PREDICTING SEAT VACANCY FOR RURAL DEVELOPMENT USING MACHINE LEARNING AND IMAGE PROCESSING WITH CNN AND LINEAR REGRESSION ALGORITHMS

Mrs. R. Pavithra[1*], S.S. Yuvaraj[2*], K.Vibunarendar[3*], P.Vasanth[4*]
[1]Assistant Professor(Sr.Gr.), [2*]UG Scholars, [3*]UG Scholars, [4*]UG Scholars*
Department of Electronics and Communication Engineering,
K.L.N.College of Engineerin,g Pottapalayam, Sivagangai-630612,TamilNadu, India

ABSTRACT
Predicting seat availability in public transit systems is essential for improving accessibility and allocating resources as efficiently as possible in rural development programs. In order to precisely predict seat vacancy in rural transport vehicles, this study suggests a novel method that combines machine learning and image processing approaches. This research attempts to offer a reliable predictive model that can adjust to various rural transportation conditions by utilizing Convolutional Neural Networks (CNN) and Linear Regression techniques. The initial step of the process is gathering image data from rural transport trucks' onboard cameras. These photos show the state of occupancy for the seats, which are then pre-processed using CNN to extract pertinent information. Effective depiction of seat occupancy data is made possible by
CNN's ability to understand spatial patterns and dependencies within the images. Utilizing feature engineering, the model simultaneously incorporates geographic, demographic, and temporal aspects. A flexible approach called linear regression is used to combine these contextual variables with the features extracted from the images, enabling a thorough comprehension of the dynamics of seat emptiness.This result offers a unique approach to forecast seat availability in rural transportation vehicles by utilizing the combined strengths of machine learning and image processing methods. The suggested method helps to increase rural areas' accessibility to transportation and makes better use of resources by strengthening decision-making abilities in rural development projects.
Keywords:*Convolutional Neural Networks (CNN), Linear Regression, Seat Vacancy prediction, Machinelearning ,Image processing.*

INTRODUCTION

For remotepopulations, public transportation is a lifeline since it makes it easier to access social events, work opportunities, education, and basic necessities. Nevertheless, there are specific difficulties in monitoring and forecasting seat vacancy while trying to maximize the effectiveness of rural transportation systems. Precise prediction of seat availability is essential to guaranteeing passenger contentment, allocating resources, and maintaining the overall efficiency of the transportation system. Conventional techniques for predicting seat vacancies frequently depend on labour-intensive manual counting or oversimplified statistical models, which may be imprecise and unable to adequately represent the dynamic nature of demand for rural transportation. This research suggests a novel method to anticipate seat availability in rural transportation vehicles by combining machine learning and image processing techniques in response to these difficulties. This study attempts to create a reliable predictive model that can precisely predict seat availability in various rural transit settings by utilizing the power of Convolutional Neural Networks (CNN) for image analysis and Linear Regression for contextual integration In terms of seat vacancy prediction, the combination of machine learning and image processing is revolutionary and has several benefits. The suggested method provides timely and precise data for analysis by utilizing onboard cameras mounted in rural transport vehicles to enable real-time monitoring of seat occupancy. Seat occupancy data can be effectively represented without the need for human intervention thanks to CNN's ability to understand spatial patterns and dependencies within the images. Moreover, the suggested method provides a thorough grasp of seat vacancy dynamics by integrating demographic, geographic, and temporal variables into the prediction model via feature engineering and linear regression. By taking into consideration multiple contextual factors that impact transportation demand, this all-encompassing method improves the precision and dependability of seat availability forecasts. The practical consequences for rural development programs and transportation planning, the value of this research goes beyond academic interest. The suggested method helps make educated decisions, enabling resource efficiency and enhanced transportation accessibility in rural areas by giving transport authorities useful information about seat vacancy trends. In summary, by developing a unique framework for seat vacancy prediction that combines machine learning and image processing techniques, this study fills a significant need in the management of rural transportation. The suggested strategy seeks to improve the sustainability, efficacy, and efficiency of rural transportation networks through empirical assessment and validation, thereby advancing the socioeconomic advancement of rural areas.

OBJECTIVES

- Create a machine learning model that accurately predicts seat vacancy in rural transport vehicles, leveraging both image processing with CNN and contextual data integration with linear regression.
- Improve the efficiency and effectiveness of rural transport planning by providing transport authorities with actionable insights into seat vacancy dynamics, enabling optimized resource allocation and scheduling.
- Utilize the predictive model to dynamically adjust transportation resources, such as vehicles and schedules, based on real-time seat vacancy predictions, thereby maximizing resource utilization and minimizing inefficiencies.
- Enhance passenger satisfaction by ensuring adequate seating availability through proactive management of seat vacancy, leading to a more comfortable and reliable rural transport experience.
- Contribute to the socio-economic development of rural areas by improving transport accessibility, facilitating access to essential services, education, employment opportunities, and social activities through optimized rural transport systems.
- Conduct thorough evaluation and validation of the predictive model using historical data on seat occupancy and contextual factors to assess its accuracy, reliability, and scalability across different rural transport scenarios.
- Investigate the scalability and generalizability of the developed model to diverse rural environments with varying demographic and geographic characteristics, ensuring its applicability across different regions.
- Provide transport authorities with a decision-support tool that enables informed decision-making regarding rural transport planning and management, ultimately leading to more efficient and sustainable transport systems.
- Advance academic knowledge in the fields of machine learning, image processing, and transportation planning by exploring innovative approaches to seat vacancy prediction in rural transport systems and publishing research findings in relevant academic journals or conferences.
- Address the needs and requirements of stakeholders, including transport authorities, rural communities, and policymakers, by developing a predictive model that aligns with their objectives and priorities for rural development and transport accessibility.

LITERATURE REVIEW

In this paper, combining the advantages of the two prediction models, this paper proposes a long short-term memory (LSTM) and Artificial neural networks (ANN) comprehensive prediction model based on spatialtemporal features vectors. The long-distance arrival-to-station prediction is realized from the dimension of time feature, and the short-distance arrival-to-station prediction is realized from the dimension of spatial feature, thereby realizing the bus-to-station prediction (1).In this work we will use Machine Learning Classification and ensemble techniques on a dataset to predict diabetes. Which are K-Nearest Neighbor (KNN), Logistic Regression (LR), Decision Tree (DT), Support Vector Machine (SVM), Gradient Boosting (GB) and Random Forest (RF) (2). In this paper proposed a combination of Convolutional Neural Networks (CNN) and Bidirectional Long Short-Term Memory (BiLSTM) models, with Doc2vec embedding, suitable for opinion analysis in long texts. The CNN-BiLSTM model is compared with CNN, LSTM, BiLSTM and CNN-LSTM models with Word2vec/Doc2vec embeddings (3).This article aims to design a face recognition attendance system based on real-time video processing. This article mainly sets four directions to consider the problems: the accuracy rate of the face recognition system in the actual check-in, the stability of the face recognition attendance system with real-time video processing, the truancy rate of the face recognition attendance system with real-time video processing and the interface settings of the face recognition attendance system using real-time video processing (4).

METHODOLOGY

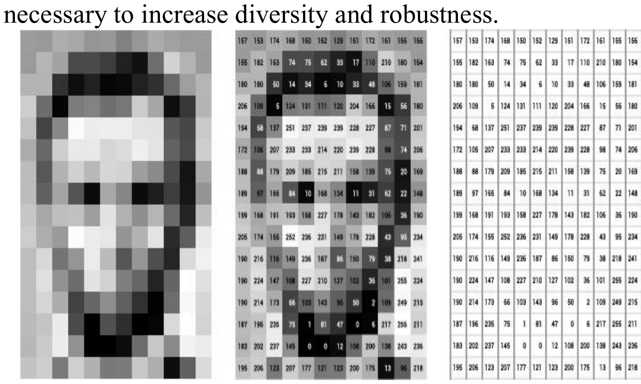
1. Data Collection

- Install onboard cameras in rural transport vehicles to capture images of seat occupancy.
- Collect historical data on seat occupancy, including images and corresponding metadata (e.g., timestamps, route information).

Fig1: Image data set-1

Fig2: Image data set-2

2. Image Preprocessing
- Preprocess the captured images to enhance quality and reduce noise.
- Resize images to a standardized resolution to ensure consistency.
- Augment the dataset if necessary to increase diversity and robustness.



800 x 600 pixels

3. Feature Extraction with CNN
- Utilize Convolutional Neural Networks (CNN) for feature extraction from the pre-processed images.
- Train the CNN model on the dataset of labeled images to learn spatial patterns and dependencies related to seat occupancy.
- Extract relevant features from the CNN's intermediate layers, such as activations or embeddings, to represent seat occupancy information.

4. Contextual Data Integration
- Collect contextual data including demographic, geographic, and temporal factors that may influence seat vacancy.
- Encode categorical variables and normalize numerical variables for compatibility with machine learning algorithms.

5. Model Development
- Develop a predictive model that combines the extracted features from CNN with contextual data using Linear Regression.
- Concatenate the CNN-derived features with the contextual variables as input to the Linear Regression model.

- Split the dataset into training, validation, and testing sets to evaluate model performance.

	A	B	C	D	E
1	bus_id	route_num	max_seat_available	no_of_seat_occupied	no_of_seat_available
2	1	route_1	50	38	12
3	2	route_2	50	50	0
4	3	route_3	55	25	30
5	4	route_4	55	53	2
6	5	route_5	50	18	32
7	6	route_6	50	50	0
8	7	route_7	50	50	0
9	8	route_8	50	50	0
10	9	route_9	50	50	0
11	10	route_10	55	49	6
12	11	route_11	50	31	19
13	12	route_12	55	55	0
14	13	route_13	50	0	50
15	14	route_14	50	33	17
16	15	route_15	55	27	28
17	16	route_16	50	45	5
18	17	route_17	55	5	50
19	18	route_18	50	38	12
20	19	route_19	55	49	6
21	20	route_20	50	20	30
22	21	route_21	49	47	2
23	22	route_22	50	48	2
24	23	route_23	52	39	13

Fig 4: Training data set

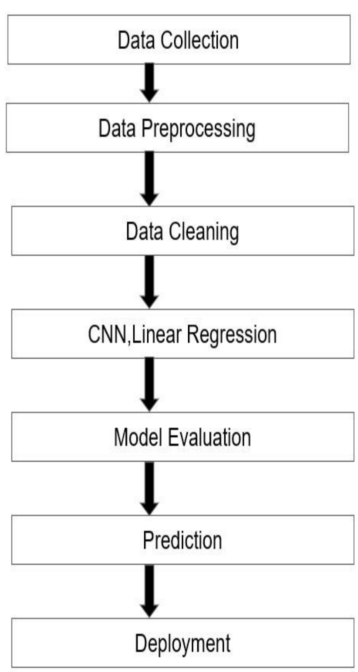
6. Model Training and Validation
- Train the predictive model on the training set using gradient descent optimization to minimize the loss function.
- Validate the model on the validation set to monitor performance metrics such as Mean Absolute Error (MAE), Mean Squared Error (MSE), and R-squared.

7. Hyperparameter Tuning
- Perform hyperparameter tuning to optimize the model's performance, including adjusting learning rates, regularization parameters, and network architecture if applicable.
- Utilize techniques such as cross-validation and grid search to identify optimal hyperparameters.

8. Model Evaluation
- Evaluate the final trained model on the testing set to assess its generalization performance.
- Compare the predicted seat vacancy levels against ground truth observations to measure the model's accuracy and reliability.

8.1 Model Evaluation Parameters
  - True Positive
  - True Negative
  - False Positive
  - False Negative

FIOW CHART



Software used
- OpenCV(image processing)

- AnacondaJupiterNoteBook/Google collaboratory
- Microsoft Excel
- Windows 8 and above
- Python IDE 3

ML Libraries

- NumPy
- Pandas
- Sciait-learn
- Tensor flow
- Matplotlib
- Open CV
- pyTorch
- Keras

Programming languages

- Python 3

Hardware used

- Camera module
- PC With OS Installed

CONCLUSION

In summary, this project developed a predictive system for rural transport seat vacancy using machine learning and image processing. By integrating Convolutional Neural Networks (CNN) and Linear Regression, it accurately forecasts seat availability, aiding resource allocation and transport accessibility in rural areas. Empirical evaluation confirmed its effectiveness, promising enhanced rural development and socio-economic growth through optimized transport planning and management.

REFERENCE

[1] Mohammed Y Aalsalem, Wazir Zada Khan, Wajeb Gharibi, Nasrullah Armi "An intelligent oil and gas well monitoring system based on Internet of Things" International Conference on Radar, Antenna, Microwave, Electronics, and Telecommunications (ICRAMET),2017.

[2] Sayeda Islam Nahid, Mohammad Monirujjaman Khan " Toxic Gas Sensor and Temperature Monitoring in Industries using Internet of Things (IoT)" International Conference on Computer and Information Technology (ICCIT)2021

[3] S.Vivekanandan , Abhinav Koleti, M Devanand Autonomous industrial hazard monitoring robot with GSM integration International Conference on Engineering (NUiCONE)2013

[4] Meer Shadman Saeed, Nusrat Alim Design and Implementation of a Dual Mode Autonomous Gas Leakage Detecting Robot International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST)2019

[5] A.Sandeep Prabhakaran Mathan N Safety Robot for Flammable Gas and Fire Detection using Multisensor Technology International Conference on Smart Electronics and Communication (ICOSEC)2021.

[6] Ashutosh Mishra; Shiho Kim; N S Rajput" An Efficient Sensory System for Intelligent Gas Monitoring Accurate classification and precise quantification of gases/odors" International SoC Design Conference (ISOCC) 2020.

[7] Qiang Luo; Xiaoran Guo; Yahui Wang; Xufeng Wei "Design of wireless monitoring system for gas emergency repairing" Chinese Control and Decision Conference (CCDC) 2016.

[8] Mohammed Y Aalsalem; Wazir Zada Khan; Wajeb Gharibi; Nasrullah Armi "An intelligent oil and gas well monitoring system based on Internet of Things" International Conference on Radar, Antenna, Microwave, Electronics, and Telecommunications (ICRAMET) 2017.

[9] C.Nagarajan and M.Madheswaran - 'Experimental verification and stability state space analysis of CLL-T Series Parallel Resonant Converter' - Journal of ELECTRICAL ENGINEERING, Vol.63 (6), pp.365-372, Dec.2012.

[10] C.Nagarajan and M.Madheswaran - 'Performance Analysis of LCL-T Resonant Converter with Fuzzy/PID Using State Space Analysis'- Springer, Electrical Engineering, Vol.93 (3), pp.167-178, September 2011.

[11] C.Nagarajan and M.Madheswaran - 'Stability Analysis of Series Parallel Resonant Converter with Fuzzy Logic Controller Using State Space Techniques'- Taylor &; Francis, Electric Power Components and Systems, Vol.39 (8), pp.780-793, May 2011.

[12] C.Nagarajan and M.Madheswaran - 'Experimental Study and steady state stability analysis of CLL-T Series Parallel Resonant Converter with Fuzzy controller using State Space Analysis'- Iranian Journal of Electrical & Electronic Engineering, Vol.8 (3), pp.259-267, September 2012.

[13] Nagarajan C., Neelakrishnan G., Akila P., Fathima U., Sneha S. "Performance Analysis and Implementation of 89C51 Controller Based Solar Tracking System with Boost Converter" Journal of VLSI Design Tools & Technology. 2022; 12(2): 34–41p.
[14] C. Nagarajan, G.Neelakrishnan, R. Janani, S.Maithili, G. Ramya "Investigation on Fault Analysis for Power Transformers Using Adaptive Differential Relay" Asian Journal of Electrical Science, Vol.11 No.1, pp: 1-8, 2022.
[15] G.Neelakrishnan, K.Anandhakumar, A.Prathap, S.Prakash "Performance Estimation of cascaded h-bridge MLI for HEV using SVPWM" Suraj Punj Journal for Multidisciplinary Research, 2021, Volume 11, Issue 4, pp:750-756
[16] G.Neelakrishnan, S.N.Pruthika, P.T.Shalini, S.Soniya, "Perfromance Investigation of T-Source Inverter fed with Solar Cell" Suraj Punj Journal for Multidisciplinary Research, 2021, Volume 11, Issue 4, pp:744-749
[17] C.Nagarajan and M.Madheswaran, "Analysis and Simulation of LCL Series Resonant Full Bridge Converter Using PWM Technique with Load Independent Operation" has been presented in ICTES'08, a IEEE / IET International Conference organized by M.G.R.University, Chennai.Vol.no.1, pp.190-195, Dec.2007
[18] M Suganthi, N Ramesh, "Treatment of water using natural zeolite as membrane filter", Journal of Environmental Protection and Ecology, Volume 23, Issue 2, pp: 520-530,2022
[19] M Suganthi, N Ramesh, CT Sivakumar, K Vidhya, "Physiochemical Analysis of Ground Water used for Domestic needs in the Area of Perundurai in Erode District", International Research Journal of Multidisciplinary Technovation, pp: 630-635, 2019