International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 14s,2025
https://theaspd.com/index.php

Smart Enquiry Chatbot Using Ann And Nlp For Locomotives

Jerin Jose J[1], Sajitha Banu S[2], Krishnamoorthy J[3], Sumathi P[4], Dr. K.Venkata Padma[5], Dr. P.Jyothi Kumari[6]

[1]Associate Professor, Department of Artificial Intelligence and Machine Learning, Sasi Institute of Technology and Engineering, Tadepalligudem, Andhra Pradesh, India.

[2]Assistant Professor, Department of Artificial Intelligence and Machine Learning, Sasi Institute of Technology and Engineering, Tadepalligudem, Andhra Pradesh, India.

[3]Assistant Professor, Department of Information Technology, Saveetha Engineering College, Chennai, Tamil Nadu, India.

[4]Assistant Professor, Department of Artificial Intelligence and Data Science, JJ College of Engineering and Technology, Trichy, Tamil Nadu, India.

[5]Asst Professor, Department of Management Studies, Aditya University, Surampalem, East Godavari District, Andhra Pradesh, India.

[6]Associate Professor, Department of Nutrition and Dietetics, Ch S D St Theresa's College for Women (A), Eluru, Andhra Pradesh, India.

Abstract

The railway sector is the major transportation network, catering. Due to the demand of efficient and personalized customer support, classical systems often don't match with people's expectations. This research targets to develop a Railway Enquiry chatbot by Artificial Neural Networks (ANN) and Natural Language Toolkit (NLTK), It aims to design an intuitive, efficient, and user-friendly interface for answering passenger queries. This chatbot uses NLTK for natural language processing tasks, such as tokenization, stemming, and intent recognition, ensuringaccurateunderstandingofuserinputs.ANN classifies and generates the responses, providing a continuous interaction. The major functions of this chatbot are ticket availability checks, train schedule inquiries, platform details, and fare estimations. The system utilizes a curated and preprocessed dataset comprising frequently asked railway-related questions and their corresponding answers, ensuring it addresses a wide spectrum of user inquiries. A supervised learning approach is applied to train the Artificial Neural Network (ANN) model, enabling it to accurately identify user intent and deliver appropriate responses. The model's performance is assessed using metrics such as accuracy, precision, recall, and F1-score, all of which affirm its reliability in providing correct information. The chatbot is implemented through a web-based interface, allowing users to conveniently obtain railway-related details. By automating standard queries, the system minimizes reliance on human customer service, thereby improving overall passenger experience. This study underscores the transformative role of ANN and Natural Language Processing (NLP) in modernizing support systems within the railway industry.

Key Words: *Stress Detection, Mental Health, Machine Learning, Deep Learning, Transfer Learning, Chatbot, Artificial Neural Network (ANN), Stress Management*

I. INTRODUCTION

A Railway Enquiry Chatbot is an advanced conversational tool developed to provide users with immediate and precise information related to railway services [1]. It delivers a wide range of details, including train timings, seat availability, ticket fares, station information, and real-time train updates. Additionally, it simplifies interactions for users seeking help with bookings, cancellations, and general queries through an intuitive interface [2]. By employing cutting-edge technologies such as Machine Learning (ML) and Deep Learning (DL), the chatbot offers fast, efficient, and reliable assistance, reducing reliance on conventional sources like helpline numbers or physical enquiry desks [3].The growing expectation for instant, personalized support in the railway domain has created a demand for intelligent solutions like chatbots.

Traditional ways of obtaining information—such as queuing or navigating complex web portals—often prove to be inefficient and frustrating. A railway chatbot overcomes these issues by offering round-the-clock availability, rapid responses, and the capacity to manage numerous queries at once [4]. It can also tailor suggestions to user preferences, creating a more engaging and seamless experience [5].ML and DL significantly enhance the functionality of railway enquiry systems. With Natural Language Processing (NLP), the chatbot can interpret user inputs, determine intent, and deliver responses that consider conversational context. Transformer-based models like BERT and GPT have further refined the chatbot's understanding of natural language, enabling it to handle even nuanced and complex queries [6]. Additionally, speech recognition powered by ML and DL converts spoken input into text using methods such as feature extraction and Recurrent Neural Networks (RNNs) [7].Beyond query resolution, these intelligent systems provide personalized recommendations using collaborative filtering techniques and embedding-based algorithms.

For live train tracking and predicting delays, the chatbot employs time-series analysis with models like Long Short-Term Memory (LSTM) networks [8]. Reinforcement learning ensures that the chatbot continues to improve with every user interaction, boosting its accuracy and dependability over time [9].The integration of Machine Learning and Deep Learning into the Railway Enquiry Chatbot fundamentally redefines how passengers access information. It streamlines communication, boosts satisfaction, and reduces the workload on railway personnel—making it an essential component of contemporary transport infrastructure [10][11].

2. RELATED WORKS

2.1 Speech-enabled railway enquiry system

"*EnhancingPassenger Convenience: An Efficient Speech Driven Enquiry System for Railway Stations*",L. Sakri et al. (2024) –The research paper introduces a speech-enabled railway enquiry system that incorporates NLP, speech recognition, and database integration. It effectively answers queries like train schedules and seat availability in under 10 seconds, improving passenger convenience.

2.2 Chatbot driven Support

" *My Rail Guide –Railway Enquiry System: An Artificial Intelligence Approach*", L. I. Sakri et al. (2024) . In a separate work, the authors present "My Rail Guide," an Android-based AI assistant for train travel. It features live scheduling, PNR tracking, QR code tickets, and chatbot-driven support, offering a hands-free and eco-friendly travel experience.

2.3 Railway Network Analyzer

"Automatic Railway Signaling Generation for Railways Systems Described on Railway Markup Language (rail ML)," M. N. Menéndez et al. (2024).The "Railway Network Analyzer" (RNA) is presented as a tool for automating the placement of railway signals based on rail ML descriptions, significantly reducing human error and improving signaling efficiency.

2.4 Rail Surface Defects

"*DeepConvolutional Neural Networks for Rail Surface Defect Perception*", Y. Wang et al. (2023)This study focuses on using deep learning—specifically CNNs, YOLO, and R-CNN—for identifying rail surface defects. It stresses the need for standardized datasets and explores challenges in detecting rare defect types to enhance railway safety.

2.5 Leveraging Web GL and GUID

"*Research on an IntegratedGraphicalModel-BasedMethodforTwo-and Three-Dimensional Model Management in Railway Engineering*",Q. Gao et al. (2023). A graphical model-based framework is introduced for managing 2D/3D railway project data. Leveraging WebGL and GUID-based association, the system strealines project oversight and enhances data management efficiency.

2.6Simplified Method for Forecasting

"*Overview of PowerElectronicsApplicationsforFixedInstallationsof Urban Railway Power Supply for Regenerative Energy Utilization,*"T. Suzuki et al. (2022).In their study on urban railway power systems, the authors examine JR(Japan Railway) East's adoption of regenerative energy equipment—including static storage and inverters—which contributed to a 2–8% reduction in substation energy consumption. A simplified method for forecasting the effectiveness of such systems is also proposed.

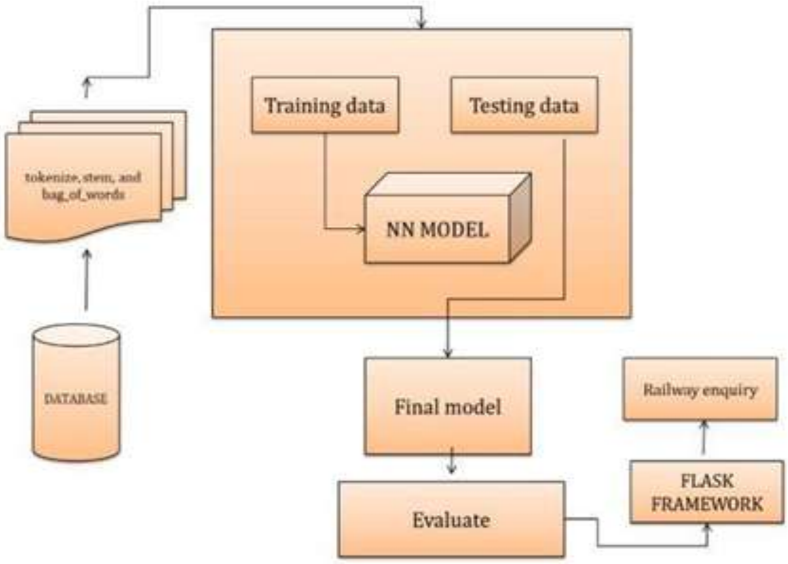
2.7 IoT based real-time tracking of railway

"The Real Time Railway Monitoring System Suitable for MultiView Object Based on Sensor Stream Data Tracking ",A. Abduvaytov et al. (2020). An IoT-driven monitoring solution is developed for real-time tracking of railway infrastructure such as tunnels and bridges. The system improves safety and operational awareness using sensor-based streaming data.

As railway networks become more complex and the need for immediate information grows, traditional enquiry channels—such as manual help desks, telephone support, and static web portals—are increasingly unable to meet the evolving expectations of passengers. These legacy systems often suffer from long response times, restricted availability, and difficulty managing a large number of varied inquiries. As a result, passengers commonly experience delays in receiving information, unclear responses, and a lack of personalized support, all of which contribute to frustration and hinder effective travel planning.

3. PROPOSED METHOD

3.1 System Design



1. Data Collection Module

This module serves as the backbone of the chatbot system, focusing on gathering all essential data required for accurate functioning. It compiles structured information such as train timetables, station codes, fare details, PNR statuses, and real-time train delay updates. Sources may include official railway databases, open APIs, and verified third-party providers.

In addition, a dataset of user queries paired with intents (e.g., "Find available trains to Delhi") is collected to train the chatbot's Artificial Neural Network (ANN) and Natural Language Processing (NLP) models. This module ensures that the chatbot operates with updated and diverse information to cater to a broad user base.

2. Data Preprocessing Module

This component prepares the raw data for integration into the chatbot's learning models. The process begins with text normalization, where input text is standardized—converted to lowercase and stripped of non-essential characters. It then undergoes tokenization to divide sentences into manageable units (words or phrases).

Irrelevant terms, known as stopwords, are filtered out, while lemmatization simplifies words to their base forms. Datasets are also tagged with intent and entity labels to facilitate supervised training. Through this structured and thorough preparation, the data becomes cleaner and more effective for training purposes, significantly boosting model performance.

3. Natural Language Processing (NLP) Module

The NLP module is responsible for interpreting user input and extracting meaningful context. It performs linguistic tasks such as tokenization and Part-of-Speech (POS) tagging to understand the grammatical structure of queries. Named Entity Recognition (NER) helps identify important entities like train names, dates, and station codes.

Intent classification determines the user's purpose—whether they're asking for a schedule, availability, or fare details. This module may employ advanced techniques like Transformer-based architectures (e.g., BERT or GPT) to enhance understanding of complex queries. The NLP layer ensures user messages are accurately understood before further processing.

4. Database Management Module

This module organizes and maintains all railway-related data for efficient retrieval. It stores critical details such as schedules, availability, fares, and PNR status within relational databases. Real-time synchronization with railway APIs ensures that data remains current.

The module is designed for high-speed access and accuracy, enabling the chatbot to answer queries reliably and in real-time. Efficient database operations are vital for maintaining system responsiveness and ensuring consistent performance.

5. Artificial Neural Network (ANN) Module

At the heart of the chatbot's intelligence is the ANN module, which uses a multi-layer perceptron (MLP) model to classify intents and trigger appropriate actions. It is trained using labelled examples of user queries linked to specific intents.

Training is optimized using algorithms such as Adam or Stochastic Gradient Descent (SGD) to enhance precision. Once the model is adequately trained, it can handle a wide range of user inputs and deliver contextually relevant responses, contributing to a smooth and engaging user experience.

6. Response Generation Module

This module is tasked with crafting clear and appropriate replies to user queries. For frequently asked questions, predefined templates are used to ensure consistency and accuracy.

In contrast, dynamic queries are addressed with responses generated using NLP methods, allowing for contextual and conversational output. The goal is to provide users with information that is not only correct but also easy to understand and relevant to their requests.

4. Implementation

This system is a full-stack web application that integrates AI-powered natural language processing with a modern, responsive user interface. It consists of four main layers: frontend, backend, database, and cloud infrastructure.

1. Backend & Database Layer

a) **Python (Flask)** is used to build the server-side logic, handling API requests, routing, and business logic.
b) The backend exposes endpoints that facilitate communication between the frontend and the AI/NLP components.
c) **MySQL** is used for persistent data storage, handling user information, queries, and application metadata.
d) **Gen API** is integrated to connect the system with external services or AI models. This could include model inference, data retrieval, or analytics functions.

2. Frontend & User Interface

a) **React.js** powers the client-side of the application, enabling a dynamic, responsive, and user-friendly interface.
b) React components communicate with the Flask backend using RESTful APIs, sending user inputs and receiving processed results.
c) The frontend is designed to work seamlessly across devices and browsers.

3. AI & NLP Processing

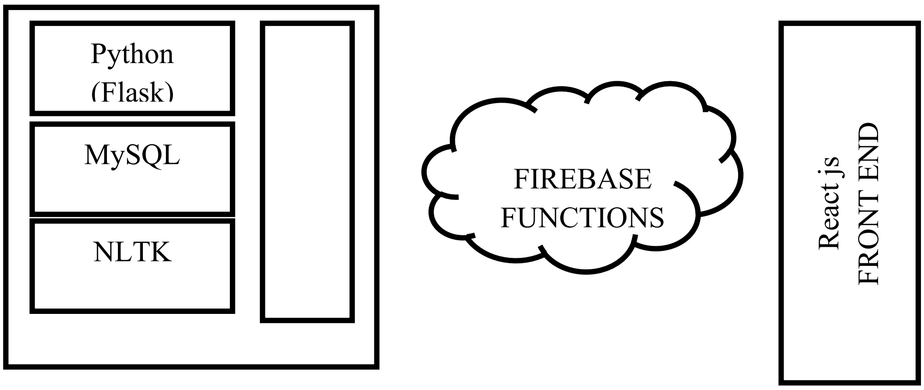
a) **NLTK (Natural Language Toolkit)** is used to process, analyze, and interpret natural language inputs.
b) It enables features such as:
    a. Tokenization

b. Named Entity Recognition (NER)
c. Sentiment analysis
d. Text classification

c) The processed results are returned to the backend, which relays them to the frontend for display.

4. Cloud & Hosting Infrastructure

- The application leverages **serverless computing** for scalability and cost-efficiency:
  - **Firebase Function** is used to run lightweight backend services or AI inference functions without maintaining dedicated servers.
- This architecture supports autoscaling and reduces latency for specific real-time operations.



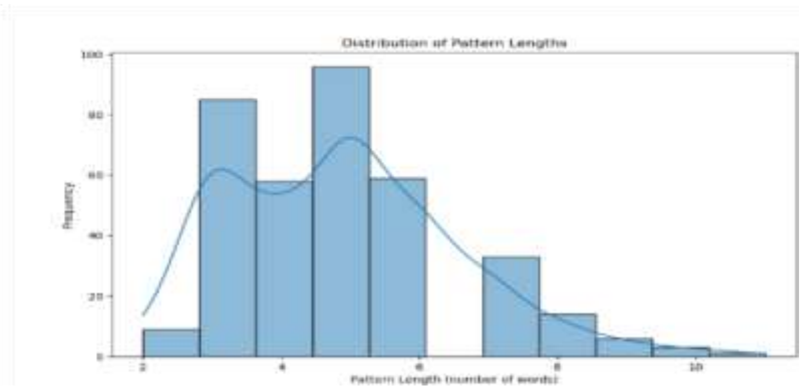
5. Training Analysis

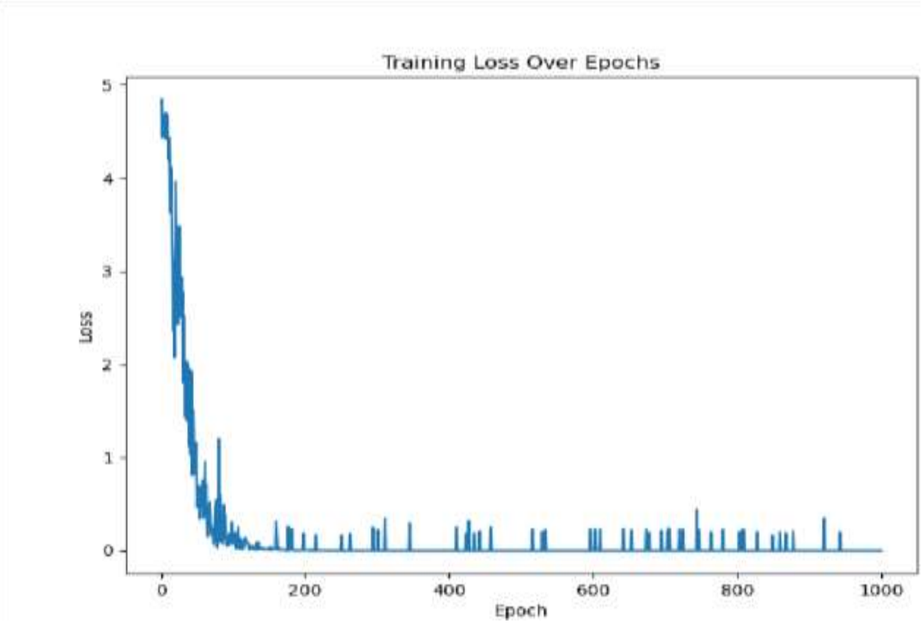
1. Distribution of Pattern Lengths

This analysis illustrates the number of words in each user input pattern from the dataset. The majority of inputs contain between **3 to 6 words**, while very short or excessively long queries are relatively uncommon. Understanding this distribution is essential for shaping the input preprocessing strategy. Maintaining consistency in input length helps prevent model bias toward unusually short or lengthy queries, ensuring balanced performance across different types of inputs.

2. Training Loss Over Epochs

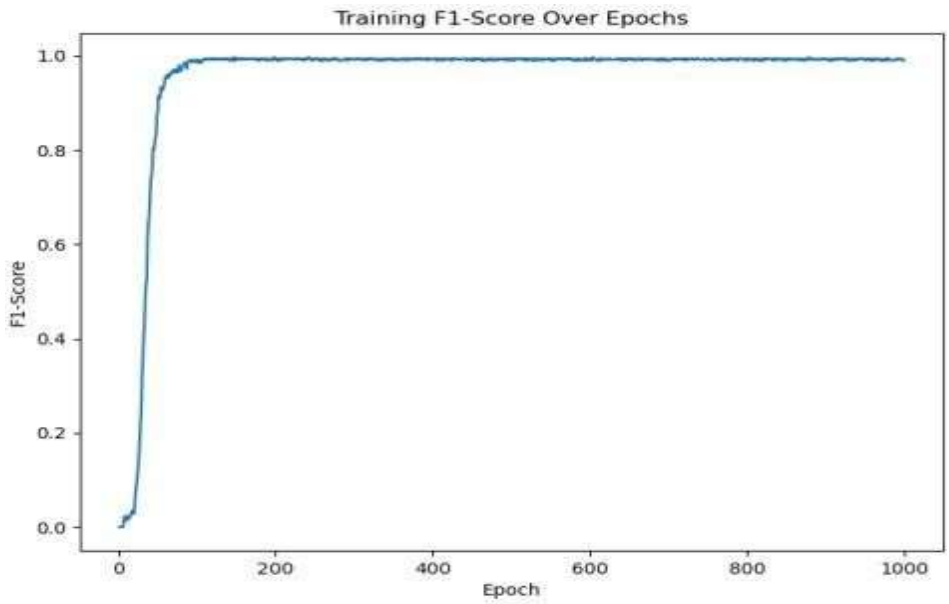
This metric tracks how much the model's predictions diverge from the true output during training. Initially, the **loss value is high** but decreases rapidly in the early training stages, eventually leveling off near zero. The **stabilization of loss** at a low level indicates that the model has effectively learned the underlying patterns. A flat curve toward the end implies that continued training no longer brings improvement, suggesting an optimal stopping point.





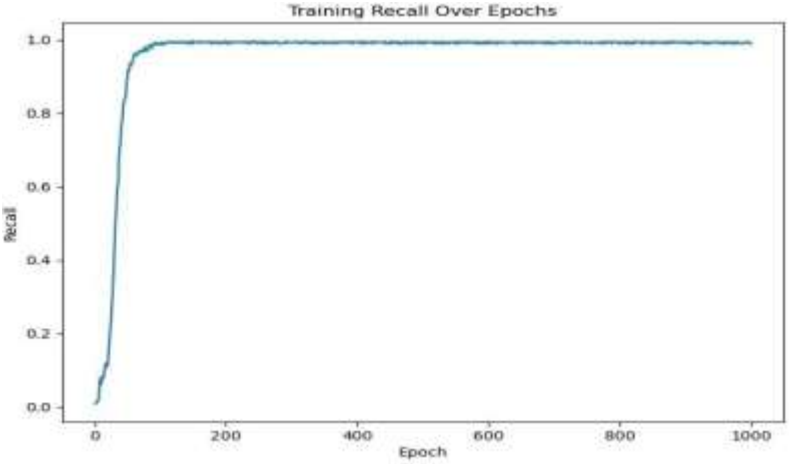
3. Training F1-Score Over Epochs

The F1-score, which is the **harmonic mean of precision and recall**, reflects a balanced evaluation of the model's classification performance. It generally mirrors trends in both precision and recall, **stabilizing close to 1**, which signifies strong overall performance. A high F1-score confirms the model's competence in handling both relevant (true positives) and irrelevant (false positives) predictions, making it highly reliable.



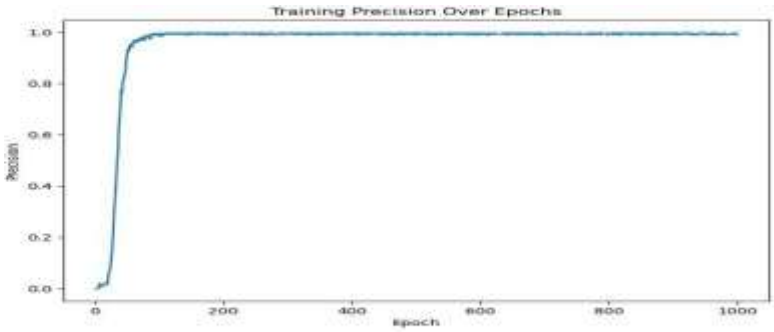
4. Training Recall Over Epochs

Recall quantifies the proportion of actual positive samples the model correctly identifies. Over training epochs, **recall rises and stabilizes near 1**, indicating the model captures most valid intents accurately. Its curve is typically aligned with those of precision and accuracy. A high recall rate ensures the chatbot does not miss meaningful queries, reflecting strong generalization over the training dataset.



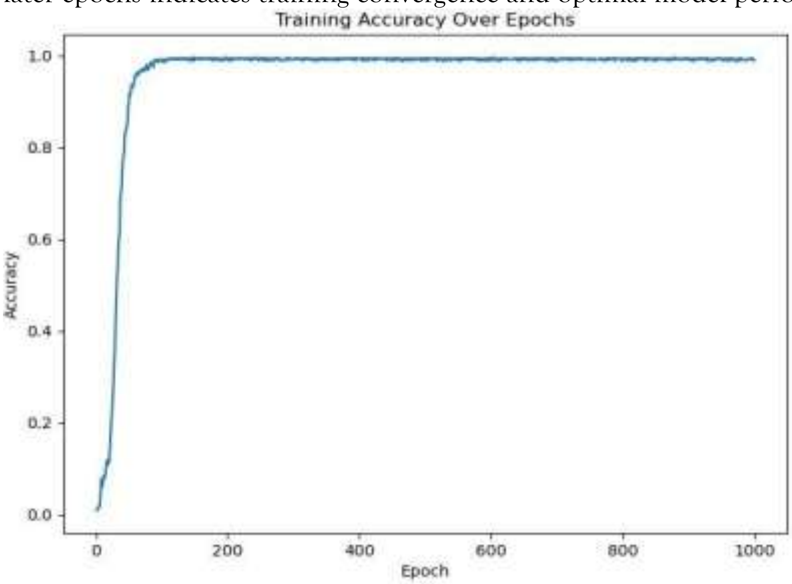
5. Training Precision Over Epochs

Precision measures the proportion of relevant results among all predicted positives. Like accuracy, it begins lower but **increases significantly over initial epochs**, eventually stabilizing close to 1. A steady precision curve suggests that the model **consistently predicts the correct intent**, minimizing false positives. This is especially important for chatbot reliability, as it reduces the likelihood of responding incorrectly to user inputs.



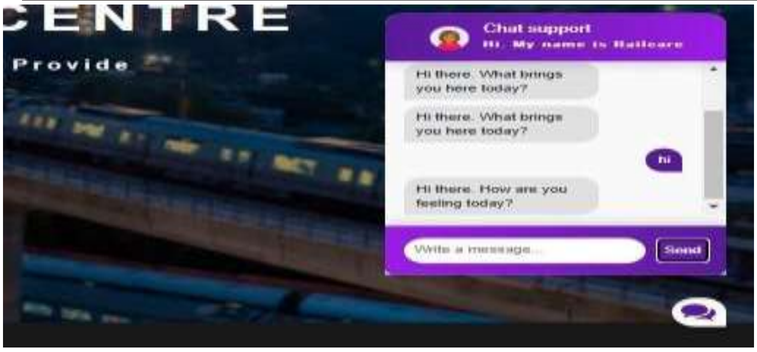
6. Training Accuracy Over Epochs

Accuracy captures the overall correctness of the model's predictions across all categories. During training, accuracy starts low and **climbs rapidly**, ultimately stabilizing near **99.9%**. This trend shows that the model has successfully learned the dataset without signs of overfitting or underfitting. The flattened curve in later epochs indicates training convergence and optimal model performance.



International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 14s,2025
https://theaspd.com/index.php

5.RESULTS



6. CONCLUSION

The Railway Enquiry Chatbot, powered by Artificial Neural Networks (ANN) and Natural Language Processing (NLP), presents a modern, efficient, and scalable alternative to conventional railway enquiry systems. By modularizing its core functionalities, the chatbot ensures smooth interaction—from interpreting user input to delivering accurate and context-aware responses. The integration of ANN for intent classification and NLP for natural query understanding allows the system to handle a wide range of user inquiries with precision. Comprehensive processes for data collection, preprocessing, and management ensure the chatbot relies on reliable and up-to-date information at all times. Furthermore, the system's feedback and learning modules support continuous performance enhancement, enabling the chatbot to adapt to changing user expectations and query trends. Overall, this solution not only enhances user experience but also demonstrates how intelligent automation can transform traditional public service infrastructures.

7. Future Enhancement: Now, our proposed research supports 4 languages Tamil, English, Hindi, Telugu. In future the no of languages supported is increased by using Large Language Models.

REFERENCES

1. Abduvaytov, R. M. Abdu Kayumbek, H. S. Jeon and R. Oh, "The Real Time Railway Monitoring System suitable for Multi-View Object based on Sensor Stream Data Tracking," 2020 International Conference on Information Science and Communications Technologies (ICISCT), Tashkent, Uzbekistan, 2020, pp. 1-4.
2. H. Hayashiyaetal.,"Necessity and possibility of smart grid technology application on railway power supply system," Proceedings of the 2011 14th European Conference on Power Electronics and Applications, Birmingham, UK, 2011, pp. 1-10.
3. J. Liu, M. Liu, H. Xu, Y. Wu and Y. Zhou, "Railway Worker Safety Analysis by the PSO Algorithm in China Railway Bureau,"202241st Chinese Control Conference (CCC), Hefei, China, 2022, pp. 1833- 1839.
4. J. Shi, Y. Wang and W. Zheng, "Correlation Analysis of Causes of Railway Accidents Based on Improved Apriori Algorithm," 2020 13th International Symposium on Computational Intelligence and Design (ISCID), Hangzhou, China, 2020, pp. 274-277.
5. J.Weiand W.Zhu,"Generating Travel Plan Sets in a High-Speed Railway Network With Complex Timetables and Transfers," IEEE Access, vol. 8, pp. 157050-157058, 2020.
6. L.I.Sakri, S. R. Biradar, V. Barge, R. Padakiand A. Hegde, "MyRailGuide – Railway Enquiry System: An Artificial Intelligence Approach," 2024 3rd International Conference for Advancement in Technology (ICONAT), GOA, India, 2024, pp. 1-5.
7. L. Sakri, S. R. Biradar, M. P. Kulkarni, S. Patilkukarni and S. K, "Enhancing Passenger Convenience: An Efficient Speech Driven Enquiry System for Railway Stations," 2024 IEEE International Conference for Women in Innovation, Technology & Entrepreneurship (ICWITE),Bangalore,India,2024,pp.686-690.
8. M. Choi, B. Yoon, D. Kim and D. Sung, "IAB- based Railway Communication Method for Stable Service Provision," 2021

Twelfth International Conference on Ubiquitous and Future Networks (ICUFN), Jeju Island, Korea, Republic of, 2021, pp. 176-178.
9. M. N. Menéndez, S. Germino, L. D. Díaz- Charris and A. Lutenberg, "Automatic Railway Signaling Generation for Railways Systems Described on Railway Markup Language (railML)," IEEE Transactions on Intelligent Transportation Systems, vol. 25, no. 3, pp. 2331-2341, March 2024.
10. M. Tomita et al., "Verification of Super conductingFeederCableinPulseCurrentand Notch Operation on Railway Vehicles," IEEE Transactions on Applied Superconductivity, Vol.31,no.1,pp.1-4,Jan.2021, Artno.4800104.
11. Q.Gao, W.Wang, Y.Xie,W.Lu,W. Liuand Y. Sun, "Research on an Integrated Graphical Model Based Method for Two and Three-Dimensional Model Management in Railway Engineering," 2023 IEEE 14th International Symposium on Parallel Architectures, Algorithms and Programming (PAAP), Beijing, China, 2023, pp. 1-6.
12. S. Ishizaki et al., "Confirmation of Correlation Between Hourly Electric Power and Instantaneous Maximum Power of Rectifiers for Railway,"2022 International Power Electronics Conference (IPEC-Himeji 2022- ECCE Asia), Himeji, Japan, 2022, pp. 1421-1426.
13. T.Suzukietal.,"OverviewofPowerElectronics Applications for Fixed Installations of Urban Railway Power Supply for Regenerative Energy Utilization," 2022 International Power Electronics Conference (IPEC- Himeji 2022- ECCEAsia),Himeji,Japan,2022, pp. 11071112.
14. T. Zhou,Y.Yang, L. Liu, C.Tao andY. Liang, "ADynamic 3-DWideband GBSM for Cooperative MassiveMIMOChannelsinIntelligent High-Speed Railway Communication Systems," IEEE Transactions on Wireless Communications, vol. 20, no. 4, pp. 2237-2250, April 2021.
15. Y. Wang, T. Wang, Z.Yang, S. Ren, X. Gouliu and J. Gao, "Deep Convolutional Neural Networks for Rail Surface Defect Perception," 2023 International Conference on Image Processing, Computer Vision and Machine Learning(ICICML), Chengdu, China, 2023, pp. 12111215.