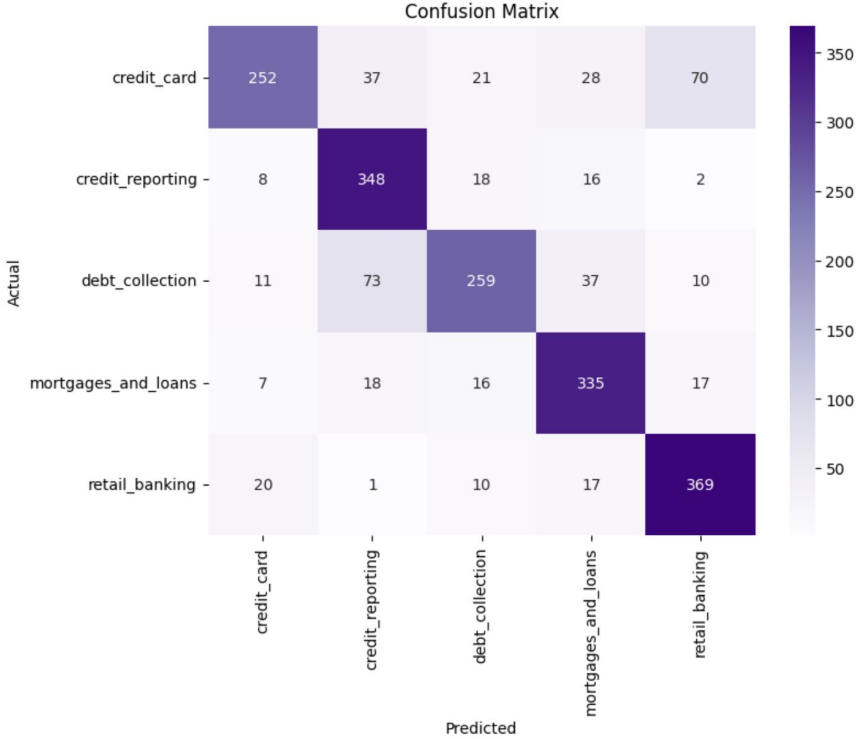
- Experimented with various parameters for the baseline models.
- Observed the model performances on various data instances.
- XGBoost had best overall accuracy, classifying narratives correctly 84% of the time.

Word Prominence per class



BERT Model





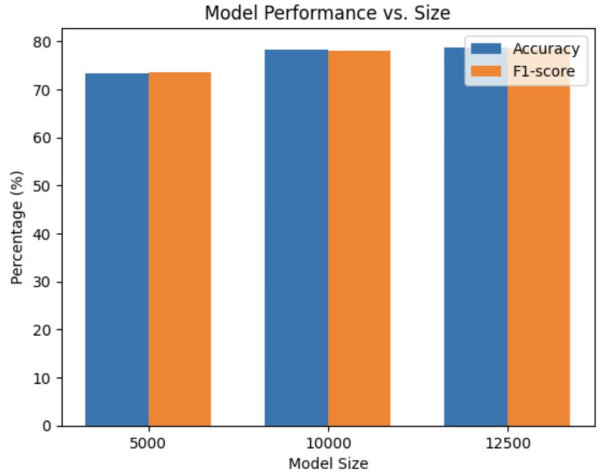
Accuracy: 0.7815
F1-Score: 0.777
Weighted Average Precision: 0.79
Weighted Average Recall: 0.78

Tokenizer Size: Max_legth = 256
Learning rate: 2e-5
Batch Size: 32
epoch : 1

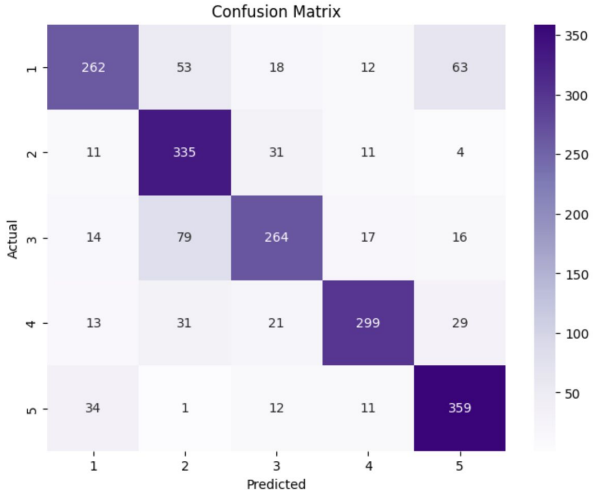
Analysis:

1. Why BERT Instead of Other Pre-trained Language Models?
   - BERT is pre trained on large corpus of text data
   - Transformer Artitecture
   - Performed better on multi-class text classification

2. What Will Happen if lr=5e-5 Instead of lr=2e-5?
   - Achieved 75% accuracy with a learning rate of 5e-5 on 5000 data.
   - Noticed that training time tripled compared to a lower learning rate



Accuracy when data has duplicate

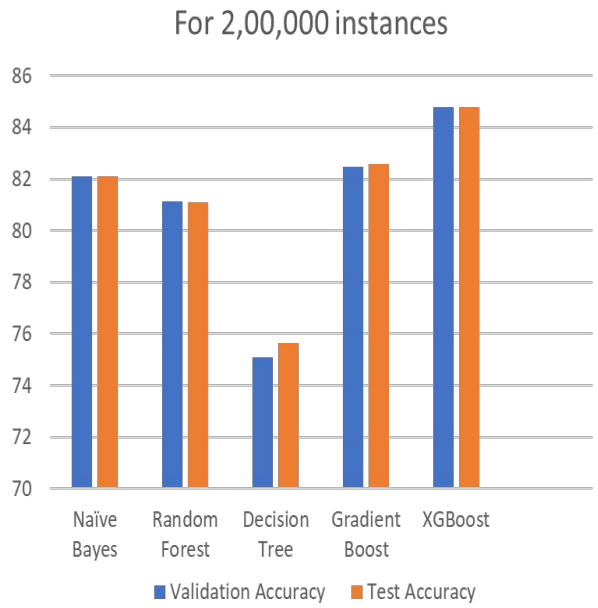


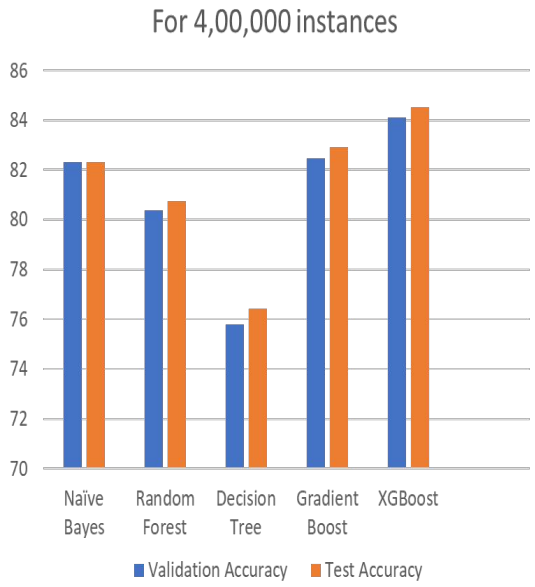
Future Analysis:

1.Does Accuracy Increase with a Larger Dataset?

2. What Will Happen if the Training Tokenizer Max Length is Changed to 512?

Results obtained by varying the number of instances





Comparison of the primary evaluation metric for 1,40,000 instances

Model	Validation Accuracy	Test Accuracy
Multinomial Naive Bayes	81.2	83.8
Random Forest	80.86	85.54
Decision Tree	75.4	85.46
Gradient Boosting	83.02	84.11
XGBoost	84.52	84.82

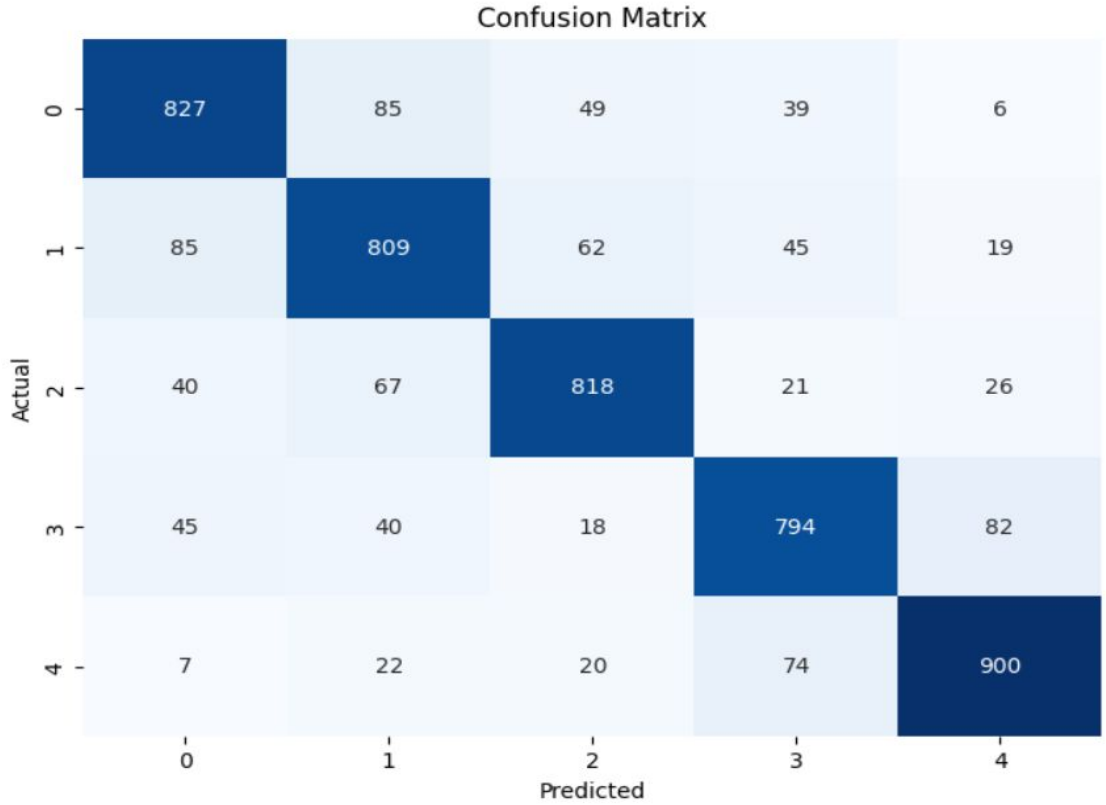
Comparison of the primary evaluation metric for 2,00,000 instances

Model	Validation Accuracy	Test Accuracy
Multinomial Naive Bayes	82.08	82.1
Random Forest	81.14	81.1
Decision Tree	75.10	75.65
Gradient Boosting	82.4	82.58
XGBoost	84.78	84.8

Comparison of the primary evaluation metric for 4,00,000 instances

Model	Validation Accuracy	Test Accuracy
Multinomial Naive Bayes	82.3	82.33
Random Forest	80.39	80.74
Decision Tree	75.78	76.41
Gradient Boosting	82.45	82.9
XGBoost	84.11	84.52

XGBOOST

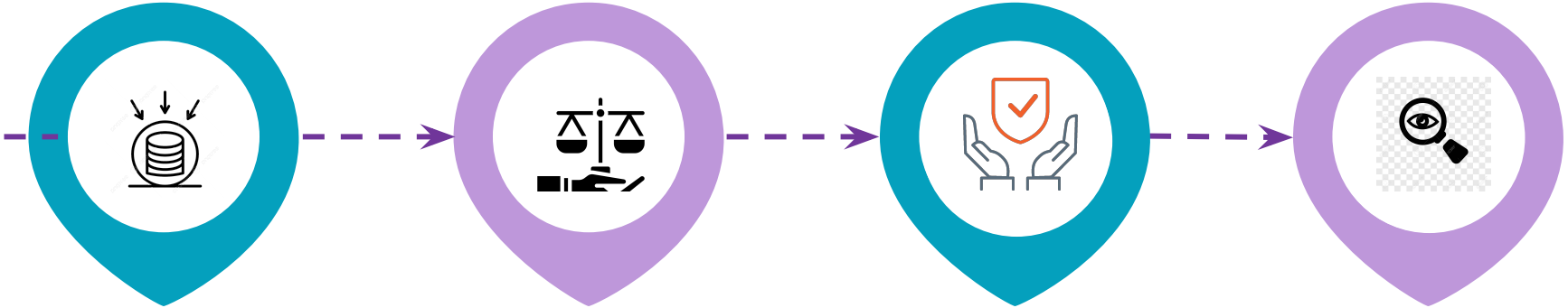


Accuracy: 0.8411
F1-Score: 0.8452
Weighted Average Precision: 0.832
Weighted Average Recall: 0.841

Challenges:

- More than twenty complaints category
- Class Imbalance
- Fine Tune the Model
- Execution time when we increase the number of instances
- Consolidating the complaints with the category of other financial service

Ethical Considerations



Data Collection Standards

Use consumer complaint data solely for the intended purpose of prediction and resolution

Fairness

Promoting fairness in representation to avoid bias in our models.

Privacy

Our project does not directly identify any individual with their personal data in the Database.

Transparency

Our model ensures transparency by clearly communicating the project's purpose, the workings of the model

Future scope

- Advancing modeling capabilities for sub-product, issue, and sub-issue categorizations.
- Refine models using advanced techniques like LLM.
- Investigate the potential for improved accuracy in real-time application to better assist the Consumer Financial Protection Bureau (CFPB).

Conclusion:

- We employed five baseline models, including Multinomial Naive Bayes, Random Forest, Decision Tree, Gradient Boosting, and XGBoost.
- With the use of real-time dataset, all of these models are trained, validated and tested. All these models shown accurate results for the large dataset.


DEMO


THANK
YOU