DEEP ADAPTIVE FEATURE FUSION FOR ORIGIN DESTINATION PASSENGER FLOW FORECASTING IN MASS EVENTS

BHUKYA RAJAKUMAR[1]
D KAVYA[2], P KEERTHI PRIYA[2], MIRIYALA ISHA SREE[2], HRITIK KUMAR[2], M S LAENIN[2]

[1] *Assistant Professor, Department of Computer Science & Information Technology, Siddharth Institute of Engineering & Technology, Andhra Pradesh, India*
[2] *Research Scholar, Department of Computer Science & Information Technology, Siddharth Institute of Engineering & Technology, Andhra Pradesh, India*

ABSTRACT

This project aims to forecast passenger flow for mass events using a Deep Adaptive Feature Fusion approach. By leveraging multiple machine learning algorithms, including Random Forest, Linear Regression, Stacking Regressor, and XGBoost Regression, the system is trained to predict both inflow and outflow of passengers. The model will be trained using historical event data, considering various influencing factors such as time, location, and event type. Based on this input data, the system will predict the number of people likely to attend the event and the estimated outflow, assisting in better crowd management and planning for large-scale events.

Keyword: - Mass Events, Passenger Flow, Inflow Prediction, Outflow Prediction, Random Forest, Linear Regression, Stacking Regressor, XGBoost Regression, Event Management, Predictive Modeling.

1. INTRODUCTION

Mass events, ranging from concerts and sports games to political rallies and festivals, attract large crowds that require meticulous planning to ensure smooth operations and safety. One of the most challenging aspects of event management is predicting passenger flow—both in terms of inflow and outflow. An accurate understanding of how many people are expected to attend and when they are likely to leave can help organizers allocate resources, manage crowd control, and plan transportation efficiently. Traditional methods, which primarily involve manual observations and rule-based predictions, are often insufficient to manage the complexity of large events. This project proposes a Deep Adaptive Feature Fusion approach that integrates multiple Machine learning algorithms, including Random Forest, Linear Regression, Stacking Regressor, and XGBoost Regression, to predict passenger flow. By leveraging historical data from past events, the system aims to forecast the number of attendees and their movement, enabling better planning and management. This introduction sets the foundation for a more advanced, data-driven solution to crowd management that can be applied to a wide range of mass events, ensuring better safety, operational efficiency, and attendee satisfaction.

The primary objective of this project is to predict passenger inflow and outflow during mass events using a Deep Adaptive Feature Fusion approach. This system is designed to forecast the number of attendees expected at an event and the flow of passengers in and out of the venue. It leverages machine learning algorithms such as Random Forest, Linear Regression, Stacking Regressor, and XGBoost Regression to analyze historical event data, including time, location, and event type. The goal is to enhance crowd management and optimize event planning by providing accurate predictions that can be used to allocate resources effectively, plan for transportation, manage crowd control, and reduce safety risks. The system aims to offer predictive insights that help event organizers understand crowd dynamics, improve security measures, and streamline the logistical operations needed to handle large crowds. By combining these models, the system will create a more accurate forecasting tool that can adapt to different event contexts and contribute to improved event outcomes.

Crowd management for large-scale events has become a critical task for organizers, authorities, and event planners. In particular, understanding passenger flow is essential for ensuring safety, minimizing congestion, and enhancing the overall experience. With the increasing number of mass events globally, it becomes more challenging to predict the number of attendees and their movement. Traditional methods relying on manual observation and basic algorithms lack precision and often fail to predict crowd behavior accurately. Predicting both inflow and outflow patterns through advanced machine learning models can offer a more proactive approach to crowd control. By implementing predictive models, such as Random Forest, Linear Regression, Stacking Regressor, and XGBoost Regression, this project aims to improve crowd management by forecasting passenger flow more accurately, providing event organizers with valuable insights for resource allocation, security planning, and logistics. The ability to accurately forecast passenger flow can significantly reduce the chances of overcrowding, improve safety, and increase operational efficiency during large-scale events.

2. LITERATURE SURVEY

[1], Wei and Chen (2012) proposed a hybrid method combining empirical mode decomposition (EMD) and neural networks for short-term metro passenger flow forecasting. Their model addresses the non-linear and non-stationary nature of flow data by decomposing it into intrinsic mode functions using EMD before prediction, resulting in improved accuracy over traditional forecasting techniques.

[2], Bai et al. (2019) introduced Stg2seq, a spatial-temporal graph-to-sequence model for multi-step passenger demand forecasting. By modeling spatial and temporal dependencies through graphs and sequence learning, their approach outperforms traditional methods in dynamic urban environments, aiding in efficient transport planning.

[3], Lim et al. (2021) presented the Temporal Fusion Transformer (TFT), a deep learning model for interpretable multi-horizon time series forecasting. TFT incorporates attention mechanisms to capture both short- and long-term dependencies and provides insights into which inputs drive predictions, making it effective for demand forecasting in transportation.

[4], Bruna et al. (2013) developed spectral networks, which apply convolution in the spectral domain using the graph Laplacian, enabling efficient learning on graph-structured data. Their work laid the foundation for modern graph neural networks, which have since been applied to transportation systems, among other domains.

3. METHODOLOGY

3.1EXISTING SYSTEM

Currently, crowd management relies heavily on manual observation, historical knowledge, and basic algorithms to predict passenger flow during mass events. Traditional methods often focus on monitoring specific event variables like time of day, weather, and transportation availability. However, these methods typically lack the ability to provide real-time, data-driven predictions and are not sophisticated enough to account for various factors affecting crowd movement. In addition, many existing systems do not integrate multiple machine learning models to predict both inflow and outflow, limiting their predictive capabilities. Event organizers often rely on approximate estimations, leading to inefficient resource allocation, delays in transportation, and potential safety risks. Some systems use simple regression models or heuristic-based methods, but they tend to be less accurate in predicting complex crowd behaviors. Given the limitations of current approaches, there is a need for more advanced solutions that combine various data sources and machine learning techniques to improve the accuracy and reliability of passenger flow predictions.

3.1.1 DISADVANTAGES OF EXISTING SYSTEM

- **Limited Predictive Accuracy**: Existing systems rely on simplistic models or manual observation, which are not always accurate in predicting passenger flow.
- **Lack of Real-Time Predictions**: Many systems cannot provide real-time predictions during events, limiting their usefulness for dynamic crowd management.
- **Inadequate Resource Allocation**: Traditional methods may lead to poor resource allocation, resulting in overcrowding or underutilized resources.
- **Limited Factors Considered**: Existing systems often focus on just a few variables (e.g., time and weather), neglecting other crucial factors that can influence crowd behavior.
- **Dependence on Past Data**: Many systems only use historical data, which may not always be representative of future events, especially in changing environments.
- **Inflexibility**: Most current systems are not adaptable to various event types, which limits their scope and versatility.

3.2 PROPOSED METHODOLOGY

The proposed system aims to enhance passenger flow prediction using a more sophisticated approach that combines multiple machine learning models, including Random Forest, Linear Regression, Stacking Regressor, and XGBoost Regression. By integrating these models, the system will provide more accurate predictions of both inflow and outflow based on historical event data. The system will consider various influencing factors such as event type, time, location, and special circumstances (e.g., weather, transport schedules) to improve prediction accuracy. Additionally, the system will be capable of providing real-time predictions, enabling dynamic adjustments to crowd management strategies as the event unfolds. By training the model on a diverse set of past event data, the system can adapt to different event types, ensuring its versatility. The proposed system will allow event organizers to make informed decisions about resource allocation, security measures, and crowd control, ultimately improving safety and operational efficiency. This approach represents a significant improvement over existing systems, providing a more reliable, data-driven solution for crowd management.

3.2.1 ADVANTAGES OF PROPOSED METHODOLOGY

- **Improved Predictive Accuracy**: By using advanced machine learning algorithms, the proposed system offers higher accuracy in forecasting passenger inflow and outflow.
- **Real-Time Predictions**: The system will provide real-time predictions, allowing event organizers to adjust their strategies dynamically.
- **Better Resource Allocation**: Accurate predictions will enable organizers to allocate resources efficiently, reducing the chances of overcrowding or underutilization.
- **Scalability**: The system can be adapted to various event types, making it versatile for different mass events.
- **Data-Driven Decision Making**: By leveraging historical event data, the system ensures that decisions are based on empirical evidence rather than guesswork.
- **Enhanced Safety**: Better crowd management leads to improved safety measures, reducing the likelihood of accidents or overcrowding.

4. SYSTEM DESIGN

System design is the process of defining the architecture, components, modules, interfaces, and data for a system to satisfy specified requirements. Systems design could be seen as the application of systems theory to product development.

4.1 SYSTEM ARCHITECTURE

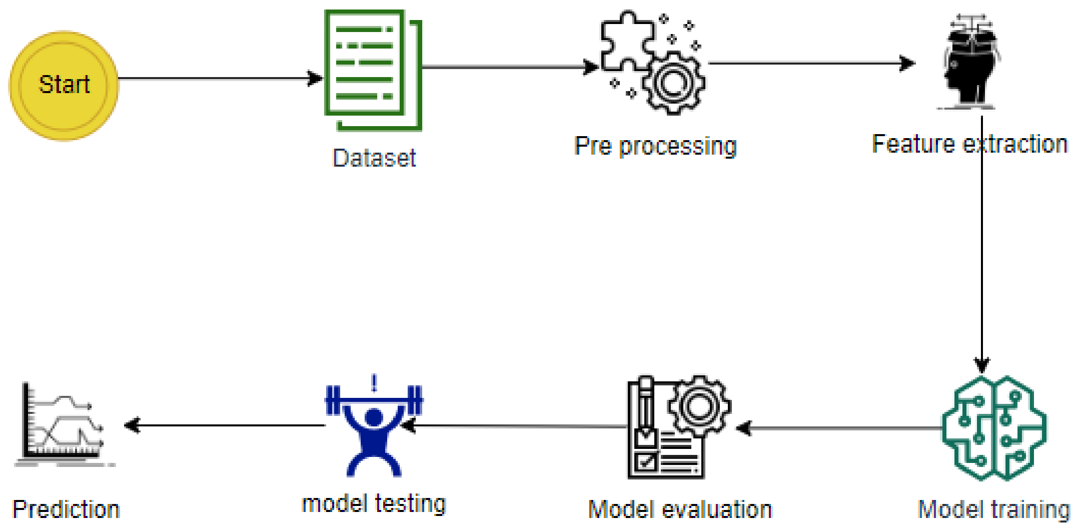


Fig. System Architecture

4.2 MODULES

In this Project , There are Two Modules. They are:

- System Provider
- User

4.2.1 MODULES DESCRIPTION

USER:

Users can upload a dataset, which is a crucial initial step for the system to work with relevant data. This dataset likely contains historical information or examples that the system will use for its predictions.

Users have the capability to view the dataset they've uploaded. This feature helps users confirm the data they've provided and ensures transparency in the process.

Users need to input specific values or parameters into the system to request predictions or results. These input values likely correspond to the variables or features in the dataset.

SYSTEM:

Take the Dataset: The system accepts and processes the dataset provided by the user. This dataset forms the foundation for building the predictive model.

Preprocessing: Before training a predictive model, the system preprocesses the dataset. This includes handling missing data, data cleaning, and feature extraction. Preprocessing Ensures that the data is in a suitable format for modeling

Training: The system uses machine learning techniques and python modules to train a model Based on the preprocessed dataset. The model learns patterns and relationships within the data, allowing it to make predictions.

Generate Results: Once the model is trained, the system can generate results based on User input values. These results typically indicate whether the input data corresponds to a specific condition, event, or prediction, such as Medical Insurance Cost.

5. RESULTS AND DISCUSSION

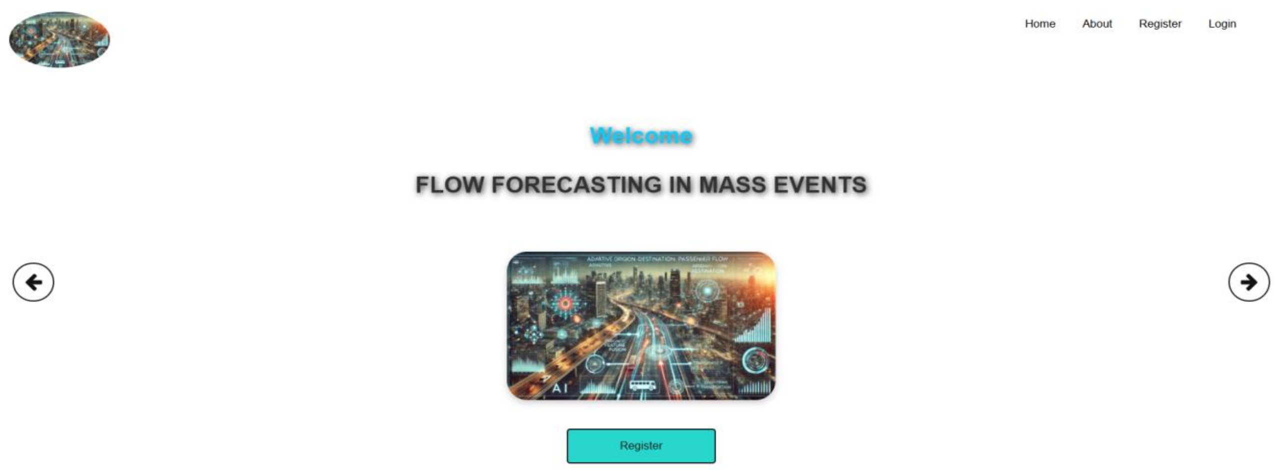


Fig. Home Page

This interface enables users to, facilitating Register and login to model home page and prediction and training page

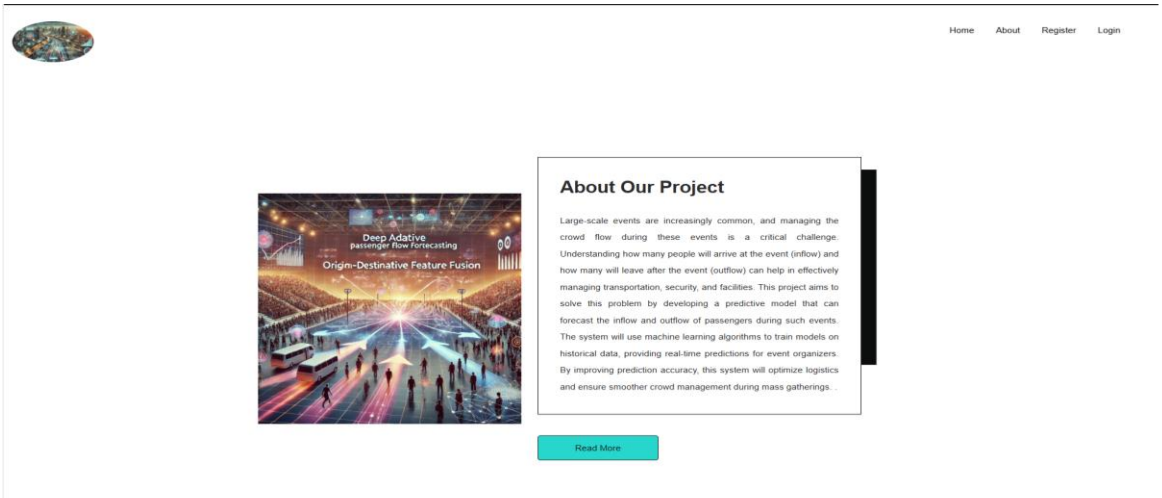


Fig. About Page

The project predicts using machine learning models, emphasizing interpretability and superior accuracy with ensemble methods.

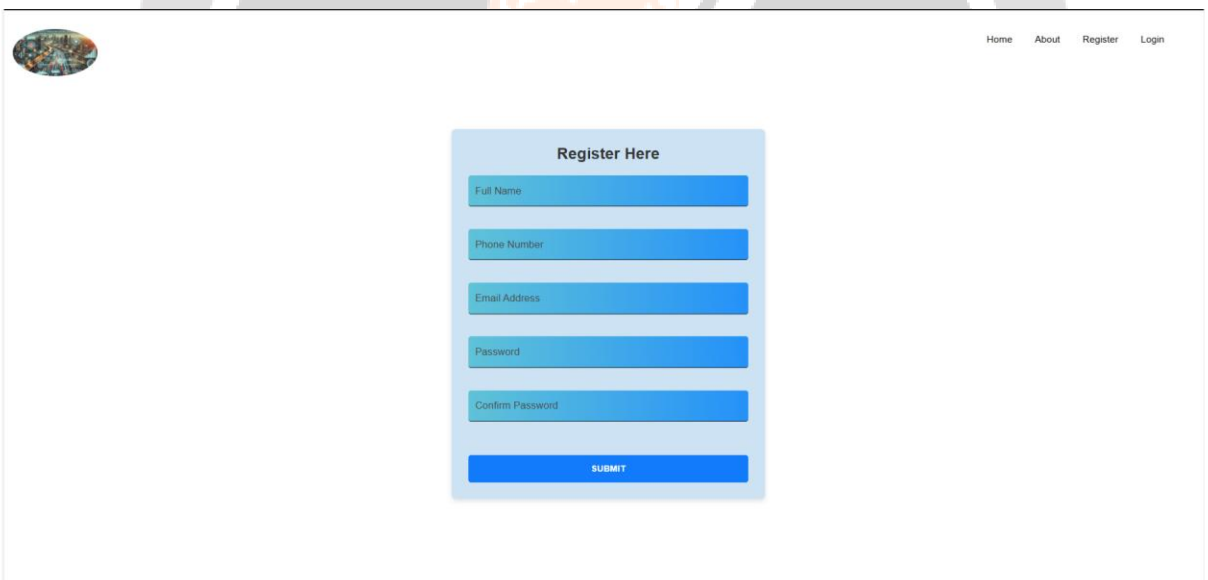


Fig. Registration Page

This page allows users to register for services, ensuring secure access by requiring personal details and password confirmation. It provides a user-friendly interface for creating a secure account.

Fig. Login Page

This page provides a secure login interface for users to access the prediction account using their email and password.

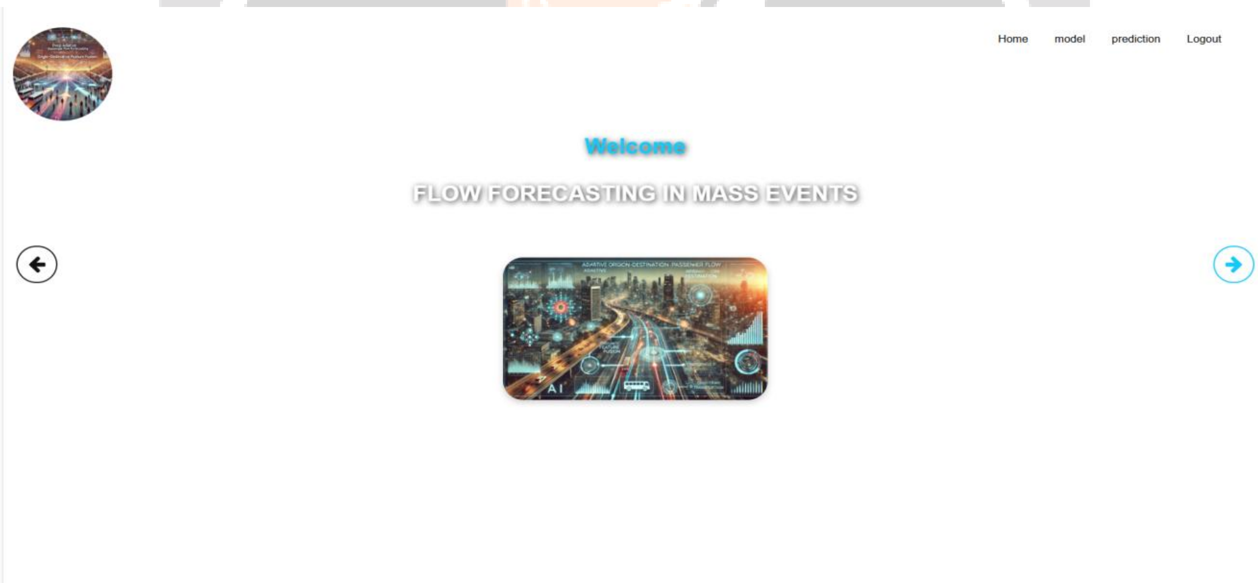


Fig. Home Page

This interface enables users facilitating Selection the dataset uploads and model training and prediction and evaluation to achieve precise predictive outcomes.

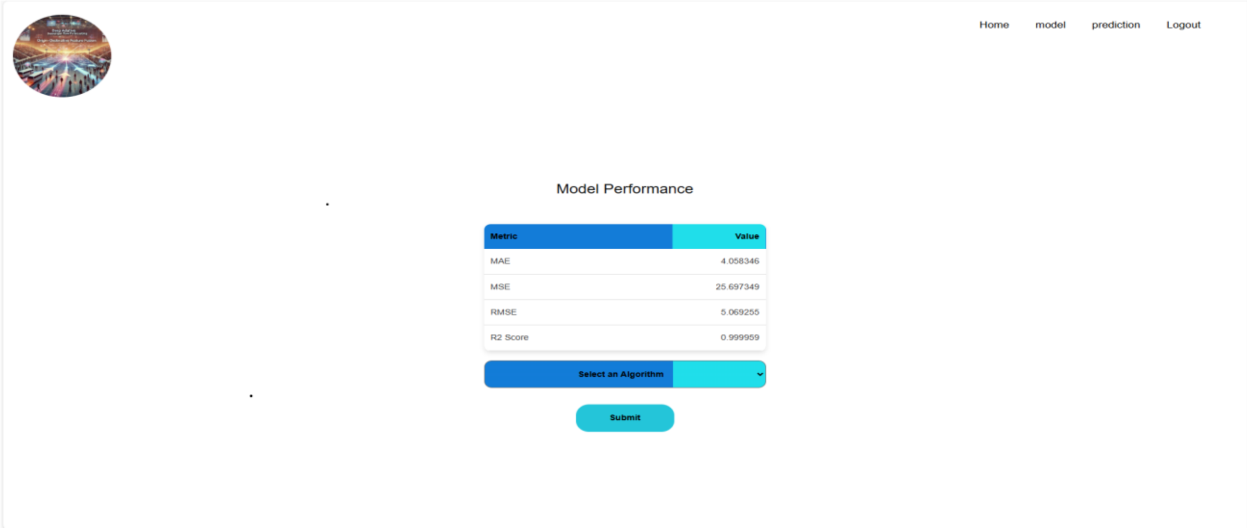


Fig. Model Selection Page

This page allows users to select an algorithm analysis, enhancing decision-making through machine learning models.

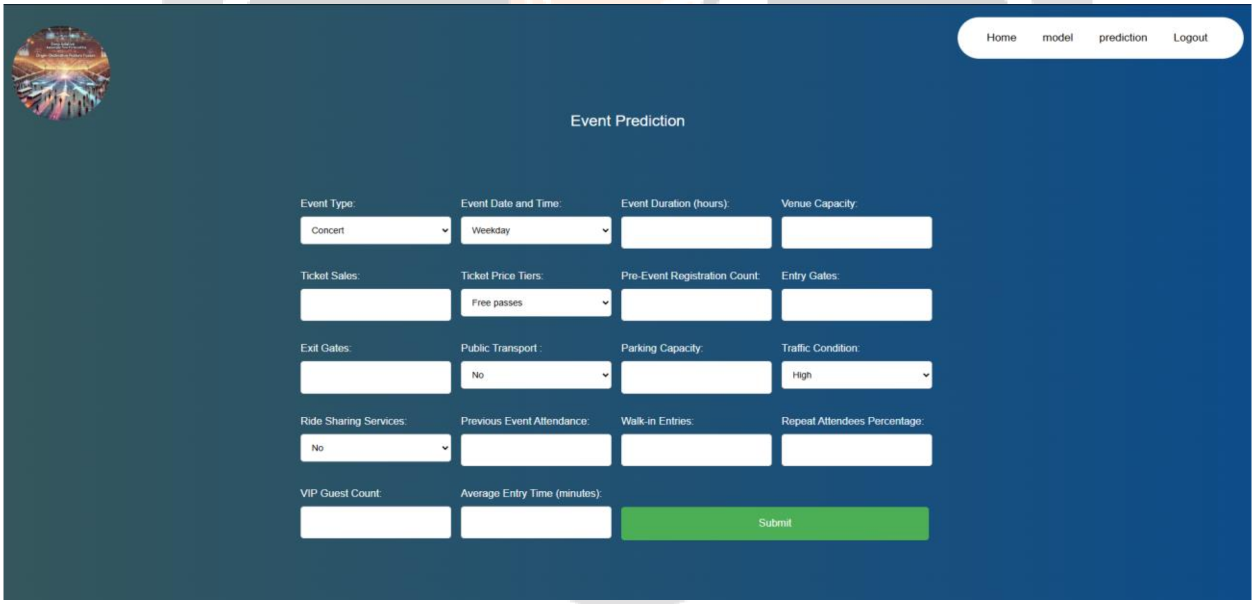
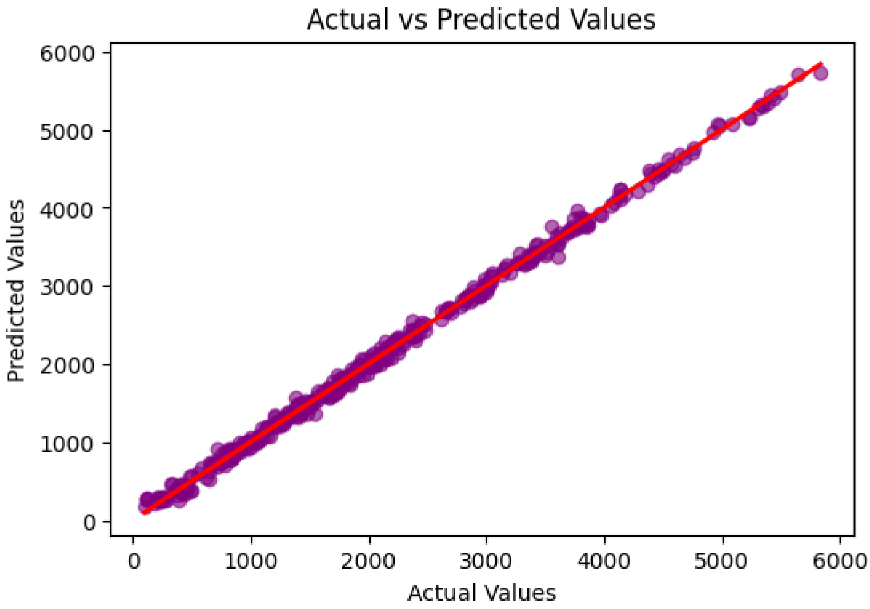


Fig. Prediction Page

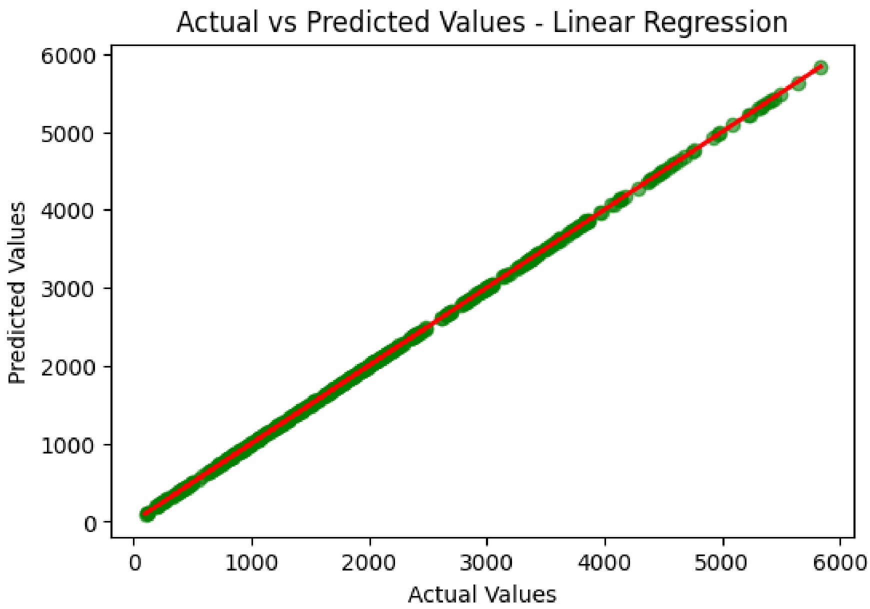
This page collects user input for various parameters to predict the New inputs

Results Algorithm Wise

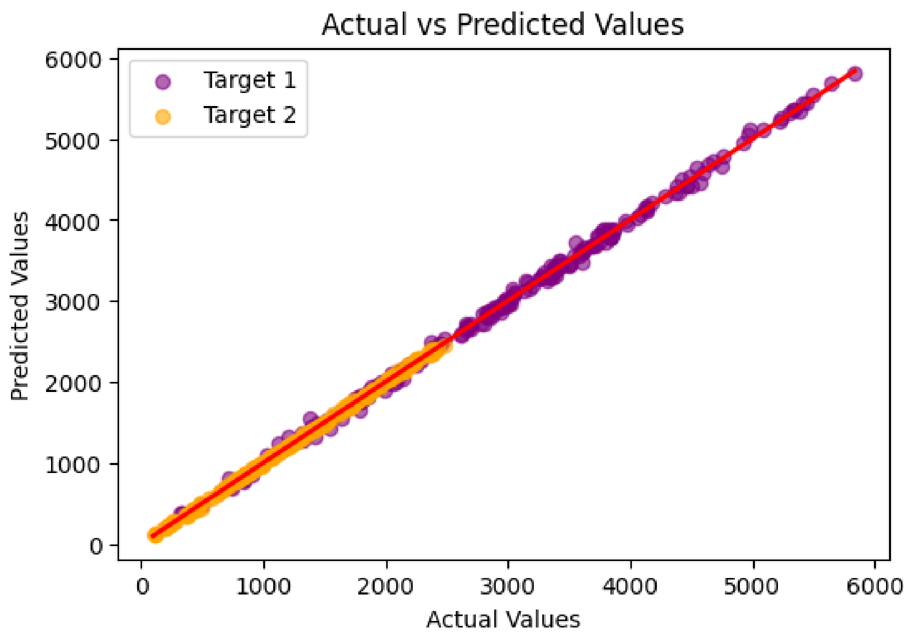
Random Forest Regression



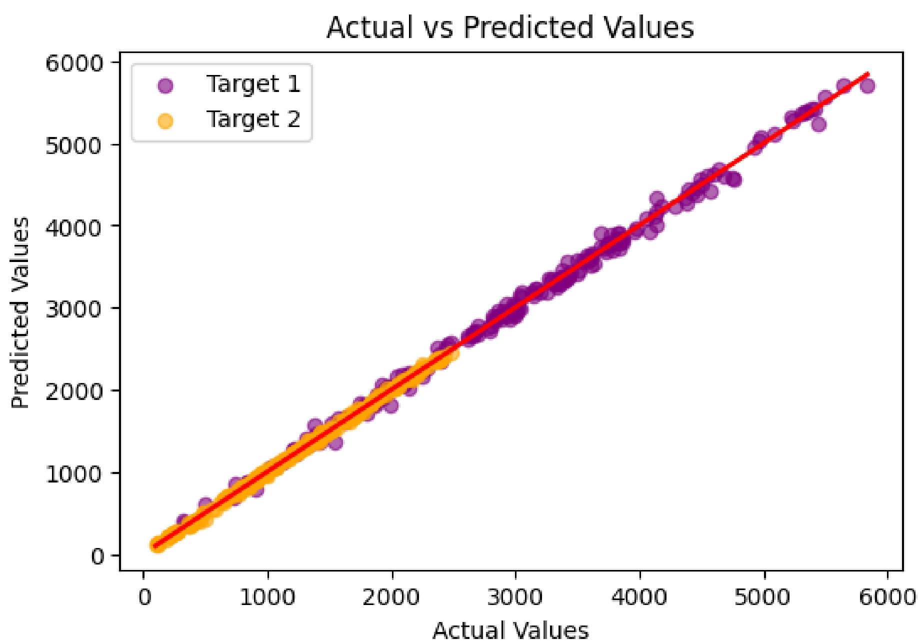
Linear Regression



Stacking Regression



XGB Regression



CONCLUSION

In conclusion, this project aims to revolutionize crowd management at mass events by using a Deep Adaptive Feature Fusion approach, incorporating machine learning algorithms such as Random Forest, Linear Regression, Stacking Regressor, and XGBoost Regression to predict passenger flow. The system will enable event organizers to make better decisions regarding resource allocation, crowd control, and safety measures, ultimately improving the event experience for attendees. By leveraging historical data, the system will provide highly accurate predictions of both inflow and outflow, helping to reduce overcrowding and ensure that appropriate safety measures are in place. The proposed approach represents a significant step forward in the field of crowd management, providing event planners with a powerful tool to optimize operations and enhance attendee safety.

FUTURE ENHANCEMENT

The future scope of this project involves several enhancements and expansions. One key area for future work is the integration of real-time data during events, such as monitoring live attendee counts and transportation status. By incorporating real-time data, the system could offer even more precise predictions and enable organizers to make dynamic adjustments as the event progresses. Another potential improvement is the incorporation of more advanced machine learning techniques, such as deep learning, to further enhance prediction accuracy. Additionally, the system could be expanded to include more granular predictions, such as forecasting crowd behavior at specific locations within the venue, providing deeper insights for crowd management. The system's scope can also be extended to accommodate different types of events, such as indoor conferences, outdoor festivals, and sports tournaments, by incorporating specialized data and features tailored to each event type. Finally, the system could be integrated with other event management tools, such as ticketing and security systems, to provide a more comprehensive solution for large-scale event planning.

7. REFERENCE

1. Wei, Y.; Chen, M.C. Forecasting the short-term metro passenger flow with empirical mode decomposition and neural networks. Transp. Res. Part C Emerg. Technol. 2012, 21, 148–162. [CrossRef]

2. Bai, L.; Yao, L.; Kanhere, S.S.; Wang, X.; Sheng, Q.Z. Stg2seq: Spatial-temporal graph to sequence model for multi-step passenger demand forecasting. In Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence {IJCAI-19}, Macao, China, 10–16 August 2019; pp. 1981–1987.

3. Lim, B.; Ark, S.; Loeff, N.; Pfister, T. Temporal fusion transformers for interpretable multi-horizon time series forecasting. Int. J. Forecast. 2021, 37, 1748–1764. [CrossRef]

4. Bruna, J.; Zaremba, W.; Szlam, A.; LeCun, Y. Spectral networks and locally connected networks on graphs. arXiv 2013, arXiv:1312.6203.

5. Defferrard, M.; Bresson, X.; Vandergheynst, P. Convolutional neural networks on graphs with fast localized spectral filtering. In Advances in Neural Information Processing Systems; ACS: Washington, DC, USA, 2016.

6. Atwood, J.; Towsley, D. Diffusion-convolutional neural networks. Comput. Sci. 2015, 29. Available online: https://proceedings. neurips.cc/paper/2016/hash/390e982518a50e280d8e2b535462ec1f-Abstract.html (accessed on 1 September 2022).

7. Velikovi, P.; Cucurull, G.; Casanova, A.; Romero, A.; Liò, P.; Bengio, Y. Graph Attention Networks. arXiv 2017, arXiv:1710.10903. Mathematics 2022, 10, 3664 29 of 30

8. Vlahogianni, E.I.; Golias, J.C.; Karlaftis, M.G.; Banister, D.; Givoni, M. Short-term traffic forecasting:

Overview of objectives and methods. Transp. Rev. 2003, 24, 533–557. [CrossRef].

9. Williams, B.; Durvasula, P.; Brown, D. Urban freeway traffic flow prediction: Application of seasonal autoregressive integrated moving average and exponential smoothing models. Transp. Res. Rec. 1998, 1644, 132–141. [CrossRef]

10. Lee, S.; Fambro, D.; Lee, S.; Fambro, D. Application of subset autoregressive integrated moving average