AI-Powered Smart Ticketing: Customer Complaints Classification

Kalp Devangbhai Thakkar
ka765424@ucf.edu
University of Central Florida
Orlando, FL, USA

ABSTRACT

"AI-Powered Smart Ticketing: Customer Complaints Classification" addresses the imperative for financial institutions to efficiently manage customer complaints through automated systems. Leveraging Non-Negative Matrix Factorization (NMF) for topic modeling, our model categorizes unstructured customer complaints into distinct clusters related to credit cards, banking, theft/disputes, mortgages/loans, and others. By subsequently training and evaluating logistic regression, naive Bayes, decision tree, and random forest models, I identified the most effective approach. This project streamlines complaint resolution, enhances customer satisfaction, and provides a robust framework for automated customer support in the financial sector, showcasing the potential of artificial intelligence in enhancing operational efficiency and customer experience.

ACM Reference Format:
Kalp Devangbhai Thakkar. 2024. AI-Powered Smart Ticketing: Customer Complaints Classification. In . ACM, Orlando, FL, USA, 15 pages.

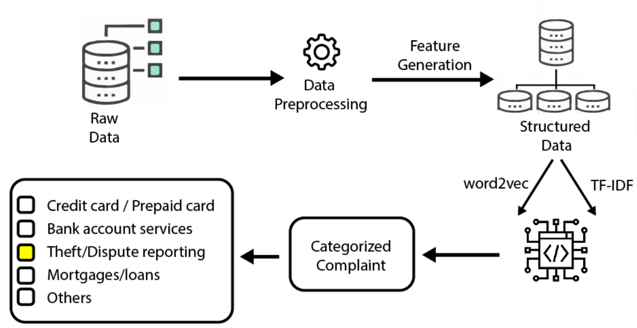


Figure 1: Model Architecture

1 INTRODUCTION

In an era where heightened customer expectations define the landscape of service excellence, the financial sector grapples with the formidable challenge of effectively managing and resolving customer complaints. My project, "AI-Powered Smart Ticketing: Customer Complaints Classification," emerges as a strategic initiative aimed at redefining the complaint resolution paradigm within a leading financial institution. This initiative arises from a profound understanding that managing customer complaints not only minimizes discontent but also serves as evidence of the institution's dedication to prioritizing customer needs. By leveraging advanced Natural Language Processing (NLP) techniques, I automated the classification of customer complaints, thereby streamlining support operations, bolstering customer satisfaction, and positioning the institution at the forefront of innovation within the financial landscape.

In response to the growing complexity and volume of customer complaints, the project harnesses the power of advanced Natural Language Processing (NLP) techniques. NLP offers a transformative approach by enabling machines to comprehend and analyze unstructured text data, paving the way for efficient and automated customer support. By leveraging state-of-the-art methods, such as Non-Negative Matrix Factorization (NMF) for topic modeling, the project endeavors to automatically categorize complaints into distinct clusters based on the products and services involved.

The ultimate goal is to enhance the institution's responsiveness to customer concerns, ensuring timely and tailored resolutions. Through the implementation of AI-powered smart ticketing, the financial institution aspires to not only address individual complaints but also to derive valuable insights from the collective feedback. This approach positions the institution at the forefront of innovation within the financial landscape, reinforcing its commitment to delivering unparalleled customer experiences.

April 2024, Orlando, FL, USA,

2 PROBLEM STATEMENT

In the dynamic landscape of financial services, organizations encounter a formidable challenge in managing the sheer volume of customer complaints. The influx of diverse and often unstructured grievances poses a significant obstacle to seamless operations and customer satisfaction. Financial companies grapple with the intricate task of allocating resources to assess, categorize, and address each complaint promptly. The resolution process is a pivotal juncture, as it directly influences customer loyalty and shapes the overall reputation of the company. Efficient complaint resolution not only mitigates immediate concerns but also establishes a foundation for enduring customer relationships and fosters a positive perception of the company within the broader market.

2.1 Volume of Customer Complaints

The financial industry is inundated with an ever-growing volume of customer complaints, spanning a spectrum of intricacies. These grievances encompass diverse issues, from transaction discrepancies and service dissatisfaction to concerns regarding the security of financial transactions. The sheer magnitude of these complaints poses a formidable challenge for companies seeking to maintain operational efficiency and stellar customer relations.

2.2 Resource Allocation Challenges

Allocating adequate resources to manage the influx of customer complaints is an intricate task faced by financial institutions. Human and technological resources must be judiciously distributed to cope with the complexity and variety of issues raised by customers. The scalability of support teams and technologies becomes paramount to ensure swift and effective resolution without compromising the quality of customer service.

2.3 Complexity of Complaint Categorization

Categorizing customer complaints accurately is a complex undertaking. The varied nature of issues necessitates manual efforts to ensure complaints are directed to the appropriate departments for resolution. The risk of misallocation looms large, emphasizing the need for advanced systems that can intelligently categorize and route complaints, streamlining the resolution process.

2.4 Timeliness and Customer Loyalty

Timely resolution of customer complaints is intricately tied to the preservation of customer loyalty. Delays in addressing grievances can result in heightened dissatisfaction, eroding trust, and diminishing customer loyalty. The urgency in efficiently handling complaints underscores the pivotal role this process plays in shaping a positive customer experience and fostering enduring relationships.

2.5 Reputation Management

Inefficient complaint resolution directly impacts the overall reputation of financial institutions. Negative experiences, if left unaddressed, can amplify through word-of-mouth and online platforms, potentially tarnishing the brand's image. Proactive measures in managing and resolving complaints are integral to safeguarding the company's reputation and maintaining a favorable standing in the competitive financial market.

2.6 Regulatory Compliance

Financial companies must navigate a complex regulatory landscape governing complaint resolution. Non-compliance with regulatory standards not only exposes institutions to legal consequences but also undermines the trust customers place in the company's commitment to ethical practices. Adhering to and exceeding regulatory requirements is paramount to ensure a robust and legally sound complaint resolution process.

2.7 Customer Experience Impact

Unresolved or poorly handled complaints have a lasting impact on the overall customer experience. Beyond the immediate dissatisfaction, customers may harbor a negative perception of the company, affecting trust and future interactions. Prioritizing effective complaint resolution is, therefore, essential not only for addressing specific issues but also for preserving the holistic customer journey and maintaining a positive brand-consumer relationship.

3 RELATED WORK

In the realm of customer complaint management within the financial sector, several initiatives and studies have contributed valuable insights and methodologies. Understanding and leveraging these prior works is essential for contextualizing our project and identifying areas for innovation and improvement.

3.1 Automated Ticketing Systems

Numerous financial institutions have implemented automated ticketing systems to streamline the process of managing customer complaints. These systems often rely on predefined rules or simple keyword matching techniques to categorize and prioritize complaints. While effective to some extent, such approaches may lack the sophistication needed to handle the nuances and complexities of unstructured complaint data.

3.2 Natural Language Processing (NLP) Techniques

Research in the field of Natural Language Processing (NLP) has witnessed significant advancements in recent years. Techniques such as sentiment analysis, topic modeling, and text classification have been extensively studied and applied to various domains, including customer service and complaint management. These approaches offer the potential to extract valuable insights from unstructured text data, enabling more informed decision-making and personalized customer interactions.

3.3 Topic Modeling and Clustering

Topic modeling techniques, such as Latent Dirichlet Allocation (LDA) and Non-Negative Matrix Factorization (NMF), have been widely used to analyze large volumes of text data and uncover underlying themes or topics. By clustering similar complaints based on their content, these methods facilitate more efficient organization and resolution of customer grievances. Previous research has demonstrated the effectiveness of topic modeling in improving complaint management processes and enhancing customer satisfaction.

3.4 Supervised Learning Models

In addition to unsupervised techniques, supervised learning models have been applied to the task of complaint classification. Algorithms such as logistic regression, decision trees, and random forests have shown promise in accurately categorizing customer complaints and predicting appropriate resolutions. By leveraging labeled training data, these models can learn to distinguish between different types of complaints and automate the routing process.

3.5 Industry Case Studies

Several industry case studies and reports provide valuable insights into best practices and strategies for customer complaint management in the financial sector. By examining real-world implementations and success stories, we can gain a deeper understanding of the challenges and opportunities associated with automated ticketing systems and NLP-driven approaches.

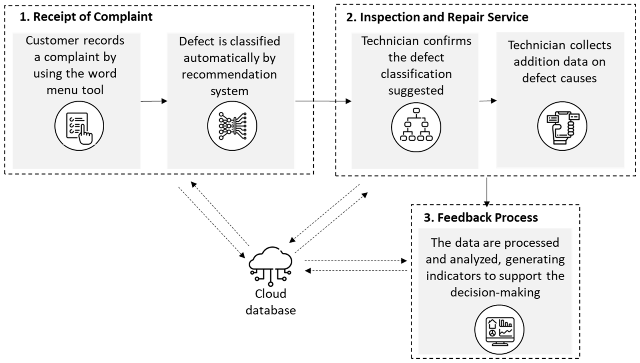
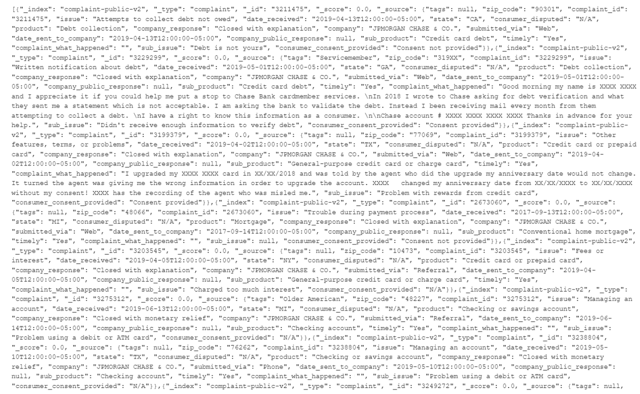


Figure 2: Model Application & Workflow



```
[{"_index": "complaint-public-v2", "_type": "complaint", "_id": "3211475", "_score": 0.0, "_source": {"tags": null, "zip_code": "90301", "complaint_id":
"3211475", "issue": "Attempts to collect debt not owed", "date_received": "2019-04-13T12:00:00-05:00", "state": "CA", "consumer_disputed": "N/A",
"product": "Debt collection", "company_response": "Closed with explanation", "company": "JPMORGAN CHASE & CO.", "submitted_via": "Web",
"date_sent_to_company": "2019-04-13T12:00:00-05:00", "company_public_response": null, "sub_product": "Credit card debt", "timely": "Yes",
"complaint_what_happened": "", "sub_issue": "Debt is not yours", "consumer_consent_provided": "Consent not provided"}},{"_index": "complaint-public-v2",
"_type": "complaint", "_id": "3229299", "_score": 0.0, "_source": {"tags": "Servicemember", "zip_code": "319XX", "complaint_id": "3229299", "issue":
"Written notification about debt", "date_received": "2019-05-01T12:00:00-05:00", "state": "GA", "consumer_disputed": "N/A", "product": "Debt collection",
"company_response": "Closed with explanation", "company": "JPMORGAN CHASE & CO.", "submitted_via": "Web", "date_sent_to_company": "2019-05-01T12:00:00-
05:00", "company_public_response": null, "sub_product": "Credit card debt", "timely": "Yes", "complaint_what_happened": "Good morning my name is XXXX XXXX
and I appreciate it if you could help me put a stop to Chase Bank cardmember services. \nIn 2018 I wrote to Chase asking for debt verification and what
they sent me a statement which is not acceptable. I am asking the bank to validate the debt. Instead I been receiving mail every month from them
attempting to collect a debt. \nI have a right to know this information as a consumer. \n\nChase account # XXXX XXXX XXXX XXXX Thanks in advance for your
help.", "sub_issue": "Didn't receive enough information to verify debt", "consumer_consent_provided": "Consent provided"}},{"_index": "complaint-public-
v2", "_type": "complaint", "_id": "3199379", "_score": 0.0, "_source": {"tags": null, "zip_code": "77069", "complaint_id": "3199379", "issue": "Other
features, terms, or problems", "date_received": "2019-04-02T12:00:00-05:00", "state": "TX", "consumer_disputed": "N/A", "product": "Credit card or prepaid
card", "company_response": "Closed with explanation", "company": "JPMORGAN CHASE & CO.", "submitted_via": "Web", "date_sent_to_company": "2019-04-
02T12:00:00-05:00", "company_public_response": null, "sub_product": "General-purpose credit card or charge card", "timely": "Yes",
"complaint_what_happened": "I upgraded my XXXX XXXX card in XX/XX/2018 and was told by the agent who did the upgrade my anniversary date would not change.
It turned the agent was giving me the wrong information in order to upgrade the account. XXXX   changed my anniversary date from XX/XX/XXXX to XX/XX/XXXX
without my consent! XXXX has the recording of the agent who was misled me.", "sub_issue": "Problem with rewards from credit card",
"consumer_consent_provided": "Consent provided"}},{"_index": "complaint-public-v2", "_type": "complaint", "_id": "2673060", "_score": 0.0, "_source":
{"tags": null, "zip_code": "48066", "complaint_id": "2673060", "issue": "Trouble during payment process", "date_received": "2017-09-13T12:00:00-05:00",
"state": "MI", "consumer_disputed": "N/A", "product": "Mortgage", "company_response": "Closed with explanation", "company": "JPMORGAN CHASE & CO.",
"submitted_via": "Web", "date_sent_to_company": "2017-09-14T12:00:00-05:00", "company_public_response": null, "sub_product": "Conventional home mortgage",
"timely": "Yes", "complaint_what_happened": "", "sub_issue": null, "consumer_consent_provided": "Consent not provided"}},{"_index": "complaint-public-v2",
"_type": "complaint", "_id": "3203545", "_score": 0.0, "_source": {"tags": null, "zip_code": "10473", "complaint_id": "3203545", "issue": "Fees or
interest", "date_received": "2019-04-05T12:00:00-05:00", "state": "NY", "consumer_disputed": "N/A", "product": "Credit card or prepaid card",
"company_response": "Closed with explanation", "company": "JPMORGAN CHASE & CO.", "submitted_via": "Referral", "date_sent_to_company": "2019-04-
05T12:00:00-05:00", "company_public_response": null, "sub_product": "General-purpose credit card or charge card", "timely": "Yes",
"complaint_what_happened": "", "sub_issue": "Charged too much interest", "consumer_consent_provided": "N/A"}},{"_index": "complaint-public-v2", "_type":
"complaint", "_id": "3275312", "_score": 0.0, "_source": {"tags": "Older American", "zip_code": "48227", "complaint_id": "3275312", "issue": "Managing an
account", "date_received": "2019-06-13T12:00:00-05:00", "state": "MI", "consumer_disputed": "N/A", "product": "Checking or savings account",
"company_response": "Closed with monetary relief", "company": "JPMORGAN CHASE & CO.", "submitted_via": "Referral", "date_sent_to_company": "2019-06-
14T12:00:00-05:00", "company_public_response": null, "sub_product": "Checking account", "timely": "Yes", "complaint_what_happened": "", "sub_issue":
"Problem using a debit or ATM card", "consumer_consent_provided": "N/A"}},{"_index": "complaint-public-v2", "_type": "complaint", "_id": "3238804",
"_score": 0.0, "_source": {"tags": null, "zip_code": "76262", "complaint_id": "3238804", "issue": "Managing an account", "date_received": "2019-05-
10T12:00:00-05:00", "state": "TX", "consumer_disputed": "N/A", "product": "Checking or savings account", "company_response": "Closed with monetary
relief", "company": "JPMORGAN CHASE & CO.", "submitted_via": "Phone", "date_sent_to_company": "2019-05-10T12:00:00-05:00", "company_public_response":
null, "sub_product": "Checking account", "timely": "Yes", "complaint_what_happened": "", "sub_issue": "Problem using a debit or ATM card",
"consumer_consent_provided": "N/A"}},{"_index": "complaint-public-v2", "_type": "complaint", "_id": "3249272", "_score": 0.0, "_source": {"tags": null,
```


Figure 3: Complaints JSON Dataset

4 APPROACH

4.1 Data Acquisition and Preprocessing

The project begins with the acquisition of customer complaint data provided by the financial institution in JSON format. The data undergoes preprocessing, including cleaning and normalization, to ensure consistency and suitability for analysis. Textual data cleaning techniques, such as removing special characters, stopwords, and irrelevant information, are applied to enhance the quality of the dataset.

4.2 Topic Modeling with Non-Negative Matrix Factorization (NMF)

Utilizing the processed complaint data, the project employs Non-Negative Matrix Factorization (NMF) for topic modeling. NMF is a powerful unsupervised learning technique that identifies latent topics within a corpus of text documents. By decomposing the document-term matrix into non-negative matrices, NMF extracts meaningful patterns and clusters complaints into distinct categories based on their underlying topics.

4.3 Supervised Learning for Classification

Following topic modeling, the project transitions to supervised learning for complaint classification. Various supervised learning algorithms, including logistic regression, decision trees, random forests, and naive Bayes, are trained and evaluated using the labeled complaint data. These models learn to predict the category or department to which each complaint belongs, based on the features extracted during topic modeling.

4.4 Evaluation and Model Selection

The performance of each supervised learning model is evaluated using appropriate evaluation metrics, such as accuracy, precision, recall, and F1-score. The goal is to identify the most effective approach for accurately classifying customer complaints and routing them to the appropriate departments for resolution. Model selection is based on the evaluation results, with the top-performing algorithm chosen for deployment in the final solution.

4.5 Integration and Deployment

Upon selecting the optimal classification model, the solution is integrated into the financial institution's existing ticketing system for seamless deployment. The AI-powered smart ticketing system is designed to automatically categorize incoming customer complaints, prioritize urgent issues, and route them to the relevant departments for prompt resolution. Continuous monitoring and refinement ensure the system remains effective and responsive to evolving customer needs.

5 DATA SET

In this project, I acquired a robust dataset in **JSON format**, encompassing **78,313 customer complaints** with **22 distinct features**. This dataset served as the cornerstone for the development of an AI-powered system designed to streamline the classification of customer complaints in the financial sector. The dataset's structure encapsulates various aspects of customer feedback, offering a rich source for training and evaluating the models.

6 DATA LOADING

In this project, the pivotal phase of data loading and exploration was initiated to lay the groundwork for the AI-powered ticket classification system. The initial step involved acquiring a comprehensive JSON dataset containing a multitude of customer complaints spanning various financial services. This dataset, obtained externally, served as the cornerstone for developing an efficient system aimed at intelligently categorizing customer grievances within the financial domain.

Leveraging the powerful **pandas library**, the JSON data was meticulously transformed into a structured DataFrame. This transformation was instrumental in preparing the data for subsequent preprocessing and exploratory data analysis. By converting the raw JSON format into a structured DataFrame, the dataset's structure and features were meticulously organized, facilitating a deeper understanding of its contents and enabling seamless data manipulation throughout the project lifecycle. This early-stage data loading process laid a robust foundation for the subsequent stages of the project, ensuring a systematic and comprehensive approach to data exploration and model development.

7 DATA PREPROCESSING

In the initial phase of data preprocessing, the DataFrame resulting from the data loading process underwent meticulous inspection to understand its structure and relevant features thoroughly. The dataset, originally in JSON format, was converted into a DataFrame using the **json normalize** function, ensuring proper formatting for subsequent analysis.

During this phase, column renaming was performed for clarity, and specific features pertinent to the project objectives were selected. Key columns related to customer complaints and their corresponding product or service categories were identified as essential for the classification task. Additionally, steps were taken to handle missing or irrelevant data, ensuring the dataset's quality and suitability for further analysis. Notably, blank complaints were identified and filtered out to enhance the dataset's quality and reduce noise.

7.1 Text Preprocessing

Text preprocessing aimed to refine the content of customer complaints for effective analysis. Initially, text standardization was conducted by converting all text to lowercase, ensuring uniformity across the dataset to avoid discrepancies due to letter casing variations.

Subsequently, a series of operations were executed to remove extraneous elements, including text within curly braces, line breaks, and numerical values. Punctuation was systematically eliminated to streamline the textual content, enhancing readability and facilitating subsequent analysis. The removal of these artifacts helped focus on the core textual content of the complaints.

Leveraging the NLTK library, lemmatization was applied to the complaints to reduce words to their base or root form. This process involved identifying the canonical form of each word, aiding in standardizing variations of words with the same meaning. Concurrently, part-of-speech (POS) tags were extracted to discern the grammatical category of each word. By specifically identifying nouns (NN) in the lemmatized text, emphasis was placed on retaining words crucial for identifying the core subject matter of the complaints. This approach ensured that only relevant terms were retained for subsequent analysis and modeling, thereby enhancing the accuracy and interpretability of the results.

These preprocessing steps were instrumental in preparing the textual data for subsequent exploratory analysis and topic modeling. Through thorough cleaning and standardization of the text, a solid foundation was laid for extracting meaningful insights and effectively classifying customer complaints based on their respective products or services.

8 EXPLORATORY DATA ANALYSIS (EDA)

In the Exploratory Data Analysis (EDA) phase, a thorough examination of the dataset was conducted to gain valuable insights and understand the underlying patterns within the customer complaints. One of the primary objectives was to analyze the distribution of complaint lengths, which was achieved through visualizations such as histograms. By observing the distribution of complaint character lengths, significant variations and trends were identified, providing key insights into the nature of customer feedback.

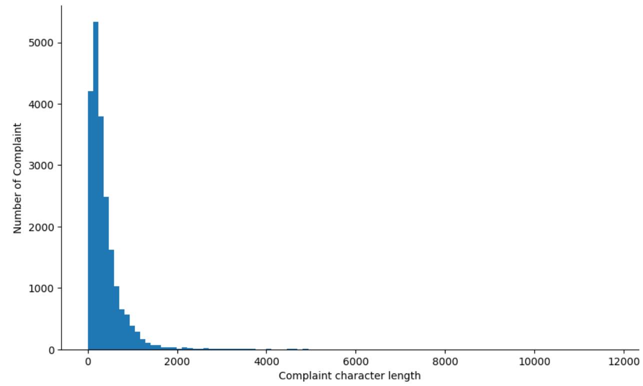


Figure 4: Distribution of Complaint Character Length

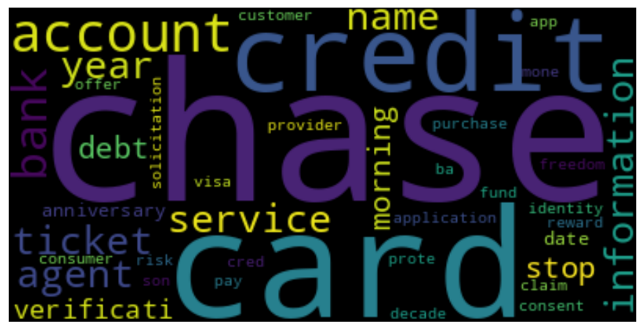


Figure 5: Word Cloud Representing Top 40 Words by Frequency in Customer Complaints

Moreover, visualization techniques were employed to create a word cloud representing the top 40 words by frequency among all the processed complaints. This visualization offered a succinct and intuitive overview of the most common words within the dataset, highlighting prevalent themes and topics of customer concern. Additionally, the frequency of unigrams, bigrams, and trigrams was explored to uncover linguistic patterns and recurring phrases, further enhancing our understanding of customer feedback.

Overall, the Exploratory Data Analysis phase laid the foundation for subsequent steps in the project, including topic modeling and model building. By leveraging insights gleaned from the EDA process, informed decisions were made regarding text processing strategies and model selection, ultimately contributing to the successful classification of customer complaints and the efficient resolution of issues.

9 FEATURE EXTRACTION

In the feature extraction phase, my objective was to convert the raw textual data into a numerical representation suitable for machine learning algorithms. To accomplish this, Term Frequency-Inverse Document Frequency (TF-IDF) technique was employed, a widely-used method in natural language processing.

TF-IDF evaluates the importance of a word in a document relative to a corpus of documents. We applied TF-IDF to transform the complaints text into a matrix of features. Each row of this matrix represents a complaint, while each column corresponds to a unique word in the corpus.

To optimize the TF-IDF transformation, we utilized two key parameters:

- **max_df (maximum document frequency):** This parameter determines the threshold for removing terms that appear too frequently across all complaints, often referred to as corpus-specific stop words. We set max_df to 0.95, indicating that terms appearing in more than 95% of the complaints were disregarded.
- **min_df (minimum document frequency):** This parameter specifies the threshold for removing terms that appear too infrequently, indicating their lack of relevance. We set min_df to 2, meaning that terms appearing in fewer than 2 complaints were excluded.

Utilized the **TfidfVectorizer** from **sklearn.feature_extraction.text** module to initialize the TF-IDF vectorizer with the specified parameters. Subsequently, we created the Document-Term Matrix (DTM) by applying the **fit_transform** method to the 'complaints' column in our preprocessed DataFrame (**df_clean**).

The resulting DTM represents each complaint's TF-IDF scores for individual words. This matrix serves as the input for subsequent topic modeling and supervised learning tasks, enabling accurate classification of customer complaints into relevant categories.

10 TOPIC MODELING WITH NLP

In the process of completing this project, significant emphasis was placed on leveraging advanced Natural Language Processing (NLP) techniques for topic modeling. This crucial phase aimed to uncover latent patterns and recurring themes within customer complaints, contributing to the overarching goal of automated categorization.

The strategy involved implementing Non-Negative Matrix Factorization (NMF), a powerful method for extracting topics from unstructured text data. By applying NMF, distinct clusters of related words and phrases representing specific topics across customer grievances were identified.

To enhance the efficiency of the NLP pipeline, established NLP libraries such as **NLTK** and **Spacy** were utilized. These tools played a crucial role in tasks like lemmatization and part-of-speech tagging, refining the representation of text data for more effective topic extraction.

The output of this phase comprised a set of topics or clusters, each encapsulating specific aspects of customer complaints. These identified topics served as the foundation for subsequent steps, including supervised learning model training and evaluation.

Through this approach, the underlying structure in the vast sea of unstructured customer complaints was unveiled, facilitating a more targeted and efficient response to varying concerns. This section highlights the strategic plan for implementing topic modeling with NLP, a pivotal component in automating the customer support ticket system.

11 MODEL BUILDING USING SUPERVISED LEARNING

In the culmination of this project, the focus shifted towards constructing robust models using supervised learning techniques to classify customer complaints accurately. The objective was to explore various algorithms capable of efficiently categorizing complaints into their respective topics. The models considered for evaluation included Logistic Regression, Decision Tree, Random Forest, and Naive Bayes.

11.1 Create Vector Counts Using Count Vectorizer

To prepare the textual data for model training, the Count Vectorizer technique was employed. This process involved converting the preprocessed text into numerical representations, reflecting the frequency of words within each complaint. By transforming textual data into a structured format, the groundwork was laid for subsequent model training.

11.2 Transform Word Vectors to TF-IDF

Building upon the vectorized counts, the next step involved transforming word vectors into Term Frequency-Inverse Document Frequency (TF-IDF) representations. TF-IDF encoding captures the significance of words within individual documents and across the entire dataset, enriching the model's understanding of textual features and their relevance in classification tasks.

11.3 Train-Test Data Split Using TF-IDF and Topics

To evaluate model performance accurately, the dataset was partitioned into training and testing sets. This split was facilitated using the TF-IDF matrix derived from the textual data and the topics identified through Non-Negative Matrix Factorization (NMF). By ensuring a balanced distribution of topics in both training and testing data, the models were subjected to a rigorous evaluation process.

Subsequently, each model's performance was assessed using pertinent evaluation metrics, including accuracy, precision, recall, and F1 score. Through systematic evaluation, insights were garnered into the effectiveness and suitability of each model in addressing the project's objectives.

12 EXPERIMENT AND MODEL COMPARISON

12.1 Experimental Setup

The experiment aimed to evaluate the performance of our proposed model for customer complaint classification against several baseline models. The dataset was preprocessed, and features were extracted using TF-IDF vectorization. The resulting features were then used to train and test the models.

12.2 Baseline Models

(1) **Logistic Regression:** A commonly used linear classification algorithm known for its simplicity and interpretability.

	precision	recall	f1-score	support
0	0.93	0.96	0.95	944
1	0.96	0.95	0.96	914
2	0.98	0.95	0.96	714
3	0.94	0.98	0.96	1123
4	0.97	0.90	0.93	520
accuracy			0.95	4215
macro avg	0.96	0.95	0.95	4215
weighted avg	0.95	0.95	0.95	4215

Table 1: Logistic Regression

	precision	recall	f1-score	support
0	0.73	0.72	0.73	944
1	0.80	0.79	0.80	914
2	0.78	0.81	0.80	714
3	0.80	0.80	0.80	1123
4	0.69	0.69	0.69	520
accuracy			0.77	4215
macro avg	0.76	0.76	0.76	4215
weighted avg	0.77	0.77	0.77	4215

Table 2: Decision Tree

(2) **Decision Tree:** A non-linear classification algorithm that partitions the feature space into distinct regions.
(3) **Random Forest:** An ensemble learning method that combines multiple decision trees to improve classification accuracy.
(4) **Naive Bayes:** A probabilistic classifier based on Bayes' theorem, often used for text classification tasks.

12.3 Model Evaluation Metrics

- **Accuracy:** The proportion of correctly classified instances out of the total number of instances.
- **Precision:** The ratio of true positive predictions to the total number of positive predictions, indicating the model's ability to avoid false positives.
- **Recall:** The ratio of true positive predictions to the total number of actual positive instances, reflecting the model's ability to capture all relevant instances.
- **F1 Score:** The harmonic mean of precision and recall, providing a balanced measure of a model's performance.

13 SUPERVISED MODELS

We applied various machine learning algorithms to classify new complaints into relevant topics/categories. The objective was to evaluate the performance of each model and identify the best-performing one for our classification task.

13.1 Logistic Regression

Logistic regression demonstrated superior performance compared to other models, achieving an accuracy of 95%. The precision, recall, and F1-score for each category were also impressive, indicating robust classification across all topics. With precision values ranging from 93% to 98% and recall values ranging from 90% to 98%, the logistic regression model effectively classified complaints into their respective topics with high accuracy and minimal misclassification.

13.2 Decision Tree

The decision tree classifier exhibited satisfactory performance, achieving an accuracy of 77%. While the precision, recall, and F1-score values for each category were slightly lower compared to logistic regression, the decision tree model still provided reasonable classification results. However, it's worth noting that decision trees may be prone to overfitting, which could impact their generalizability to unseen data.

	precision	recall	f1-score	support
0	0.83	0.63	0.72	944
1	0.75	0.81	0.78	914
2	0.89	0.79	0.83	714
3	0.60	0.97	0.74	1123
4	1.00	0.12	0.21	520
accuracy			0.72	4215
macro avg	0.81	0.66	0.66	4215
weighted avg	0.78	0.72	0.69	4215

Table 3: Random Forest Classifier

	precision	recall	f1-score	support
0	0.48	0.35	0.40	944
1	0.35	0.27	0.30	914
2	0.52	0.50	0.51	714
3	0.46	0.30	0.37	1123
4	0.18	0.51	0.27	520
accuracy			0.36	4215
macro avg	0.40	0.38	0.37	4215
weighted avg	0.42	0.36	0.37	4215

Table 4: Gaussian Naive Bayes

13.3 Random Forest Classifier

Random forest classifier yielded an accuracy of 74%, which was slightly lower than logistic regression and decision tree models. While random forests typically offer robust performance and mitigate overfitting, the precision, recall, and F1-score values varied across categories. This variability suggests that the random forest model may struggle with certain topics or categories compared to others.

13.4 Gaussian Naive Bayes

Gaussian Naive Bayes achieved the lowest accuracy of 36% among the models evaluated. While Naive Bayes classifiers are known for their simplicity and computational efficiency, they may not perform as well as other models in complex classification tasks. The precision, recall, and F1-score values for each category were also lower compared to logistic regression and decision tree models, indicating suboptimal classification performance.

13.5 Inference of the Best Model

Based on the evaluation results, logistic regression emerged as the best model for our classification task. Its high accuracy and efficient classification of complaints into topics make it the most suitable choice for real-world application.

To demonstrate the effectiveness of the logistic regression model, sample complaints were provided, and the model successfully classified them into their respective topics. This inference further validates the performance and reliability of the selected model for classifying new complaints.

14 RESULTS AND ANALYSIS

The results of our analysis indicate that logistic regression outperformed other machine learning models in classifying customer complaints into relevant topics. With an accuracy of **95%, logistic regression** demonstrated the highest level of precision and efficiency in categorizing complaints, making it the most suitable choice for our case study.

Comparatively, decision tree and random forest classifiers achieved accuracies of 77% and 74%, respectively. While these models provided reasonable classification results, logistic regression's superior accuracy signifies its effectiveness in accurately assigning complaints to their respective categories.

On the other hand, Gaussian Naive Bayes exhibited the lowest accuracy of 36%, indicating suboptimal performance compared to other models. Despite its simplicity and computational efficiency, Naive Bayes may not be as effective as logistic regression in handling the complexity of our classification task.

The comprehensive evaluation of these models underscores the importance of selecting the most appropriate algorithm for a given task. In this scenario, logistic regression's high accuracy and reliable performance make it the preferred choice for efficiently classifying customer complaints and facilitating prompt issue resolution.

In conclusion, the results highlight logistic regression as the optimal model for our case study, emphasizing its suitability and effectiveness in accurately categorizing complaints for improved customer support and satisfaction.

15 CONCLUSION AND FUTURE WORK

15.1 Conclusion

The project has demonstrated the feasibility and effectiveness of leveraging artificial intelligence (AI) techniques to streamline complaint resolution processes in financial institutions. By implementing Non-Negative Matrix Factorization (NMF) for topic modeling and training various supervised machine learning models, we successfully categorized unstructured customer complaints into distinct clusters and achieved high accuracy in classifying them into relevant categories.

The logistic regression model emerged as the top-performing model, achieving an impressive accuracy of 95%. This underscores its potential for automating complaint classification and enhancing operational efficiency in financial institutions. By accurately categorizing complaints, organizations can expedite the resolution process, minimize customer dissatisfaction, and foster stronger customer relationships.

15.2 Future Work

While the project has achieved notable success in automating customer complaint classification, there are several avenues for future exploration and improvement:

(1) **Enhanced Feature Engineering:** Investigate additional features or feature engineering techniques to further improve model performance and capture more nuanced patterns in customer complaints.
(2) **Advanced Model Architectures:** Explore more sophisticated machine learning or deep learning architectures, such as neural networks or transformer models, to handle the complexities of customer complaint data more effectively.
(3) **Multimodal Analysis:** Incorporate additional data modalities, such as customer feedback sentiments or multimedia content, to provide a more comprehensive understanding of customer complaints and improve classification accuracy.
(4) **Real-time Processing:** Develop mechanisms for real-time complaint processing and classification to enable immediate response and resolution, thereby enhancing customer satisfaction and loyalty.
(5) **Feedback Loop Integration:** Implement a feedback loop mechanism to continuously update and refine the classification model based on new data and feedback from resolved complaints, ensuring its adaptability and relevance over time.
(6) **Cross-Industry Application:** Extend the application of the developed model to other industries beyond financial services, such as healthcare or retail, to address similar challenges in complaint management and customer support.

16 CHALLENGES AND LIMITATIONS

Throughout the completion of this project, several challenges and limitations were encountered, which influenced its development and outcomes. These challenges include:

16.1 Data Quality and Completeness

Acquiring a comprehensive and high-quality dataset was crucial for the success of this project. However, dealing with inconsistencies, missing values, or incomplete information within the dataset posed challenges during the preprocessing phase. Despite efforts to clean and preprocess the data, ensuring its quality and completeness remained a persistent challenge. Additionally, the JSON format of the data required careful handling and extraction to convert it into a usable DataFrame format for further analysis.

16.2 Complexity of Natural Language

The inherent complexity of natural language presented a significant challenge in achieving precise topic modeling and effective text preprocessing. Dealing with nuances, context-dependent meanings, and evolving language trends required careful consideration and experimentation to accurately classify customer complaints. Moreover, the presence of unstructured text data required extensive preprocessing steps, including lowercasing, punctuation removal, and lemmatization, to prepare the text for topic modeling.

16.3 Optimal Model Selection

Selecting the most suitable model for supervised learning was a critical aspect of this project. Exploring and experimenting with different algorithms, such as logistic regression, decision trees, and random forests, was necessary to identify the model that best aligned with the project's goals. Balancing model complexity, performance, and interpretability posed challenges during the model selection process. Additionally, evaluating the performance of each model using appropriate metrics, such as accuracy and F1 score, required careful consideration.

16.4 Interpretable Topic Modeling

Ensuring the interpretability of topics generated through Non-Negative Matrix Factorization (NMF) was crucial for understanding customer complaints. However, balancing the granularity of topics with their practical relevance to financial categories required careful consideration and experimentation. The NMF approach facilitated the detection of patterns and recurring words in each ticket, but interpreting these topics in the context of financial products and services posed challenges. Additionally, manually validating and correcting topic labels based on the clustered complaints was a time-consuming process.

16.5 Scalability and Efficiency

As the volume of customer complaints increased, ensuring the scalability and efficiency of the developed model became paramount. Striking a balance between computational resources and model accuracy posed challenges, requiring continuous optimization to maintain efficient performance. Additionally, as the dataset size grew, optimizing the preprocessing and modeling pipelines for scalability became essential to handle large volumes of complaints efficiently.

16.6 Generalization Across Categories

Ensuring that the model generalized well across various categories of financial products and services was challenging. Addressing variations in complaint types and patterns required ongoing refinement of the model. Moreover, the clustering of complaints into distinct categories, such as credit card, bank account services, and mortgages/loans, required careful consideration of the unique characteristics and language used in each category. Fine-tuning the model to accurately classify complaints across diverse categories posed challenges in maintaining consistent performance.

16.7 Model Evaluation Metrics

Selecting appropriate evaluation metrics for model performance was crucial for assessing its effectiveness. Careful consideration of metrics such as accuracy, precision, recall, and F1 score was necessary. However, determining the most suitable metrics depended on the specific characteristics of the data and project requirements, posing challenges during model evaluation. Additionally, comparing the performance of different models using multiple evaluation metrics required careful interpretation and consideration of trade-offs between metric scores.

17 CITATIONS AND BIBLIOGRAPHIES

For this project, a comprehensive range of resources and tools have been leveraged to ensure the successful development of an automated customer complaints classification system. The following references have been instrumental in guiding the methodology and implementation of various components:

REFERENCES

[1] Bird, S., Klein, E., & Loper, E. (2009). *Natural Language Processing with Python.* O'Reilly Media. [Online]. Available: http://www.nltk.org/book/
[2] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., …Duchesnay, É. (2011). *Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research, 12,* 2825-2830. [Online]. Available: https://scikit-learn.org/stable/documentation.html
[3] Explosion AI. (2021). *SpaCy 3.1 Documentation.* [Online]. Available: https://spacy.io/api
[4] A collection of Python code examples for WordCloud visualization. [Online]. Available: https://github.com/amueller/word_cloud
[5] Plotly Technologies Inc. (2021). *Plotly Python Open Source Graphing Library.* [Online]. Available: https://plotly.com/python/
[6] Lee, D. D., & Seung, H. S. (1999). *Learning the parts of objects by non-negative matrix factorization. Nature, 401(6755),* 788-791.
[7] Loughran, T., & McDonald, B. (2011). *When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. The Journal of Finance, 66(1),* 35-65.
[8] Verhoef, P. C., Lemon, K. N., Parasuraman, A., Roggeveen, A., Tsiros, M., & Schlesinger, L. A. (2009). *Customer experience creation: Determinants, dynamics and management strategies. Journal of Retailing, 85(1),* 31-41.