



B.Tech Project Synopsis

Civic Sentiment Analysis Platform

Sakshi Sharad Mane	22610047
Vaibhavi Balaji Hipparkar	23620009
Radhika Mahadev Bhosale	22610084
Tamnna Khalil Malgave	23620007

Guide: Mrs. M. B. Shinde

Department: Information Technology

Academic Year: 2025-26

1. Introduction

Citizens frequently use social media and civic portals to express opinions, complaints, and suggestions about essential services such as traffic management, water supply, waste collection, and sanitation. This feedback, if analyzed properly, can provide valuable insights to policymakers and municipal authorities. However, the unstructured nature and high volume of such data make manual analysis time-consuming, inconsistent, and ineffective in real time.

The proposed project, *Civic Sentiment Analysis Platform*, aims to address this gap by leveraging Artificial Intelligence (AI), Natural Language Processing (NLP), and Explainable AI (XAI) techniques. The platform will preprocess and analyze citizen feedback, detect sentiment (positive, neutral, negative), group issues into civic domains, and present results in an interactive dashboard. A key highlight is the use of XAI, which ensures transparency by explaining why the AI made specific predictions. This project is scoped as a **Minimum Viable Product (MVP)** for the semester, focusing on core components while laying the foundation for future enhancements such as forecasting and chatbot integration.

3. Problem Statement

Citizens share civic complaints on social media and municipal portals, but authorities lack an automated, real-time system to process and interpret this feedback. Manual analysis is slow and prone to oversight, leading to delays in addressing urgent civic problems. There is a need for an AI-based platform that can analyze sentiments, categorize issues, and provide explainable and actionable insights through dashboards to assist authorities in timely decision-making.

4. Objectives

1. **Collect and preprocess civic feedback data** from social media and civic portals to create a clean, structured dataset.
2. **Implement sentiment analysis** using pretrained NLP models (e.g., BERT/DistilBERT) and apply **Explainable AI (XAI)** techniques (e.g., SHAP, LIME, attention visualization) to make predictions interpretable.
3. **Apply topic categorization methods** (e.g., LDA or BERTopic) to group feedback into major civic domains such as traffic, water supply, and sanitation.

-
4. **Develop an interactive dashboard** to visualize sentiment trends, categorized issues, and explainability insights for municipal stakeholders.
-

5. Literature Review / Related Work

Existing civic complaint platforms, such as municipal helplines and portals, depend heavily on manual logging and categorization, which makes them slow and unable to handle large-scale data. NLP research has advanced significantly in recent years, particularly with transformer-based models such as BERT, RoBERTa, and DistilBERT, which have set new benchmarks for sentiment classification tasks. Topic modeling approaches like Latent Dirichlet Allocation (LDA) and BERTopic are widely applied to discover hidden patterns and themes in large text datasets.

Recent work also emphasizes the importance of **explainability** in AI. Black-box models often lack transparency, which reduces trust among decision-makers. Explainable AI techniques such as SHAP and LIME allow stakeholders to understand which features influenced predictions, making AI systems more interpretable and trustworthy. This is particularly crucial in governance contexts where accountability is essential.

Studies such as “Social Media, Topic Modeling and Sentiment Analysis in Municipal Decision Support” (2023) demonstrate the utility of combining sentiment and topic analysis for civic data. Similarly, newer works highlight the role of explainability in enhancing adoption of NLP tools for policymaking. Our project extends these directions by building an integrated MVP that unifies **sentiment analysis, topic modeling, and XAI-based interpretability** into a single dashboard for civic governance.

6. Proposed Methodology / Approach

The methodology follows a pipeline approach:

1. Data Collection & Preprocessing

- Collect citizen feedback from civic datasets, municipal complaint records, or social media posts.
- Preprocess data by removing noise, special characters, and irrelevant content.

- Normalize text through tokenization, lemmatization, and multilingual handling using NLP libraries like SpaCy and NLTK.

2. Sentiment Analysis with XAI

- Implement sentiment classification using pretrained transformer models (BERT/DistilBERT).
- Classify feedback into positive, neutral, or negative sentiments.
- Apply Explainable AI methods (e.g., SHAP, LIME, attention visualization) to show why the model predicted a certain sentiment, improving transparency.

3. Topic Categorization

- Apply topic modeling techniques such as LDA or BERTopic to discover recurring civic issues.
- Group issues into categories like traffic, sanitation, water supply, electricity, etc.
- Evaluate coherence of discovered topics for interpretability.

4. Dashboard Development & Visualization

- Build a web-based dashboard using Flask/Django (backend) and Chart.js/Plotly/Dash (visualization).
- Display key insights including sentiment trends, distribution of civic issue categories, and **XAI-based explanations**.
- Provide filters for timeline, location (if available), and issue type.

5. Testing & Validation

- Evaluate sentiment model accuracy against test data.
 - Assess topic modeling quality using coherence scores.
 - Collect feedback on dashboard usability and interpretability.
-

7. Expected Outcomes / Deliverables

- A functional AI-based MVP capable of processing and analyzing civic feedback.
 - **Sentiment classification with interpretability** (positive, neutral, negative + XAI explanations).
 - Topic categorization into key civic domains.
 - An interactive dashboard for visualizing results, including explainability insights.
 - Documentation and demonstration of the system.
-

8. Timeline / Work Plan

Phase	Activity	Duration
Phase 1	Literature review & dataset collection	~1 week
Phase 2	Data preprocessing & NLP pipeline setup	~2 weeks
Phase 3	Sentiment analysis model + Explainable AI integration	~2 weeks
Phase 4	Topic categorization (LDA/BERTopic)	~2 weeks
Phase 5	Dashboard development & integration	~2 weeks
Phase 6	Testing, validation & documentation	~1.5 weeks

9. References

1. Social Media, Topic Modeling and Sentiment Analysis in Municipal Decision Support
<https://arxiv.org/abs/2308.04124> (Aug 2023)
 2. [Explaining Sentiments: Improving Explainability in Sentiment Analysis Using Local Interpretable Model-Agnostic Explanations and Counterfactual Explanations | IEEE Journals & Magazine | IEEE Xplore](#) (July 2023)
 3. [Sentiment Analysis Meets Explainable Artificial Intelligence: A Survey on Explainable Sentiment Analysis | IEEE Journals & Magazine | IEEE Xplore](#) (April 2025)
-