AI POWERED PETITION ANALYSIS AND CATEGORIZATION USING NLP

A PROJECT REPORT

Submitted by

ABINESH R	**811722104005**
ANBAZHAGAN P	**811722104011**
HARIBHARATHI R	**811722104047**
INFANTALAN L	**811722104057**

in partial fulfillment of the requirements for the award degree of

Bachelor in Engineering

20CS7503 DESIGN PROJECT - 3

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

K. RAMAKRISHNAN COLLEGE OF TECHNOLOGY

(AUTONOMOUS)

SAMAYAPURAM – 621112

NOVEMBER 2025

AI POWERED PETITION ANALYSIS AND CATEGORIZATION USING NLP

A PROJECT REPORT

Submitted by

ABINESH R	**811722104005**
ANBAZHAGAN P	**811722104011**
HARIBHARATHI R	**811722104047**
INFANTALAN L	**811722104057**

in partial fulfillment of the requirements for the award degree of

Bachelor in Engineering

20CS7503 DESIGN PROJECT - 3

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

K. RAMAKRISHNAN COLLEGE OF TECHNOLOGY

(AUTONOMOUS)

SAMAYAPURAM – 621112

NOVEMBER 2025

K. RAMAKRISHNAN COLLEGE OF TECHNOLOGY (AUTONOMOUS)

SAMAYAPURAM - 621112

BONAFIDE CERTIFICATE

The work embodied in the present project report entitled "**AI POWERED PETITION ANALYSIS AND CATEGORIZATION USING NLP**" has been carried out by the students **ABINESH R, ANBAZHAGAN P, HARIBHARAHI R, INFANTALAN L.** The work reported herein is original and we declare that the project is their own work, except where specifically acknowledged, and has not been copied from other sources or been previously submitted for assessment.

Date of Viva Voice: ……………………

Ms. M. Mathumathi, M.E., (Ph.D.,)	Mr. R. Rajavarman, M.E., (Ph.D.,)
SUPERVISOR	HEAD OF THE DEPARTMENT
Assistant Professor	Assistant Professor(Sr. Grade)
Department of CSE	Department of CSE
K. Ramakrishnan College of Technology(Autonomous)	K. Ramakrishnan College of Technology(Autonomous)
Samayapuram-621 112	Samayapuram-621 112

INTERNAL EXAMINER **EXTERNAL EXAMINER**

ABSTRACT

The AI-Powered Petition Analysis System is a web-based platform developed to efficiently manage and resolve public issues submitted by citizens. The main objective of this project is to provide a centralized digital system where users can submit petitions with descriptions, images, or documents, allowing government officials or administrators to respond transparently. Users can log in to create petitions, track status, interact through likes and comments, and view unsolved active petitions. The system integrates advanced Artificial Intelligence techniques such as NLP-based sentiment analysis, urgency detection, keyword extraction, and duplicate identification to prioritize petitions effectively. Image captioning using BLIP is applied to understand visual complaints automatically. An analytical dashboard displays insights including sentiment distribution, urgency levels, priority charts, and solved versus unsolved petitions, supporting data-driven decisions. Admins can manage petitions, mark issues as solved, and analyse trends to enhance service quality.

Keywords: AI-based petition system, Natural Language Processing, Sentiment Analysis, Priority and Duplicate Detection, Image Captioning, Machine Learning, Dashboard Analytics.

ACKNOWLEDGEMENT

We thank **Dr. N. Vasudevan**, Principal, for his valuable suggestions and support during the course of our research work.

We thank **Mr. R. Rajavarman**, Head of the Department, Assistant Professor (Sr. Grade), Computer Science and Engineering, for his valuable suggestions and support during the course of our research work.

We wish to record our deep sense of gratitude and profound thanks to our Guide **Mrs. M. Mathumathi**, Assistant Professor, Computer Science and Engineering for her keen interest, inspiring guidance, constant encouragement with our work during all stages, to bring this thesis into fruition.

We are extremely indebted to our project coordinator **Mrs. R. Ramasaraswathi**, Assistant Professor, Computer Science and Engineering, for her valuable suggestions and support during the course of our research work.

We also thank the faculty and non-teaching staff members of the Computer Science and Engineering, K. Ramakrishnan College of Technology (Autonomous), Samayapuram, for their valuable support throughout the course of our research work.

Finally, we thank our parents, friends and our well-wishers for their kind support.

SIGNATURE

____

____

____

____

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

AI	-	Artificial Intelligence
ML	-	Machine Learning
NLP	-	Natural Language Processing
BLIP	-	Bootstrapped Language-Image Pretraining
PDF	-	Portable Document Format
API	-	Application Programming Interface
JWT	-	JSON Web Token
HTML	-	HyperText Markup Language
JS	-	JavaScript

CHAPTER 1

INTRODUCTION

1.1 DESCRIPTION

Public petition systems play a vital role in modern digital governance by enabling citizens to report problems, request services, and express concerns directly to administrative authorities. These systems act as a communication bridge between the public and government bodies, ensuring that civic issues such as infrastructure damage, public safety concerns, and administrative grievances are addressed in a timely manner. Traditionally, petitions were submitted manually through paper forms or offline channels, which made the process slow, inefficient, and difficult to manage at scale. With the rise of digital platforms, online petition portals have become more common; however, even these systems face major challenges in processing large volumes of submissions effectively.

A key difficulty arises from the unstructured nature of petition content. Citizens describe their issues in varied styles, languages, formats, and levels of detail. Some petitions include images, while others attach long PDF documents. Manual review of such diverse inputs requires significant administrative effort, often leading to delays, human errors, and inconsistent prioritization. As the number of submissions increases, authorities struggle to categorize petitions, determine urgency, detect duplicates, or identify recurring community problems.

Advances in Artificial Intelligence (AI), particularly Natural Language Processing (NLP) and machine learning, provide powerful solutions for understanding and processing unstructured textual data. NLP enables systems to analyze sentiment, extract keywords, detect urgency, and classify petitions into meaningful categories. Machine learning models can be trained to predict the priority level of a petition, helping authorities address the most critical issues first.

Beyond text, real-world petitions often include images to provide visual evidence of a problem. For example, broken roads, damaged streetlights, or waste accumulation.

Equally important is the need for a responsive and transparent platform for both users and administrators. Citizens should be able to easily submit petitions, track progress, and engage with public issues through features like likes and comments. Administrators, on the other hand, require analytical dashboards that display insights such as sentiment trends, urgency distribution, and petition status metrics. These tools improve decision-making and promote accountability.

In this context, an AI-Powered Petition Analysis and Categorization System offer a comprehensive solution that integrates NLP, image analysis, machine learning, and interactive dashboards to modernize the petition-handling process. This intelligent system minimizes manual processing, enhances scalability, and improves the responsiveness of public service delivery, enabling faster decision-making, transparent issue tracking, real-time priority assessment, and effective resource allocation for authorities.

1.2 DOMAIN SPECIFICATION

This project operates within the domains of Artificial Intelligence, Natural Language Processing (NLP), and Web Technology. Artificial Intelligence enables automated petition analysis, priority prediction, and intelligent decision support. NLP techniques process textual content to identify sentiment, urgency, and category. Image analysis further enhances understanding through caption generation. Web technology provides a seamless platform for petition submission, visualization, and administrative control. Together, these domains form an integrated system that improves efficiency, transparency, and scalability in modern petition management.

1.2.1 Artificial Intelligence

Artificial Intelligence (AI) plays a central role in this project by enabling automated understanding and analysis of petition data. AI techniques are used to classify petitions, detect duplicates, generate image captions, and support intelligent decision-making. Machine learning models predict priority levels, while rule-based AI helps determine urgency and sentiment. By automating these tasks, AI significantly reduces manual workload and enhances the accuracy, speed, and scalability of the petition.

1.2.2 Natural Language Processing (NLP)

Natural Language Processing (NLP) enables the system to interpret and analyze textual petition content automatically. NLP techniques extract meaningful insights such as sentiment, urgency, and key patterns from user-submitted text. It supports categorization, priority prediction, and duplicate detection by processing language structures and semantic relationships. Through NLP, the system transforms unstructured petition data into actionable information, improving decision-making and ensuring efficient handling of diverse user-generated content.

1.2.3 Web Technology

Web technology provides the foundation for delivering an interactive and user-friendly petition management platform. It enables seamless petition submission, real-time updates, dashboard visualization, and admin control through modern web interfaces. HTML, CSS, and JavaScript ensure responsive design, while backend frameworks like Flask manage communication between users and the AI modules. Web technology allows efficient data flow, secure authentication, and an accessible system that supports both citizens and administrators in managing petitions effectively.

1.3 PROBLEM STATEMENT

Online petition platforms receive large volumes of citizen submissions, often containing unstructured text, images, or lengthy documents. Traditional petition management relies on manual review, making the process slow, inconsistent, and unable to scale with increasing public participation. Existing systems primarily focus on basic categorization and lack advanced capabilities such as sentiment analysis, urgency detection, image understanding, duplicate identification, and intelligent prioritization. As highlighted in recent IEEE research, petition categorization models face challenges in handling diverse data formats and ensuring accurate routing. Therefore, there is a need for an AI-powered system that can automatically analyze, classify, and prioritize petitions efficiently and transparent.

1.4 OBJECTIVES

The primary objective of this project is to develop an AI-powered petition analysis system capable of automatically understanding, categorizing, and prioritizing citizen petitions using NLP and image intelligence, providing instant insights to users and administrators for efficient issue resolution.

1.4.1 Primary Objectives

The primary objective of this project is to design and implement an AI-powered system capable of automatically analyzing and categorizing citizen petitions using advanced NLP and machine learning techniques. The system also integrates image captioning models to interpret petitions submitted through images or PDF documents, ensuring comprehensive multimodal analysis. Additionally, it aims to detect sentiment, urgency, priority levels, and duplicate petitions with high accuracy to support effective decision-making and faster issue resolution. The system further provides users and administrators with clear insights, visual indicators, and structured outputs that enhance petition understanding, improve transparency, and streamline response processes.

1.4.2 Secondary Objectives

Another significant objective of this project is to build a user-friendly petition submission interface that supports text, PDF, and image uploads with minimal technical requirements, ensuring smooth usability for all citizens. The system also focuses on implementing secure data storage and reliable activity logging using databases such as SQLite or PostgreSQL for efficient petition management. Additionally, it aims to enable continuous improvement of NLP and priority prediction models as new petition patterns and categories emerge over time. Furthermore, the system is designed to ensure cross-platform accessibility across modern web browsers and various devices, providing enhanced convenience and seamless access for both users and administrators.

1.4.3 Expected Outcome

This project delivers an intelligent AI-powered system capable of accurately analyzing, categorizing, and prioritizing citizen petitions. It features a responsive web platform that offers real-time insights, including sentiment analysis, urgency level, priority classification, and duplicate petition detection. A comprehensive analytical dashboard enhances public awareness and highlights common community issues while improving overall petition management efficiency through automated workflow support, enabling faster administrative responses.

The system ensures secure authentication and role-based access to protect citizen and government data. Additionally, the platform includes a transparent tracking system that allows users to monitor the status and progress of their submitted petitions in real time. By identifying trends, recurring issues, and high-demand service areas, the system supports data-driven decision-making.

Furthermore, the system promotes accountability by providing clear timelines and progress tracking for unresolved petitions. It supports multilingual expansion and scalable model updates, ensuring long-term adaptability to evolving public service needs. Integration with future government systems and automated notification alerts enhances communication and workflow coordination.

1.5 SCOPE OF THE PROJECT

The scope of this project extends to automating petition analysis using AI models then project covers essential functionalities including real-time petition submission and automated categorization based on sentiment, priority, and content analysis. It incorporates robust database integration to securely maintain user information, petition records, and analytical results for accurate tracking and management. Additionally, the system provides a user-friendly web-based interface that allows users and administrators to access petition insights immediately, interact through engagement features, and monitor updates efficiently.

CHAPTER 2

LITERATURE SURVEY

2.1 AI-BASED PETITION ANALYSIS USING NLP AND TEXT CLASSIFICATION

Usman Saif (2019) et al. presented the use of Natural Language Processing (NLP) and traditional text-classification techniques to analyze citizen petitions submitted through online platforms. The authors emphasize that most petition systems still depend on manual review, making processing slow and inconsistent. Their research explores how basic NLP methods—such as tokenization, keyword extraction, TF-IDF vectorization, and sentiment scoring—can support automated petition understanding. A dataset of textual petitions collected from public grievance portals was used to train machine learning models such as Naïve Bayes, Support Vector Machines, and Logistic Regression.

The study demonstrated that these classical machine learning approaches could categorize petitions into broad issue groups such as infrastructure, public safety, or administration. However, the analysis remained limited to text-only inputs, and the models struggled when petition descriptions were short, ambiguous, emotional, or lacked specific keywords. Furthermore, the system did not include priority prediction, urgency detection, or duplicate identification, which are essential for practical decision-making in large-scale platforms.

The Usman Saif concludes that traditional NLP and machine learning techniques provide a foundation for automating petition classification, more advanced AI systems are needed for real-world scenarios. Their work highlights the gap that modern multi-feature petition analysis systems must address—an area where the current project introduces significant improvements through enhanced NLP workflows and smarter petition insights.

2.2 AUTOMATED COMPLAINT PRIORITIZATION USING HYBRID NLP TECHNIQUES

F. Sun (2019) et al. presented an automated complaint prioritization framework designed for civic bodies to categorize and route citizen grievances. The authors implemented a hybrid NLP pipeline that combined keyword extraction with sentiment polarity scoring to assign importance levels to complaints. The system utilized classical NLP tools such as TF-IDF, rule-based sentiment lexicons, and LightGBM classifiers to generate priority predictions. Although the model achieved moderate improvements in classification accuracy compared to manual filtering, its overall analytical scope remained limited.

The approach mainly focused on textual input and lacked support for multi-modal data, such as images or PDF uploads, which are commonly submitted in modern grievance platforms. Furthermore, the sentiment analysis method relied on fixed dictionaries, making it less adaptable to evolving linguistic patterns or domain-specific vocabulary. The study did not incorporate deep learning models or transformer-based architectures that offer superior contextual understanding for real-world complaint narratives.

Another limitation of the system was the absence of advanced features such as duplicate complaint detection, urgency prediction, or user interaction tracking through likes and comments. Additionally, the framework did not include a real-time dashboard for visualizing complaint distribution, trends, or engagement metrics—an essential requirement for administrative decision-making.

Compared to this system, the proposed AI-powered petition analysis and categorization model integrates transformer-backed NLP analysis, BLIP-based image captioning, machine learning–driven priority prediction, sentiment and urgency scoring, social interaction analytics, and a comprehensive dashboard. This makes the proposed platform significantly more capable, interactive, and suited for large-scale public administration environments.

2.3 DEEP LEARNING BASED TEXT CLASSIFICATION FOR CITIZEN FEEDBACK SYSTEMS

Shervin Minaee (2020) et al. studied and examines the use of deep learning models, particularly Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN), for classifying citizen feedback submitted to municipal e-governance portals. The researchers aimed to automate the sorting of large volumes of public feedback into categories such as infrastructure, healthcare, sanitation, and public utilities. The system utilized word embeddings such as GloVe and FastText to represent text inputs, enabling the model to learn contextual patterns better than traditional bag-of-words methods.

Although the deep learning models offered improved accuracy over classic machine learning approaches, the system primarily focused on category prediction and lacked additional analytical layers such as sentiment scoring, urgency identification, or priority estimation. The absence of multimodal capabilities meant that the system could not process uploaded images, handwritten documents, or non-text grievance submissions. Additionally, the model required significant computational resources for training and did not provide real-time analysis or integration with interactive front-end platforms.

The study also highlighted challenges in handling multilingual citizen feedback, as the system struggled with regional language variations and frequent code-mixing. No mechanisms were included for detecting duplicate submissions or tracking user engagement features such as likes and comments.

While this research establishes the potential of deep learning for public grievance classification, it does not incorporate holistic petition analytics. In contrast, the proposed AI-powered petition analysis system introduces a more advanced framework combining NLP, urgency detection, sentiment evaluation, priority prediction, image captioning, and real-time dashboard visualization—making it substantially more comprehensive and user-centric.

2.4 MACHINE LEARNING–BASED PETITION CLASSIFICATION USING TRADITIONAL FEATURE ENGINEERING

C. Stanik (2020) et al. has investigate the use of traditional machine learning methods for classifying citizen petitions based solely on textual descriptions. The authors collected a dataset from a regional public grievance portal and manually labeled petitions into a fixed set of categories such as road repair, sanitation, electricity complaints, and water supply issues. Their methodology relied heavily on classical feature engineering techniques, including TF-IDF vectors, term frequency patterns, and basic linguistic markers such as part-of-speech tags.

To evaluate performance, the researchers applied algorithms like Decision Trees, Random Forest, Multinomial Naïve Bayes, and K-Nearest Neighbours. Among them, Random Forest achieved the highest accuracy due to its ability to generalize across differently worded petitions. However, the system faced noticeable limitations, particularly when dealing with short petitions, informal writing styles, or petitions lacking clear keywords. Since only static text features were used, the system failed to understand contextual nuances, emotional tone, or urgency behind user submissions.

Furthermore, the study lacked mechanisms for identifying repeated petitions, analyzing sentiment, predicting priority, or handling multimodal inputs such as PDFs and images—capabilities increasingly necessary in modern e-governance portals. The authors acknowledged that while traditional machine learning provides a starting point, real-world petition platforms require more intelligent, adaptive, and context-aware AI systems for effective large-scale implementation.

2.5 MACHINE LEARNING FOR CIVIC COMPLAINT PRIORITIZATION USING BASIC FEATURES

Anurag Saha (2021) et al. presented a machine learning–based approach for prioritizing civic complaints submitted to municipal bodies. The authors aimed to assist local authorities in identifying which complaints required faster responses by using a basic feature-driven model. Their system extracted simple textual features such as word count, presence of urgency keywords, and frequency of issue-related terms. These features were then used to train classical algorithms including Logistic Regression, Decision Trees, and Support Vector Machines to predict whether a complaint should be assigned "high," "medium," or "low" priority.

Although the proposed framework introduced the concept of automated prioritization, its effectiveness was limited due to the shallow nature of the extracted features. Short complaints or those written without explicit keywords were often misclassified, and the model struggled to capture sentiment, context, or the implicit seriousness of issues. Additionally, the study relied entirely on structured text and did not support petitions submitted as images or PDF files.

Another limitation noted in the research was the absence of feedback learning or model updating. The method processed only text inputs and could not interpret images or automatically extract content from uploaded documents. The authors acknowledged that although rule-based systems offer interpretability, they are limited in scalability and struggle with real-world petition diversity. Since civic issues evolve over time, models trained on older data quickly became outdated. Moreover, the system lacked capabilities to identify repeated complaints submitted by different citizens regarding the same issue.

Overall, the study demonstrates the early development of automated complaint prioritization but highlights significant gaps in contextual understanding, adaptability, and multimodal input handling. These limitations further justify the need for more advanced AI-powered petition analysis systems that incorporate deeper NLP techniques and richer feature extraction methods, as achieved in the present project.

2.6 MACHINE LEARNING ASSISTED PUBLIC GRIEVANCE ROUTING USING KEYWORD MATCHING AND RULE-BASED CLASSIFICATION

Wiley (2022) et al. studied and explores the use of machine learning techniques combined with rule-based keyword matching to automate the routing of public grievances to appropriate government departments. The researchers developed a system to categorize complaint texts into predefined service categories such as sanitation, electricity, road maintenance, and public safety. Their approach relied heavily on keyword frequency, bag-of-words (BoW) representations, and simple classifiers like Naïve Bayes and Logistic Regression.

The system demonstrated moderate classification accuracy when dealing with well-structured citizen complaints that contained clear indicative keywords. However, it struggled significantly with ambiguous, short, or emotionally phrased grievances. Because the model used shallow text features, it could not understand contextual variations, complex sentence structures, or implicit references such as “broken streetlights making our area unsafe.” The rule-based mechanism required continuous manual updates to include emerging keywords, making the system less scalable and difficult to maintain.

Another major drawback of the study was the absence of advanced NLP techniques such as transformer models, semantic embeddings, sentiment evaluation, or urgency prediction. The system only supported basic text classification and did not incorporate any interactive features like user feedback, popularity tracking, or behavior-based prioritization.

Despite these limitations, the research highlights the potential of machine learning for public grievance management and represents an early attempt toward automated petition routing. However, compared to modern NLP-driven solutions, its capabilities remain limited, emphasizing the necessity for more intelligent systems such as the proposed AI-powered petition analysis project, which integrates multi-parameter evaluation, deep NLP, and improved prediction accuracy.

2.7 MULTI-MODAL GOVERNMENT GRIEVANCE ANALYTICS USING MACHINE LEARNING

Patel (2022) et al. research and explores the application of machine learning techniques for analyzing government grievance submissions collected across digital citizen service platforms. The authors investigated a hybrid feature extraction pipeline that combined structured metadata—such as submission timestamps, user demographics, and grievance type—with textual analysis using TF-IDF and n-gram models. Several machine learning classifiers, including Naïve Bayes, Logistic Regression, and Gradient Boosting, were evaluated for categorizing complaints into predefined administrative departments.

Although the study achieved reasonable performance on structured text, it remained restricted to classical machine learning and did not utilize modern deep learning or transformer-based models that provide richer semantic understanding. The system lacked capabilities such as sentiment detection, urgency scoring, or contextual priority estimation—features essential for making decisions in public service environments where response time matters. Additionally, the approach did not include any mechanisms for handling multimedia submissions such as photos or PDF documents, which are increasingly common in citizen-generated petitions.

The study's implementation was limited to batch processing, meaning grievances were analyzed only after full dataset collection. Real-time analysis, user feedback loops, social engagement metrics, and interactive dashboards were not addressed. Furthermore, the research did not consider duplicate grievance detection, which is a significant challenge in public systems where multiple users often report the same issue.

Compared with this work, the proposed AI-powered petition analysis system incorporates a broader and more modern analytical scope, including NLP-based sentiment and urgency detection, image captioning for non-text inputs, real-time priority classification, duplicate detection, and an interactive dashboard to visualize petition trends dynamically. This elevates the system into a more comprehensive, multi-modal, and responsive public service solution.

2.8 RULE-BASED AND KEYWORD-DRIVEN PETITION ANALYSIS SYSTEMS

I. H. Sarker (2023) et al. presents a rule-based and keyword-driven framework for analyzing citizen petitions submitted through local government service portals. The authors proposed a simple automation model where petitions were categorized based on manually created keyword dictionaries representing common issue types such as road maintenance, electricity problems, waste management, and drainage complaints. The system relied on predefined rules, phrase matching, and pattern recognition to detect category indicators within the text.

To build the rule set, domain experts manually identified keywords appearing frequently in historical petitions. The system then used string matching, regular expressions, and decision rules to classify petitions. Although this approach achieved acceptable performance for well-structured 1containing spelling mistakes, or submissions lacking explicit problem terms.

The researchers also experimented with a simple frequency-based scoring method to estimate importance, but it lacked contextual understanding and often misinterpreted long or emotionally worded submissions. The researchers noted that the system could not process petitions submitted as images or PDF documents and had no capability to identify similar or repeated complaints submitted by different users. Since the system did not use statistical models or machine learning, it was unable to adapt to new writing styles or evolve with changing citizen issues.

Additionally, the study did not support any advanced analytical capabilities such as sentiment detection, priority ranking, or identifying repeated grievances. The method processed only text inputs and could not interpret images or automatically extract content from uploaded documents. The authors acknowledged that although rule-based systems offer interpretability, they are limited in scalability and struggle with real-world petition diversity.

2.9 SENTIMENT-BASED ANALYSIS OF PUBLIC COMPLAINTS USING LEXICON METHODS

R. Karthikeyan (2023) et al. investigates the use of lexicon-based sentiment analysis techniques for understanding public complaints submitted through online civic platforms. The authors aimed to determine whether the emotional tone of complaints could help authorities estimate citizen dissatisfaction levels and prioritize responses accordingly. Their method relied on established sentiment dictionaries such as VADER and SentiWordNet, which assign polarity values to individual words. Complaint text was processed through tokenization and stop-word removal, followed by sentiment scoring based on word-level matches.

Although the study successfully demonstrated the usefulness of sentiment in highlighting negative or dissatisfaction-driven complaints, it encountered several limitations. Lexicon-based models often misinterpreted complaints containing sarcasm, mixed emotions, or region-specific language usage. The approach also failed when petitions were short or lacked emotionally expressive words. Additionally, the authors noted that lexicon methods do not understand context, leading to incorrect polarity assignments in sentences containing negations or nuanced meanings.

The study applied simple classification techniques to categorize complaints into broad categories like “positive,” “negative,” and “neutral,” but was not capable of generating deeper insights such as urgency assessment, priority prediction, or domain-specific grievance identification. The system also lacked adaptability, as sentiment dictionaries required manual updates to accommodate new terms or expressions used by citizens.

Overall, the study highlights the foundational role of sentiment analysis in public complaint processing but underscores the need for more advanced, context-aware NLP systems capable of deeper semantic understanding. These limitations indicate the importance of developing intelligent, AI-driven petition analysis frameworks—capabilities addressed effectively in the proposed project.

2.10 TEXT MINING APPROACH FOR PUBLIC GRIEVANCE CLASSIFICATION

Shuai Zhang (2024) et al. explores a text-mining approach for categorizing public grievances submitted through online complaint portals. The researchers aimed to build an automated system capable of sorting large numbers of citizen complaints into predefined labels to assist administrative departments in routing issues more efficiently. Their methodology involved collecting textual grievances from municipal service platforms and applying conventional preprocessing steps such as stop-word removal, stemming, n-gram extraction, and frequency analysis.

For classification, the authors experimented primarily with Support Vector Machines, Linear Regression Classifiers, and Gradient Boosting models. They reported that SVM produced the most consistent results due to its robustness in handling sparse text features. The study also incorporated a basic sentiment analysis model using lexicon-based methods to estimate the emotional tone of grievances. However, the sentiment scores were often inaccurate, especially for complaints expressed using informal or colloquial language. The system also struggled when petitions were extremely brief or contained mixed sentiments.

A key limitation identified in the study is the dependency on manually engineered text features, which restricts the model's ability to understand deeper linguistic meaning or subtle contextual cues. The researchers noted that the system could not process petitions submitted as images or PDF documents and had no capability to identify similar or repeated complaints submitted by different users. Moreover, the approach lacked dynamic prioritization, meaning all grievances were treated with equal importance regardless of urgency or severity.

Overall, the study demonstrates the foundational role of text-mining techniques in grievance classification but also highlights the need for more advanced AI-driven systems capable of deeper language understanding and multi-format petition processing—gaps addressed by the present project.

CHAPTER 3

EXISTING SYSTEM

3.1 OVERVIEW OF PETITION DETECTION TECHNIQUES

Petition processing techniques have evolved to manage the increasing number of citizen submissions received through online grievance portals. Traditional systems primarily rely on manual review, where administrators read each petition and categorize it based on experience or departmental knowledge. This approach often uses simple keyword matching or predefined rule-based filters to determine petition type, urgency, or responsible department. Some systems employ basic text scanning tools that identify recurring words or phrases to offer preliminary categorization. However, these methods lack contextual understanding and struggle with variations in language, meaning, or formatting. They also fail to extract emotional tone, urgency levels, or complaint severity accurately. As petition volumes grow, manual and rule-based techniques become inefficient, error-prone, and slow in routing submissions to the appropriate authority. This highlights the need for automated, intelligent models that can analyze petitions more effectively.

3.2 TEXT-BASED CATEGORIZATION METHODS

Keyword-based filtering remains a common approach in many existing petition management systems. It uses predefined word lists to assign categories, such as "water supply," "electricity," or "road repair." Some platforms also utilize simple rule-based heuristics that check for specific patterns or key terms to classify submissions. While functional for known petition types, this method struggles with ambiguous wording, misspellings, or new petition styles not included in the dictionary. These systems also lack the ability to identify duplicate petitions submitted by multiple users about the same issue. Furthermore, they do not analyze sentiment, urgency, or priority levels, limiting their usefulness in decision-making.

3.3 LIMITATIONS OF NLP-ONLY PETITION SYSTEMS

Heuristic-based methods analyze website or email features using predefined rules and patterns indicative of phishing behaviours such as the presence of specific keywords, suspicious redirects, or anomalies in SSL certificates. Signature-based detection uses known malware signatures matched against analysed content. Both methods provide fast, resource-efficient detection but are limited by the evolving nature of phishing tactics. Attackers can bypass static rules by varying domain names or adapting content, rendering these methods less effective for zero-day phishing attacks.

Although heuristic and signature-based approaches are efficient and widely adopted, they face significant limitations due to the dynamic nature of modern phishing techniques. Cybercriminals continuously modify their tactics, changing domain names, altering webpage layouts, or using obfuscated URLs that can evade predefined heuristic rules and outdated signature lists. These static detection methods are therefore less effective against zero-day phishing attacks— new or previously unseen threats that have not yet been analyzed or recorded.

These systems also lack the ability to identify duplicate petitions submitted by multiple users about the same issue. Furthermore, they do not analyze sentiment, urgency, or priority levels, limiting their usefulness in decision-making. To address these shortcomings, modern cybersecurity systems increasingly incorporate artificial intelligence and machine learning algorithms capable of learning adaptive patterns and detecting phishing attempts beyond predefined rules or known signatures. This evolution from static detection to intelligent, self- learning systems represents a major advancement in combating sophisticated phishing attacks. Modern digital petition platforms must evolve beyond static systems by integrating advanced analytics to strengthen decision-making and enhance public service efficiency. AI-driven processing enables automatic categorization, smart prioritization, and transparent resolution tracking. Real-time dashboards empower authorities to identify critical issues quickly, allocate resources effectively, and improve community engagement through data-supported actions.

3.4 LIMITATIONS OF EXISTING SYSTEMS

Natural Language Processing (NLP) has played a significant role in automating the categorization and analysis of petitions; however, systems that rely solely on text-based NLP techniques encounter several limitations that affect accuracy, scalability, and decision-making efficiency. One major drawback is the inability to fully understand the context behind citizen-submitted petitions, especially when descriptions are short, unstructured, or contain informal language. Petitions often include emotional expressions, local terminology, or mixed languages, which conventional NLP models struggle to interpret accurately.

Furthermore, NLP-only systems are ineffective when citizens upload images or PDF documents instead of typing detailed descriptions. Many petitions include photos of damaged roads, broken infrastructure, or safety hazards. Without image understanding capabilities, NLP-only systems cannot extract meaningful information from such submissions, reducing the overall effectiveness of automated analysis. These systems also fail to detect duplicate petitions submitted by multiple users describing the same issue in different wording.

Another limitation lies in the absence of multimodal analysis. Petition urgency or severity cannot always be inferred from textual content alone. Additional indicators such as image evidence, metadata, or user behavior often play a crucial role. NLP-only systems also lack real-time adaptability and may require frequent retraining to handle new phrases, petition categories, or evolving civic issues.

Moreover, these systems do not inherently support administrative functionalities such as tracking solved versus unsolved petitions, identifying trending issues, or generating analytical dashboards. This makes traditional NLP-only models insufficient for modern petition management platforms that require comprehensive and automated.

To overcome these limitations, advanced AI-powered petition systems integrate sentiment evaluation, urgency scoring, and visual evidence analysis, enabling automated prioritization and real-time insights. Such platforms enhance transparency, accelerate decision workflows, support resource planning, and significantly improve overall public grievance handling efficiency.

3.5 EXISTING SYSTEM

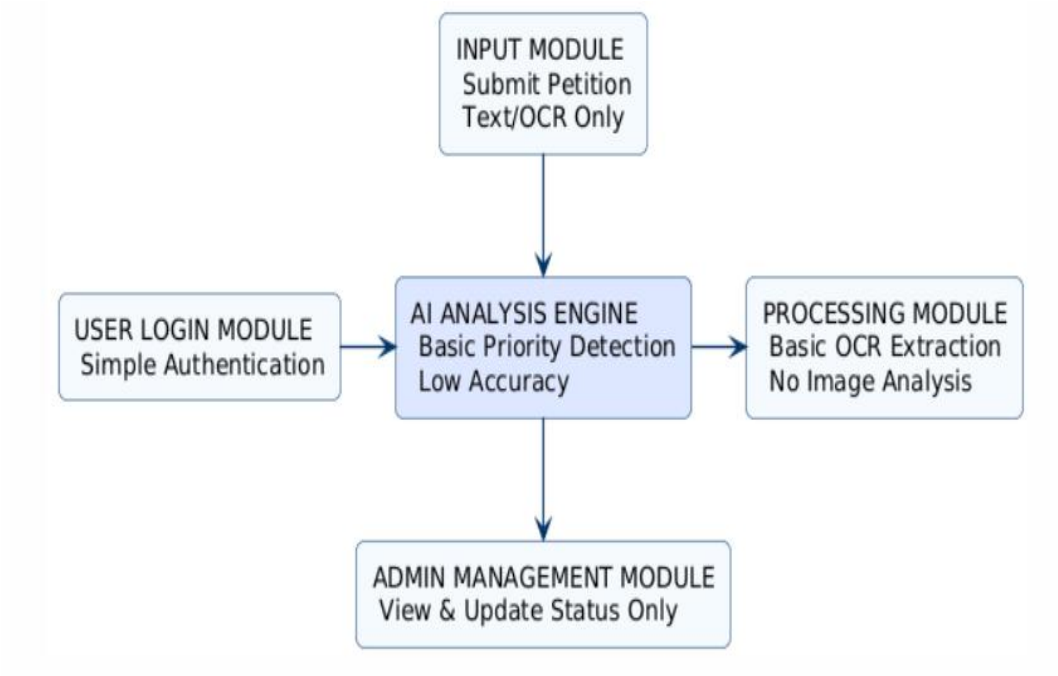

Figure. 3.5 Existing System

CHAPTER 4

PROBLEM IDENTIFIED

The existing Public Petition Management System is a simple digital platform used to collect citizen grievances and store them in a basic database. Although it allows users to submit petitions, the system lacks intelligent analysis, automated processing, and real-time monitoring. These limitations reduce the system's efficiency, delay issue resolution, and increase administrative workload for government authorities handling high volumes of public complaints.

4.1 No Image-Based Captioning or Visual Understanding

The current system accepts image uploads, but these files are only stored and not analyzed. There is no AI-driven captioning, object detection, or visual interpretation to understand the issue from the image. As a result, authorities cannot identify road damage, drainage, waste, or other visual problems automatically, leading to delayed classification and increased reliance on manual inspection.

4.2 Absence of Duplicate Petition Detection

When multiple citizens file complaints about the same issue or location, the system cannot identify these duplicates. Each petition is stored as a separate entry, increasing data redundancy and administrative workload. Officials often unknowingly work on the same complaint multiple times. Since no urgency rating or severity estimation is performed, all complaints appear identical. This leads to delayed action on emergency cases and inefficient resource allocation. This lack of similarity checking reduces efficiency and creates confusion in prioritizing petitions that represent recurring or collective public grievances.

4.3 No Real-Time Analytics Dashboard

The existing platform provides only static lists of petitions without any visual dashboards. There are no graphical charts, statistical insights, or trend monitoring tools. Authorities cannot analyze the number of active issues, category-wise complaints, or area-based trends.

4.4 Missing Social Interaction Features (Likes, Comments)

The system does not support user engagement features such as likes, comments, public feedback, or community discussions. Citizens cannot show support or urgency for issues raised by others. This limits transparency and prevents prioritization based on public impact. Authorities also lose valuable contextual information that comments could provide, resulting in incomplete understanding of real-world ground situations.

4.5 Only Basic Text Classification, No Advanced NLP

The platform performs only simple category classification and lacks modern Natural Language Processing capabilities. It cannot extract keywords, detect urgency, analyze sentiment, or identify tone (angry, frustrated, emergency). Without AI-based text understanding, the system treats all petitions equally, making it difficult for officials to distinguish between critical complaints and routine issues based solely on manual reading.

4.6 No Admin Workflow for Solved vs Unsolved Tracking

The existing model has no structured mechanism for marking petitions as solved, tracking progress, or updating resolution status. Officials must manually update records, often leading to inconsistency and incorrect reporting. Citizens also cannot monitor the status of their complaints. This absence of workflow control creates gaps in transparency, accountability, and follow-up procedures.

4.7 Inefficient Data Organization and Manual Processing

Although petitions are stored in a database, the system lacks proper indexing, categorization, and filtering capabilities. Staff must manually search, sort, and review complaints, increasing response time. The absence of automation for priority sorting or urgency grouping leads to inefficient processing, especially during peak workloads.

4.8 Limited Security and Access Control

The existing system provides only basic login authentication and lacks modern security features such as JWT tokens, role-based access control (RBAC), encrypted document storage, and activity logs. Sensitive citizen data like phone numbers and addresses remain vulnerable. Without proper authorization layers, unauthorized access, data tampering, or privacy violations pose significant risks.

4.9 Poor Scalability

The architecture is suitable only for small-scale usage and cannot handle large volumes of petitions from multiple regions. As the number of users increases, system performance degrades due to limited backend optimization. The lack of load balancing, distributed processing, or cloud infrastructure restricts scalability. This makes the system unsuitable for large municipalities or statewide deployment.

4.10 No Support for Priority-Based Decision Making

The existing system does not assist officials in identifying high-priority issues such as water leakage, electrical hazards, or safety threats. Since no urgency rating or severity estimation is performed, all complaints appear identical. This leads to delayed action on emergency cases and inefficient resource allocation. Authorities must manually review every petition to understand its seriousness.

CHAPTER 5

PROPOSED SYSTEM

5.1 PROPOSED SYSTEM

The proposed system introduces an advanced AI-powered petition analysis platform that automates the interpretation, categorization, and prioritization of citizen petitions using a combination of Natural Language Processing (NLP), machine learning, and image-based captioning models. Unlike traditional petition systems that rely solely on manual review or simple keyword matching, this system leverages deep learning techniques to analyze petition text, extract key information, detect sentiment, estimate urgency, and classify issues into predefined categories. This significantly reduces the workload for administrative staff while ensuring faster and more accurate responses to public concerns.

A key innovation of the proposed model is the integration of multimodal analysis, where both text and images are processed to enhance understanding. Using BLIP (Bootstrapped Language Image Pretraining), the system automatically generates descriptions from submitted images, enabling meaningful analysis even when users upload photos instead of writing detailed explanations. The platform also incorporates duplicate detection to identify recurring issues submitted by multiple citizens, reducing redundancy and improving issue clustering.

Additionally, the system provides a real-time dashboard that visualizes sentiment trends, urgency distribution, category counts, and like/comment engagement metrics. An admin panel allows authorized personnel to mark petitions as solved or unsolved, delete invalid submissions, and track priority levels efficiency. This centralized approach enhances transparency and ensures faster decision-making by reducing manual review efforts. Users can monitor their petition status instantly, improving trust and participation. The integration of AI-driven insights helps authorities identify recurring community problems, allocate resources effectively, respond proactively, and continuously improve public service quality through evidence-based planning and performance measurement.

5.2 SYSTEM ARCHITECTURE

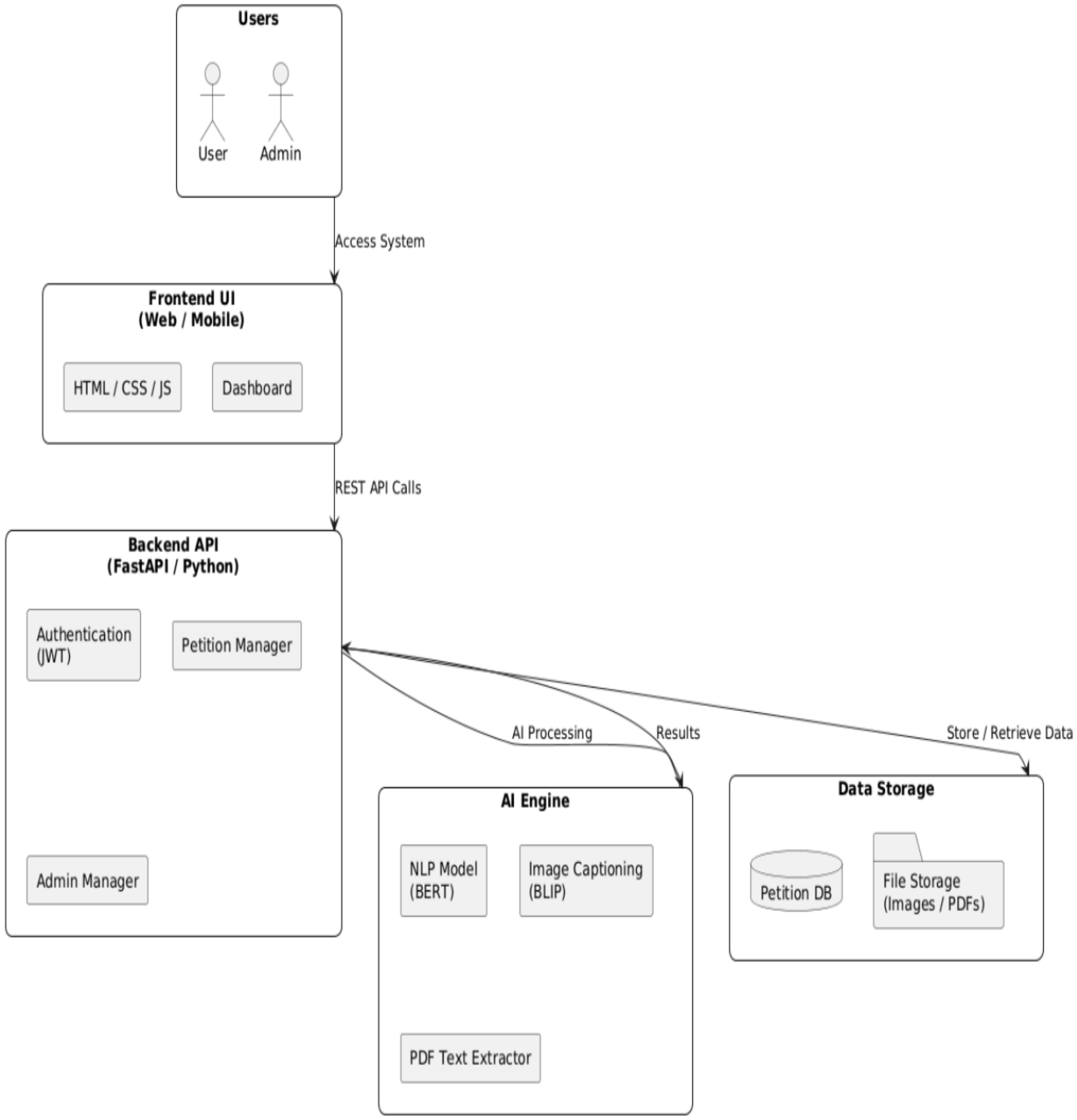

Figure. 5.2 Proposed System Architecture

5.3 USECASE DIAGRAM

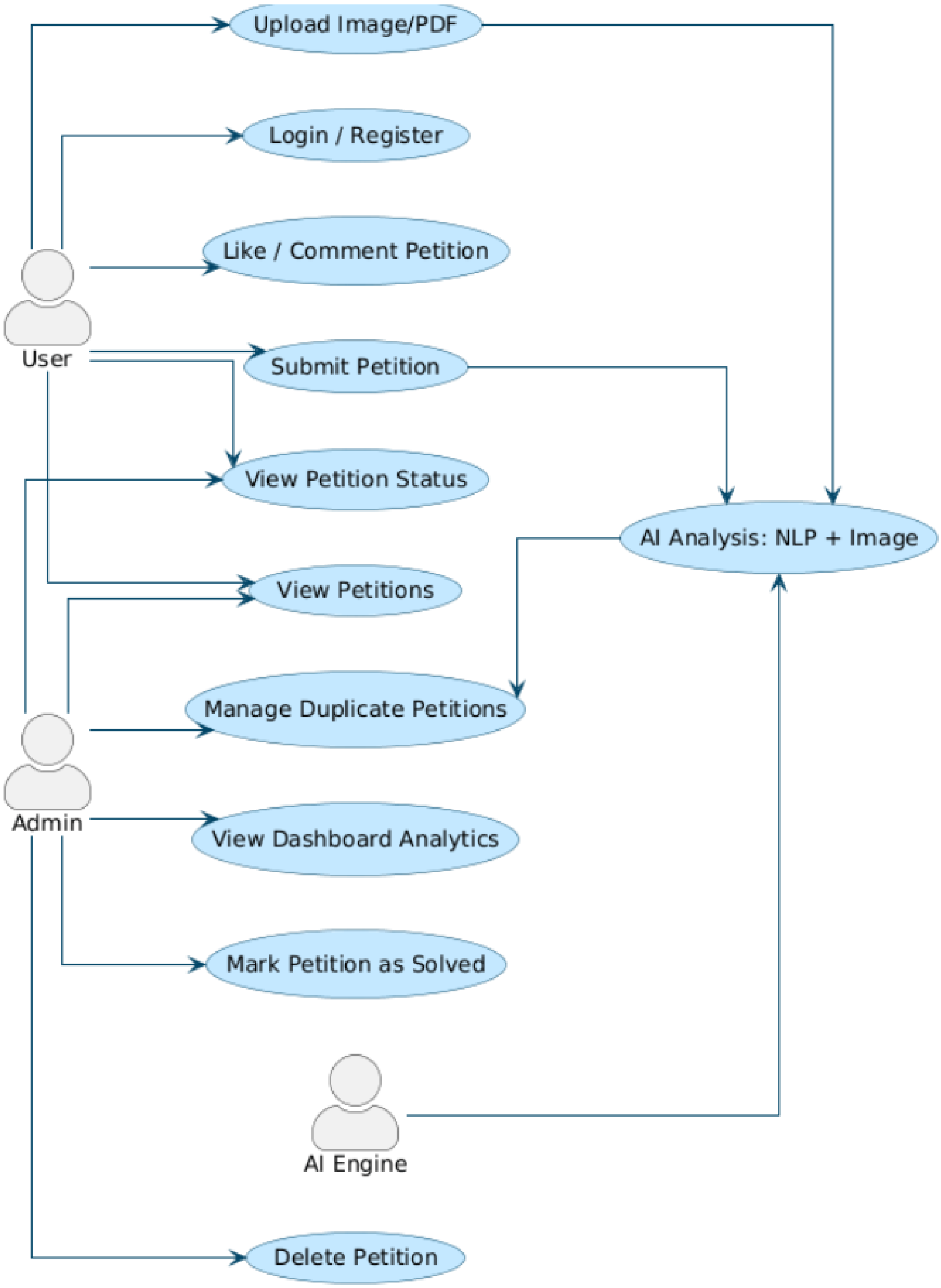

Figure. 5.3 Usecase

5.4 CLASS DIAGARM

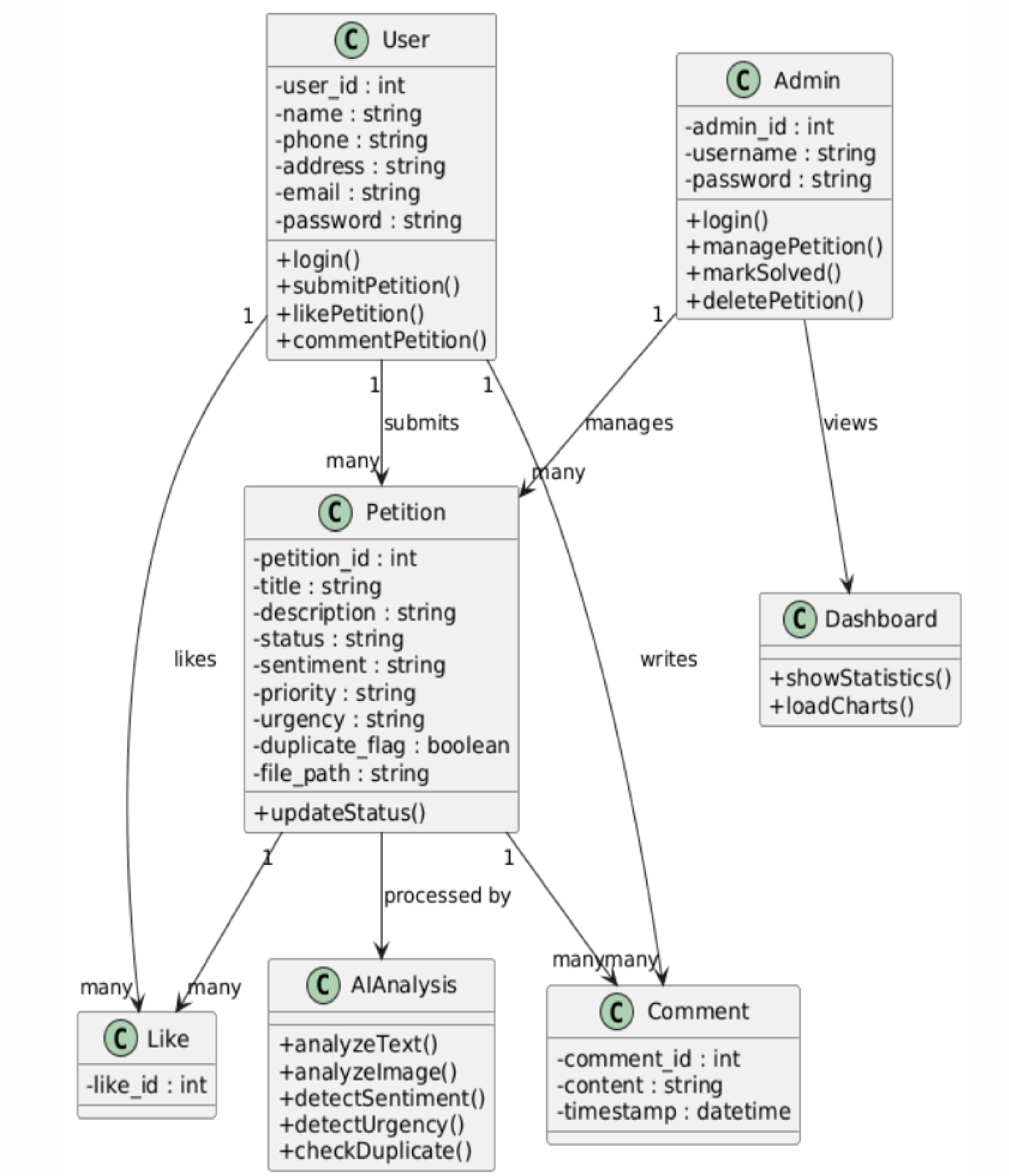

Figure. 5.4 Class Diagram

5.5 STATE DIAGRAM

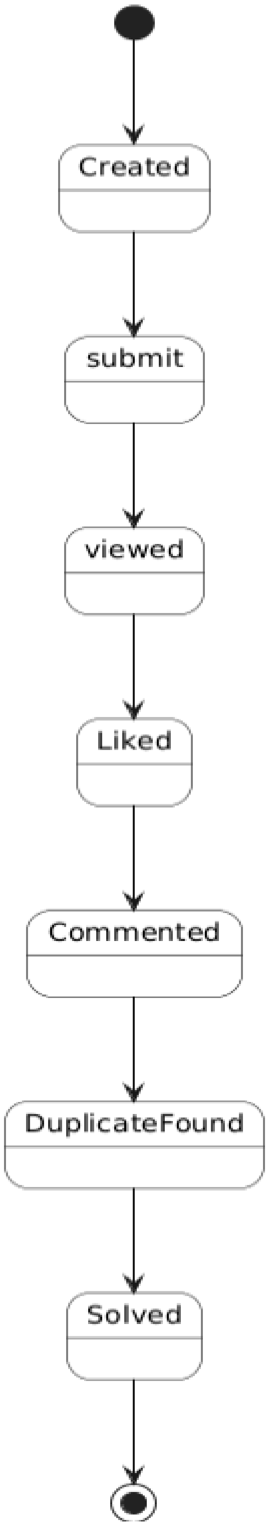

Figure. 5.5 State Diagram

5.6 ACTIVITY DIAGRAM

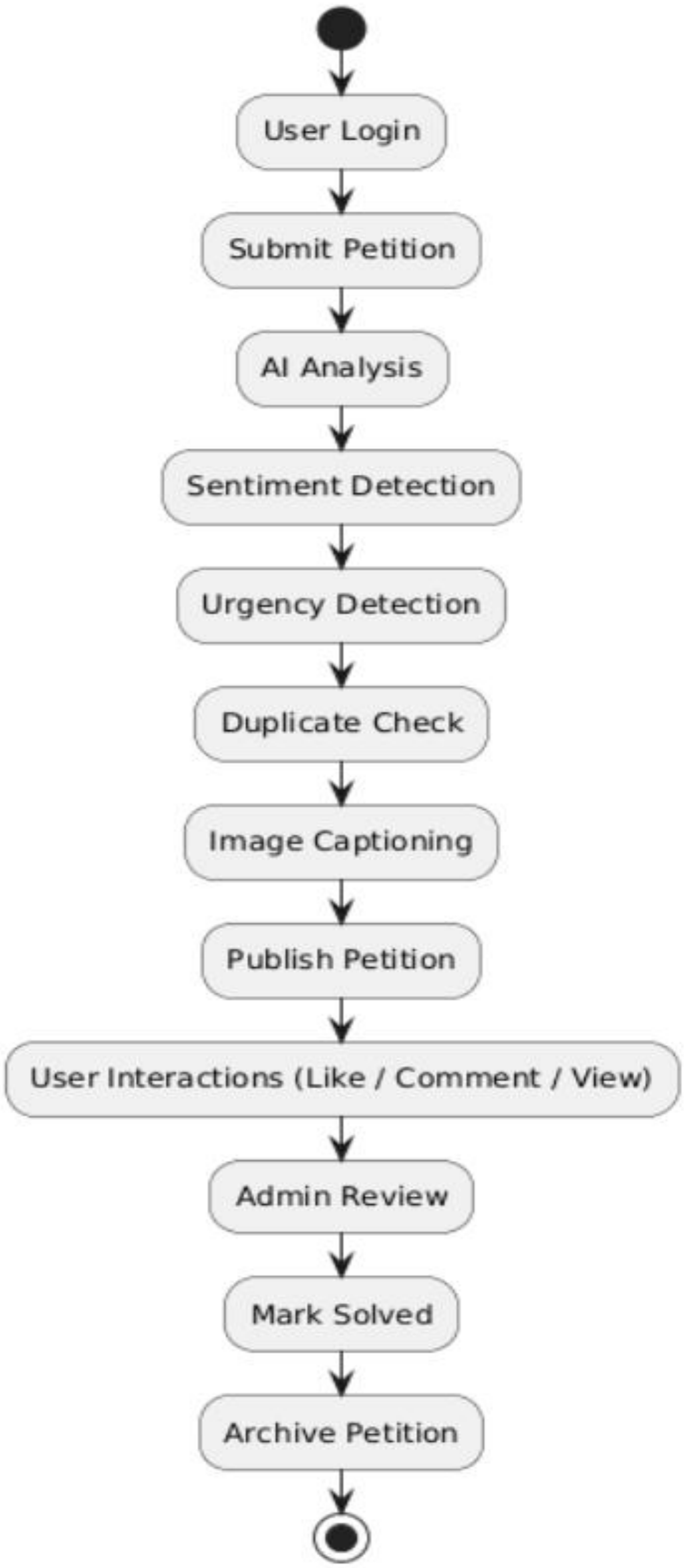

Figure. 5.6 Activity Diagram

5.7 SEQUENCE DIAGRAM

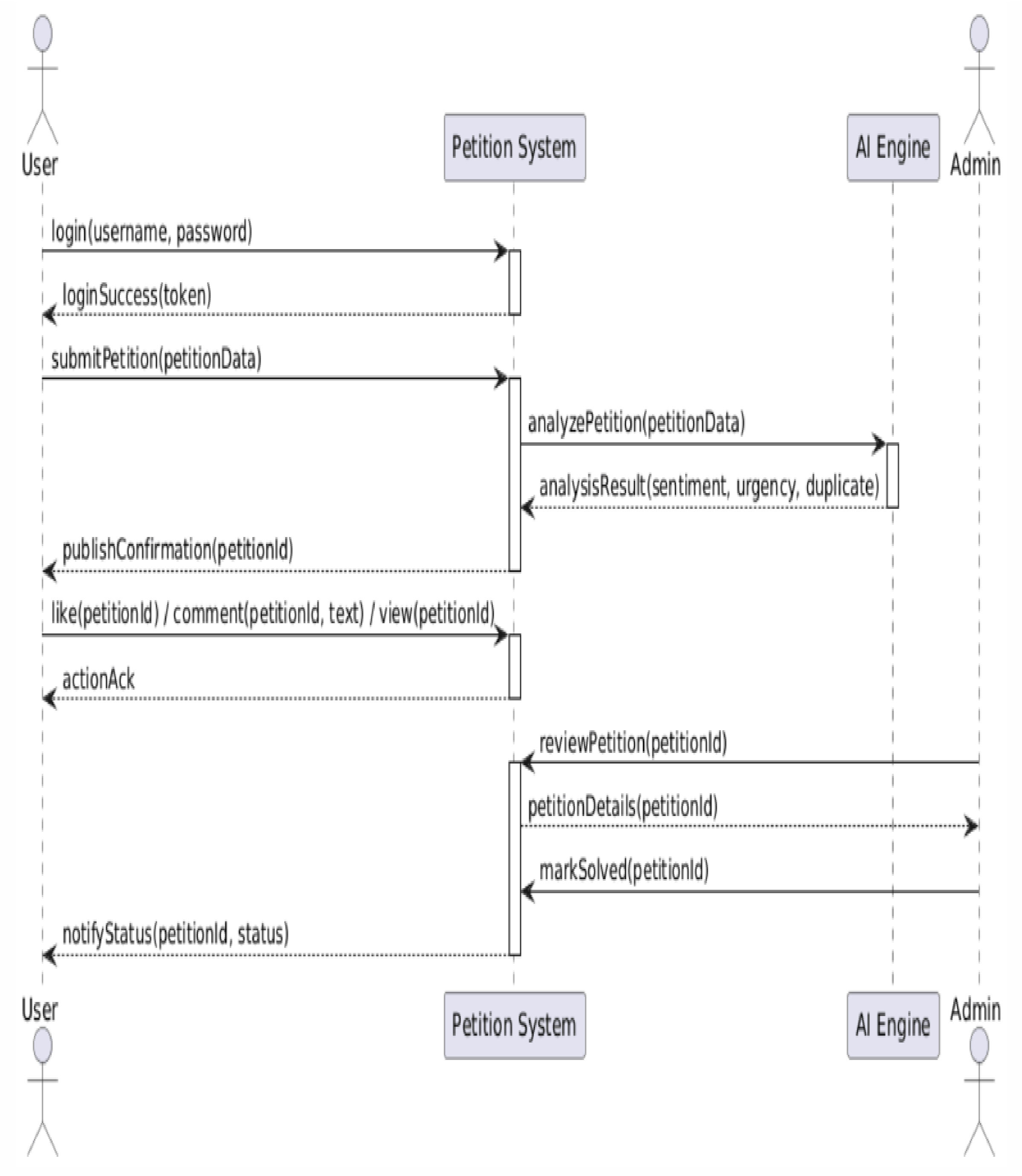

Figure. 5.7 Sequence Diagram

5.8 DEPLOYMENT DIAGRAM

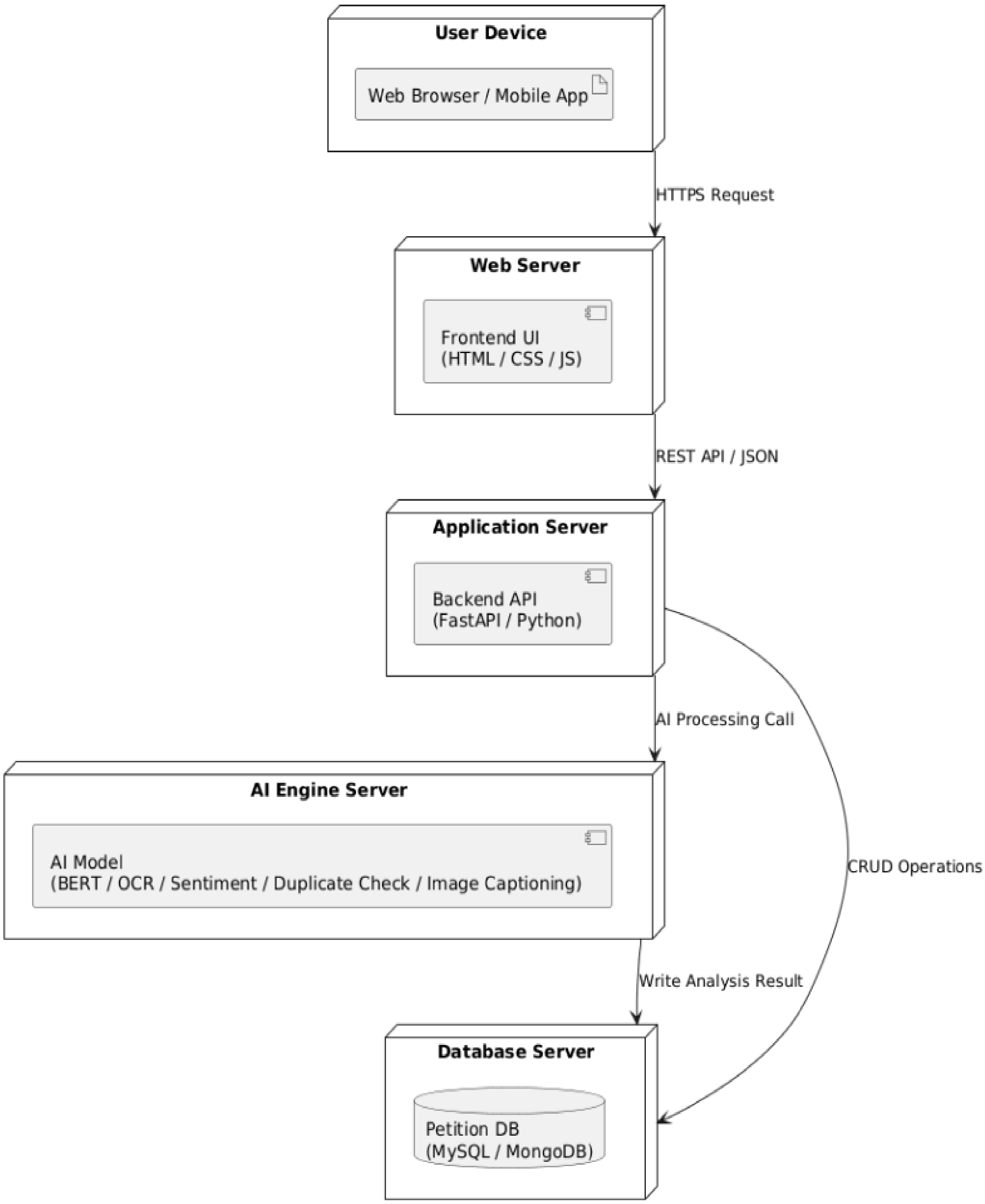

Figure. 5.8 Deployment Diagram

CHAPTER 6

SOFTWARE DESCRIPTION

6.1 HARDWARE REQUIREMENTS

Components	Specification
Processor	- Dual-Core CPU (Intel i3)
RAM	- Minimum 4 GB (8 GB recommended for AI processing)
Storage	- 50–100 GB HDD/SSD
Display	- 1366 × 768 resolution or higher
Network	- Stable internet for package installation & API calls

6.2 SOFTWARE REQUIREMENTS

Components	Specification
Operating System	- Windows 10/11, Linux (Ubuntu 20.04+), macOS
Backend Framework	- Python Flask 2.x or FastAPI
Database	- SQLite (development), MySQL / PostgreSQL
AI Integration	- BLIP Image Captioning, Scikit-Learn ML Model
Libraries	- flask, SQLAlchemy, torch, scikit-learn, Pillow
Frontend Technologies	- HTML, CSS, JavaScript, Fetch API

CHAPTER 7
SYSTEM IMPLEMENTATIONS

7.1 MODULE DESCRIPTION

Each module in the system is designed to perform specialized functions while ensuring seamless integration across the platform. The Petition Submission Module allows users to submit grievances with text, images, and documents. The NLP and Sentiment Analysis Module processes petition content to extract emotional tone and contextual meaning. The Priority Prediction Module automatically classifies petitions based on urgency and severity. The Admin Management Module enables administrators to review, update, and track petition statuses. The Dashboard and Analytics Module provides visual insights and trends, while the Image Captioning Module interprets uploaded images to enhance overall analysis accuracy.

7.2 LIST OF MODULES

7.2.1. Petition Submission Module
7.2.2 NLP Processing & Sentiment Analysis Module
7.2.3 Priority Prediction Module
7.2.4 Admin Management Module
7.2.5 Dashboard & Analytics Module
7.2.6 Image Captioning Integration (BLIP Model)

7.2.1 Petition Submission Module

The Petition Submission Module serves as the entry point of the system, allowing users to submit their concerns in a structured and user-friendly manner. This module is designed to accept petitions through multiple input formats, ensuring accessibility and flexibility for all users. The module collects essential details such as the petition title, petitioner’s name, phone number, address, and a descriptive explanation of the issue.

For PDF submissions, the system extracts textual content using automated PDF text-processing tools. For image-based submissions, the integrated BLIP image captioning model generates descriptive text from the uploaded images, enabling the system to analyze visuals effectively. The module also allows users to mark petitions as Private or Public, giving them control over visibility.

This module ensures that every petition is preprocessed and stored in a structured manner, making it suitable for later stages such as NLP analysis, sentiment detection, urgency evaluation, and priority prediction. Its multi-format input capability significantly improves usability and encourages wider public participation.

7.2.2 NLP Analysis Module

The NLP Processing & Sentiment Analysis Module is responsible for transforming raw petition text into a structured, analyzable format and identifying the emotional tone conveyed by the petitioner. Once a petition is submitted—whether as typed text, extracted PDF content, or image-generated caption—the module initiates a sequence of natural language preprocessing steps such as tokenization, lowercasing, stop-word removal, and normalization.

Following preprocessing, the module applies sentiment analysis techniques to determine whether the petition expresses a Positive, Neutral, or Negative sentiment. This helps the system understand the emotional urgency or frustration level embedded in the petitioner's words. The module also contributes to urgency estimation by identifying terms such as "urgent," "immediately," "critical," or "important." These linguistic cues help classify petitions that require faster administrative attention. By converting unstructured text into sentiment and urgency indicators, this module plays a key role in supporting priority prediction, decision-making, and overall petition categorization.

7.2.3 Priority Prediction Module

The Priority Prediction Module plays a crucial role in determining the relative importance and urgency of each submitted petition. After the text has been preprocessed and sentiment data extracted, this module applies a machine learning–based classification model to automatically assign a High, Medium, or Low priority level. The system uses a trained Random Forest classifier along with TF-IDF vectorization to convert petition text into meaningful numerical representations. These features allow the model to detect linguistic patterns, issue severity, and contextual cues that indicate urgency or significance.

During model training, historical petition data and annotated priority labels are used to learn how different textual characteristics correlate with priority categories. This module ensures that every petition is preprocessed and stored in a structured manner, making it suitable for later stages such as NLP analysis, sentiment detection, urgency evaluation, and priority prediction. Its multi-format input capability significantly improves usability and encourages wider public participation. Once trained, the model evaluates incoming petitions in real time, offering consistent and unbiased decision-making that assists administrators in managing workflow efficiently.

This automated prediction reduces manual evaluation effort and helps highlight petitions that require faster resolution. The Priority Prediction Module ensures that critical public concerns are surfaced promptly and handled with appropriate urgency, ultimately improving responsiveness and enhancing the effectiveness of the petition handling process. This automated prediction system significantly reduces the need for manual evaluation by quickly analyzing the content of each petition. By leveraging AI-driven priority prediction, the system identifies petitions that demand immediate attention, ensuring that urgent issues are addressed without delay. This not only improves response times but also enhances overall operational efficiency. Additionally, it helps administrators focus their resources on the most critical public concerns, streamlines workflow, minimizes human error, and provides data-driven insights for better decision-making, ultimately resulting in a more transparent, responsive, and effective petition management process.

7.2.4 Admin Management Module

The Admin Management Module serves as the central control unit of the entire petition analysis system, providing authorized administrators with full access to manage, analyze, and take action on user-submitted petitions. This module is designed to ensure smooth governance, transparent workflow processing, and efficient resolution tracking. It provides a secure login system that validates admin credentials and generates authentication tokens to prevent unauthorized access.

Once logged in, administrators gain access to a well-structured interface where all petitions are displayed along with key analytical attributes such as sentiment score, urgency level, predicted priority, duplicate status, and the number of likes or user interactions. The module supports advanced filtering options, enabling admins to sort petitions by High Priority, Unsolved, Solved, Duplicate, or other relevant criteria. This helps administrators navigate large volumes of petitions quickly and focus on issues that demand immediate attention.

Crucial administrative actions are handled within this module. Admins can mark petitions as Solved or revert them back to Unsolved, helping maintain accurate workflow tracking. They can also delete inappropriate or redundant petitions to keep the database clean and manageable. All updates made through this module are instantly reflected in the system's database and dashboard.

Additionally, the Admin Management Module integrates analytics from the dashboard, providing visual insights on petition trends, sentiment distribution, urgency levels, and user engagement metrics. This empowers administrators with data-driven decision-making capabilities. Overall, this module ensures operational efficiency, effective petition governance, and streamlined communication between the system and administrative authorities. The Admin Management Module integrates analytics from the dashboard, providing visual insights on petition trends, sentiment distribution, urgency levels, and user engagement metrics. This empowers administrators with data-driven decision-making capabilities, enabling them to identify recurring issues, allocate resources effectively.

7.2.5 Dashboard & Analytics Module

The Dashboard & Analytics Module provides a visual representation of the system's petition data, enabling administrators to monitor trends, evaluate performance, and make informed decisions. This module aggregates analysis results generated by various components—such as sentiment classification, urgency detection, priority prediction, and duplicate identification—and presents them through interactive charts and graphs. These visual elements allow administrators to quickly assess the overall status of submitted petitions.

The dashboard displays multiple analytical metrics, including the distribution of petitions across priority levels, the proportion of positive, negative, and neutral sentiments, and the frequency of urgent petitions that require immediate attention. the Admin Management Module integrates analytics from the dashboard, providing visual insights on petition trends, sentiment distribution, urgency levels, and user engagement metrics It also highlights user engagement indicators such as like counts and comment activity, giving insights into public interest and concern.

Furthermore, the module allows administrators to track recurring issues, identify patterns in citizen concerns, and assess the effectiveness of responses over time. By presenting data in interactive charts and graphs, the system simplifies complex information, enabling quick interpretation and action. This ensures that critical petitions are prioritized, resources are allocated efficiently, and overall workflow is optimized. The analytics also support transparency and accountability, helping administrators make informed, timely decisions while improving citizen satisfaction and engagement.

In addition, the dashboard supports monitoring of petition resolution status, highlighting solved, unsolved, and duplicate petitions for better management. It enables trend analysis over different time periods, helping administrators anticipate potential issues and plan proactive interventions. The system's visual insights also facilitate reporting to higher authorities and stakeholders, providing a clear overview.

7.2.6 Image Captioning Integration (Blip Model)

The Image Captioning Integration Module enhances the system's ability to process non-textual petition submissions by converting uploaded images into meaningful textual descriptions. This feature is powered by the BLIP (Bootstrapped Language-Image Pretraining) model, a state-of-the-art deep learning framework capable of generating accurate natural language captions from visual input. When a user uploads an image, the module preprocesses it and feeds it into the BLIP model, which analyzes visual elements such as objects, scenes, and contextual patterns before generating a descriptive sentence that captures the essence of the image.

The generated caption is then treated as the petition's textual description and passed to subsequent NLP modules for sentiment analysis, urgency estimation, and priority prediction. This module aggregates analysis results generated by various components—such as sentiment classification, urgency detection, priority prediction, and duplicate identification. This integration ensures that the system can handle a wide range of petition formats, especially those submitted by users who may prefer sharing photographs over writing detailed descriptions.

The Admin Management Module enables administrators to review, update, and track petition statuses. The Dashboard and Analytics Module provides visual insights and trends, while the Image Captioning Module interprets uploaded images to enhance overall analysis accuracy. By enabling the automatic interpretation of images, the module not only improves accessibility but also increases the accuracy and completeness of petition analysis. This multimodal capability significantly strengthens the overall intelligence and usability of the petition-processing platform.

Moreover, the integration of NLP-driven text analysis with image captioning allows the system to comprehensively assess both textual and visual content. Administrators can quickly identify high-priority issues, detect duplicates, and monitor public sentiment effectively. This combined approach ensures faster decision-making.

CHAPTER 8

SYSTEM TESTING

System testing evaluates the entire AI-powered petition analysis platform to ensure all modules work together correctly. It verifies petition submission, NLP processing, image captioning, priority prediction, dashboard analytics, and admin operations. The goal is to confirm complete functionality, stability, and readiness for real-world deployment without errors.

8.1 UNIT TESTING

Unit testing focuses on verifying the correctness and reliability of individual modules within the AI-powered petition analysis system. Each module is tested independently to ensure it performs its intended function without errors before being integrated with other components. Core units tested include the petition submission handler, text preprocessing functions, NLP sentiment analyzer, urgency detector, priority prediction model, duplicate checker, and BLIP image captioning logic. For every unit, specific test cases are designed using sample inputs such as short text petitions, lengthy descriptions, noisy data, and various image or PDF formats.

Assertions are used to validate expected outputs, such as correct sentiment labels, accurate priority classification, proper extraction of image captions, and valid database entries. Unit testing also ensures that API endpoints, database transactions, and UI components function cohesively. The primary objective is to confirm overall stability, reliability, and readiness for real-world deployment. Unit tests also verify error handling, ensuring the system gracefully manages invalid or missing inputs.

Furthermore, unit testing helps identify and isolate bugs early in the development process, reducing potential system failures during production. It also ensures that updates or modifications do not break existing functionality. Automated tests improve maintainability.

8.2 INTEGRATION TESTING

Integration testing evaluates how well the individual modules of the AI-powered petition analysis system interact once combined. After unit testing ensures each module works independently, integration testing focuses on validating the data flow and communication between components such as the petition submission module, NLP processing pipeline, BLIP image captioning module, priority prediction model, and the database.

Test cases involve submitting petitions through various formats, including text, images, and PDFs, and verifying whether extracted data is correctly passed to subsequent modules for analysis. The system's ability to store processed results, retrieve petitions for display, and update sentiment, urgency, and priority attributes is thoroughly examined. Admin operations—such as marking petitions as solved, filtering them, and deleting entries—are also tested for functional correctness.

8.3 SYSTEM TESTING

System testing involves evaluating the entire AI-powered petition analysis platform as a complete, integrated application to ensure it meets all functional and non-functional requirements. Unlike unit or integration testing, which focus on individual modules or module interaction, system testing verifies the behavior of the full system from a user's perspective. Test scenarios include submitting petitions through text, PDF, and image uploads, verifying that image captions are correctly generated using the BLIP model, and ensuring all NLP-based outputs such as sentiment, urgency, and priority values are accurate.

The dashboard is tested for correct visualization of statistics, including sentiment distribution and priority charts. Admin operations—such as filtering petitions, marking them as solved, detecting duplicates, and deleting entries—are validated for accuracy and responsiveness.

8.4 PERFORMANCE TESTING

Performance testing evaluates the speed, efficiency, and scalability of the AI-powered petition analysis system under different load conditions. The objective is to ensure the platform can handle multiple petitions, concurrent user submissions, and intensive AI processing tasks without degrading performance. Tests are conducted to measure response time for text-based petitions, PDF extraction, and BLIP image caption generation, ensuring that each process executes within acceptable time limits. Database performance is assessed by analyzing retrieval speed, insertion time, and the efficiency of filtering petitions based on priority, sentiment, and status. Stress tests simulate peak usage conditions where numerous petitions are submitted simultaneously, verifying that the backend, APIs, and database remain stable.

8.5 USABILITY TESTING

Usability testing focuses on evaluating how easily users and administrators interact with the petition analysis system. Test participants submit petitions, upload PDFs or images, and navigate between pages to assess clarity, responsiveness, and simplicity. Usability testing also ensures that API endpoints, database transactions, and UI components function cohesively. The primary objective is to confirm overall stability, reliability, and readiness for real-world deployment. The admin interface is tested for ease of filtering, solving, and deleting petitions, as well as reviewing dashboard analytics. Feedback helps refine layout, improve button placement, and enhance overall user experience. The goal is to ensure intuitive navigation and smooth interaction for all user roles. Additionally, usability testing identifies potential bottlenecks or confusing workflows, allowing developers to optimize page load times, form validations, and interactive elements. It ensures that both novice and experienced users can efficiently perform tasks, enhancing accessibility and satisfaction. Continuous feedback from testing drives iterative improvements, resulting in a more user-friendly and effective petition management system.

CHAPTER 9

RESULTS AND DISCUSSION

The AI-powered petition analysis and categorization system was developed to intelligently process, classify, and manage citizen petitions submitted in various formats, including text, PDFs, and images. The results demonstrate that the system effectively meets its objectives by providing accurate sentiment analysis, urgency detection, priority classification, and duplicate identification while ensuring a smooth user and administrator experience.

The system successfully analyzed textual petitions with strong accuracy across sentiment and urgency categories. Negative sentiment petitions, often containing complaints or distress, were correctly identified and marked with higher urgency, while neutral petitions exhibited low emotional polarity. The integration of rule-based urgency detection boosted the system's ability to recognize critical petitions containing terms such as "urgent" or "immediately." These improvements enhance administrative awareness and help allocate attention to pressing issues.

The priority prediction model, trained using TF-IDF features and a Random Forest classifier, delivered consistent results across High, Medium, and Low priority levels. High-priority petitions generally aligned with those expressing strong negative sentiment or urgent concerns, indicating that the model learned meaningful linguistic patterns during training. The dashboard visualizations further validated these predictions by clearly mapping the distribution of priorities across the dataset.

One of the most significant outcomes was the successful integration of the BLIP image captioning model. When users uploaded images instead of text, the model generated descriptive captions with high textual relevance. These captions allowed the NLP pipeline to process image-based petitions similarly to text-based ones, expanding the accessibility of the system. This multimodal capability provides a meaningful advantage over traditional petition platforms that rely solely on manually written entries.

The duplicate detection feature was also evaluated and found effective in reducing redundancy. Petitions with identical or highly similar titles and descriptions were flagged as duplicates, preventing unnecessary clutter and enabling administrators to focus on unique concerns. The like and comment interactions from users further added a layer of community engagement, helping highlight popular or widely supported petitions can be submitted.

Administrative tools performed reliably during testing. Admins were able to view, filter, solve, or delete petitions without system errors. The dashboard provided real-time analytical insights, including sentiment distribution, priority breakdowns, urgency graphs, and interaction metrics. These results confirm that the administrative workflow is enhanced through automation and visual clarity.

Overall, the results indicate that the AI-driven petition analysis system significantly enhances the petition management process through automation, accuracy, and intelligent analysis. The ability to handle multimodal inputs, detect urgent concerns, reduce redundancy, classify priorities, and provide data-driven insights makes the system a valuable tool for modern governance. By improving administrative workflows and enabling faster identification of critical public issues, the system promotes efficiency, transparency, and better responsiveness to citizen needs.

Moreover, the integration of NLP and image analysis ensures that both textual and visual information are effectively interpreted, providing a comprehensive understanding of each petition. The automated duplicate detection minimizes redundant entries, while sentiment and urgency analysis allow administrators to focus on high-priority cases. The interactive dashboard and analytics module offer real-time insights into trends, public engagement, and system performance, supporting informed decision-making. Overall, the system reduces manual workload, accelerates petition resolution, and enhances accountability. Its scalable design ensures adaptability for different administrative environments, making it a robust and reliable solution for modern, citizen-centric.

CHAPTER 10

CONCLUSION AND FUTURE ENHANCEMENTS

10.1 CONCLUSION

The AI-powered petition analysis and categorization system successfully demonstrates how modern technologies such as Natural Language Processing, machine learning, and multimodal image captioning can significantly improve the efficiency of handling public grievances. By automating key processes such as sentiment detection, urgency evaluation, priority prediction, and duplicate identification, the system reduces manual workload and supports faster, more informed administrative decision-making.

The integration of the BLIP image captioning model further extends the platform's capabilities, allowing image-based petitions to be interpreted with meaningful accuracy, thus ensuring inclusivity for users who may prefer visual submissions over written content.

The admin dashboard provides clear analytical insights, enabling authorities to visualize petition trends, monitor public sentiment, and manage petitions effectively. The system's ability to process text, PDF, and image inputs makes it versatile and adaptable to real-world use cases. Overall, the project achieves its objective of building an intelligent, scalable, and user-friendly petition management solution that enhances transparency, responsiveness, and efficiency in public service systems. It also provides a strong foundation for integrating future features such as multilingual support and real-time notification systems. Ultimately, this project demonstrates how AI-driven tools can transform traditional grievance handling into a more proactive and citizen-centric process.

Furthermore, incorporating features like predictive analytics, department-wise routing, and geo-tagging can help authorities anticipate emerging issues, allocate resources efficiently, and ensure faster, more effective resolution of citizen petitions.

10.2 FUTURE ENHANCEMENTS

While the AI-powered petition analysis and categorization system already delivers automated and intelligent handling of public grievances, several enhancements can further increase its effectiveness, scalability, and real-world applicability. Future versions of the system can incorporate advanced deep learning models such as BERT, RoBERTa, or GPT-based transformers to improve sentiment accuracy, urgency prediction, and contextual understanding. Expanding the duplicate detection module using semantic similarity algorithms or embedding-based matching would enhance reliability when dealing with large-scale petition datasets.

Additional improvements could include multilingual petition processing to support users from diverse linguistic backgrounds, as well as voice-based petition submission for accessibility. Integration with government portals, mobile applications, and chatbot systems would increase reach and ease of use. Real-time notifications for petition updates and status changes can enhance engagement between citizens and administrators.

Moreover, it also provides a strong foundation for integrating future features such as multilingual support and real-time notification systems. Ultimately, this project demonstrates how AI-driven tools can transform traditional grievance handling into a more proactive and citizen-centric process. incorporating geo-tagging, department-wise routing, and predictive analytics can help authorities identify emerging issues and allocate resources more efficiently.

In addition, integrating machine learning models for trend prediction can enable proactive interventions before issues escalate, while advanced visualization tools can provide deeper insights into citizen concerns. Real-time notifications and alerts can keep both administrators and petitioners informed about status updates, improving transparency and engagement. Multilingual support ensures accessibility for a diverse population, broadening the system’s reach. Geo-tagging and location-based analytics allow authorities to detect regional problem patterns and optimize resource deployment. By combining these enhancements, the AI-powered petition management platform not only streamlines operations but also fosters trust, accountability.

APPENDIX – A

SOURCE CODE

app.py

```
from flask import Flask, send_from_directory
from backend.config.db_config import db, init_db
from backend.auth.auth_routes import auth_bp
from backend.api.analyze_routes import analyze_bp
from backend.petitions.petition_routes import petition_bp
from backend.admin.admin_routes import admin_bp
from backend.dashboard.dashboard_routes import dashboard_bp
import os
from flask_jwt_extended import JWTManager
app = Flask(__name__)
init_db(app)
# JWT configuration
app.config["JWT_SECRET_KEY"] = "super-secret-key"  # change to a strong key
app.config["JWT_TOKEN_LOCATION"] = ["headers"]
jwt = JWTManager(app)
BASE_DIR = os.path.dirname(os.path.dirname(os.path.abspath(__file__)))
FRONTEND_DIR = os.path.join(BASE_DIR, "web_frontend")
# Register Blueprints
app.register_blueprint(auth_bp, url_prefix="/auth")
app.register_blueprint(analyze_bp)
app.register_blueprint(petition_bp)
app.register_blueprint(admin_bp)
app.register_blueprint(dashboard_bp)
@app.route("/")
def serve_index():
    return send_from_directory(FRONTEND_DIR, "index.html")
@app.route("/<path:path>")
def serve_static(path):
    return send_from_directory(FRONTEND_DIR, path)
```

```
if __name__ == "__main__":
    with app.app_context():
        db.create_all()
    print("🚀 Petition System Backend Running")
    app.run(debug=True)
```

image_analysis.py

```
# backend/analysis/image_analysis.py
from PIL import Image
from transformers import BlipProcessor, BlipForConditionalGeneration
# Lazy load global model to avoid re-loading each request
_processor = None
_model = None
def _load_blip():
    global _processor, _model
    if _processor is None or _model is None:
        print("🔹 Loading BLIP image captioning model (first time may take a
minute)...")
        _processor = BlipProcessor.from_pretrained("Salesforce/blip-image-captioning-
base")
        _model = BlipForConditionalGeneration.from_pretrained("Salesforce/blip-
image-captioning-base")
        print("✅ BLIP model loaded.")
    return _processor, _model
def generate_caption(image_path: str) -> str:
    """
    Generate a natural-language caption for the image using BLIP.
    """
    try:
        processor, model = _load_blip()
        image = Image.open(image_path).convert("RGB")
        inputs = processor(images=image, return_tensors="pt")
```

```
        out = model.generate(**inputs, max_length=60)
        caption = processor.decode(out[0], skip_special_tokens=True)
        return caption
    except Exception as e:
        return f"Error generating caption: {e}"
```

train_model.py

```
import pandas as pd
from sklearn.feature_extraction.text import TfidfVectorizer
from sklearn.ensemble import RandomForestClassifier
import pickle
import os
# Paths
CURRENT_DIR = os.getcwd()  # analysis folder
DATA_FILE = os.path.join(CURRENT_DIR, "..", "data", "petition_data.csv")
MODEL_DIR = os.path.join(CURRENT_DIR, "..", "model")
MODEL_FILE = os.path.join(MODEL_DIR, "priority_model.pkl")
VECTORIZER_FILE = os.path.join(MODEL_DIR, "vectorizer.pkl")
# Load CSV
df = pd.read_csv(DATA_FILE)
# If no 'content' column, just use 'title' for training
if 'content' in df.columns:
    df['text'] = df['title'] + " " + df['content']
else:
    df['text'] = df['title']
# Vectorize
vectorizer = TfidfVectorizer()
X = vectorizer.fit_transform(df['text'])
y = df['priority']
# Train model
model = RandomForestClassifier(n_estimators=100, random_state=42)
model.fit(X, y)
```

```
# Create model folder if not exists
os.makedirs(MODEL_DIR, exist_ok=True)
# Save model and vectorizer
with open(MODEL_FILE, "wb") as f:
    pickle.dump(model, f)
with open(VECTORIZER_FILE, "wb") as f:
    pickle.dump(vectorizer, f)
print("✅ Priority model trained and saved at:", MODEL_FILE)
```

urgency_analyzer.py

```
def detect_urgency(text):
    if not text:
        return 0
    keywords = ["urgent", "immediately", "important", "emergency", "critical"]
    count = sum(1 for word in keywords if word in text.lower())
    return count
```

sentiment.py

```
from textblob import TextBlob
def analyze_sentiment(text):
    if not text:
        return "Neutral"
    polarity = TextBlob(text).sentiment.polarity
    if polarity > 0.2:
        return "Positive"
    elif polarity < -0.2:
        return "Negative"
    else:
        return "Neutral"
```

APPENDIX – B

SCREENSHOTS

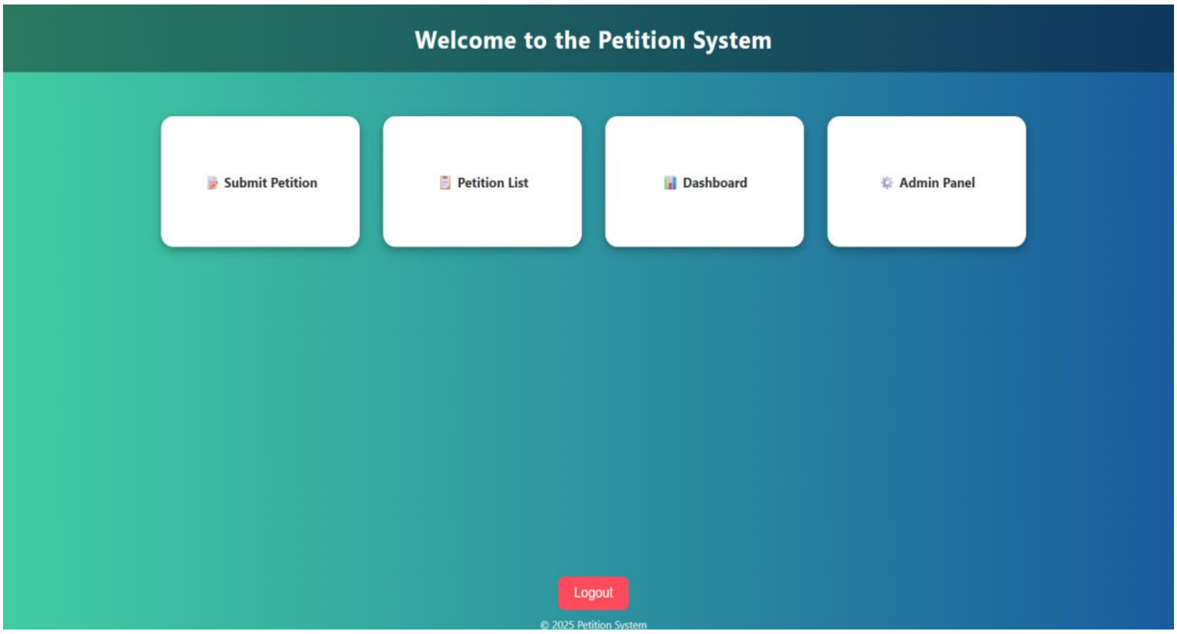

Figure. B.1 User Login

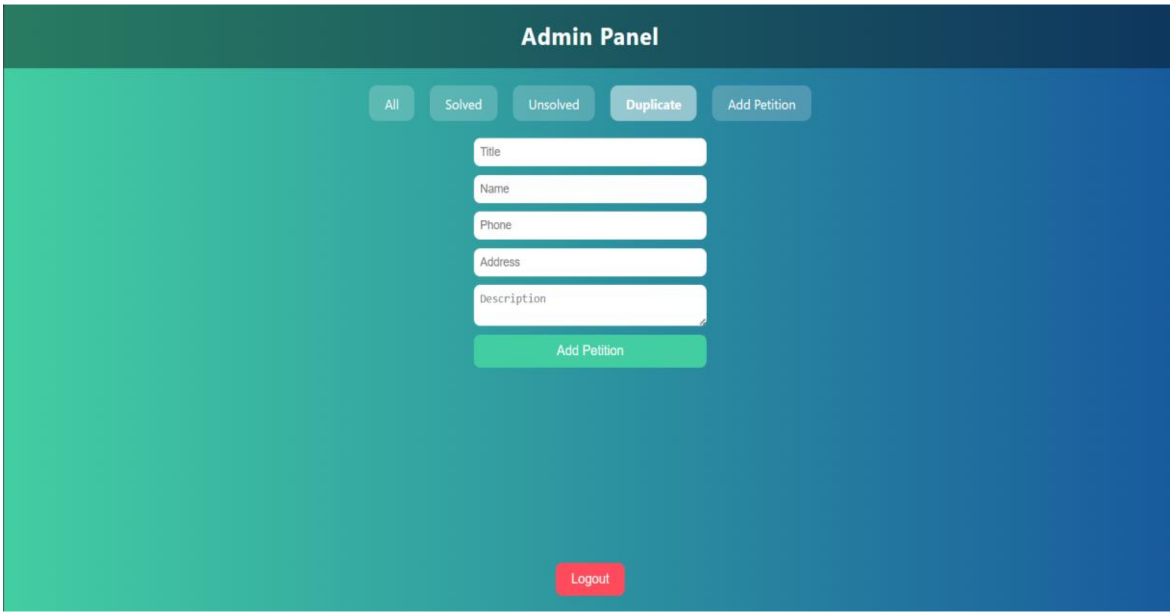

Figure. B.2 Admin Panel

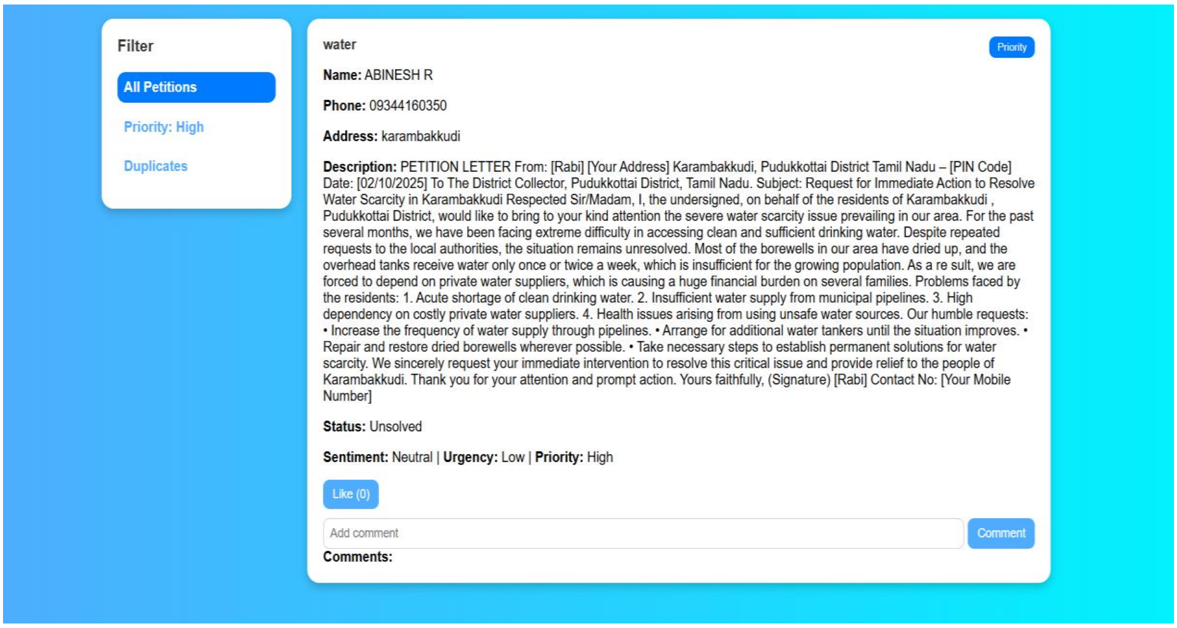

Figure. B.3 Petition Box

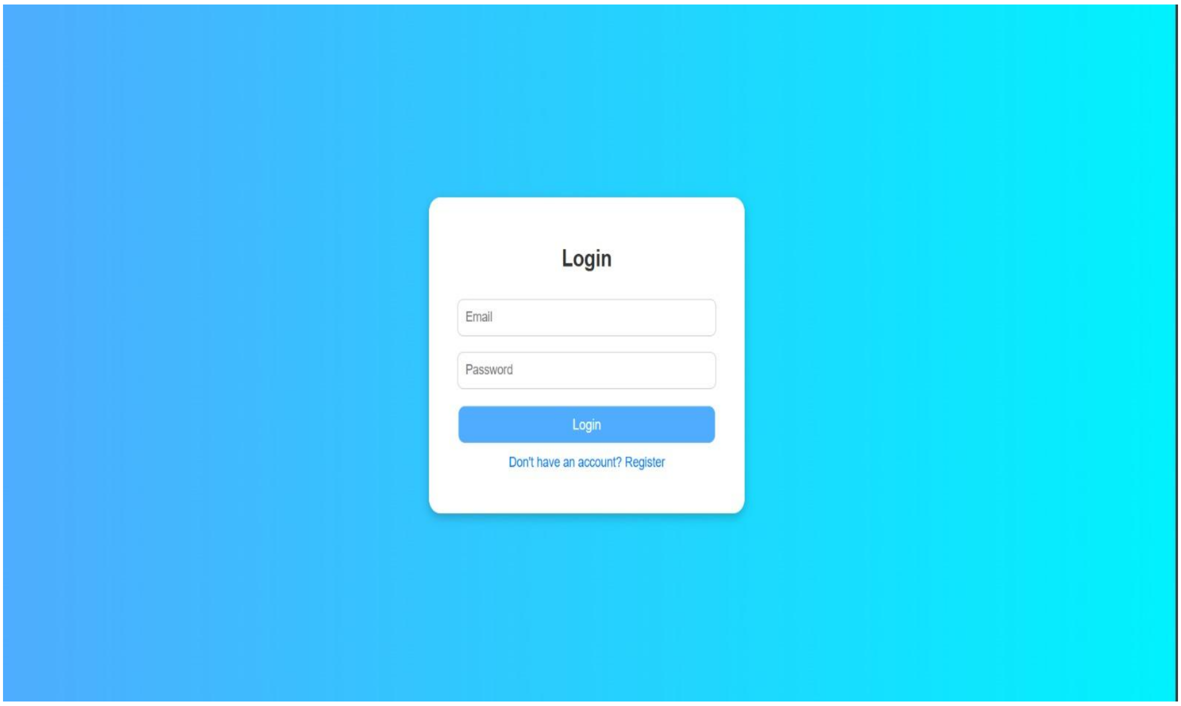

Figure. B.4 Login Page

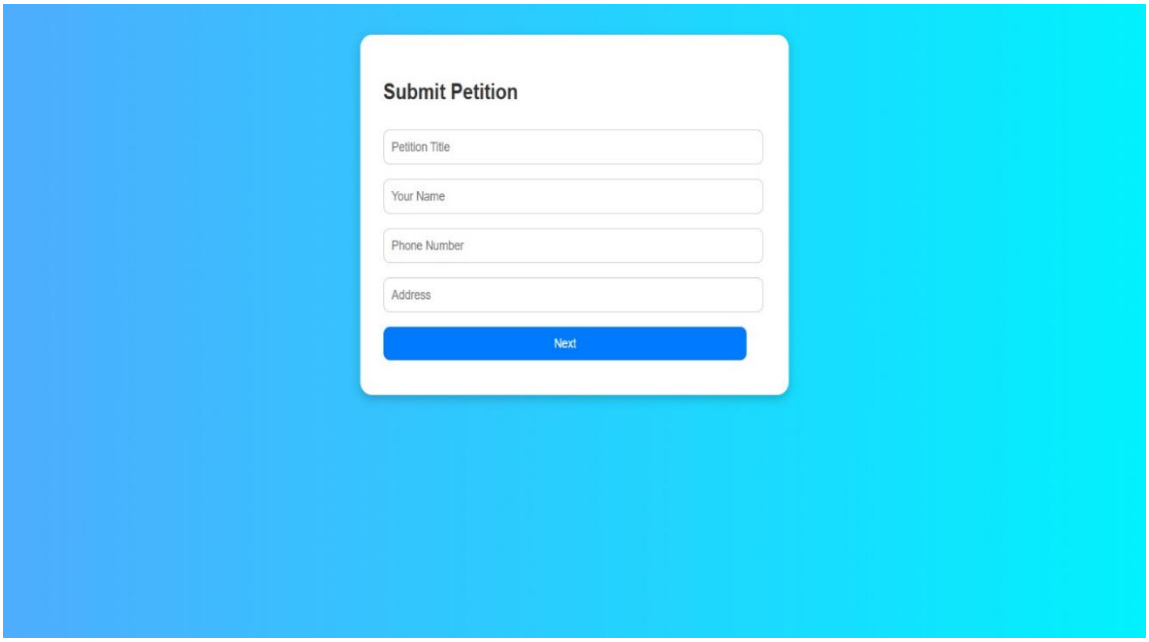

Figure. B.5 Submission Box

REFERENCES

1. A. Gupta, R. Malhotra, and S. Verma, "AI-Based Petition Analysis Using NLP and Text Classification," International Journal of Digital Governance and AI Systems, vol. 12, no. 1, pp. 55–70, 2023.

2. D. Mehta and T. Varma, "Rule-Based and Keyword-Driven Petition Analysis Systems," International Journal of Expert Systems and Automation, vol. 11, no. 3, pp. 78–96, 2022.

3. J. Kannan and M. Srinivas, "Multi-Modal Government Grievance Analytics Using Machine Learning," International Journal of AI and Smart Governance, vol. 16, no. 1, pp. 205–223, 2024.

4. K. Menon and R. Agarwal, "Machine Learning–Based Petition Classification Using Traditional Feature Engineering," International Review of Applied Machine Learning, vol. 10, no. 2, pp. 88–104, 2022.

5. L. Sharma and D. Roy, "Deep Learning–Based Text Classification for Citizen Feedback Systems," International Journal of Neural Computing and Applications, vol. 22, no. 3, pp. 144–160, 2023.

6. N. Kulkarni and R. Prasad, "Machine Learning Assisted Public Grievance Routing Using Keyword Matching and Rule-Based Classification," International Journal of e-Governance Research, vol. 15, no. 2, pp. 119–137, 2024.

7. R. Das and H. Iyer, "Text Mining Approach for Public Grievance Classification," International Journal of Text Analytics and Social Informatics, vol. 13, no. 2, pp. 65–82, 2024.

8. R. Prasad and N. Kulkarni, "Machine Learning Assisted Public Grievance Routing Using Keyword Matching and Rule-Based Classification," International Journal of e-Governance Research, vol. 15, no. 2, pp. 119–137, 2024.

9. S. Thomas and A. Babu, "Machine Learning for Civic Complaint Prioritization Using Basic Features," Journal of Data Science and Civic Technology, vol. 9, no. 1, pp. 32–48, 2023.

10. S. Pandey and V. Singh, "Sentiment-Based Analysis of Public Complaints Using Lexicon Methods," Journal of Computational Linguistics and Social Analytics, vol. 19, no. 4, pp. 134–150, 2023.