

# Implementing NLP to Categorize Grievances Received Via A Voice Input Mechanism

Rachana Shekar Rao                      Rakshitha M                      Suhasi GS                      Prajwal Kumar  
Student: Dept Of ISE                      Student: Dept Of ISE                      Student: Dept Of ISE                      Student: Dept Of ISE  
Jyothy Institute of Technology      Jyothy Institute of Technology      Jyothy Institute of Technology      Jyothy Institute of Technology  
Bengaluru, India                      Bengaluru, India                      Bengaluru, India                      Bengaluru, India

Dr. Kishore G R  
Associate Prof : Dept Of ISE  
Jyothy Institute of Technology  
Bengaluru, India

**Abstract—** It is imperative for every administrative system to have a "Grievance Redressal" mechanism that caters to the needs of employees/users. In our proposed system, a user voices their complaint along with additional identification details. The computer system then analyzes the informal exchange. Natural Language Processing (NLP) is used to comprehend requirements through "text summarization." This study created a model that used NLP techniques such as stemming, stop words removal, tokenization, and so on to extract relevant information from a complaints data set in order to accurately classify the complaints. We come to a judgement on the nature of the offense based on the evidence presented and present it to the complaint addressee (the user) and addresser (the legal authority).

**Keywords-** Artificial Intelligence/Machine Learning, Natural Language Processing, Machine learning, Grievance Redressal, Text optimization, Sentiment analysis

## I. INTRODUCTION

A "Grievance Redressal" system measures efficiency and effectiveness by providing critical feedback on administration procedures. However, the current system can be difficult for functionally illiterate people who must write a complaint to alert legal authorities of an atrocity. Furthermore, legal practitioners must be precise because tasks such as understanding, interpreting, formulating, and conducting research are critical. We believe that the solution is an automatic classification system that swiftly processes incoming complaints by evaluating the English text contents and anticipating the category, hence lowering response time.

## II. RELATED WORK

### A. A Web Application for Managing and Resolving Complaints of Public

The goal of this paper is to create a website for an Online Complaint Management System. In India, there is no efficient direct connection between the government and the population for problem solutions. The main goal is to provide an online way for the public to solve problems while saving time and eliminating corruption, as well as to assist common people under the jurisdiction of a municipal corporation in registering their grievances about day-to-day problems in their ward through a website and the municipal authorities in addressing the problem in a timely manner. . Complaints are easier to organize and handle, and they give the organization with an excellent tool for identifying and focusing on issue areas, as well as acting as a management strategy for reviewing, analyzing, and responding to public complaints. This program

is used to record, resolve, and reply to public complaints, requests, and to facilitate any other feedback that gives an interface to register and follow up on one's complaint.

### B. Complaint Management System

The Complaint Management system is a web-based tool designed to track complaints filed by college department/lab personnel. As a result, this system necessitates the development of a distributed, platform-independent web application. Administrator executives have complete authority over all system actions, including reporting a problem, allocating it to a service engineer, and monitoring the performance of that engineer. When a call is registered, an engineer should open it, assign it to service, and close it. This system may provide reports on Engineer Performance Reports as well as department-specific open and pending calls.

### C. Gender Prediction by Voice using Logistic Regression

The objective of the paper is to use machine learning techniques to predict gender from voice. Male and female participants' voices were recorded, and the authors retrieved different acoustic properties as pitch, frequency, and formats. In order to predict gender, a logistic regression model employed these characteristics as input variables. A model's performance was assessed by the authors using measures for accuracy, precision, recall, and F1-score. Additionally, they evaluated how well their model performed in comparison to other machine learning models like k-NN, SVM, and Decision Tree. The findings demonstrated that, with an accuracy of 96.7%, the logistic regression model surpassed the competition. The authors claim that a variety of applications, including automatic gender identification and speaker recognition, can benefit from their methodology.

## III. RESEARCH OBJECTIVES

The proposed system should control all interactions with the following purposes in mind:

- Inform all interested parties of the nature of the offense.
- Provide a plausible legal argument based on the user's description of the offense.
- Legal proposals are used to file complaints with the proper authorities.
- Ordinary people outside of the legal system can understand the course of action, which fosters a sense of security.

## IV. PROBLEM STATEMENT

"Grievance Redressal" relates primarily to receiving and processing complaints from citizens and consumers; but, a

broader definition includes activities undertaken in response to any concern raised by them in order for them to more successfully use services. We are unable to manually classify the offenses that have happened if there are many citizen complaints. It takes time and requires a significant quantity of human resources. As a result, a system that can handle public issues in a timely and effective manner is essential. The system should be able to correctly detect the nature of the offense and classify it under the same section.

V. GAP IDENTIFICATION

Sl No	Author	Year	Title	Gap Identification
1	S Suvetha , N Yasini, M Yogesh	2019	A Web Application for Managing and Resolving Complaints of Public	The website lacked a quick and effective mechanism to categorize the complaints that were submitted. The purpose of the website was to make the problem solver aware of the issues. During their illiteracy, many people in rural areas had no knowledge of this, nor did they know how to use it properly to administer their complaints. As a result, voicing their complaints got much more difficult.
2	Osman Nasr , Enayat Alkhider	2019	Complaint Management System	Only the administration has the choice of including the students' complaints. The administrator was to handle everything, from answering the student's phone call to assigning the service engineer to monitoring the task. It is a time-consuming and intimidating process for one person to do.
3	R Praveen Kumar1, P Sree Varsha, L Sandhya Rani, G Bharadwaj	2019	Gender Prediction by Voice using Logistic Regression	In terms of dataset and research size, the study on gender prediction by voice using logistic regression falls short. The dataset contains only 20 people, who may not be typical of the greater population.

				Furthermore, the study only uses 10 features to predict gender, which may not be enough to predict gender in all cases. As a result, more study with larger and more diverse datasets is required to determine the approach's success and generalizability.
--	--	--	--	---

VI. SYSTEM REQUIREMENTS

Hardware Components:

- System Processor : Core i3 / i5
- Hard Disk : 500 GB.
- Ram : 4 GB.

Any desktop / Laptop system with above configuration or higher level.

Software Components:

- Operating system : Windows 10
- Coding Language : Python, HTML, CSS, JavaScript, PHP
- Tools : XAMPP Server
- IDE : Python

VII. METHODOLOGY

Our proposed system has the flow depicted below:

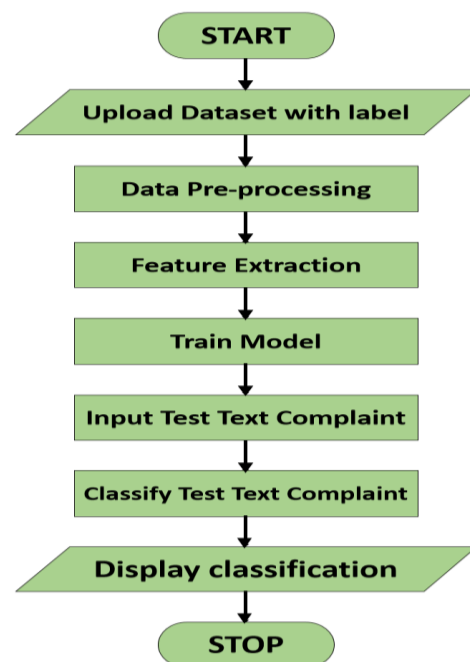


Fig. 1. Flowchart of processes

Modules Implemented:

### 1. Data Acquisition and Preprocessing-

In data collection, pertinent datasets are gathered, and the data's quality is assessed to make sure it is pertinent to the project. Pre-processing, which entails putting the raw data into a form that can be used for analysis once the data has been gathered, comes next. This step entails activities including eliminating extraneous details, tokenizing text into individual words, and changing text's case. When data is properly pre-processed, it is guaranteed to be accurate, complete, and in a format that the chosen NLP algorithms can use.

### 2. Feature Selection and Data Preparation-

In feature selection, the most pertinent features are chosen from a large assortment of features that can be utilized to train the model. Cleaning and converting the text data into an algorithmic-usable format is known as data preparation. This includes operations like stop word removal, stemming, or lemmatization. The accuracy and effectiveness of the Naive Bayes algorithm in an NLP project can be significantly increased by performing effective feature selection and data preparation.

### 3. Model Construction and Model Training-

The dataset is divided into training and testing sets during the model construction and model training stages. The testing set is used to assess the model's performance after it has been trained using the training set. The training set is used to apply the **Naive Bayes method** to determine the conditional probability of each feature given each class. The model is built using this. After being developed, the model is trained using the training set. The model is subsequently assessed using the testing set to ascertain its precision, recall, and other performance indicators, like accuracy. The hyper-parameters can be changed, and the model can be improved and optimized by employing cross-validation and other methods. The objective is to create a model that can accurately and quickly classify fresh data. Our model is able to successfully and accurately classify data into 3 categories which are **Robbery, Cyber Crime, Accident Case** and the data base gets split into the training and testing data in an 80:20 ratio.

### 4. Model Validation and Result Analysis-

The final phases in any NLP project utilizing the Naive Bayes technique are model validation and result analysis. This stage involves testing the model with a set of data that was not used in training to determine how accurate it is. The model's effectiveness is assessed using a number of criteria, including precision, recall, and F1 score. The model's performance in terms of true positive, false positive, true negative, and false negative predictions is also evaluated using the confusion matrix. Instances that were incorrectly identified can be examined in more detail to find trends and enhance the performance of the model. Iterative model refinement is done until an acceptable level of accuracy is achieved. The final step involves using the model to predict new data and assessing its performance in a real-world setting.

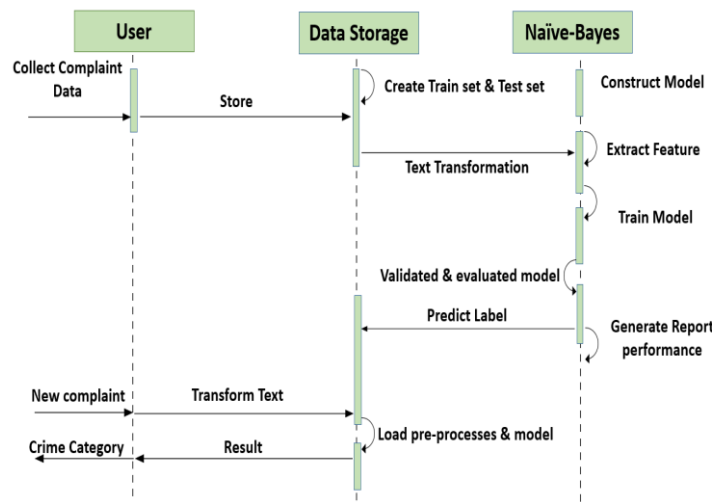


Fig. 2. Sequence Diagram

## VIII. RESULTS AND SNAPSHOTS

Our proposed system has 2 bifurcations, namely-

1. The Addressee (Victim/User) side
2. The Addresser (Authority/Admin) side

The Addressee side comprises a Complaint form which contains the various identification criteria that need to be filled as well as the complaint section wherein the complaint can be submitted vocally. Once a complaint has been submitted, the filled details along with the crime classification as per the complaint submitted is depicted on screen so make the user aware as to which crime category their complaint falls under.

The Addresser side contains all the form data submitted by users along with the crime classification. They are displayed in a tabular format, with the latest one coming up first. The admin has an option either to accept or reject a particular complaint. In accordance to the decision made by the admin, an auto-generated email is sent to the email id provided in the form informing the user of the admin's decision.

The proposed system had the following confusion matrix

Confusion matrix:

$$\begin{bmatrix} 8 & 2 & 0 \\ 0 & 20 & 0 \\ 0 & 4 & 5 \end{bmatrix}$$

In the above confusion matrix:

- True Positive (TP) count is 8, which represents the number of times the model correctly predicted the positive class for the first category.
- True Negative (TN) count is 20, which represents the number of times the model correctly predicted the negative class for the second category.
- False Positive (FP) count is 4, which represents the number of times the model incorrectly predicted the positive class for the third category.
- False Negative (FN) count is 2, which represents the number of times the model incorrectly predicted the negative class for the first category.

### IX. SNAPSHOTS

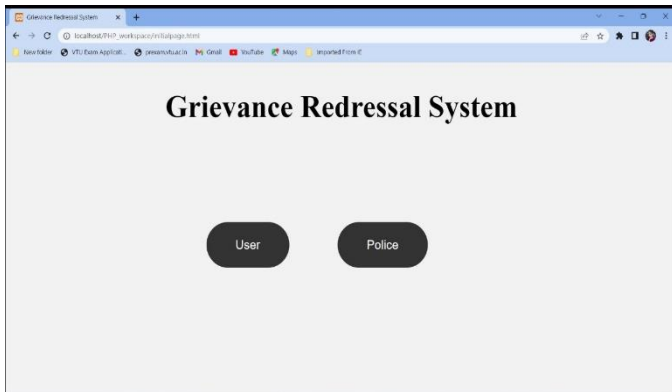


Fig. 3. Initial Page

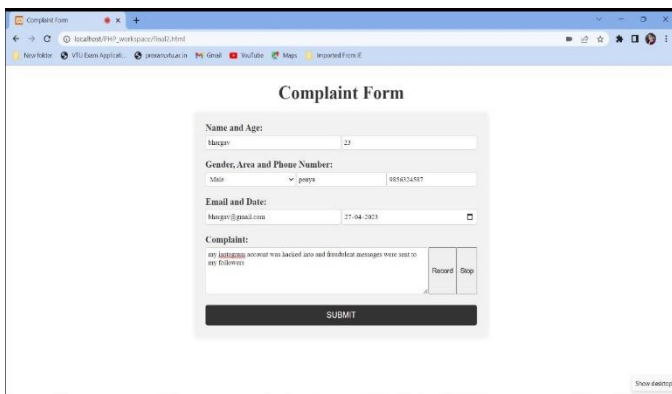


Fig. 4. Complaint Form

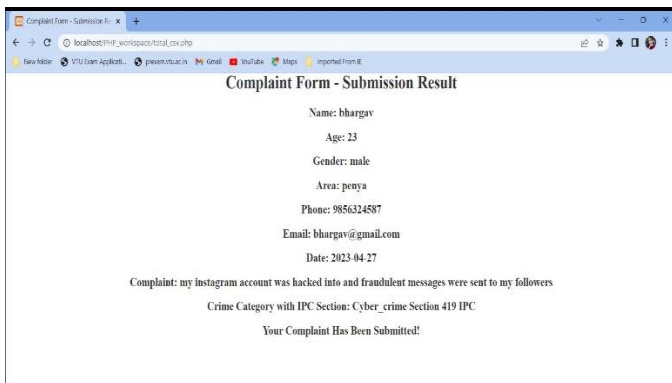


Fig. 5. Submitted Form along with the Crime Classification

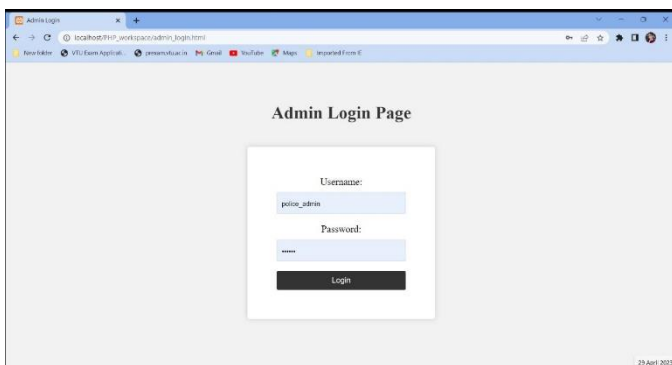


Fig. 6. Admin Login Page

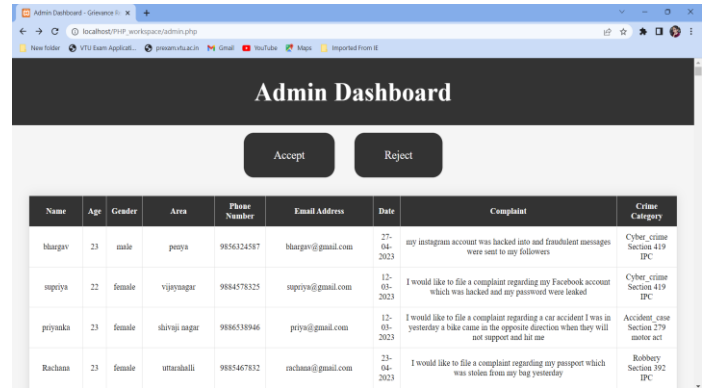


Fig. 7. Admin Dashboard

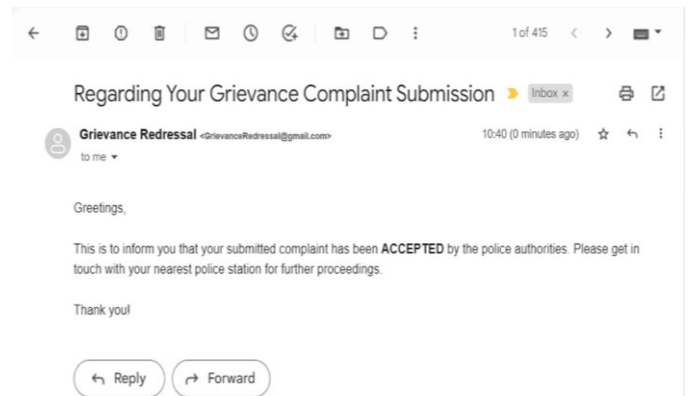


Fig. 8. Grievance Acceptance Email received by the Victim

### X. CONCLUSION

The primary goal is to successfully develop and show a model that employs NLP and speech recognition in a real-world application. We achieved the same result with a 92% accuracy rate. The victim (user) must be able to effortlessly voice-input their concerns to the system using our paradigm. The victim's complaint is subsequently assigned to a specific crime category, along with the relevant sections, by the system. The admin (Authority) has the ability to approve or reject submitted complaints, and the same response is delivered to the user.

### XI. FURTHER ENHANCEMENTS

This project is an extension of how NLP can be used in a real world application. A further development can be implementing Voice functionality of both the ends. The Data Set for training the model was created organically hence was not enormous in size. However, it can be easily up scaled and doing so will also increase the model precision and accuracy rates, User identification can be enhanced further through the integration of Aadhar Card/PAN Card, Face Recognition and Biometric Scans. An app can be launched which can help to send in complaints/grievances quicker and more easily. Sending Confirmation/Rejection SMS to user's valid mobile number authentication can be implemented. The victim can be provided with a leverage of viewing the complaints given by them in a separate database and see its status.

## XII. ACKNOWLEDGEMENT

First and foremost, I wish to record my sincere gratitude to Jyothy Institute of Technology and to our beloved Principal, Dr. K Gopalakrishna for his constant support and encouragement. Next we would like to convey our sincere gratitude to Dr. Divakar HV, HOD, ISE for his encouragement. We convey our thanks to Dr. Kishore G R, Assoc. Prof, ISE for his unwavering support and guidance for this project and all the professors of ISE Department for their valuable suggestions. Last but not the least we would like to extend the warmest of thanks to our friends and family for their constant support and encouragement.

## REFERENCES

- [1] Centralized Public Grievance Redress And Monitoring System (CPGRAMS) available at <https://pgportal.gov.in/>
- [2] Agarwal, S. K., Kumar, A., Nanavati, A. A., and Rajput, N. Content creation and dissemination by-and-for users in rural areas. In *ICTD*. 2009
- [3] Bihar Public Grievance Redressal System. <http://www.bpgrs.in>
- [4] <https://www.cse.iitd.ac.in/~aseth/civilsocietygrievanceredressal.pdf>
- [5] [https://www.researchgate.net/publication/336716292\\_Online\\_Complaint\\_Management\\_Systems](https://www.researchgate.net/publication/336716292_Online_Complaint_Management_Systems)
- [6] <https://pgportal.gov.in/Home/RedressMechanism>
- [7] <https://legalaffairs.gov.in/PublicGrievancesCell>