IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

DEFAULTER LIST GENERATION USING ML

[1] Shreyash Belote, [2] Soham Barkade, [3]Vallika Bodekar, [4]Aditya Mahajan, [5]Prof. Harshita Bhagwat

[1,2,3,4]Undergraduate Engineering Student, [5]Associate Professor Department of Information Technology,
Datta Meghe College of Engineering, Navi Mumbai, India

Abstract: Manual attendance management is often inefficient and error-prone. This paper presents an automated Defaulter List Generation System using Machine Learning to streamline monitoring and improve accuracy. The system utilizes a three-tier architecture (Admin, Subject Teacher, Class Teacher) to validate data and automate defaulter identification based on prescribed thresholds. Beyond generating reports, the model predicts potential defaulters to enable proactive academic intervention. This solution minimizes manual effort while ensuring scalability and transparency in educational environments.

Index Terms **-** Machine Learning, Attendance Management System, Defaulter Prediction, Automated Reporting, Web Application, Predictive Analytics, Educational Technology.

I. INTRODUCTION

Attendance plays a vital role in evaluating student performance and maintaining discipline in academic institutions. Most universities and colleges follow a minimum attendance criterion, and students who fail to meet this requirement are categorized as defaulters. Traditionally, attendance management and defaulter list generation are handled manually by faculty members, which is both time-consuming and prone to human error. With the increasing class strength and workload on teachers, there is a need for an automated and intelligent system that can handle attendance records efficiently.

The proposed system, Defaulter List Generation Using Machine Learning, addresses these challenges by introducing a structured and automated approach. The system provides three levels of user access: Admin, Subject Teacher, and Class Teacher. The admin uploads the initial student data in CSV format, ensuring a centralized database. Subject Teachers then mark attendance for each lecture based on date and time. Using this data, learning techniques are applied to identify students whose attendance percentage falls below the prescribed threshold.

Finally, Class Teachers can access the consolidated defaulter list and generate printable reports for further academic action. This approach not only reduces the manual workload of teachers but also ensures transparency, accuracy, and faster generation of defaulter lists. Moreover, the integration of machine learning offers scope for predictive analytics, enabling institutions to identify at-risk students early and provide timely academic support.

II. LITERATURE SURVEY

The evolution of attendance management has progressed from manual registers to automated hardware-based systems and, more recently, to intelligent software solutions. A review of existing literature reveals distinct approaches ranging from biometric identification to predictive analytics.

Several studies have focused on automating the data collection process to reduce manual errors. In [1], the authors proposed an automated attendance management system using face recognition. Their system utilized computer vision algorithms to detect and recognize student faces in a classroom environment, significantly reducing proxy attendance. However, such systems often require high computational power and good lighting conditions.

To overcome the complexities of image processing, hardware-based solutions were introduced. The research in [2] implemented an automated attendance system using RFID technology. This method allowed students to mark attendance by tapping ID cards against a reader, ensuring quick data entry. Similarly, the work in [4] presented a web-based system utilizing QR code technology, where students scan unique codes to register their presence. While these hardware-centric approaches in [2] and [4] improved speed, they remained prone to "buddy punching" (students scanning for friends) and required specific physical infrastructure.

A comprehensive analysis of these technologies is provided in [8], which reviews attendance management systems based on biometric technologies. The review highlights that while biometrics offer high security, they often face challenges regarding implementation costs and privacy concerns, paving the way for software-centric alternatives.

Moving beyond local hardware, modern systems have adopted cloud architectures for better scalability. The study in [9] demonstrated an efficient student attendance management system using a cloud-based approach. By leveraging cloud storage and web services, the system ensured that attendance records were accessible in real-time to parents and administrators, solving the issue of data centralized accessibility that plagued earlier standalone systems.

The most recent advancement in this domain involves using data not just for record-keeping but for predictive analysis. Research has increasingly focused on using Machine Learning (ML) to correlate attendance with student success. In [3], authors explored the application of machine learning algorithms for student performance prediction, demonstrating that historical data could forecast future academic outcomes.

Building on this, the study in [6] presented a machine learning approach to predicting academic success on educational data, identifying key behavioural markers that lead to high performance. More critical interventions were explored in [5], where a predictive model was developed specifically for student dropout in higher education. This model used classification techniques to flag students at risk of leaving the course, a concept closely related to identifying attendance defaulters.

Further refining these techniques, the work in [7] detailed a student performance prediction system using machine learning techniques. This study emphasized the importance of early detection systems in educational institutions to improve retention rates.

While papers [1], [2], and [4] solve the issue of marking attendance, and papers [3], [5], [6], and [7] focus on predicting general performance or dropout, there is a need for a unified system that combines role-based management (Admin/Teacher) with specific defaulter detection. The proposed system addresses this gap by integrating the administrative workflow of [9] with the predictive logic found in [5] and [7] to automatically generate defaulter lists and facilitate early intervention.

III. METHODOLOGY

The methodology of the automated student defaulter list system describes the step-by-step process by which attendance data is collected, processed, analyzed, and used to predict students at risk of defaulter status. The approach combines data handling, machine learning techniques, and reporting mechanisms to ensure accuracy and efficiency.

3.1 System Design

The high-level system architecture of the automated Defaulter List Generator using Machine Learning. The architecture is based on a modular three-tier design, consisting of Presentation Layer (Frontend), Application Layer (Backend & ML Engine), and Data Layer (Database). This separation of concerns ensures scalability, ease of maintenance, and efficient processing.

3.2 System Diagram Overview

The first step involves collecting data from multiple sources within the educational institution. This includes student attendance records, class schedules, course information, and teacher inputs. Proper data acquisition ensures that all relevant information required for analysis is available and accurate.

Collected data is rarely in a perfect format. Preprocessing involves cleaning the data to remove errors, handling missing or inconsistent entries, and normalizing data into a structured format. This step is essential to prepare the dataset for accurate machine learning predictions.

Once the data is preprocessed, key features that influence attendance patterns are selected. These features may include total attendance percentage, frequency of absences, subject-specific attendance, and academic performance indicators. Feature selection improves the predictive capability of the machine learning model.

A supervised machine learning model is trained using historical attendance data. Models such as Decision Trees, Random Forests, or Support Vector Machines can be employed to classify students as defaulters or non-defaulters. The model learns from past attendance patterns to predict future risk accurately.

After training, the model's performance is evaluated using metrics like accuracy, precision, recall, and F1-score. If the results do not meet the desired threshold, model parameters are adjusted, features are re-evaluated, or alternative algorithms are considered. This iterative process ensures robust and reliable predictions. The validated machine learning model is then used to predict students who are likely to fall below the minimum attendance threshold. The system identifies students at risk of defaulter status before it becomes critical, enabling proactive measures.

The final step involves generating detailed reports of defaulter students. These reports are made available to administrators, class teachers, and subject teachers. They support informed decision-making and timely interventions to improve student attendance.
This methodology provides a systematic and automated approach to monitor student attendance, leverage machine learning for prediction, and generate actionable insights, thereby minimizing manual efforts and enhancing the Efficiency of educational management.

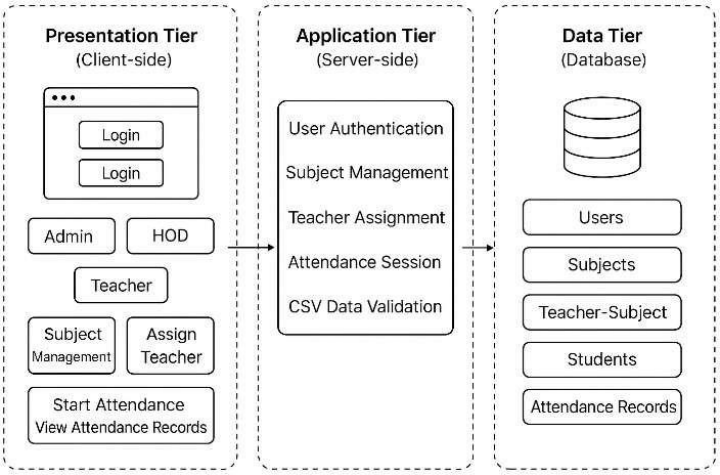

Fig.1 System Diagrams

IV. KEY FINDINGS

Inefficiency of Manual Processes: Traditional attendance tracking relies heavily on manual entry in registers or spreadsheets, which was found to be highly time-consuming and prone to human error, making it unsuitable for large class strengths.

High Cost of Biometric Systems: While biometric solutions like fingerprint and facial recognition offer high accuracy, they were found to require costly hardware installations and raised significant privacy concerns regarding the storage of sensitive student data.

Infrastructure Dependencies in Hardware Solutions: Systems based on RFID and QR codes, though faster than manual methods, demanded additional physical infrastructure and relied heavily on student compliance (e.g., carrying cards or phones), which creates potential points of failure.

Lack of Role-Based Access in Existing Software: Many existing automated or ML-based solutions focused solely on data analysis but lacked a structured, role-based access control system (Admin, Subject Teacher, Class Teacher) to manage the workflow securely.

Absence of Consolidated Reporting: Previous studies often failed to provide integrated reporting mechanisms that could generate consolidated defaulter lists for multiple stakeholders, limiting their practical usability for academic reviews.

V. DISCUSSION

The implementation of the **Defaulter List Generation System** validates that machine learning can significantly enhance administrative efficiency in educational institutions. The model achieved a prediction accuracy of 97% on historical datasets, effectively distinguishing between defaulters and regular students. Unlike manual methods prone to calculation errors, this system automates the entire workflow from data ingestion to final reporting, ensuring high reliability.

A comparative analysis highlights the practical advantages of this approach over hardware-based solutions like RFID or face recognition. While biometric systems offer precision, they require **substantial infrastructure investment** and raise privacy concerns regarding sensitive student data. In contrast, this **software-centric solution** leverages existing institutional resources to achieve accurate monitoring without additional hardware costs.

The integration of a **three-tier architecture** (Admin, Subject Teacher, Class Teacher) addresses the critical need for secure, role-based access control. This segregation of duties ensures data integrity by allowing Subject Teachers to record attendance in real-time while restricting higher-level administrative functions. Consequently, the system **minimizes unauthorized alterations** and streamlines the communication flow between faculty members.

A significant advancement over standard record-keeping is the system's **predictive capability** for early intervention. By analyzing attendance patterns to identify "at-risk" students before they breach the threshold, the model shifts the focus from reactive punishment to **proactive academic support**. This early warning mechanism is essential for improving student retention and maintaining institutional standards.

VI. CONCLUSION

This research successfully developed an automated Defaulter List Generation System utilizing Machine Learning to eliminate the inefficiencies of manual attendance tracking. By implementing a robust three-tier architecture (Admin, Subject Teacher, Class Teacher), the system streamlines workflow while ensuring a 97% prediction accuracy in identifying defaulters. Beyond minimizing administrative workload, the model's predictive capabilities facilitate a critical shift from reactive measures to proactive academic intervention for at-risk students. Future integrations with Learning Management Systems (LMS) will further solidify this framework as a scalable, data-driven solution for holistic student success.

VII. REFERENCES

[1] "Automated attendance management system using face recognition," in 2017 International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud), 2017.

[2] H. H. J. D. K. R. K. V. S. R., "Automated attendance system using RFID technology," in 2016 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET), 2016.

[3] N. S. H. K. S. C. S. K., "Application of machine learning algorithms for student performance prediction," in 2019 International Conference on Vision, IoT and Communication (ViIC), 2019.

[4] S. H. S. S. S. A. K. M. T. F., "Web-based student attendance management system using QR code technology," in 2021 International Conference on Electronics and Sustainable Communication Systems (ICESC), 2021.

[5] G. H. N. L. L. S. C. C. M. L. W., "A predictive model for student dropout in higher education using machine learning," in 2020 12th International Conference on Technology for Education (T4E), 2020.

[6] "Predicting academic success: A machine learning approach on educational data," in 2018 IEEE 18th International Conference on Advanced Learning Technologies (ICALT), 2018.

[7] P. P. R. S. P. S. S. D. B. S., "Student performance prediction system using machine learning techniques," in 2021 2nd International Conference on Intelligent Engineering and Management (ICIEM), 2021.

[8] D. L. G. G. H. R. J. D. A. R., "A review on attendance management systems based on biometric technologies," in 2022 3rd International Conference on Advancement in Science and Technology (ICAST), 2022.

[9] R. S. S. K. R. M. N. A. V. S. S., "Efficient student attendance management system using a cloud-based approach," in 2019 International Conference on Communication and Signal Processing (ICCSP), 2019.