

# **Improving Southwest Airlines Customer Service**

Building a text analytics model to automatically classify customer inquiries on Twitter


# Problem Statement

- Southwest Airline received large volume of customer inquiries due to their system breakdown in Dec 2022
- Not responded to customers' inquiries effectively resulted in poor customer experience and loss of business and trust

The New York Times

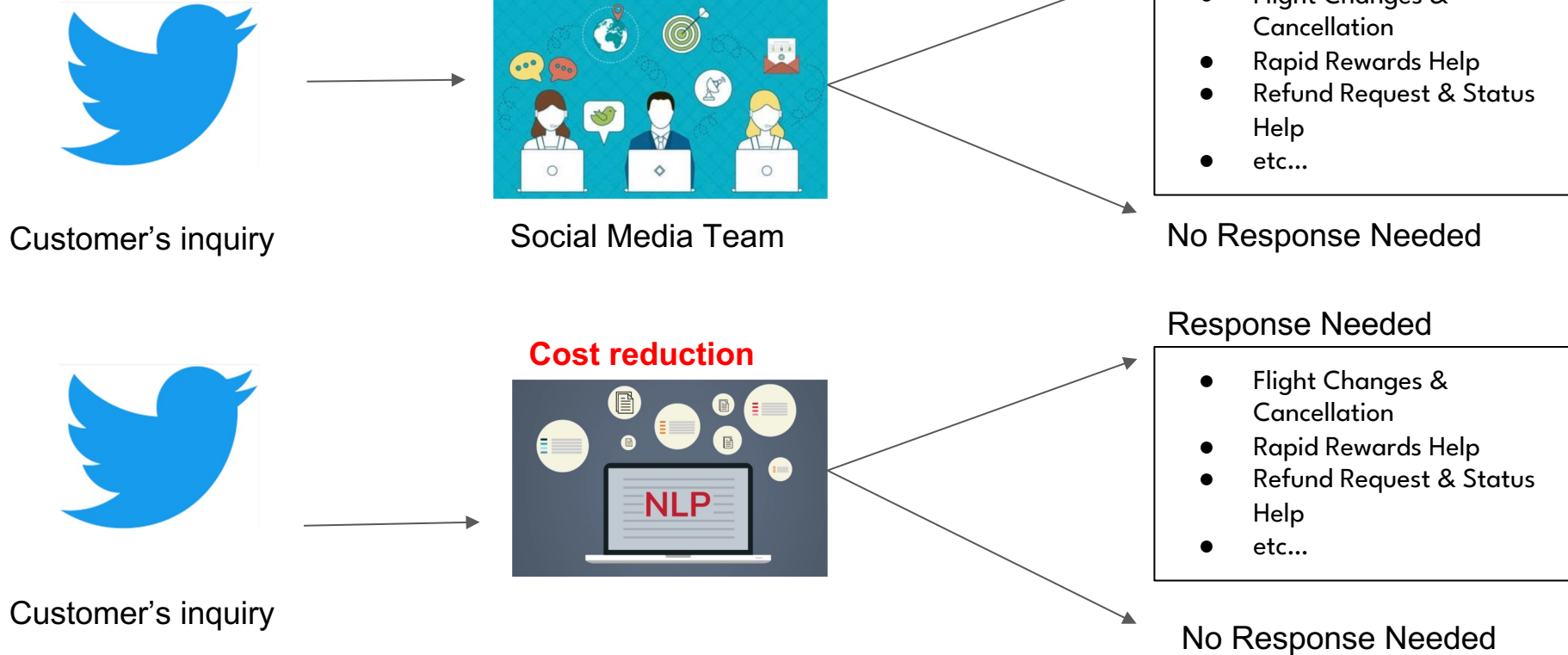
## ***Southwest Says Holiday Meltdown Will Cost It More Than \$1 Billion***

The airline, which canceled 16,700 flights before and after Christmas, said it lost \$220 million in the last three months of the year.

 Give this article



# Business Goal Analysis

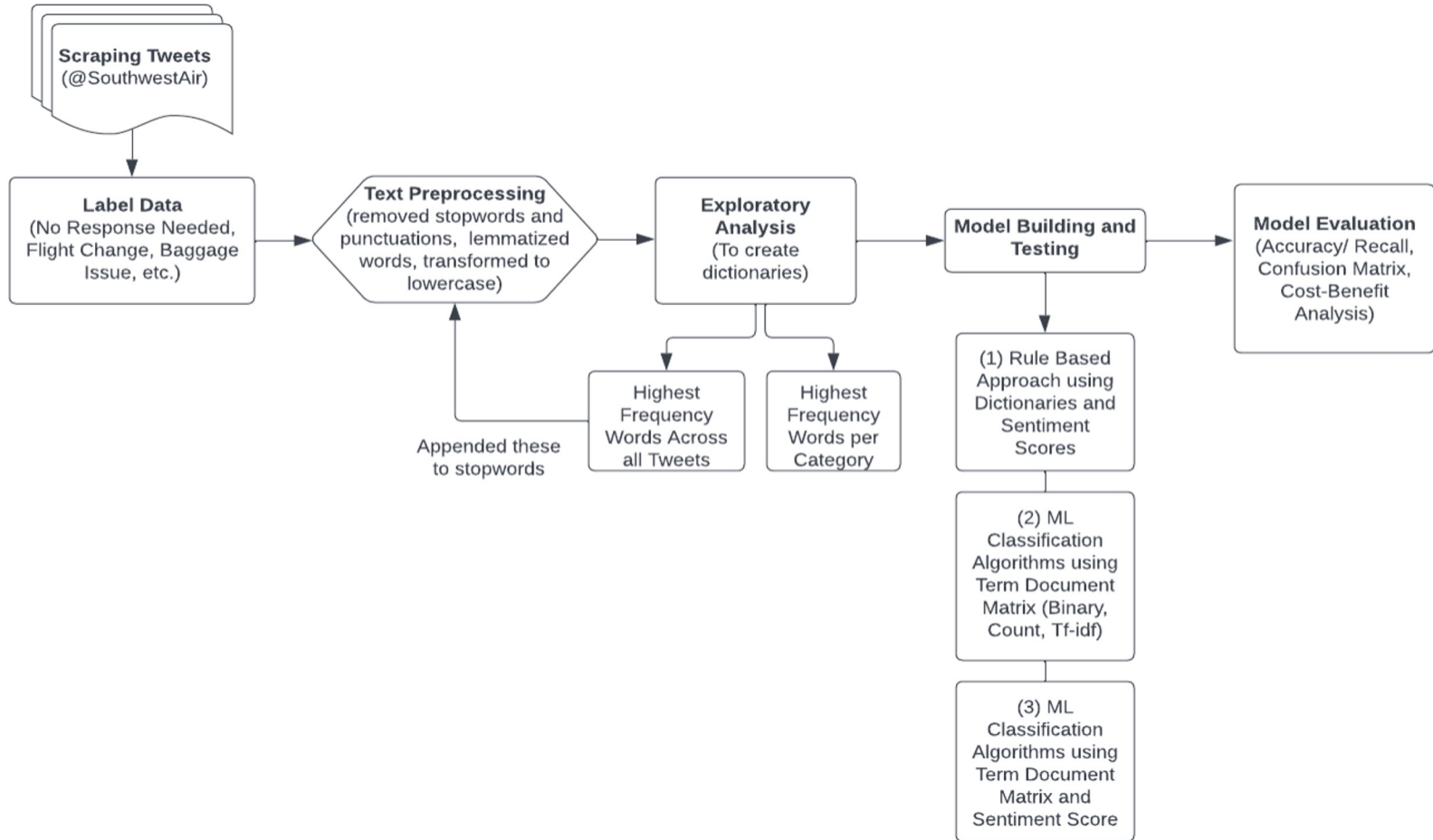


# Data Description

- Approx. 3000 tweets collected by scraping Twitter using the Twitter API with a @SouthwestAir 'tag'
- Jan 2023 - Feb 2023
- Stored in a CSV file
- Comprised of tweets pertaining to
  - Customer inquiries
  - Positive and negative reviews



# System Design



# System Implementation: Extraction to Preprocessing

## /Data Extraction: Python

Web Scrapping to Extract Tweets from Twitter API

## /Annotation: Excel

Manually Categorize Tweets into 8 Classes

- Flight Changes & Cancellation
- Rapid Rewards Help
- Refund Request & Status Help
- Others - help needed
- Baggage Help
- Seats Help
- Covid 19
- **No Response Needed**

## /Text Pre-Processing: Python

### Exploratory Analysis

/Top 30 Frequent Words

- Stopwords Elimination
- Special Character Elimination
- Frequent Words Elimination
- Tokenization
- Lemmatization

### Exploratory Analysis

/Top 30 Frequent Nouns

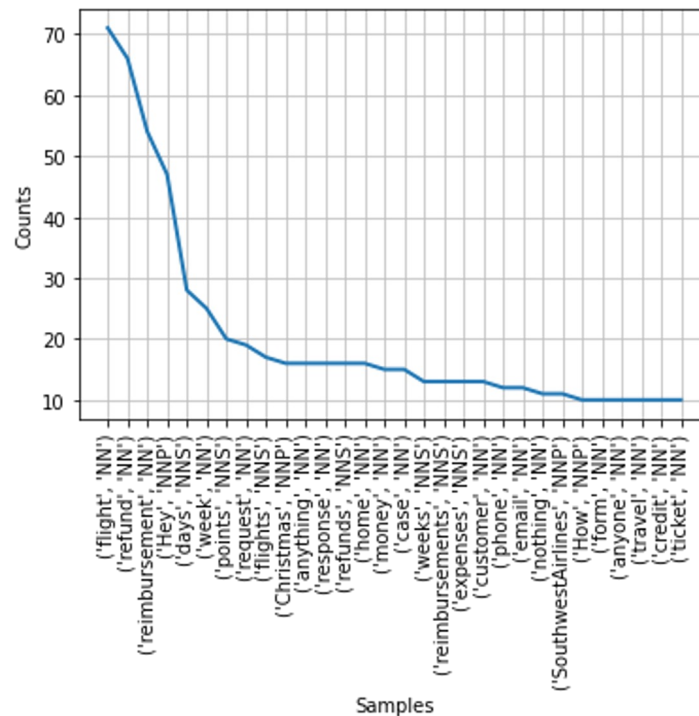
- Stopwords Elimination
- Special Character Elimination
- Frequent Words Elimination
- Tokenization

### Sentiment Analysis

- Special Character Elimination
- Adding “?” Back to Text
- Tokenization

## System Implementation: Exploratory Analysis and Dictionary Building

# /Exploratory Data Analysis: Python

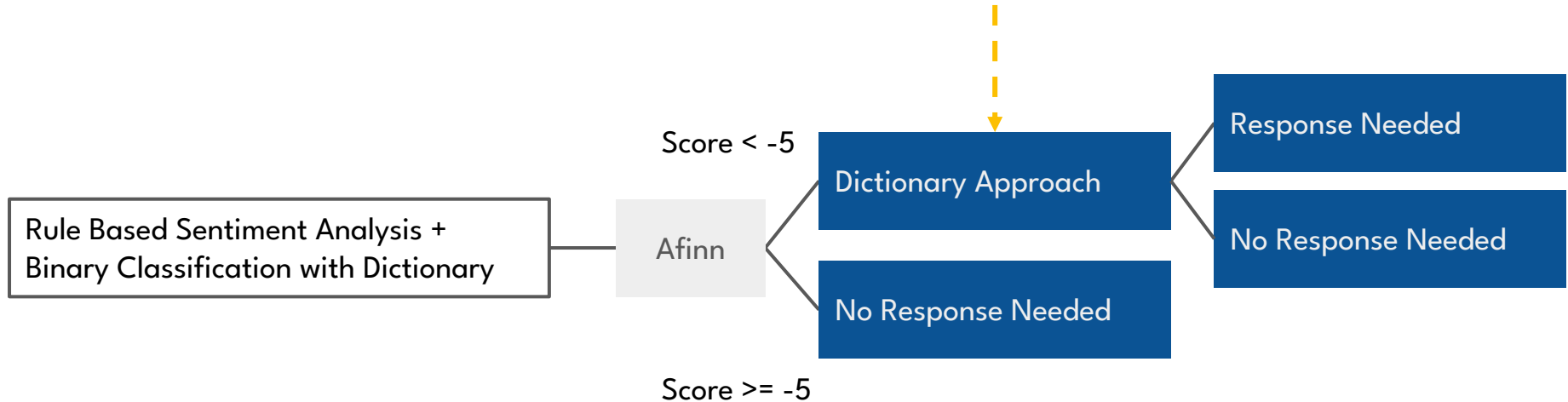


# System Implementation: Method 1 (Rule-Based Approach)

{"luggage", "bag", "baggage",  
"suitcase", "bags"}

{"find", "lost", "lose", "missing",  
"damaged", "tracking"}

**Baggage Help = luggage\_related AND baggage\_issues**





# System Implementation: Methods 2 (Term Document Matrix) and Method 3 (Term Document Matrix + Sentiment Score)

Method2:

Feature Engineering + Machine Learning

- Bag of Words - Binary
- Bag of Words - Frequency
- Tf-idf



- Naive Bayes
- Decision Tree
- Random Forest
- Logistic Regression
- SVM
- XGBoost

Method3:

Feature Engineering + Machine Learning  
+ Sentiment Score

Sentiment  
Score



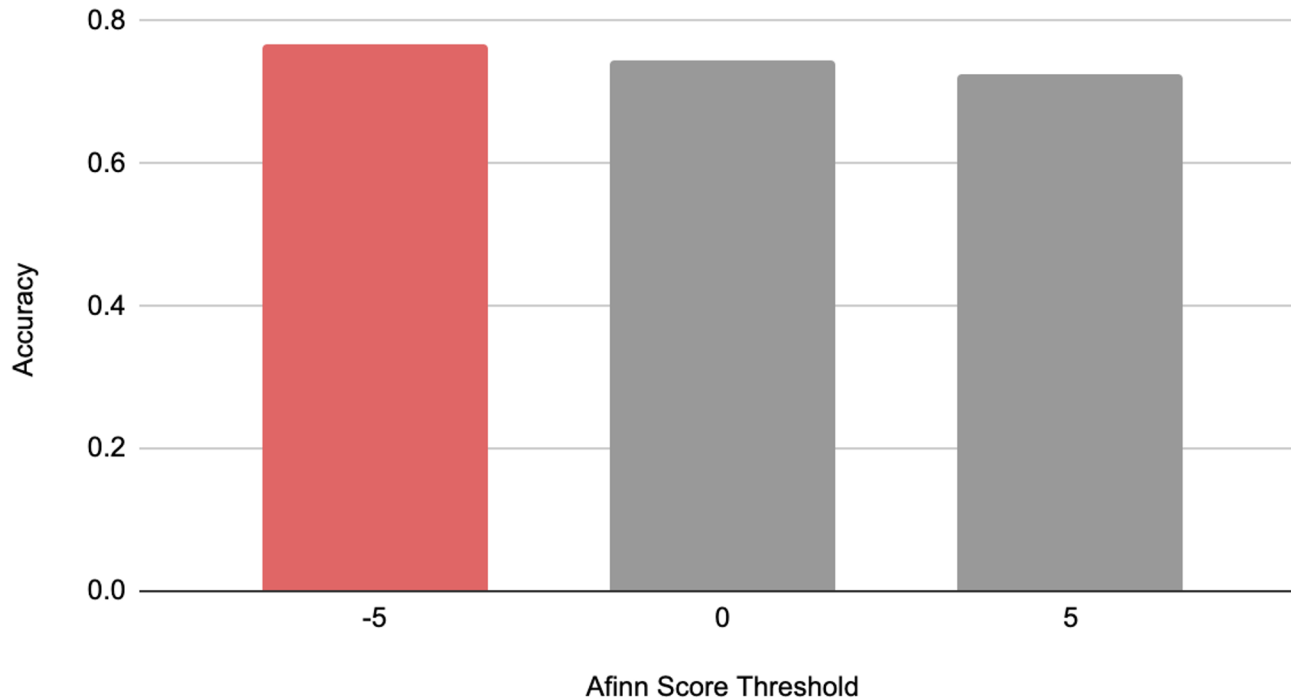
- Bag of Words - Binary
- Bag of Words - Frequency
- Tf-idf



- Naive Bayes
- Decision Tree
- Random Forest
- Logistic Regression
- SVM
- XGBoost

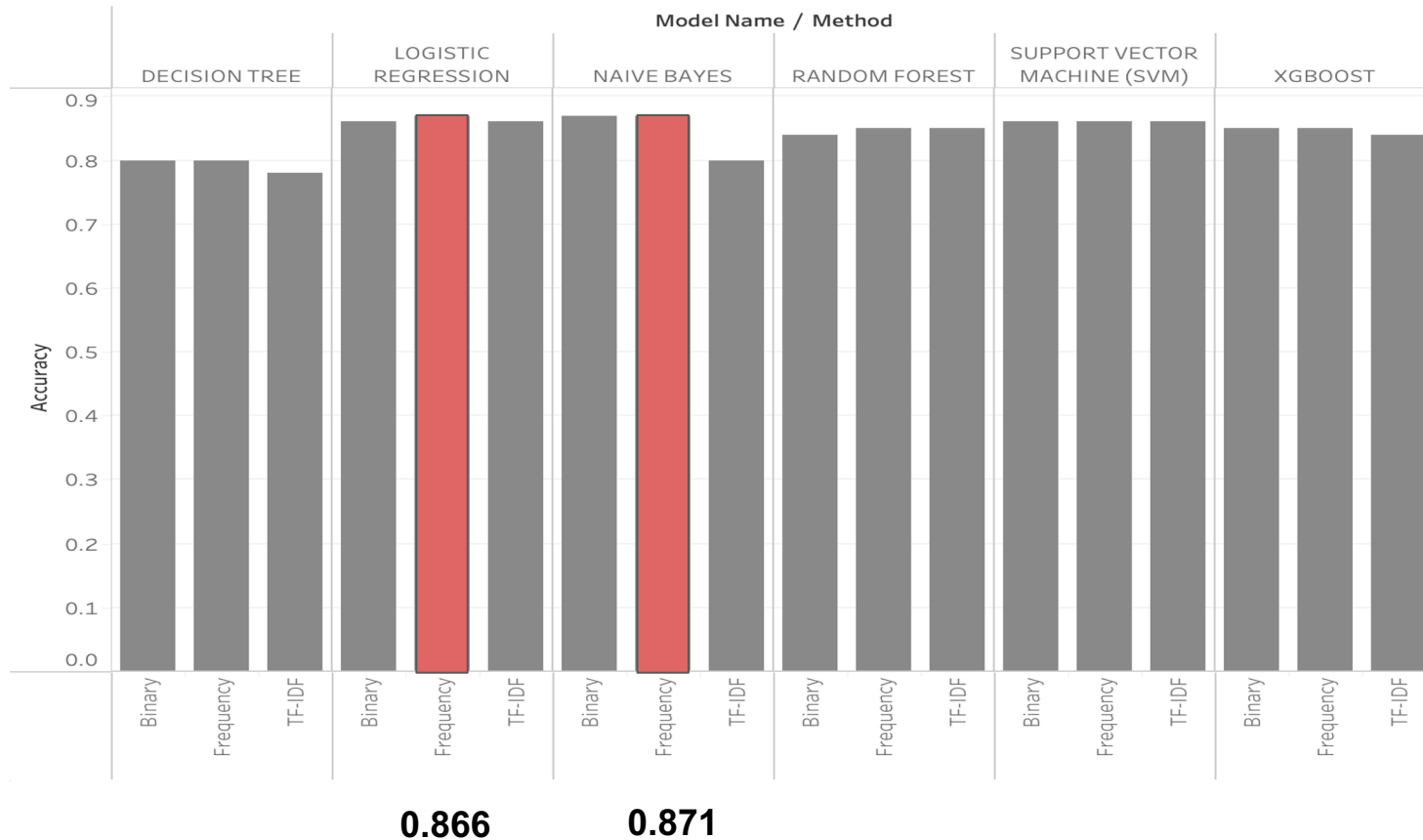
# Method 1 (Sentiment Analysis with Dictionary Approach) :

(Binary) Accuracy vs. AFINN Score Threshold

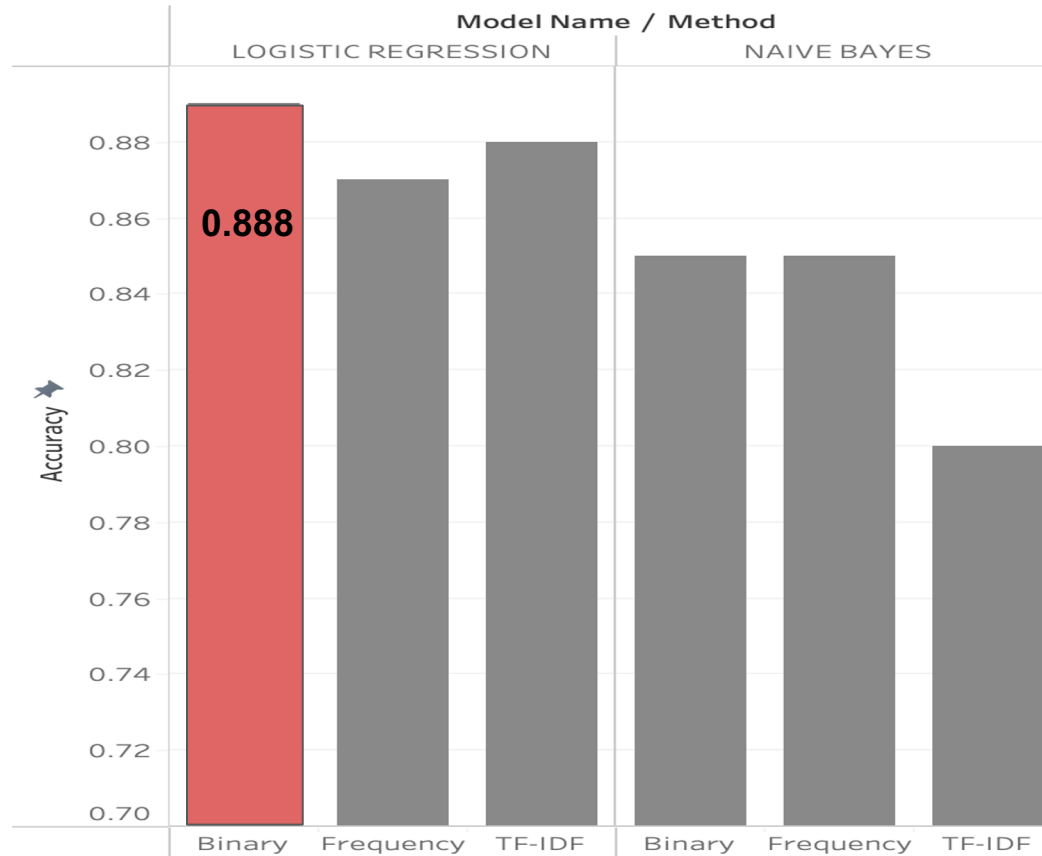


# Methods 2 (Term Document Matrix) :

Overall Accuracy



## Methods 3 (Method 2 + Sentiment Score) :



## Conclusion & Findings :

- Our logistic regression model with term document matrix and sentiment score features outperformed the other two methods
- Relying solely on keywords is not sufficient for accurately classifying customer inquiries.
- Customizing the text pre-processing steps based on the chosen text analytics approach is important for text analytics project
- The urgency of ticket inquiries can be changed over time, which adds another layer of complexity to the classification task

	Accuracy
Method1	0.767
Method2	0.871
<b>Method3</b>	<b>0.888</b>

## **Future Direction :**

- Expanding the model to include multiple classes that represent specific inquiry topics
- Incorporating more advanced natural language processing (NLP) techniques such as transformer-based models (e.g. BERT)
- Exploring the use of ensemble methods to combine the predictions from multiple models
- Using multiple evaluation metrics such as precision, recall, and F1-score, in addition to accuracy