

# **PROJECT REPORT**

## **Project Title : AI-Powered Public Service Feedback Management**

- **Team Members :**

1. KAPIL SINGHAL (23MIP10111)
2. MUDIT AGRAWAL (22BCE11128)
3. UTKARSH (24BSA10039)
4. VAISHNAVI (23MIP10125)

- **Room no. :- 017**
- **Team no. :- 26**

## **Introduction**

Public service departments handle numerous queries daily, but manual processing often leads to delays and errors. To address this, we propose an **AI-based solution** that automates query management by:

- Detecting spam,
- Classifying query domains,
- Analyzing sentiment, and
- Assigning priority levels.

This system improves response efficiency, reduces workload, and enhances user satisfaction by ensuring timely query resolution.

## **Problem Statement**

- ☐ High volume of unstructured feedback.
- ☐ Manual processing is inefficient.
- ☐ Difficulty in prioritizing urgent queries.
- ☐ Lack of automated spam detection.

# Objective

- Develop an AI-based solution to **automate query management**.
- Improve response time and enhance user satisfaction.
- Categorize and prioritize queries efficiently.

# Proposed Solution

An AI-driven model with four core functionalities:

## 1. Spam Detection :

- **Step 1:** Receive user query.
- **Step 2:** Apply **Spam Detection Model** to classify the query:
  - **Spam:** Discard and no further action.
  - **Not Spam:** Proceed to next steps for detailed analysis:
    - **Domain Classification:** Identify if query is related to electricity, water, or gas.
    - **Sentiment Analysis:** Gauge user satisfaction (Positive, Neutral, Negative).
    - **Priority Assignment:** Set urgency levels (Low, Medium, High).

## DATASET :

Query	SPAM/NOT SPAM
traveling from.amarpali express no water in toilet.called so many time 182 BT call nt s7	NOT SPAM
My train has been delayed for more than 5 hours.	NOT SPAM
very dirty toilet in coach B5 of 12509 exp. Pls pls pls do something. Pnr 4120563680. Divyanshu Mob.7847932526	NOT SPAM
Train No. 15107 Coach BE1, AC not working from Barabanki & no one is listening to our complain.	NOT SPAM
...13134 coach no. B1.....20-30 people without ticket are in this coach....jbrdsti baith rhe...tt ki bt v nhi man rhe...	NOT SPAM
- train delays and very dirty lavatory area. expected better after 2 years in ministry	NOT SPAM
The washrooms in the train are extremely dirty and unusable.	NOT SPAM
- Water pipes in Ramanathapuram Railway Station (RMD) is damaged . Passengers suffer without drinking water supply	NOT SPAM
My reserved seat was occupied by someone else and no action was taken.	NOT SPAM
I'm a passenger in Aleppey-Chennai Superfast train. Hoping to catch at least last bus. :(#indianrailwaydelay	NOT SPAM
my uncle travelling to train no 15656 coach no-B2. He found that the toilet of this coach is very dirty.Plz do something.	NOT SPAM
Petition all you like, if continually cut PSO and under invest in public transport the travellâ	NOT SPAM
Maximum train stoppage in bharuch.. We want also stoppage of more train in evening time.. Aâ	NOT SPAM
There is no announcement about the change in platform.	NOT SPAM
How it possible in Tatkal that all the 256 Available Garib Rath seats r booked within 1.5 min.? Is there any hidden tricks?	NOT SPAM
The AC is not train cooling.	NOT SPAM

## Preprocessing Summary

- **Duplicate Removal:** Identified and removed duplicate queries.
- **Null Value Handling:** Checked for null values and filled them with statistical analysis (mean/mode/median).
- **Text Cleaning:** Lowercase conversion, punctuation removal, etc.

## Model Selection for Spam Detection

### Applied 4 ML Algorithms:

- **Logistic Regression:** 95.63%
- **Naive Bayes:** 95.63%
- **SVM (Support Vector Machine):** 96.37%
- **Random Forest:** 96.09%

### Best Model for Spam Detection:

- **SVM (Linear Kernel)** remains the best choice with **96.37% accuracy**

## 2. Domain Classification

- **Step 1:** Input passes through Domain Classification Model.
- **Step 2:** Classify query into relevant domains:
  - **Electricity** – Queries about power outages, billing issues, etc.
  - **Water** – Complaints about water supply, leakage, etc.
  - **Gas** – Queries related to gas leaks, connection issues, etc.

### DATASET :

Query	Sub-topic	Domain
traveling from.amarpali express no water in toilet.called so many time 182 BT call nt s7	other	Railway Management
My train has been delayed for more than 5 hours.	reservation problem	Railway Management
very dirty toilet in coach B5 of 12509 exp. Pls pls pls do something. Pnr 4120563680. Divyanshu Mob.7847932526	ticket issue	Railway Management
Train No. 15107 Coach BE1, AC not working from Barabanki & no one is listening to our complain.	coach issue	Railway Management
...13134 coach no. B1.....20-30 people without ticket are in this coach....jbrdsti baith rhe...tt ki bt v nhi man rhe...	other	Railway Management
- train delays and very dirty lavatory area. expected better after 2 years in ministry	punctuality issue	Railway Management
The washrooms in the train are extremely dirty and unusable.	crowd management	Railway Management
- Water pipes in Ramanathapuram Railway Station (RMD) is damaged . Passengers suffer without drinking water supply	other	Railway Management
My reserved seat was occupied by someone else and no action was taken.	train cancellation	Railway Management
I'm a passenger in Aleppey-Chennai Superfast train. Hoping to catch at least last bus. :(#indianrailwaydelay	punctuality issue	Railway Management
my uncle travelling to train no 15656 coach no-B2. He found that the toilet of this coach is very dirty.Plz do something.	other	Railway Management
Petition all you like, if continually cut PSO and under invest in public transport the travellâ	other	Railway Management
Maximum train stoppage in bharuch.. We want also stoppage of more train in evening time.. Aâ	other	Railway Management
There is no announcement about the change in platform.	crowd management	Railway Management
How it possible in Tatkal that all the 256 Available Garib Rath seats r booked within 1.5 min.? Is there any hidden tricks	ticket issue	Railway Management
The AC is not train cooling.	Maintenance issue	Unknown
Sir plz clarify if No Berth opted for a child passenger but half fare paid..Is the child entitled to meals like copassengers	other	Railway Management
No proper water supply in wash basin. Also the area is unclean PNR 84 03 995416 kindly assist.	ticket issue	Railway Management
The station staff is very unhelpful and rude.	crowd management	Railway Management

## Preprocessing Summary

- **Duplicate Removal:** Removed duplicate domain-labeled queries.
- **Null Value Handling:** Filled missing values using mode for categorical data.
- **Text Cleaning:** Lowercase conversion, stopword removal, and vectorization.

## Model Selection for Domain Classification

- Applied 4 ML Algorithms:
  - **Logistic Regression:** 62%
  - **Naive Bayes:** 64.2%
  - **SVM:** 64%
  - **Random Forest:** 65.38%

## Best Model for Domain Classification

- **Random Forest** with **65.38% accuracy** performs best for classifying query domains.

## 3. Sentiment Analysis

- **Step 1:** Analyze query feedback to identify sentiment:
  - **Positive:** User is satisfied.
  - **Neutral:** Feedback is neutral or unclear.
  - **Negative:** User is dissatisfied.

**DATASET :**

A	B	C	D
Word/Query	Good	Bad	Extreme Bad
a+	1	0	0
abound	1	0	0
abounds	1	0	0
abundance	1	0	0
abundant	1	0	0
accessable	1	0	0
accessible	1	0	0
acclaim	1	0	0
acclaimed	1	0	0
acclamation	1	0	0
accolade	1	0	0
accolades	1	0	0
accommodative	1	0	0
accomodative	1	0	0
accomplish	1	0	0
accomplished	1	0	0
accomplishment	1	0	0
accomplishments	1	0	0
accurate	1	0	0
accurately	1	0	0

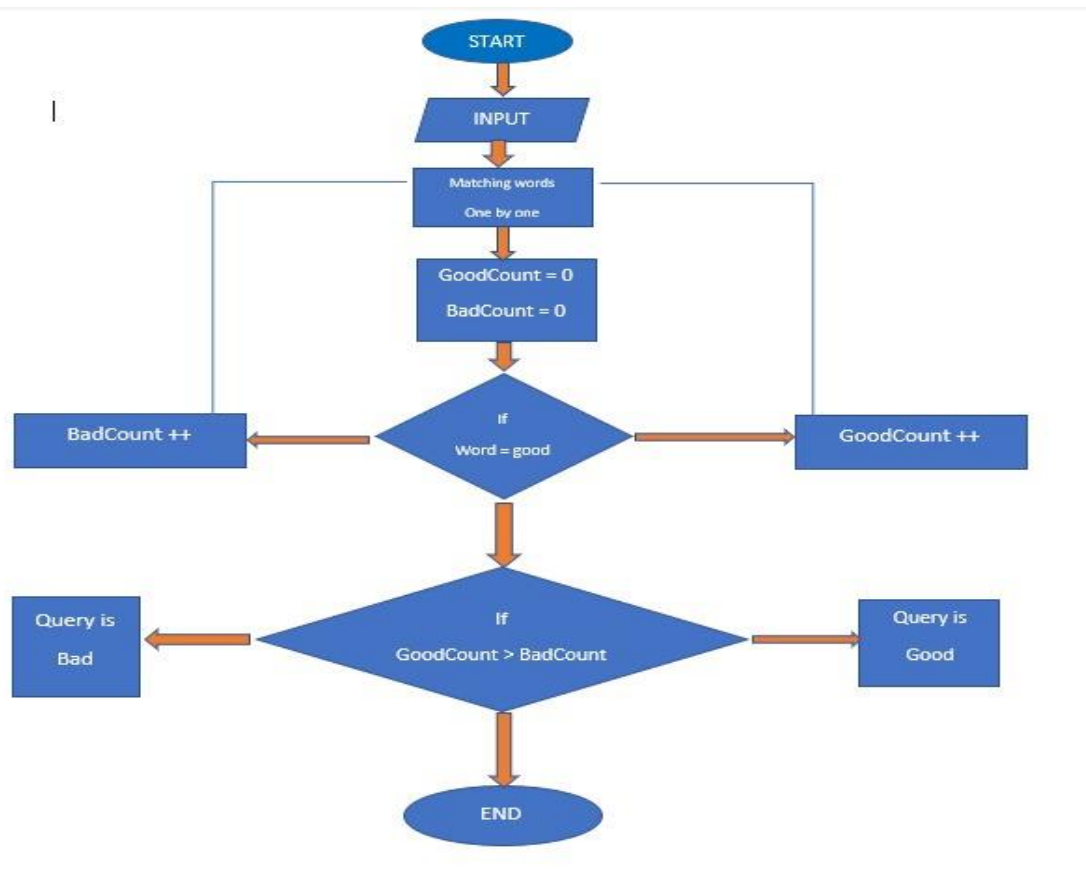
## Heuristic Approach

- **Step 1:** Analyze query feedback to identify sentiment using a **heuristic-based NLP approach**.
- **Step 2:** Count occurrences of **positive and negative words**:
  - **Positive Words:** Includes terms like good, satisfied, excellent, etc.
  - **Negative Words:** Includes terms like bad, poor, dissatisfied, etc.

## Heuristic Methodology

- **Word Matching:** Compare query text with predefined positive and negative word lists.
- **Score Calculation:**
  - Goodcount will increase 1 for each positive word.
  - Badcount will increase 1 for each negative word.
- **Sentiment Decision:**
  - **Positive Sentiment:** If Goodcount > badcount.
  - **Negative Sentiment:** If Badcount > Goodcount.
  - **Neutral Sentiment:** If scores are balanced or inconclusive.

## FLOWCHART :



## 4. Priority Assignment

- **Step 1:** Assign priority levels to queries:
  - **Low Priority:** Non-critical issues.
  - **Medium Priority:** Requires attention but not urgent.
  - **High Priority:** Emergency situations requiring immediate action.

### Preprocessing Summary

- **Duplicate Removal:** Removed duplicate priority labels.
- **Null Value Handling:** Filled missing priority values using median.
- **Text Cleaning:** Standard text preprocessing applied.

### Model Selection for Priority Assignment

- Applied 4 ML Algorithms:
  - **Logistic Regression:** 60.9%
  - **Naive Bayes:** 62.6%
  - **SVM:** 65%
  - **Random Forest:** 68%

### Best Model for Priority Assignment

- **Random Forest** achieves the highest accuracy with **68%** for priority prediction.

## Core Components Used

- **Natural Language Processing (NLP):** Extracts and interprets textual data from user queries.
- **Spam Detection Model:** Identifies and removes irrelevant or fake queries.
- **Domain Classifier Model:** Categorizes queries into relevant public service domains (electricity, water, gas).
- **Sentiment Analysis Engine:** Assesses user satisfaction and emotional tone.
- **Priority Assignment System:** Ranks query urgency to ensure timely resolution.
- **Machine Learning (ML) Algorithms:** Trains models for classification, sentiment analysis, and priority assignment.

## Potential Impacts

- ❑ **Improved Service Efficiency:** Faster query resolution reduces manual workload.
- ❑ **Enhanced User Satisfaction:** Timely responses and better prioritization boost public trust
- ❑ **Reduced Spam and Noise:** Filters out irrelevant queries, focusing on real issues.
- ❑ **Resource Optimization:** Allows service departments to allocate resources based on priority.

## Future Scope

- **Municipal Services:** Automate complaints related to sanitation, waste management, and civic issues.
- **Healthcare:** Handle patient queries, appointment scheduling, and emergency responses.
- **Law Enforcement:** Analyze public reports, detect suspicious patterns, and prioritize emergency responses.
- **Transportation and Traffic Management:** Manage complaints about delays, route issues, and public transport services.
- **Disaster Management:** Prioritize queries during floods, earthquakes, or other emergencies.
- **Education Sector:** Process student queries, feedback, and admission-related inquiries.
- **Utility Services:** Resolve complaints related to electricity, water, and gas supply.
- **E-Governance Platforms:** Improve citizen engagement by automating feedback and grievance redressal.

## Results and Evaluation

```
Enter Query:My train have been delayed for more than 5 hours
```

```
Menu:
```

1. Spam Detection
2. Query Classification
3. Query Analysis
4. Query Prioritization
5. Execute All
6. Exit

```
Enter your choice: 3
```

```
Experience: Bad
```

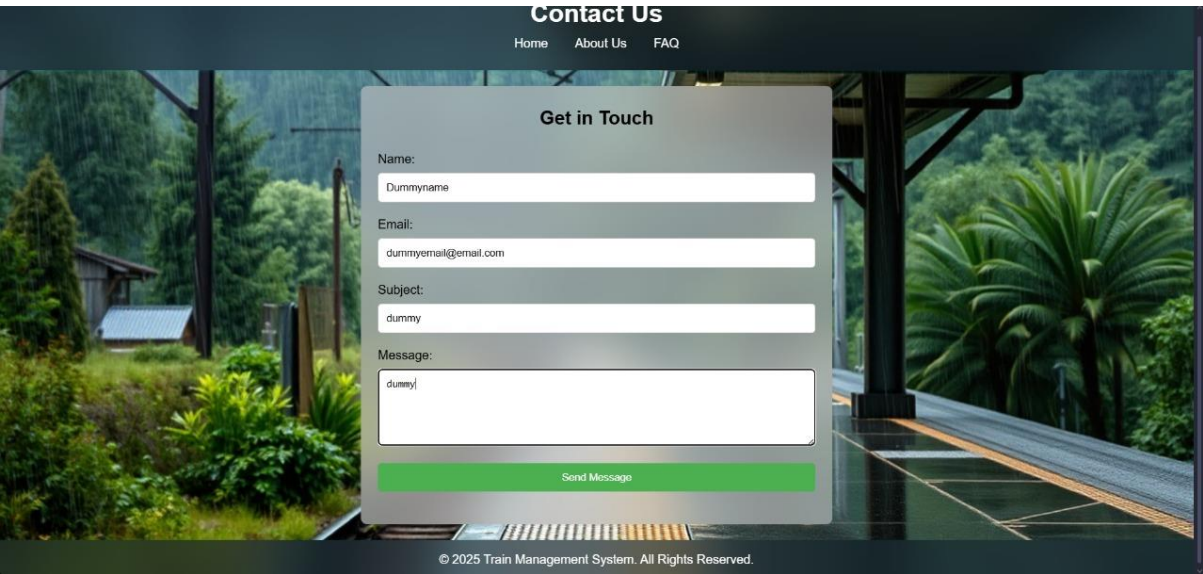
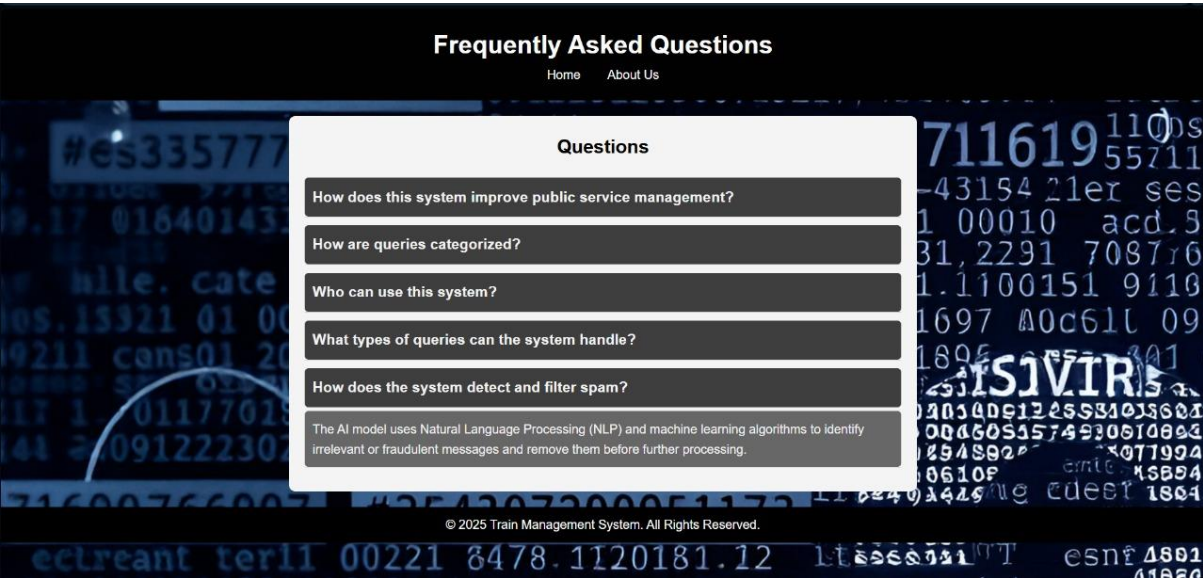
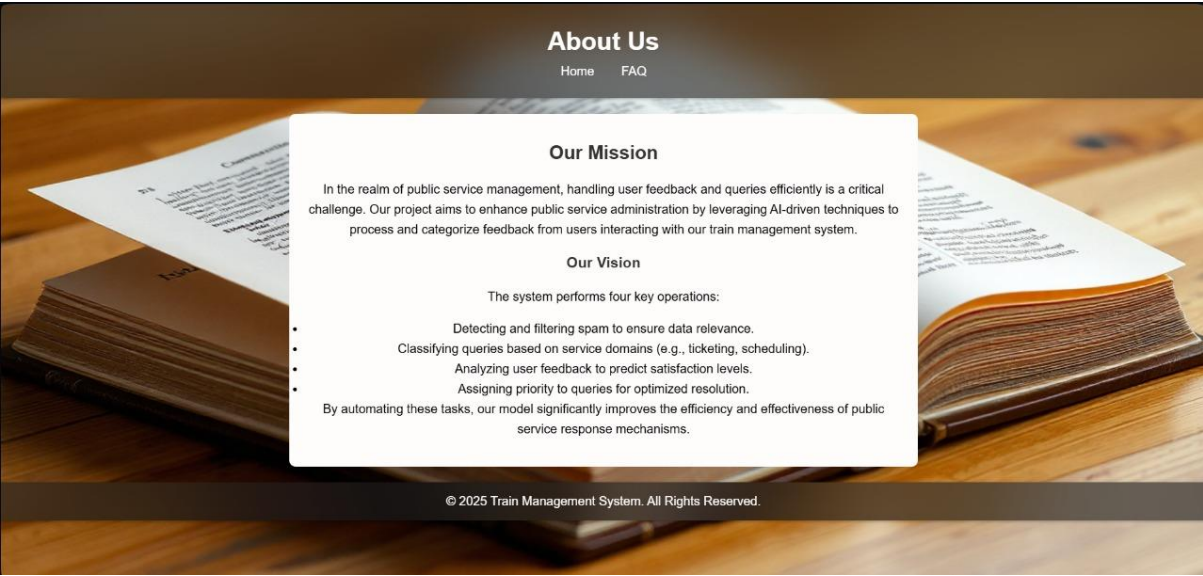
```
Menu:
```

1. Spam Detection
2. Query Classification
3. Query Analysis
4. Query Prioritization
5. Execute All
6. Exit

```
Enter your choice: 1
```

```
The Query is not Spam
```







## References

1. Breiman, L. (2001). *Random Forests*. Machine Learning Journal, 45(1), 5–32.
  - <https://link.springer.com/article/10.1023/A:1010933404324>
2. Pang, B., & Lee, L. (2008). *Opinion Mining and Sentiment Analysis*. Foundations and Trends in Information Retrieval, 2(1-2), 1–135.
  - <https://www.cs.cornell.edu/home/llee/omsa/omsa.pdf>
3. Kotsiantis, S. B., Zaharakis, I. D., & Pintelas, P. E. (2007). *Supervised Machine Learning: A Review of Classification Techniques*. Informatica, 31(3), 249–268.
  - [https://www.researchgate.net/publication/228344975\\_Supervised\\_Machine\\_Learning\\_A\\_Review\\_of\\_Classification\\_Techniques](https://www.researchgate.net/publication/228344975_Supervised_Machine_Learning_A_Review_of_Classification_Techniques)

## Conclusion

The proposed AI-based solution effectively automates query management by detecting spam, classifying domains, analyzing sentiment, and assigning priority levels. This system enhances efficiency, reduces manual workload, and ensures timely query resolution, ultimately improving user satisfaction and service quality.