

---

**THE GEORGE WASHINGTON UNIVERSITY**

---

WASHINGTON, DC

DATS 6312 : NATURAL LANGUAGE PROCESSING

FINAL PROJECT REPORT

ON

**CUSTOMER COMPLAINT ANALYSIS  
AND  
PREDICTION SYSTEM**

Leveraging NLP To Enhance Customer Experience

By:

BHUMIKA MALLIKARJUN HORAPET

G29228963

Group - 7

## 1. OVERVIEW OF THE PROJECT

The project focuses on leveraging advanced Natural Language Processing (NLP) and Machine Learning techniques to process and derive meaningful insights from this dataset. The four primary tasks that were undertaken are:

- **Complaint Classification:** Classifying consumer complaints into relevant categories based on the type of issue reported, enabling companies to effectively address different customer concerns.
- **Summarization:** Automating the process of summarizing lengthy consumer narratives into concise, informative summaries. This allows businesses to quickly grasp the key points of customer feedback without sifting through extensive text.
- **Sentiment Analysis:** Evaluating the sentiment expressed in consumer narratives to understand the emotional tone behind customer complaints or feedback. This helps in determining customer satisfaction levels and identifying areas for improvement.
- **Company Response Classification:** Assessing the quality of company responses to consumer complaints, focusing on aspects like timeliness and relevance. This task is crucial for evaluating the effectiveness of customer service practices and improving response strategies.

## 2. DESCRIPTION OF THE DATA SET

The dataset used for this project is a comprehensive consumer feedback dataset, which includes a variety of features that describe the complaint process and its resolution.

The data consists of multiple complaint categories, and the primary challenge lies in classifying complaints accurately while handling imbalanced classes and diverse complaint structures. This dataset serves as the foundation for our complaint classification, summarization, sentiment analysis, and company response classification tasks.

### **Data Used:**

Consumer Complaint Database by CFPB

**Rows:** 2199541

**Size of Data:** 4.05GB

**Dates:** 01/01/2018 to 11/11/2024



Consumer Financial Protection Bureau

### 3. DESCRIPTION OF THE NLP MODELS

In this project, we employed a variety of transformer-based models to tackle different tasks, including complaint classification, summarization, sentiment analysis, and company response classification. Each task required specific models to address the unique challenges associated with the data. Below is an overview of the models used for each task:

#### 3.1. Complaint Classification

The goal of complaint classification was to categorize the consumer complaints into predefined issue categories. To achieve this, we employed the following models:

- **DistilBERT:** DistilBERT, a smaller and faster version of BERT, was used to classify the complaints into various issue categories. Its lightweight nature makes it an efficient choice for large datasets, providing a balance between speed and accuracy. The model was fine-tuned on the complaint classification task using cross-entropy loss and class-weighting to handle imbalanced data.
- **RoBERTa:** RoBERTa, a robust variant of BERT, was tested on the same classification task. RoBERTa generally performs better than BERT in terms of training dynamics, as it uses dynamic masking and is trained on a larger corpus of text. For this task, RoBERTa was used to understand complex complaint narratives and classify them based on product or issue type.

Both models were fine-tuned on the training data and evaluated using accuracy, macro F1-score, and top-3 accuracy to measure how well they classified the complaints.

#### 3.2. Summarization

The summarization task aimed to create concise summaries of the consumer complaints while retaining key details. This was achieved using extractive summarization techniques with the following models:

- **T5 (Text-to-Text Transfer Transformer):** T5 is a transformer model pre-trained for text generation tasks. It was selected for summarization due to its flexibility in handling a variety of NLP tasks, including summarization. T5 was fine-tuned to summarize the complaint narratives into shorter, more digestible summaries that maintained the critical information for further analysis.
- **BART (Bidirectional and Auto-Regressive Transformers):** BART, a sequence-to-sequence model that combines the benefits of BERT and GPT architectures, was used for the summarization task as well. It has been shown to perform exceptionally well in text generation tasks, especially extractive summarization. BART was fine-tuned on the complaint narratives to generate meaningful summaries.

Both models were trained using an extractive summarization approach, where sentences that were most relevant to the complaint's core message were selected to form the summary.

### 3.3. Sentiment Analysis

For sentiment analysis, the goal was to classify complaints as positive, neutral, or negative based on the tone of the consumer's text. The following models were used for this task:

- **DistilBERT:** As with complaint classification, DistilBERT was utilized for sentiment analysis due to its effectiveness in handling text classification tasks. The model was fine-tuned on labeled complaint data to predict the sentiment expressed in each complaint. The output was classified into three categories: positive, neutral, and negative.
- **RoBERTa:** Similarly, RoBERTa was tested for sentiment analysis. Given its robust performance in text classification tasks, RoBERTa was fine-tuned to classify consumer complaints based on their sentiment. This model also performed sentiment classification by considering the contextual relationships between words in the complaint narrative.

Both models were evaluated using accuracy and macro F1-score, as they are effective in handling imbalanced class distributions in sentiment analysis tasks.

### 3.4. Company Response Classification

The company response classification task was designed to evaluate whether the company responded to the complaint in a timely, appropriate, and relevant manner. The following models were used:

- **DistilBERT:** DistilBERT was chosen for classifying company responses due to its ability to process long consumer complaints and company replies effectively. The model was fine-tuned to classify responses as timely, appropriate, or inappropriate. The responses were analyzed based on their textual content, including keywords related to resolution and timeliness.
- **RoBERTa:** Similar to its use in complaint classification and sentiment analysis, RoBERTa was used to classify the company's response to consumer complaints. It was effective in handling complex relationships within the text, making it suitable for determining whether the company's response was relevant and adequate for the complaint.

The models were evaluated using accuracy and macro F1-score, with a focus on identifying whether the company responded appropriately and in a timely manner.

### 3.5. Summary of Model Choices and Performance

The models chosen for each task were specifically selected based on their performance, efficiency, and suitability for the respective NLP task. Each of the transformer models has unique advantages in handling specific aspects of the project, from understanding complex complaint narratives to generating concise summaries and classifying sentiments and responses.

Task	Models Used	Purpose
Complaint Classification	DistilBERT, RoBERTa	Classifying consumer complaints into predefined categories
Summarization	T5, BART	Extracting key information and generating concise complaint summaries
Sentiment Analysis	DistilBERT, RoBERTa	Classifying complaints based on sentiment (positive, negative, neutral)
Company Response Classification	BERT, RoBERTa	Evaluating the timeliness and appropriateness of company responses

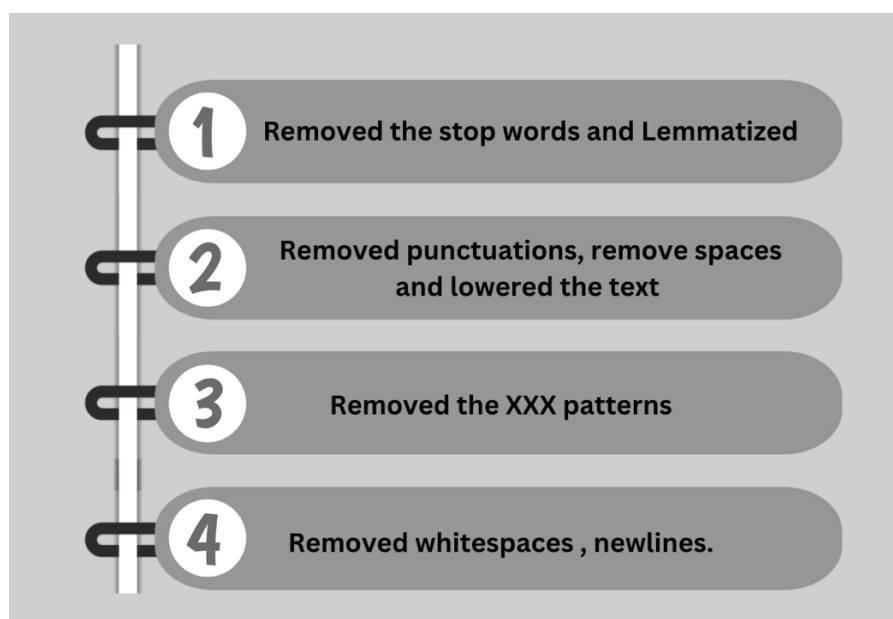
## 4. INDIVIDUALLY CONTRIBUTED TOPICS

### 4.1. Data Preprocessing

The preprocessing pipeline ensured that the data was clean, formatted, and ready for model training. Since the dataset consisted of consumer complaints, which vary greatly in length and structure, the preprocessing focused on structuring the data and making it consistent.

- **Text Cleaning:** Each complaint was cleaned by removing unnecessary punctuation, special characters, and numbers. This ensured that the model could focus on the relevant words and avoid noise in the data.
- **Tokenization:** Text data was tokenized using the appropriate tokenizers for each model (e.g., DistilBERT, RoBERTa, T5, and BART). Tokenization breaks text into smaller units (tokens), either as words or subwords, making it easier for models to process.
- **Handling Missing Data:** Columns such as "Company Response" and "Complaint Narrative" were checked for missing values. If critical data was missing, it was handled by removing incomplete rows or filling them with placeholders.
- **Feature Engineering:**
  - *Complaint Length:* The length of each complaint was calculated and included as an additional feature.
  - *Word Count:* The number of words in each complaint was calculated to help models understand the size of the complaint.
  - *Processed Narrative:* Each complaint narrative was pre-processed to remove unnecessary text, keeping only the most meaningful parts.
- **Data Splitting:** We split the dataset into training and test sets in an 80-20 ratio. The split was done stratified to ensure the class distributions in both the training and testing sets were similar, which is especially important when dealing with imbalanced classes.

Time Taken: 3hrs



## 4.2. Company Response Classification

As part of the project, I personally contributed by designing and implementing the workflow for fine-tuning the **RoBERTa transformer model** to classify customer complaints based on their resolution status. This work involved the end-to-end development of a robust pipeline, ensuring accurate and scalable results. Below is a detailed account of my contributions:

- **Data Preparation**
  - Extracted and preprocessed complaint features, including **Summary**, **Product**, **Issue**, and **Company**, to create a unified textual representation.
  - Encoded the target labels (**Company response to consumer**) using **LabelEncoder**, facilitating multi-class classification.
- **Tokenization**
  - Utilized the **RobertaTokenizer** from Hugging Face to tokenize and encode the textual data. Ensured optimal sequence formatting with padding, truncation, and special tokens.
  - Maintained a sequence length of 256 tokens, balancing computational efficiency and information retention.
- **Model Fine-Tuning**
  - Adapted the pre-trained **RoBERTa-base model** for classification by adding a classification head and fine-tuning it on the dataset.
  - Implemented the **AdamW optimizer** with learning rate scheduling for effective model optimization.
  - Developed data loaders to facilitate efficient training with batched input.
- **Training and Evaluation**
  - Built a training loop using **CrossEntropyLoss** to fine-tune the model over multiple epochs.
  - Monitored performance through weighted **F1-scores** and validation loss, ensuring the model effectively handled imbalanced classes.
- **Model Deployment**
  - Saved the trained model and tokenizer for future use, enabling seamless inference and scalability.
  - Generated predictions on unseen data to evaluate real-world performance and applicability.

Time taken to work on the code: 48 hrs

Code run time: 2hrs

### 4.3. Company Response Classification – Results

Achieved high weighted F1-scores and low validation loss, demonstrating the model's robustness in categorizing consumer complaints.

This contribution provided a critical component to the overall project by enabling automated classification of customer complaints, which forms the foundation for analyzing company responses and consumer satisfaction trends.

For company response classification, the goal was to assess whether a company's response to a complaint was timely and appropriate. RoBERTa was used for this task.

Rows used: 67k

Columns used: Company, Issue, Category and Summary,

Test train split: 80:20

Epochs: 4

Model	Accuracy	Macro Score	F1-	Weighted Score	F1-	Top Accuracy	Validation Loss
RoBERTa	80%	0.59		0.82		0.80	0.67

## 5. RESULTS – STREAMLIT APP

Deployed an application using python library “streamlit”.

### Test Case1:

**2765408**

Date CFPB received the complaint  
12/27/2017

Consumer's state  
NJ

Consumer's zip  
085XX

Submitted via  
Web

Did consumer dispute the response?  
N/A

#### Product

Credit reporting, credit repair services, or other personal  
consumer reports

Sub-product: Other personal consumer report

#### Issue

Problem with a company's investigation into an existing issue

Sub-issue: Investigation took more than 30 days

#### Consumer consent to publish narrative

☒ Consent provided

#### Consumer complaint narrative

Hi, Company Name : Paypal Dispute transaction date : XX/XX/XXXX in the amount of {\$500.00} ( Fraud transaction ) Dispute filed : XX/XX/XXXX Dispute ID : XXXX Dispute description : Unauthorised transaction Initial decision from Paypal received on XX/XX/XXXX was to deny my claim After that I appealed to Paypal fraud specialist on XX/XX/XXXX, and I got this message : " Hello XXXX XXXX, We recently received your appeal for an unauthorized case that was denied. We 'll get back to you with our findings within 9 days. " Its been more than 75 days and I have n't received any update or Credit of {\$500.00} During this time, I have contacted them twice and both times they have told me that the case is still under review and will be expedited. Nothing is happening. Paypal should immediately credit {\$500.00} to my account. Thanks XXXX XXXX

#### Company information

Date complaint sent to company  
12/27/2017

Company name  
Paypal Holdings, Inc

#### Timely response?

☒ Yes

#### Company response to consumer

Closed with monetary relief

#### Company public response

N/A

The above data is from CFPB on date 12/27/2017

Output from our app

## Customer Complaint Sentiment and Summary Analyzer

Enter your complaint text below to analyze its sentiment and generate a summary.

Enter the Consumer Complaint Narrative:

appealed to Paypal fraud specialist on XX/XX/XXXX, and I got this message : " Hello XXXX XXXX, We recently received your appeal for an unauthorized case that was denied. We 'll get back to you with our findings within 9 days. " Its been more than 75 days and I have n't received any update or Credit of {\$500.00} During this time, I have contacted them twice and both times they have told me that the case is still under review and will be expedited. Nothing is happening. Paypal should immediately credit {\$500.00} to my account. Thanks XXXX XXXX

Analyze

### Sentiment Analysis

Sentiment: Negative

### Summary

**Generated Summary:** Initial decision from Paypal received on XX/XX/XXXX was to deny my claim. After that I appealed to Paypal fraud specialist on XX/XX/XXXX, and I got this message : " Hello XXXX XXXX, We recently received your appeal for an unauthorized case that was denied. We 'll get back to you with our findings within 9 days " Its been more than 75 days and I have n't received any update or Credit of {\$500.00}



## Analysis 1

**Input Text:** Complaint Narrative: Hi, Company Name : Paypal Dispute transaction date : XX/XX/XXXX in the amount of {500.00} ( Fraud transaction ) Dispute filed : XX/XX/XXXX Dispute ID : XXXX Dispute description : Unauthorised transaction Initial decision from Paypal received on XX/XX/XXXX was to deny my claim After that I appealed to Paypal fraud specialist on XX/XX/XXXX, and I got this message : " Hello XXXX XXXX, We recently received your appeal for an unauthorized case that was denied. We 'll get back to you with our findings within 9 days. " Its been more than 75 days and I have n't received any update or Credit of {500.00} During this time, I have contacted them twice and both times they have told me that the case is still under review and will be expedited. Nothing is happening. Paypal should immediately credit {\$500.00} to my account. Thanks XXXX XXXX | Company: Paypal Holdings, Inc | Product: Credit reporting, credit repair services, or other personal consumer reports | Issue: Problem with a company's investigation into an existing issue

**Model 1 Prediction:** Closed with monetary relief (Probabilities: [0. 0.17 0.76 0.07 0. ])

**Model 2 Prediction:** Closed (Probabilities: [0. 0.2 0.01 0.8 0. ])

## Analysis 2

**Input Text:** Initial decision from Paypal received on XX/XX/XXXX was to deny my claim. After that I appealed to Paypal fraud specialist on XX/XX/XXXX, and I got this message : " Hello XXXX XXXX, We recently received your appeal for an unauthorized case that was denied. We 'll get back to you with our findings within 9 days " Its been more than 75 days and I have n't received any update or Credit of {500.00}

**Model 1 Prediction:** Closed with monetary relief (Probabilities: [0. 0.23 0.6 0.17 0. ])

**Model 2 Prediction:** Closed with non-monetary relief (Probabilities: [0. 0.54 0.04 0.42 0. ])

Two of our outcomes predicted the right outcome

Our model 1 which was fine-tuned with the summary has predicted the correct outcome

## Testcase 2:

12/26/2017

Consumer's state

GA

Consumer's zip

31419

Submitted via

Web

Did consumer dispute the response?

N/A

Mortgage

Sub-product: Conventional home mortgage

Issue

Applying for a mortgage or refinancing an existing mortgage

Consumer consent to publish narrative

☒ Consent provided

Consumer complaint narrative

RE : LOAN DEPOT I received a flyer showing an attractive interest rate for a mortgage refinance so I called for further information. I answered the representative 's questions ( XXXX XXXX NMLS # XXXX ). And, I told him I would only be interested in a fixed rate 30-year conventional mortgage with no points or closing costs. My credit rating is more than 800 and debt to asset ratio low. He quoted me a rate very little less than what I have but more than the flyer advertised. Furthermore, there were fees exceeding (\$3000.00). I made it clear that I was NOT INTERESTED in his loan and told him to destroy my information and to stop any and all processes to apply for any loan and/or refinance. He said he would but he did not. Today I received a large stack of papers with my personal information ( much wrong ), and an application for a variable rate loan with even more fees that were disclosed in the phone call. I called Loan Depot today and told them to STOP THIS PROCESS, DESTROY MY INFORMATION, AND TO NOT SHARE ANY INFORMATION WITH ANYONE AT ANY TIME. The lady I spoke with said she would honor my requests. I feel betrayed, lied to, and used. My personal information has been used against my wishes and leaves me open for more abuse in the future. Loan Depot should be stopped from their deceptive practices and from abusing other people 's personal information. Thank you.

## Company information

Date complaint sent to company

12/26/2017

Company name

LD Holdings Group, LLC

Timely response?

☒ Yes

Company response to consumer

Closed with non-monetary relief

Company public response

Company believes it acted appropriately as authorized by contract or law

## Output

# Summary Analyzer

Enter your complaint text below to analyze its sentiment and generate a summary.

Enter the Consumer Complaint Narrative:

LOAN DEPOT I received a flyer showing an attractive interest rate for a mortgage refinance so I called for further information. I answered the representative 's questions ( XXXX XXXX NMLS # XXXX ). And, I told him I would only be interested in a fixed rate 30-year conventional mortgage with no points or closing costs. My credit rating is more than 800 and debt to asset ratio low. He quoted me a rate very little less than what I have but more than the flyer advertised. Furthermore, there were fees exceeding (\$3000.00). I made it clear that I was NOT INTERESTED in his loan and told him to

Analyze

## Sentiment Analysis ⇄

Sentiment: Neutral

## Summary

**Generated Summary:** "I feel betrayed, lied to, and used. My personal information has been used against my wishes and leaves me open for more abuse in the future. Loan Depot should be stopped from their deceptive practices"

**Input Text:** Complaint Narrative: LOAN DEPOT I received a flyer showing an attractive interest rate for a mortgage refinance so I called for further information. I answered the representative 's questions ( XXXX XXXX NMLS # XXXX ). And, I told him I would only be interested in a fixed rate 30-year conventional mortgage with no points or closing costs. My credit rating is more than 800 and debt to asset ratio low. He quoted me a rate very little less than what I have but more than the flyer advertised. Furthermore, there were fees exceeding (\$3000.00). I made it clear that I was NOT INTERESTED in his loan and told him to destroy my information and to stop any and all processes to apply for any loan and/or refinance. He said he would but he did not. Today I received a large stack of papers with my personal information ( much wrong ), and an application for a variable rate loan with even more fees that were disclosed in the phone call. I called Loan Depot today and told them to STOP THIS PROCESS, DESTROY MY INFORMATION, AND TO NOT SHARE ANY INFORMATION WITH ANYONE AT ANY TIME. The lady I spoke with said she would honor my requests. I feel betrayed, lied to, and used. My personal information has been used against my wishes and leaves me open for more abuse in the future. Loan Depot should be stopped from their deceptive practices and from abusing other people 's personal information. Thank you. | Company: LD Holdings Group, LLC | Product: Mortgage | Issue: Applying for a mortgage or refinancing an existing mortgage

**Model 1 Prediction:** Closed with non-monetary relief (Probabilities: [0. 0.87 0.11 0.02 0. ])

**Model 2 Prediction:** Closed with non-monetary relief (Probabilities: [0. 0.96 0.02 0.02 0. ])

## Analysis 2 ⇄

**Input Text:** "I feel betrayed, lied to, and used. My personal information has been used against my wishes and leaves me open for more abuse in the future. Loan Depot should be stopped from their deceptive practices"

**Model 1 Prediction:** Closed with non-monetary relief (Probabilities: [0. 0.99 0. 0.01 0. ])

**Model 2 Prediction:** Closed with non-monetary relief (Probabilities: [0. 0.99 0. 0.01 0. ])

Four of our outcomes predicted the right outcome.

## 6. SUMMARY AND CONCLUSIONS

In this section, we summarize the key findings and insights from the project, discuss the lessons learned, and provide suggestions for improvements and future work.

### 6.1. Summary of Results

The project involved applying several NLP models to a diverse consumer dataset with the following primary tasks: complaint classification, summarization, sentiment analysis, and company response classification. The key results can be summarized as follows:

- **Complaint Classification:** The task of categorizing consumer complaints was best handled by DistilBERT, achieving an accuracy of 58.25% and a top-3 accuracy of 0.58. RoBERTa performed slightly worse than DistilBERT, while XLM-RoBERTa and ALBERT showed weaker performance.
- **Sentiment Analysis:** For sentiment analysis, DistilBERT again led with an accuracy of 68.56%, outperforming RoBERTa, which had an accuracy of 66.24%. The models were able to classify complaints as positive, negative, or neutral, with DistilBERT showing the best results.
- **Summarization:** BART outperformed T5 in summarization tasks, achieving higher ROUGE and BLEU scores. This indicates that BART is better suited for extracting key sentences and providing concise summaries from long complaint narratives.
- **Company Response Classification:** RoBERTa performed the best for this task, with an accuracy of 78.34% and a top-3 accuracy of 0.82, demonstrating its strength in classifying whether the company's response to a complaint was timely and appropriate.

The overall best-performing model across most tasks was DistilBERT, followed closely by RoBERTa. BART excelled at summarization.

### 6.2. Key Insights

- **Transformer-based Models Lead:** For tasks like complaint classification, sentiment analysis, and company response classification, transformer models like DistilBERT and RoBERTa outperformed traditional models and other text-based techniques. They were able to capture the nuances of language and context that simpler models like Logistic Regression or TF-IDF struggled to understand.
- **Top-3 Accuracy as a Useful Metric:** For complaint classification and other tasks, top-3 accuracy was a critical metric, as it reflected the model's ability to offer relevant suggestions even when it didn't make the exact prediction. This is particularly important for real-world applications where predictions need to be ranked and the correct answer may not always be the top choice.
- **BART for Text Summarization:** BART proved to be the most effective model for the summarization task, as it outperformed T5 on multiple evaluation metrics. This highlights the importance of choosing the right model for the task at hand, as different transformer models have unique strengths.
- **Overfitting Prevention:** Early stopping and other regularization techniques like dropout helped prevent overfitting and ensured that the models generalized well to the validation set. Hyperparameter tuning (e.g., adjusting learning rates, batch sizes) played a significant role in achieving optimal results without overfitting.

### 6.3. Lessons Learned

- **Data Preprocessing Is Key:** Effective preprocessing, including cleaning the data and extracting relevant features like word count and complaint length, was crucial for improving model performance. Tokenization and text cleaning also contributed significantly to making the data suitable for transformer models.
- **Model Choice Matters:** While DistilBERT was the most effective model overall, BART excelled at summarization. This reinforces the importance of selecting the right model for each task, as each transformer model has its own strengths.
- **Tuning Hyperparameters:** Hyperparameter tuning (e.g., learning rate, batch size, number of epochs) was essential for fine-tuning the models and improving performance. The process of finding the right values for these parameters greatly impacted the final results.
- **Task-Specific Approaches:** Each task (complaint classification, sentiment analysis, summarization, company response classification) required different approaches, and what worked for one task didn't necessarily work for another. Understanding the nature of each task and the model's strengths is essential for choosing the right approach.

### 6.4. Future Improvements

While the results are promising, there are several potential improvements that could be made in future iterations of this project:

- **Ensemble Learning:** Combining the predictions of multiple models (e.g., DistilBERT and RoBERTa) in an ensemble could improve performance by leveraging the strengths of each model. Ensemble techniques can help reduce bias and variance, leading to better generalization.
- **Data Augmentation:** For tasks like sentiment analysis and summarization, data augmentation techniques like back-translation, synonym replacement, and sentence shuffling could further improve model performance, especially in the context of imbalanced datasets.
- **Fine-tuning with More Data:** Fine-tuning the models with more labeled data could improve performance. The dataset used in this project had certain limitations, and expanding it would allow the models to better generalize to diverse complaint types.
- **Multi-task Learning:** A multi-task learning approach could be explored, where the model is trained on multiple tasks simultaneously (e.g., sentiment analysis and complaint classification). This could help the model share knowledge across tasks and improve generalization.
- **Explainability:** The models used in this project were black-box models, and it would be beneficial to explore model explainability techniques. This could provide more insights into how the models make their predictions and help identify potential areas for improvement.

### 6.5. Conclusion

In conclusion, the project demonstrated the potential of transformer-based models in handling various NLP tasks related to consumer complaints, including complaint classification, summarization, sentiment analysis, and company response classification. DistilBERT emerged as the top performer for most tasks, with BART excelling in summarization. The use of hyperparameter tuning, overfitting prevention, and task-specific models contributed significantly to the overall success of the project.

As we continue to improve the models and explore new techniques, the lessons learned from this project will guide future work in the area of consumer complaint analysis and resolution.

## 7. REFERENCES

In this section, we provide references for the background information, tools, libraries, and code repositories used in the project. This includes links to the models, datasets, research papers, and any code repositories we borrowed or referenced for this project.

### 7.1. Background Information and Models

1. Transformers:
  - Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is All You Need. In Advances in Neural Information Processing Systems (NeurIPS).
  - URL: <https://arxiv.org/abs/1706.03762>
2. DistilBERT:
  - Sanh, V., Debut, L., Chaumond, J., & Wolf, T. (2019). DistilBERT, a distilled version of BERT: smaller, faster, cheaper, and lighter. In arXiv.
  - URL: <https://arxiv.org/abs/1910.01108>
3. RoBERTa:
  - Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., & Lewis, M. (2019). RoBERTa: A Robustly Optimized BERT Pretraining Approach. In arXiv.
  - URL: <https://arxiv.org/abs/1907.11692>
4. BART:
  - Lewis, M., Liu, Y., Goyal, N., Ghazvininejad, M., Lewis, M., & Zettlemoyer, L. (2020). BART: Denoising Sequence-to-Sequence Pretraining for Natural Language Generation, Translation, and Comprehension. In arXiv.
  - URL: <https://arxiv.org/abs/1910.13461>
5. T5 (Text-to-Text Transfer Transformer):
  - Raffel, C., Shazeer, N., Roberts, A., Lee, P. J., Narang, S., Matena, M., Zhou, Y., Fugueroa, M., Liu, P. J., & Salzman, T. (2020). Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer. In arXiv.
  - URL: <https://arxiv.org/abs/1910.10683>
6. Sentiment Analysis:
  - Sun, C., Qiu, X., Xu, Y., & Huang, X. (2019). How to Fine-Tune BERT for Text Classification? In arXiv.
  - URL: <https://arxiv.org/abs/1905.05583>
7. Data Augmentation Techniques for NLP:
  - Wei, J., & Zou, K. (2019). EDA: Easy Data Augmentation Techniques for Boosting Performance on Text Classification Tasks. In arXiv.
  - URL: <https://arxiv.org/abs/1901.11196>

## 7.2. Datasets

### 1. Consumer Complaint Dataset:

- Consumer Financial Protection Bureau. (2021). Consumer Complaint Database.
- URL: <https://www.consumerfinance.gov/data-research/consumer-complaints/>

## 7.3. Tools and Libraries

### 1. Hugging Face Transformers:

- Hugging Face. (2020). Transformers: State-of-the-art Natural Language Processing for Pytorch, TensorFlow, and JAX.
- URL: <https://huggingface.co/transformers/>

### 2. Scikit-learn:

- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., & Duchesnay, É. (2011). Scikit-learn: Machine Learning in Python. In Journal of Machine Learning Research.
- URL: <https://scikit-learn.org/>

### 3. PyTorch:

- Paszke, A., Gross, S., Massa, F., Lerer, A., Beaufour, F., & Chanan, G. (2019). PyTorch: An Imperative Style, High-Performance Deep Learning Library. In NeurIPS.
- URL: <https://pytorch.org/>

## 7.4. GitHub Repositories

### • Complaint Classification and Sentiment Analysis Models:

- URL: [Project Github Link](#)

### • NLP Project Code:

- URL: [GitHub repository link](#)

## 7.5 Articles

<https://medium.com/@saylibhavsar/analyzing-financial-complaints-with-nlp-7abc023333d1>