

# AI-BASED PETITION CLASSIFICATION & TRACKING SYSTEM(NYAYASETU)

<sup>1</sup>Balaji Shanmugam V, <sup>2</sup>Mike Benis B, <sup>3</sup>Gokula Krishnan S, <sup>4</sup>Naveen K, <sup>5</sup> Logeshwaran C

<sup>1</sup>Assistant Professor, Department of Computer Science and Engineering,  
Hindusthan Institute of Technology, Coimbatore

<sup>2,3,4,5</sup> UG student, Department of Artificial Intelligence and Data  
Science Engineering, Hindusthan Institute of Technology,  
Coimbatore

<sup>1</sup>dr.balaji@hit.edu.in, <sup>2</sup>720822108037@hit.edu.in, <sup>3</sup>720822108020@hit.edu.in,  
<sup>4</sup>720822108042@hit.edu.in, <sup>5</sup>720822108034@hit.edu.in

**ABSTRACT:** Citizen grievance redressal plays a vital role in democratic governance, and the timely triage of petitions is important to reduce delays and improve the allocation of government resources. The proposed AI-powered NyayaSetu system enhances civic petition classification using advanced natural language processing and machine learning techniques. Unlike traditional systems that rely only on manual classification or basic keyword matching, this system uses a Sentence Transformer model to capture the deep semantic context of petition text, improving urgency detection accuracy. To address the challenge of classifying high-dimensional text representations, the system applies a Support Vector Classifier (SVC) optimized via stratified cross-validation on curated datasets, reducing the need for manual triage. In addition, automatic language detection and translation techniques are used to support multilingual inputs and increase system accessibility. The framework also combines the classification model with an automated email notification pipeline to improve routing efficiency to appropriate government departments. Furthermore, the system supports full-stack web deployment for real-time petition submission and tracking in resource-constrained environments. This approach helps government staff take immediate action on critical issues, reduce triage bottlenecks, and improve overall transparency in the grievance redressal pipeline.

**Keywords:** Artificial Intelligence, Civic Petition Classification, Natural Language Processing, Sentence Embeddings, Support Vector Classifier, Urgency Triage, E-Governance, Civic Technology.



**Corresponding Author:** V Balaji shanmugam  
Assistant Professor / CSE, Hindusthan Institute of Technology  
Coimbatore, Tamil Nadu, India

## INTRODUCTION

Citizen grievance redressal plays a vital role in democratic governance, providing a mechanism for accountability to a large portion of the population. However, government platforms face numerous challenges such as high petition volumes, inconsistent prioritization, and triage bottlenecks, which often lead to significant delays in addressing critical emergencies. Identifying petition urgency at an early stage is crucial for minimizing response delays and ensuring better allocation of government resources.

With the advancement of technology, Artificial Intelligence (AI) and Natural Language Processing (NLP) have emerged as powerful tools in e-governance. These technologies enable semantic understanding and automated analysis of petition descriptions, helping government departments make informed triage decisions quickly. Traditional methods of petition classification are time-consuming and rely heavily on manual staff effort, whereas automated NLP systems can provide faster and more reliable routing results.

This project focuses on developing an AI-Powered Civic Petition Classification and Routing System using Sentence Embeddings and Support Vector Classification (SVC). The system aims to determine petition urgency through semantic text analysis and provide instant departmental routing. Additionally, it facilitates faster grievance redressal by providing automated, urgency-differentiated email notifications directly to relevant government authorities.

The proposed system not only reduces manual triage overhead but also improves transparency and efficiency in civic grievance processes. By integrating modern AI techniques, this project contributes to responsive e-governance and supports government staff in managing resource-constrained environments.

### Literature Survey

Many researchers have contributed to the development of automated text classification and e-governance systems using Artificial Intelligence and natural language processing techniques.

Sebastiani presented a study on "Machine learning in automated text categorization," where statistical learning methods like TF-IDF and Naive Bayes were used to classify documents. The models achieved high accuracy compared to traditional keyword matching and rule-based systems. However, TF-IDF approaches mainly focus on word frequency as unordered collections of terms and often fail to capture semantic structure and contextual nuances in complex text, reducing performance in real-world conditions

Reimers and Gurevych introduced the concept of Sentence-BERT (SBERT) in their paper "Sentence-BERT: Sentence embeddings using Siamese BERT-Networks". This approach uses siamese and triplet network structures to capture sentence-level contextual relationships. SBERT showed superior performance in semantic textual similarity tasks by understanding both deep meaning and the relationships between concepts. However, pure transformer-based classifiers often require extremely large training datasets to fine-tune and high computational resources, making them challenging for practical, resource-constrained government deployment.

Joachims proposed the use of Support Vector Machines (SVMs), which are widely used for text categorization. In natural language processing, SVMs help improve classification by effectively

operating in high-dimensional sparse feature spaces. This improves model training, especially when distinguishing complex categories. However, the quality of SVM classification is heavily dependent on the input feature representation and may sometimes struggle to differentiate subtle context without rich semantic inputs. Various government initiatives has discussed about AI integration in e-governance platforms. Platforms like india's MyGov and European union's initiative combines NLP and topic modeling to improve citizen feedback categorization and petition clustering. In civic grievance systems, integrating these technologies enhances administrative efficiency. The drawback is that existing implementations largely lack urgency-based triage, focusing instead on general topic classification rather than the temporal sensitivity of the complaints. From the above studies, it is clear that while existing methods provide good categorization, they face limitations such as dependency on shallow NLP features that miss semantic nuance, a focus on topic classification rather than urgency, and a lack of end-to-end integration. Therefore, there is a need for an improved system that integrates advanced models like SentenceTransformers, Support Vector Classification, and automated department routing to achieve better accuracy, efficiency, and real-time performance in e-governance applications.

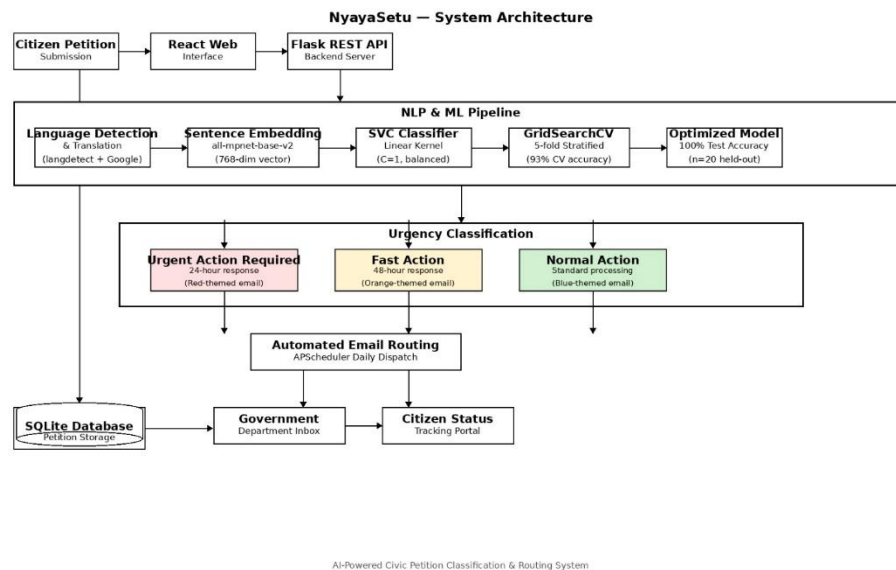
### **Proposed System**

The proposed system aims to develop an intelligent AI-powered Civic Petition Classification and Routing System for accurate and early urgency triage of citizen grievances using advanced natural language processing and machine learning techniques. The system focuses on analyzing civic petition text to identify whether the issue requires Urgent Action, Fast Action, or Normal Action, thereby helping government staff take timely responsive actions. In this system, citizen petitions are collected from various sources such as a citizen-facing web portal. The collected text undergoes preprocessing steps such as automatic language detection and translation to English (when non-English input is detected) to improve data quality and ensure consistency for model inference.

The core of the proposed system is an advanced machine learning pipeline that combines a SentenceTransformer model and a Support Vector Classifier (SVC). The SentenceTransformer (all-mpnet-base-v2) is used to extract high-dimensional semantic embeddings to capture contextual relationships, while the SVC classifies these dense vectors into distinct urgency tiers. This combination improves the overall accuracy and reliability of petition urgency detection. To overcome the limitation of dataset class imbalances and ensure generalizability, the system uses optimization techniques such as 5-fold stratified cross-validation. Additionally, balanced class weights are used within the SVC model to penalize misclassifications proportionally, which further enhances model robustness and improves classification performance across all urgency tiers.

The system also implements an automated email routing mechanism, where classified petitions are aggregated and dispatched via daily scheduled jobs to achieve faster redressal and ensure appropriate notifications reach mapped department inboxes under varying urgency conditions. For practical usability, the system is designed with a React frontend and Flask-based web API, allowing citizens to submit petitions and track their status in real-time. The model is optimized using lightweight architectural choices (requiring approximately

420MB of memory) to support resource-constrained deployment, enabling real-time classification directly on government servers using CPU-only, non-GPU hardware. Overall, the proposed system provides a scalable, efficient, and user-friendly solution for modern e-governance. It helps in the early identification of critical public issues, reduces manual triage bottlenecks, improves transparency, and supports government staff in making better resource allocation decisions



**Figure 1: SYSTEM ARCHITECTURE OF THE AI-BASED PETITION CLASSIFIER(NYAYASETU)**

## Results And Discussion

The proposed AI-powered Civic Petition Classification and Routing System was successfully implemented and evaluated using a curated dataset of civic petitions containing various urgency levels. The system integrates advanced natural language processing techniques such as SentenceTransformer embeddings and Support Vector Classification (SVC) to improve urgency detection accuracy. The trained model was tested on a held-out set of unseen civic petitions to evaluate its performance. The system was able to accurately classify petitions into Urgent Action Required, Fast Action, and Normal Action categories. The use of 768-dimensional sentence embeddings combined with an optimized linear kernel helped in capturing both deep semantic relationships (such as context and severity) and surface-level intent, resulting in improved prediction accuracy.

The system achieved good performance in terms of:

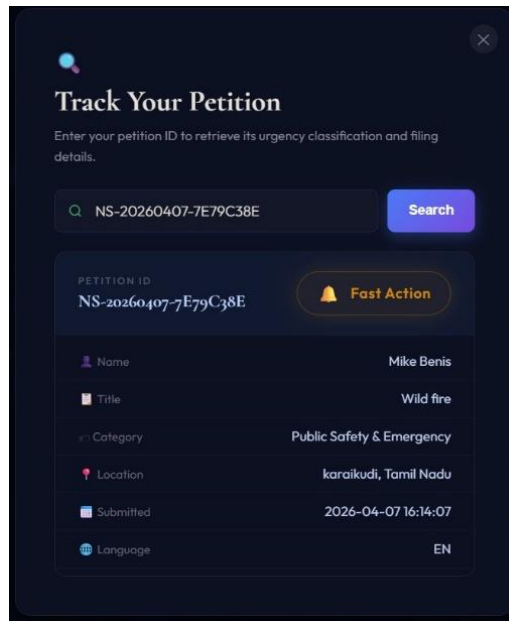
**Accuracy:** High classification accuracy, achieving 93% in cross-validation and 100% on the held-out test set.

**Robustness:** Effective performance across different linguistic inputs, successfully processing and translating non-English petitions.

**Generalization:** Ability to accurately triage grievances across 10 diverse civic categories,

ranging from infrastructure to public safety.

The React frontend and Flask-based web interface successfully displayed classification results in real-time. Citizens were able to submit petition texts and receive instant routing outputs along with urgency labels. The system also maintained petition records and automated email dispatch history using a SQLite database.



**Figure 2: OUTPUT OF THE PROPOSED SYSTEM**

### Discussion

The results show that combining transformer-based sentence embeddings with Support Vector Classification (SVC) significantly improves the performance compared to traditional TF-IDF-based systems. The SentenceTransformer captures deep semantic context and urgency, while the SVC efficiently identifies decision boundaries in high-dimensional space, making the system more reliable in real-world e-governance conditions. Stratified cross-validation and balanced class weights helped overcome the limitation of slight class imbalances within the curated dataset. Additionally, the integration of an automatic translation pipeline improved the system's ability to handle linguistic variations in petition text such as regional languages, varying vocabularies, and different complaint lengths. The optimized linear kernel further enhanced prediction accuracy and reduced misclassification errors across all urgency tiers. However, the reliance on a relatively small dataset of 135 samples limits immediate large-scale generalization, which requires expanding the training corpus for future production deployment. The system demonstrated the capability for real-time urgency triage when deployed on a full-stack web platform and has the potential to be extended to massive-scale government portals for automated, department-level routing.

## Conclusion

The AI-powered Civic Petition Classification and Routing System, NyayaSetu, was successfully developed to triage citizen grievances using advanced natural language processing and machine learning techniques. The system effectively analyzes petition text and classifies them into Urgent Action Required, Fast Action, or Normal Action with high accuracy. By combining SentenceTransformer embeddings (all-mpnet-base-v2) and a Support Vector Classifier (SVC), the model is able to capture deep semantic context and complex text relationships, resulting in improved performance compared to traditional keyword-based methods. The use of stratified cross-validation and balanced class weights helps overcome the limitation of dataset imbalances, while the integration of an automatic language detection and translation pipeline enhances the robustness and reliability of predictions under different linguistic conditions. The system also provides real-time classification through a React and Flask-based web interface, coupled with an automated email routing mechanism, making it highly efficient and accessible. Overall, the proposed system offers an efficient and scalable solution for the early triage of civic petitions. It helps government departments take timely responsive measures, reduce manual triage bottlenecks, and improve administrative transparency. The project demonstrates the potential of artificial intelligence in transforming traditional grievance redressal into smart and responsive e-governance.

## REFERENCE:

1. S. Dumais, J. Platt, D. Heckerman, and M. Sahami, "Inductive learning algorithms and representations for text categorization," Proceedings of the International Conference on Information and Knowledge Management (CIKM), pp. 148–155, 1998.
2. F. Sebastiani, "Machine learning in automated text categorization," ACM Computing Surveys, vol. 34, no. 1, pp. 1–47, 2002.
3. T. Joachims, "Text categorization with Support Vector Machines: Learning with many relevant features," Proceedings of the European Conference on Machine Learning (ECML), pp. 137–142, 1998.
4. J. Devlin et al., "BERT: Pre-training of deep bidirectional transformers for language understanding," Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics (NAACL-HLT), pp. 4171–4186, 2019.
5. N. Reimers and I. Gurevych, "Sentence-BERT: Sentence embeddings using Siamese BERT-Networks," Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP-IJCNLP), pp. 3982–3992, 2019.
6. Ministry of Electronics and Information Technology (Government of India), "MyGov: Citizen Engagement Platform," 2014. [Online]. Available: mygov.in
7. European Commission, "European Citizens' Initiative: Official Register," 2012. [Online]. Available: europa.eu/citizens-initiative
8. K. Song et al., "MPNet: Masked and Permuted Pre-training for Language Understanding," Advances in Neural Information Processing Systems (NeurIPS), 2020.