

QR Code-Based Ticketing System for Heritage Sites

Paras Wane¹, Sankalp Kamble², Rutuja Pol³, Sanika Bidkar⁴, Nandinee Mudegol⁵

¹ Walchand College of Engineering Sangli, Department of Computer Science and Engineering, Vishrambag, Sangli, Maharashtra, India. Email: paraswane15@gmail.com

² Walchand College of Engineering Sangli, Department of Computer Science and Engineering, Vishrambag, Sangli, Maharashtra, India. Email: sankalp.kamble.work@gmail.com

³ Walchand College of Engineering Sangli, Department of Computer Science and Engineering, Vishrambag, Sangli, Maharashtra, India. Email: rutujapol20@gmail.com

⁴ Walchand College of Engineering Sangli, Department of Computer Science and Engineering, Vishrambag, Sangli, Maharashtra, India. Email: sbidkar127@gmail.com

⁵ Walchand College of Engineering Sangli, Department of Computer Science and Engineering, Vishrambag, Sangli, Maharashtra, India. Email: nandinee.mudegol@walchandsangli.ac.in

Abstract

This research presents a modern QR-based ticketing system for heritage sites, designed to streamline operations and enhance visitor experiences. By integrating machine learning, the system predicts visitor numbers, enabling efficient resource management and improved operational planning. The proposed solution automates ticket booking, validation, and entry using secure QR codes, effectively reducing long queues and eliminating the environmental impact of paper tickets. Additionally, predictive analytics provide heritage site administrators with valuable insights into crowd patterns, facilitating better decision-making. This paper explores the system's architecture, predictive model, experimental results, and its broader implications for cultural heritage management.

Keywords: QR-based ticketing, heritage sites, machine learning, visitor prediction, crowd management, cultural heritage management, smart tourism.

1. Introduction

Heritage sites, as significant cultural landmarks, face operational challenges due to outdated ticketing methods that rely heavily on manual processes. These traditional systems often lead to inefficient resource allocation, long visitor queues, and environmental concerns due to excessive paper usage. Additionally, the absence of real-time analytics and predictive capabilities results in suboptimal visitor management and planning.

1.1 Current Challenges

Key issues with conventional ticketing systems at heritage sites include:

- Manual Processes – Increased staff workload, operational inefficiencies, and human error.
- Limited Accessibility – Fragmented and outdated booking systems reduce user convenience.
- Environmental Impact – Paper-based tickets contribute to ecological concerns.

- Lack of Predictive Insights – Absence of data-driven analytics leads to ineffective resource planning

1.2 Contribution of the Proposed System

The proposed QR-based ticketing system introduces:

- Automated Ticket Booking and Entry – Secure QR codes streamline visitor management.
- Real-time Visitor Forecasting – Machine learning predicts footfall for better resource allocation.
- Scalability and Security – Built using React, Flask, and MongoDB to ensure efficiency and adaptability

2. Related Work

Various works have investigated QR-code usage in e-ticketing and public transportation. QR-codes, initially designed for the automobile sector, are currently extensively utilized for encoding user data on numerous platforms [1]. Singh et al. suggested a Dynamic Seat Allocation (DSA) system with QR-code processing for equitable seat booking in Indian Railways

[2]. Shinde et al. presented an Android app for QR-code tickets but pointed towards user challenges resulting from unawareness [3].

Babar and Dongare came up with a GPS-based ticket booking model but encountered weak signal accuracy problems [4]. Trebar and Finzgar applied QR-code and RFID-based ticketing in public transport to ensure smooth passenger verification [5]. Another work presented a smart card with QR-code verification for digital bus passes to ease the process of verification for conductors [6]. Shaikh et al. highlighted QR-code advantages for mobile-ticket validation [7].

Mumbai railway ATVMs enhanced ticketing efficiency but were plagued by scalability [8]. An NFC-based ticketing pilot project was tested in Germany, with dynamic pricing but high costs that impeded adoption [9]. Jadhav et al. suggested a real-time QR-code system to enhance convenience in ticketing [10]. Chavan et al. created an Android-based train ticket reservation system but highlighted issues of user adaptability [11].

More research investigated the integration of QR-code with wireless technologies for enhanced security and ease of use in e-ticketing [12]. More research probed user awareness and adoption issues in QR-code ticketing [13]. GPS-based ticket reservation systems were improved with real-time tracking but faced connectivity problems [14]. Another research suggested a better QR-code validation system, enhancing ticket scanning efficiency [15].

2.1 Research Gaps

Despite these advancements, existing solutions fail to address:

1. The unique challenges of managing high visitor volumes in museums.
2. The integration of predictive analytics with ticketing systems.
3. The need for scalable, eco-friendly systems that cater to diverse audiences.

3. Proposed Methodology

This project implements an integrated QR-based ticketing system alongside a machine learning (ML) model for visitor prediction, ensuring seamless ticket management and optimized crowd control. The methodology follows a structured approach, incorporating both automated ticketing and predictive

analytics to enhance visitor experience and operational efficiency.

3.1 General Methodology

The methodology for this project consists of two key components: ML-based crowd prediction and QR-based ticketing. The ML model analyzes historical and real-time data to forecast visitor trends, enabling efficient crowd management and resource allocation. Meanwhile, the QR-based system ensures secure and seamless ticket generation, reducing manual intervention and streamlining the entry process. Together, these components enhance operational efficiency, minimize errors, and improve the overall user experience.

3.1.1 Crowd Prediction ML Model

3.1.1.1 Dataset Generation

The dataset for the ML model was compiled by collecting historical visitor records from museum authorities. This dataset, termed visitors.csv, includes key attributes affecting visitor count predictions:

- **Date:** The specific date of the record.
- **No. of Visitors:** The target variable representing the number of visitors.
- **Special Day:** A categorical feature indicating if a special event influenced attendance.
- **Weekday Name:** The designated day of the week (e.g., Monday, Tuesday) to analyze trends in visitor behavior.

The dataset covers a significant period, capturing seasonal fluctuations and event-based attendance trends to enhance the accuracy of predictions.

Date	No. of Visitors	Special Day	Weekday Name
1/1/2021	800	New Year's	Friday
1/2/2021	700	None	Saturday
1/3/2021	900	None	Sunday
1/4/2021	400	None	Monday
1/5/2021	300	None	Tuesday
1/6/2021	600	None	Wednesday
1/7/2021	500	None	Thursday
1/8/2021	850	None	Friday
1/9/2021	750	None	Saturday
1/10/2021	950	None	Sunday
1/11/2021	420	None	Monday
1/12/2021	320	None	Tuesday
1/13/2021	620	Lohri	Wednesday
1/14/2021	520	Makar Sankranti	Thursday
1/15/2021	880	Pongal	Friday
1/16/2021	780	None	Saturday
1/17/2021	960	None	Sunday
1/18/2021	430	None	Monday
1/19/2021	330	None	Tuesday

Figure 1 Screenshot of Sample Dataset

3.1.1.2 Data Preprocessing

To preprocess the dataset for training, several steps were undertaken:

1. **Date Transformation:** The Date column was converted into a date - time format, allowing the extraction of temporal features such as day, month, and year.
2. **Feature Engineering:**
 1. Extracted features: Day, Month, Year, and Weekday from the Date column.
 2. Created a new binary feature, IsWeekend, to distinguish weekends from weekdays.
 3. One-hot encoding was applied to Special Day and Weekday Name for categorical handling.
3. **Dropping Redundant Columns:** The original Date column was removed after feature extraction to prevent redundancy.

These steps ensure that the dataset is clean, structured, and suitable for machine learning algorithms.

3.1.1.3 Data Splitting

The dataset was systematically split into training and testing sets for model training and evaluation:

1. **Feature Selection and Target Definition**
 - **Features (X):** All columns except "No. of Visitors."
 - **Target (y):** "No. of Visitors," representing the predicted variable.
2. **Train-Test Split**
 - The dataset was partitioned into 80% training and 20% testing using Scikit-Learn's train_test_split() function."
 - A fixed random seed was utilized to maintain result reproducibility."

3.1.1.4 Model Creation

The predictive model was built using machine learning techniques to estimate future visitor counts:

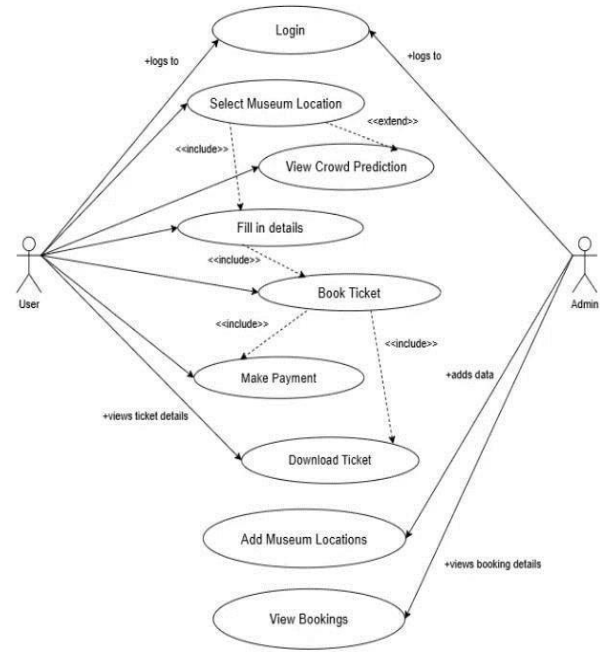


Figure 2 Use Case Diagram

1. **Pipeline Setup:** A machine learning pipeline was established with:
 - **Preprocessor:** A ColumnTransformer for data preprocessing.
 - **Model:** A RandomForestRegressor, chosen for its ability to capture complex relationships.
2. **Training:** The model was trained using the fit() function on the training dataset.
3. **Prediction and Evaluation:**

The trained model was used to generate predictions on the test dataset [16].

Performance was assessed using standard regression metrics.

- **Mean Absolute Error (MAE)** – Calculates the average difference between predicted and actual visitor counts.

$$1) \text{ MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

- **Mean Squared Error (MSE)** – Assigns greater weight to larger prediction errors, emphasizing significant deviations. [17].

$$2) \text{ MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

- **R² Score** – Measures how effectively the model explains the variance in visitor data.

$$3) R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2}$$

Where,

n = Total number of observations

y_i = Actual Value

$\hat{y}_i$ = Predicted Value

3.1.1.5. Model Performance Results

The model was validated using the test dataset, yielding the following performance metrics

Table 1. Calculated values for Random Forest Regressor

Metric	Value
MAE	11.57
MSE	221.50
R ² Score	0.9967 (99.67% accuracy)

These results confirm that the model is highly reliable for predicting visitor traffic and optimizing crowd management.

3.1.2 QR Code-Based Ticket Generation

3.1.2.1 Ticket Creation and Management

- Users submit booking details (name, email, visit date, number of visitors).
- A unique Ticket ID is generated and stored in a MongoDB database with the booking details.

3.1.2.2 QR Code and PDF Generation

- The **qrcode** library is used to generate a QR code that securely encodes ticket details.
- A PDF ticket is generated, incorporating the QR code for seamless validation.
- The ticket is saved with the **Ticket ID** to ensure easy access and retrieval.

3.1.2.3 Email Delivery

- The system sends the PDF ticket via email using nodemailer.
- The email includes booking details and a scannable QR code for entry validation.

3.1.2.4 Ticket Validation & Retrieval

- Visitors present the QR code at the entry point for verification.

- The system scans and validates the QR code, ensuring secure access.
- Users can retrieve lost tickets by providing the Ticket ID.

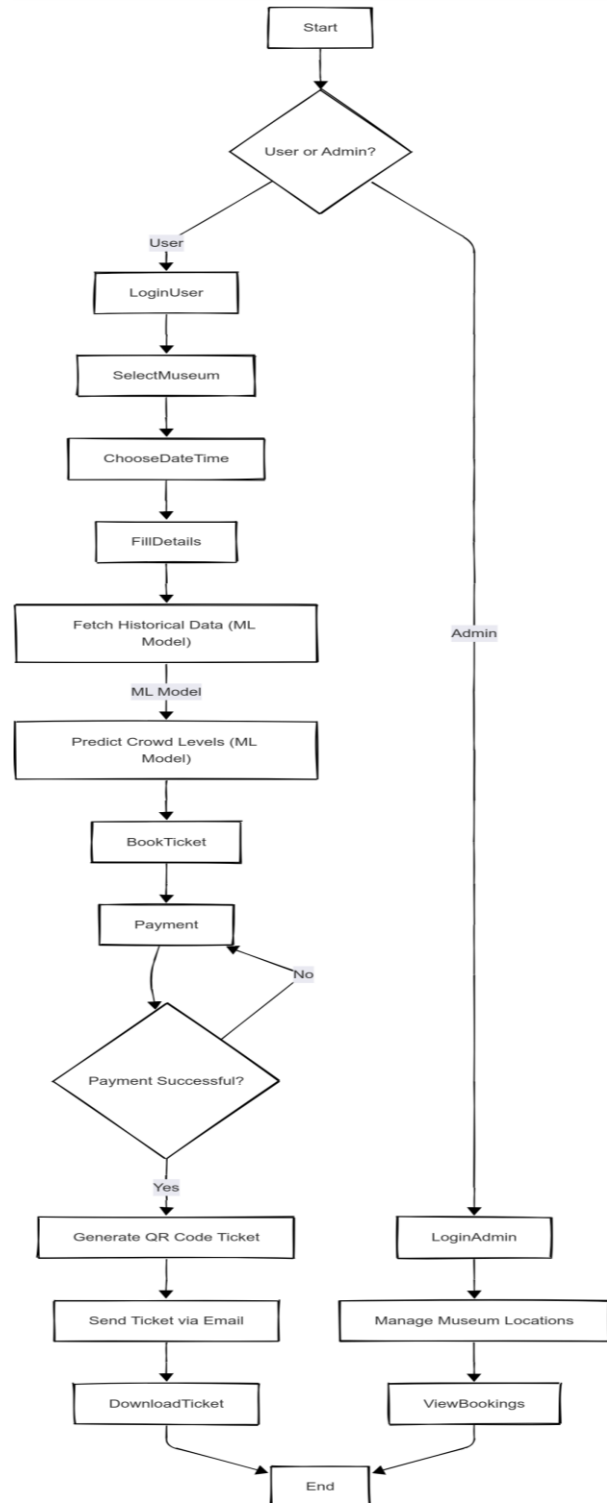


Figure 3 Flow Diagram

This **automated, secure, and paperless** ticketing system improves the booking experience and enhances operational efficiency.

4. Dataset Description and Technology Stack

This section provides a comprehensive overview of the dataset and technology stack utilized in developing the QR Code-Based Ticketing System. The dataset consists of historical visitor records, which are essential for training the machine learning model to predict visitor numbers accurately. The technology stack encompasses both hardware and software components necessary for the system's design, implementation, and deployment. It includes tools for frontend and backend development, data handling, and machine learning frameworks, ensuring a robust, efficient, and scalable solution.

4.1 Dataset Details

The dataset used in this research contains extensive information on museum visitors, which is essential for analyzing visitor trends and optimizing operational strategies. It includes key attributes such as visitor counts, timestamps, and booking types (e.g., general admission, exhibition passes, and group tours). These data points form the basis for predicting high-traffic days and effectively allocating resources.

Key Attributes of the Dataset:

- **Date:** The recorded date of visitor data.
- **Visitor Count:** The total number of visitors on a given day.
- **Booking Type:** Classification of bookings, including general admission, exhibition passes, and group tours.
- **Peak Hours:** Time periods with the highest visitor influx, aiding in operational planning and resource management.

Preprocessing and Data Preparation:

To enhance the accuracy and reliability of the predictive model, the dataset underwent several preprocessing steps:

- **Handling Missing Values:** Ensuring data consistency by filling in or removing incomplete records.
- **Normalizing Data Formats:** Standardizing date formats and numerical values for consistency.
- **Data Transformation:** Restructuring raw data into a standardized format optimized for analysis and model training.

These preprocessing steps were critical in developing an accurate visitor forecasting model, minimizing inconsistencies, and improving model performance.

4.2 Technology Stack

The development of the QR Code-Based Ticketing System relied on a robust and scalable technology stack to enable real-time ticket booking, visitor data processing, and machine learning-driven forecasting. Carefully selected hardware and software components were integrated to ensure high performance, security, and efficiency.

4.2.1 Hardware Specifications

To ensure smooth data processing, efficient ticket generation, and effective predictive model training, the system was deployed on hardware with the following specifications:

- **Processor:** Intel Core i5 or an equivalent, offering adequate computational power for system operations and data processing.
- **RAM:** At least 8 GB, enabling seamless multitasking and efficient handling of large datasets during model training.
- **Storage:** 256 GB SSD, providing faster data access, processing, and retrieval compared to HDDs.
- **GPU (Optional):** NVIDIA GTX series or an equivalent, assisting in accelerating machine learning tasks such as model training.

4.2.2 Software Stack

The software stack comprises various tools and frameworks that support frontend and backend development, database management, and machine learning model training.

Frontend Development

- **React.js:** Used to develop a dynamic and responsive user interface, enabling seamless interactions and real-time updates.
- **HTML5 & CSS3:** Employed to structure and style webpage components, ensuring an intuitive and user-friendly experience.

Backend Development

- **Node.js & Express.js:** Implemented for handling server-side logic, managing user booking requests, processing data exchanges between the frontend

and backend, and ensuring smooth system performance.

- **Nodemailer:** Integrated for sending QR-coded e-tickets to users via email upon successful booking, automating the ticket distribution process.

Database Management

- **MongoDB:** A NoSQL database selected for its scalability and efficiency in handling unstructured data, including booking records, user information, and chatbot interactions. Its flexible schema allows rapid modifications and supports large datasets.

Machine Learning

- **Scikit-learn:** A Python-based machine learning library utilized for data preprocessing, analysis, model training, and evaluation.
- **Algorithm:** The Random Forest Regressor, an ensemble learning method, was chosen for predicting visitor numbers. By combining predictions from multiple decision trees, this algorithm improves accuracy and robustness, making it well-suited for forecasting visitor trends.

5. Results and Analysis

This section provides a detailed analysis and evaluation of the QR Code-Based Ticketing System, focusing on its core functionalities such as QR code generation, ticket processing, and visitor prediction. The system's performance is assessed based on efficiency, real-time monitoring capabilities, and user feedback, demonstrating its effectiveness in optimizing the ticketing process.

5.1 Model Evaluation

The model's performance was analyzed using statistical methods and data visualizations. A correlation matrix provided insights into the relationships between different features and visitor numbers.

Key Observations from the Correlation Analysis

- **Weekday and Visitor Numbers:** A strong positive correlation of 0.79 was observed, indicating an increase in visitor traffic as the week progresses, likely due to more weekday visits than weekends.
- **Weekend Effect:** The "IsWeekend" feature showed a moderate positive correlation of 0.55, suggesting higher visitor numbers on weekends.

Impact of Specific Weekdays:

- **Sunday:** Moderate correlation of **0.50**, indicating high foot traffic.
- **Saturday:** Correlation of 0.22, suggesting moderate weekend activity.

- **Month Influence:** A positive correlation of 0.36 suggests that certain months attract more visitors, possibly due to favorable weather or holiday seasons.

Special Days:

- Children's Day, New Year's Eve, and Christmas Eve showed a slight positive correlation (0.09 - 0.10), indicating a modest rise in visitor numbers.
- Republic Day, Makar Sankranti, and Buddha Purnima exhibited negative correlations (-0.08 to -0.10), suggesting a lower influence on visitor attendance.

Weekday Influence:

- Tuesday (-0.66), Monday (-0.29), and Thursday showed negative correlations, indicating lower visitor counts on these days.

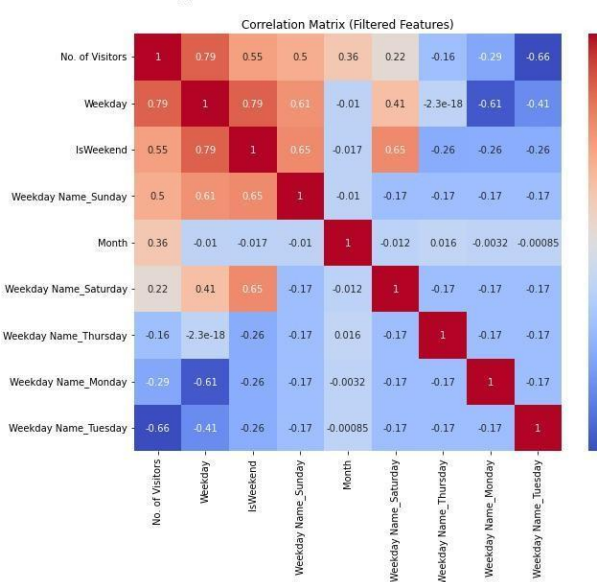


Figure 4 Correlation Matrix

6. Figures and Tables

Key Insights from the Plot:

- **Shape of the Distribution:** The histogram, combined with the KDE curve, suggested a right-skewed distribution, where most days experience

fewer visitors, with a few days exhibiting exceptionally high visitor numbers.

- **Peak at Lower Visitor Numbers:** The histogram revealed a concentration of data on the lower end of visitor counts, indicating that a majority of the days had moderate or low traffic.
- **Presence of Outliers:** Longer tails on the right side of the distribution signaled days with extremely high visitor numbers, which could be attributed to special events or holidays.
- **Spread of Data:** The variability in the number of visitors was wide, suggesting that the number of visitors fluctuates greatly from day to day. This variability is crucial for forecasting and resource allocation.

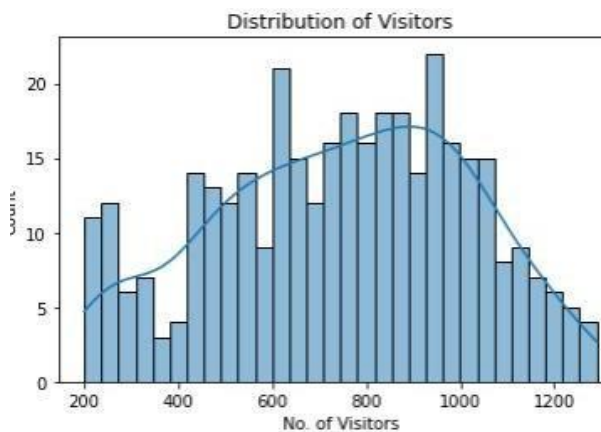


Figure 5 Distribution of visitors graph

Key Data Visualizations:

6.1. Area Chart (Monthly Visitors with Cumulative Data):

This chart tracks monthly visitor numbers and cumulative data throughout the year. It reveals steady growth in visitor numbers, with December showing the highest cumulative count, useful for identifying annual visitor trends.

6.2. Radar Chart (Compare Visitor Counts Across Weekdays):

The radar chart highlights the distribution of visitors across weekdays. Peaks on weekends suggest higher traffic, while troughs on weekdays like Tuesday and Monday indicate lower visitor numbers.

6.3. Line Chart (Daily Visitors Over Time):

The line chart displays fluctuations in daily visitor numbers, reflecting the impact of events, holidays, or weather conditions.

6.4. Bar Chart (Number of Visitors by Month):

This chart summarizes visitor numbers by month, helping administrators plan resources based on seasonal traffic patterns.

6.5. Pie Chart (Distribution of Visitors on Special Days):

The pie chart categorizes visitor distribution on special days (e.g., holidays, cultural events). It reveals which special days attract the most visitors, assisting in event planning and promotions.



Figure 6 Data Visualization



Figure 7 ML Model Prediction

7. Conclusion

The The QR code-based ticketing system developed in this study provides an efficient and user-friendly solution for managing visitor flow and streamlining entry processes across various venues. By incorporating machine learning algorithms for visitor forecasting, the system not only optimizes resource allocation but also improves crowd management. The seamless integration of an intuitive booking interface, secure QR code generation, and robust backend support illustrates the system's efficacy in improving the overall visitor experience.

7.1 Major Achievements:

- **Automated Ticketing & QR Code Generation:** Streamlined ticket issuance with secure QR codes, reducing wait times and enhancing security.
- **ML-Based Visitor Forecasting:** Implemented a predictive model to anticipate peak visitor times, optimizing resource allocation.
- **User-Friendly Interface:** Designed an intuitive booking system, enhancing user experience and simplifying ticketing.
- **Real-Time Monitoring:** Enabled live tracking for administrators, ensuring system reliability and efficiency during peak hours.

References

- [1] Vinay Maheshwar, Kalpesh Patil, Azim Maredia, Apeksha Waghmare "Android Application on ETicketing Railway System Using Qr-Code", IOSRJEN, ISSN (e): 2250-3021, ISSN (p): 2278- 8719 Volume 13, PP 33-38
- [2] Durgesh Jaiswal, Prof. Jayashri Sonawane, Rushikesh Badgujar, Chetan Shivade, "Revolutionizing Museum Experiences: Implementing Ticketless Entry Systems Using QR Code." *International Research Journal of Modernization in Engineering, Technology and Science (IRJMETS)*, e-ISSN: 2582-5208, Volume 06, Issue 05, May 2024, Impact Factor: 7.868. Available at: https://www.irjmets.com/uploaded/files/paper/issue_5_may_2024/57248/final/fin_irjmets1720011603.pdf.
- [3] Aaroohi Rathore, Aayush Gupta, Abhay Gour, Ankur Nagar, "E-Ticketing System for Indian Museums & Heritage Sites." *International Research Journal of Modernization in Engineering, Technology and Science (IRJMETS)*, e-ISSN: 2582-5208, Volume 04, Issue 11, November 2022, Impact Factor: 6.752. Available at: https://www.irjmets.com/uploaded/files/paper/issue_11_november_2022/31115/final/fin_irjmets1668090963.pdf.
- [4] Avantika Jirapure, Tanaya Joshi, Tanushri Kakas, Prachiti Panchbhai, Prof. S. C. Shirbhate, "Ticketless Entry System for Heritage/Museums." Available at: <https://doi.org/10.22214/ijraset.2024.60107>.
- [5] Xucai Zhang, Yeran Sun, Fangli Guan, Kai Chen, Frank Witlox, Haosheng Huang, "Forecasting the Crowd: An Effective and Efficient Neural Network for Citywide Crowd Information Prediction at a Fine Spatio-Temporal Scale." *Transportation Research Part C: Emerging Technologies*, Volume 141, 2022. Available at: <https://www.sciencedirect.com/science/article/abs/pii/S0968090X2200273X>.
- [6] Karbovskii, V., Lees, M., Presbitero, A., Kurilkin, A., Voloshin, D., Derevitskii, I., Karsakov, A., & Slood, P. M. A. (2021). "Utilizing Ensemble Learning for Large-Scale Crowd Flow Prediction." *Engineering Applications of Artificial Intelligence*, 104. Retrieved from ScienceDirect.
- [7] Bhiwgade, M., Dambale, R., Ambadkar, S., Raja, S., Sawalakhe, K., Gudadhe, S. R., & Shandilya, V. K. (2023). "Ticketless Entry System for Monuments and Museums." Available at DOI.
- [8] Yuan, X., Han, J., Wang, X., He, Y., Xu, W., & Zhang, K. "An Innovative Learning Approach for Predicting Citywide Crowd Flow." Retrieved from IEEE Xplore.
- [9] A. M. Ansari, A. Alam, M. M. Barga, "Next Generation E-ticketing system" Available at: 10.48175/IJARST-13654.
- [10] Jogekar, R., Wasnik, R., Supare, P., Gawande, N., Chopkar, H., & Ukeybondre, R. *An Overview of QR Code-Based Ticketing Systems: A Review*. Published in the *International Journal of Scientific Research in Science and Technology (IJSRST)*. Available at: IJSRST.
- [11] Gautam Anand, Jatin Suthar, Harshal Jain, Devanshi Minda, Nandani Dalsaniya, Jyoti Kaushal, "Ticketless Entry in Heritage Museums." *International Advanced Research Journal in Science, Engineering and Technology (IARJSET)*, Vol. 10, Special Issue 2, May 2023, International Conference on Multi-Disciplinary Application & Research Technologies (ICMART-2023), Geetanjali Institute of Technical Studies. Available at: <https://iarjset.com/wp-content/uploads/2023/06/IARJSET-ICMART-18.pdf>.
- [12] Maheshwar, V., Patil, K., Maredia, A., & Waghmare, A. (2019). Android Application on E-Ticketing Railway System Using QR-Code. IOSRJEN, 13, 33-38.
- [13] Swarup, M. M., Dwivedi, A., Sonkar, C., Prasad, R., & Bag, M. (2012). *QR-Code-Based Processing for Dynamic and Transparent Seat Allocation in Indian Railways. IJCSI International Journal of Computer Science Issues*, 9(3), 1-6.

- [14] Sheikh, S., Shinde, G., Potghan, M., & Shaikh, T. (2020). *Urban Railway Ticketing Application. International Journal of Scientific Research in Science and Technology (IJSRST)*, 7(1), 103-107.
- [15] Dongare, T., & Babar, A. (2014). *Android-Based Ticket Reservation System Utilizing GPS for Ticket Validation. International Journal of Emerging Research in Management and Technology (IJERMT)*, 3(3).
- [16] Dagur, A., Shukla, D. K., Makhmadiyarovich, N. F., Rustamovich, A. A., & Sindorovich, J. J. *Advanced Artificial Intelligence Algorithms*.
- [17] Dey, S. K., & Rahman, M. M. *Impact of Machine Learning Techniques on Flow-Based Anomaly Detection in Software-Defined Networking*.