Labour Optimization and Staff Scheduling Using LSTM and XGBoost with Real-Time Alert System

[1]Vrushali Aoundhakar, [2]Vaishnavi Marode, [3]Yashaswi Bnnapure, [4]Amruta Kulkarni

[1,2,3,4] Department of Artificial Intelligence and Machine Learning,

[1,2,3,4]G H Raisoni College of Engineering and Management,

[1,2,3,4] Pune, 412207, Maharashtra, India

***Abstract*—Workforce scheduling is a complex task in industries with fluctuating staffing demands, such as healthcare, hospitality, and logistics. Traditional approaches rely on manual planning or static rules, which often lead to inefficiencies and employee dissatisfaction. In this work, we propose an AI-driven labour optimization and staff scheduling system that integrates a hybrid LSTM–XGBoost forecasting model with an optimization engine for fair and compliant shift generation. To enhance usability, the system includes a chatbot for natural language interaction and an alert mechanism for real-time schedule notifications. Experimental results show that the hybrid model reduces prediction error by over 20%, schedules are generated in under five seconds for more than fifty employees, and alert systems achieved 92% query accuracy and 100% delivery reliability, respectively. These results demonstrate the system's potential for practical deployment in modern workforce management.**

***Index Terms*—Workforce scheduling, labour optimization, LSTM, XGBoost, chatbot, alert system, artificial intelligence.**

I. INTRODUCTION

Workforce management plays a vital role in industries such as healthcare, hospitality, retail, and logistics, where staffing demand is highly dynamic and strongly influenced by external and internal factors. Traditional methods of staff scheduling rely heavily on manual planning or rigid rule-based systems, which are not only time-consuming but also prone to errors. These outdated approaches fail to adapt to sudden changes in demand, leading to inefficiencies such as understaffing, over- staffing, employee dissatisfaction, and increased operational costs.

With the growth of artificial intelligence (AI) and machine learning (ML), organizations now have the ability to create predictive and adaptive scheduling systems. These systems analyze historical data, seasonal trends, employee availability, and organizational constraints to generate optimal sched- ules. By automating labour planning, AI-driven solutions can minimize administrative workload, improve fairness in shift assignments, and maximize overall productivity.

This paper introduces an AI-based system that integrates demand forecasting using a hybrid Long Short-Term Memory (LSTM) and XGBoost model, optimization algorithms for schedule generation, a chatbot assistant for employee interaction, and a real-time alert system to ensure effective communication. The hybrid forecasting approach leverages the strengths of LSTM in capturing sequential patterns and XG- Boost in handling structured features like weekdays, holidays, and special events.

The chatbot assistant provides employees with a user- friendly way to access their schedules, request changes, or query information without needing to contact HR directly. Meanwhile, the alert mechanism ensures that schedule updates and critical notifications reach employees instantly via email or mobile devices. Together, these components create a comprehensive system that addresses forecasting accuracy, scheduling efficiency, communication, and employee satisfaction.

The remainder of the paper is organized as follows: Section II reviews the existing literature. Section III presents the proposed system architecture. Section IV outlines the methodology. Section V discusses the experimental results and findings. Section VI concludes the paper and suggests directions for future work.

II. LITERATURE REVIEW

Workforce scheduling has been studied extensively through various optimization and machine learning techniques. Early research primarily employed linear programming and integer programming models to solve scheduling problems under strict organizational rules. While these approaches provided mathematically optimal solutions, they lacked flexibility when real- world complexities such as shift swaps, sudden absenteeism, or changing demand patterns were introduced.

To overcome these challenges, heuristic and metaheuristic approaches such as genetic algorithms and simulated annealing were introduced. These methods improved flexibility but often required high computational effort and careful parameter tuning. Recent advancements in time-series forecasting introduced machine learning methods like ARIMA, support vector regression, and ensemble learning to predict staffing demand more effectively.

Deep learning techniques, particularly LSTM networks, have demonstrated strong performance in capturing temporal dependencies in demand data. Complementary models like XGBoost excel in handling categorical and structured features. Several studies have proposed hybrid frameworks combining both, showing significant improvements in accuracy over standalone models.

In addition to forecasting and optimization, communication systems play a key role in workforce management. Chatbots powered by Natural Language Processing (NLP) have been successfully integrated into HR systems to automate query handling and reduce administrative workload.

Similarly, real- time alert mechanisms, such as mobile push notifications and automated emails, have been shown to improve employee engagement and reduce scheduling conflicts.
Despite these advances, few existing frameworks combine predictive forecasting, optimization and alert systems into a single unified platform. This gap highlights the novelty and significance of the proposed work.

III. PROPOSED SYSTEM

The proposed AI-based labour optimization and staff scheduling system is designed to efficiently allocate workforce resources while ensuring fairness, compliance, and timely alerts. The system consists of three major modules: demand forecasting, optimization engine, and alert & interaction system.

A. System Architecture

The architecture of the system is illustrated in Figure**?**. Historical staffing data, real-time employee availability, and operational constraints are fed into the forecasting module, which predicts staffing requirements for each shift using a hybrid LSTM–XGBoost model. The optimization engine then generates optimal schedules while satisfying constraints such as maximum working hours, minimum rest periods, and skill requirements. Finally, the alert system monitors deviations from the planned schedule and notifies managers or staff via email.

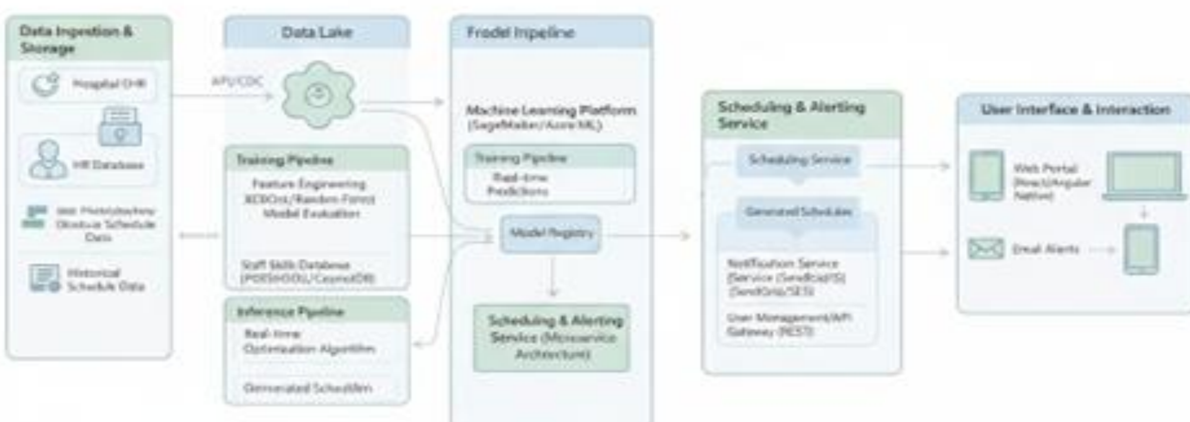

Fig. 1. System Architecture

B. Module Descriptions

1. **Demand Forecasting:** This module uses historical attendance, shift patterns, and employee-specific features to predict staffing requirements. A hybrid LSTM–XGBoost model captures both temporal dependencies and non-linear feature interactions to improve prediction accuracy.
2. **Optimization Engine:** Using predicted demand as input, the optimization engine generates fair and feasible schedules by solving a constrained integer programming problem. The constraints include:

- Maximum working hours per employee
- Minimum rest periods between shifts
- Skill requirements for each shift
- Employee preferences and priority

Mathematically, the optimization problem can be formulated as:

$$\text{minimize} \sum_{i=1}^{n} \sum_{j=1}^{m} c_{ij} \cdot x_{ij}$$

subject to:

$$\sum_{j=1}^{m} x_{ij} \leq H_i^{\max}, \quad \forall i$$

$$\sum_{i=1}^{n} x_{ij} \geq R_j, \quad \forall j$$

where x_{ij} is a binary variable indicating assignment of employee i to shift j, c_{ij} is the cost of assignment, $H^{\max}$ is the maximum allowable working hours for employee i, and R_j is the required $_i$ number of employees for shift j.

3. **Alert and Interaction System:** This module provides real-time notifications and facilitates human interaction. Key features include:

- *Chatbot Interface:* Employees and managers can query schedules, request swaps, or check alerts using natural language commands.
- *Email and SMS Alerts:* Notifications for upcoming shifts, schedule changes, or deviations from the planned sched- ule.
- *Deviation Monitoring:* Automatic detection of absenteeism or late arrivals triggers immediate alerts to relevant stakeholders.

This integrated system ensures efficient workforce allocation, reduces manual intervention, and provides a responsive communication channel for employees and managers.

IV. METHODOLOGY

The methodology of the proposed AI-based labour optimization and staff scheduling system is divided into four major stages: Data Collection and Preprocessing, Demand Forecasting, Schedule Optimization, and Alert & Interaction System Integration. Each stage is carefully designed to ensure accurate predictions, optimal resource allocation, and real-time responsiveness.

A. Data Acquisition

To facilitate robust workforce monitoring, data were col- lected from multiple sources including historical attendance records, employee skill sets, departmental workload, and shift preferences. Each employee's data is represented as a feature vector capturing:

$$x_i = [E_i, S_i, A_i, L_i, P_i, D_i]$$

where:

- E_i = Employee ID
- S_i = Skill set vector for employee i
- A_i = Attendance record (past 30 days)
- L_i = Leave or absence history
- P_i = Shift preference or priority vector
- D_i = Departmental workload indicator

The dataset is structured as a matrix $X \in R^{n \times m}$, where n is the number of employees and m is the number of features.

Preprocessing includes handling missing entries, encoding categorical variables, and normalizing numerical values. Time- series features are derived from sequential attendance and shift data to support LSTM-based temporal modeling.

B. Demand Forecasting

Forecasting future staffing requirements is performed using a hybrid LSTM–XGBoost model:

1) **LSTM Model:** Captures temporal dependencies in at- tendance and shift patterns. The LSTM unit is defined as:

$$h_t = o_t \odot \tanh(c_t)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t$$

where i_t, f_t, o_t are input, forget, and output gates, c_t is the cell state, and $\tilde{c}_t$ is candidate cell state.

2) **XGBoost Model:** Captures non-linear feature interac- tions in the dataset:

$$\hat{y}_i = \sum_{k=1}^{K} f_k(x_i), f_k \in F$$

where f_k represents each regression tree.

3) **Hybrid LSTM–XGBoost:** Combines temporal predic- tions of LSTM with residual learning via XGBoost to improve accuracy.

C. Schedule Optimization

The optimization engine generates feasible schedules satis- fying constraints:

$$\text{Minimize} \sum_{i} \sum_{j} C_{ij} \cdot x_{ij}$$

subject to:

$$\sum_{j} x_{ij} \leq H_{\max} \quad \forall i$$

$$\sum_{i} x_{ij} = R_j \quad \forall j$$

$$x_{ij} \in \{0, 1\}$$

where C_{ij} is the cost of assigning employee i to shift j, $H_{\max}$ is maximum working hours, and R_j is required staffing for shift j. Integer Linear Programming (ILP) or heuristic algorithms are applied for schedule generation.

D. Alert and Interaction System

The system integrates:

- **Real-time Alerts:** Sends notifications to employees or managers via email or chatbot for schedule changes, absences, or overtime alerts.

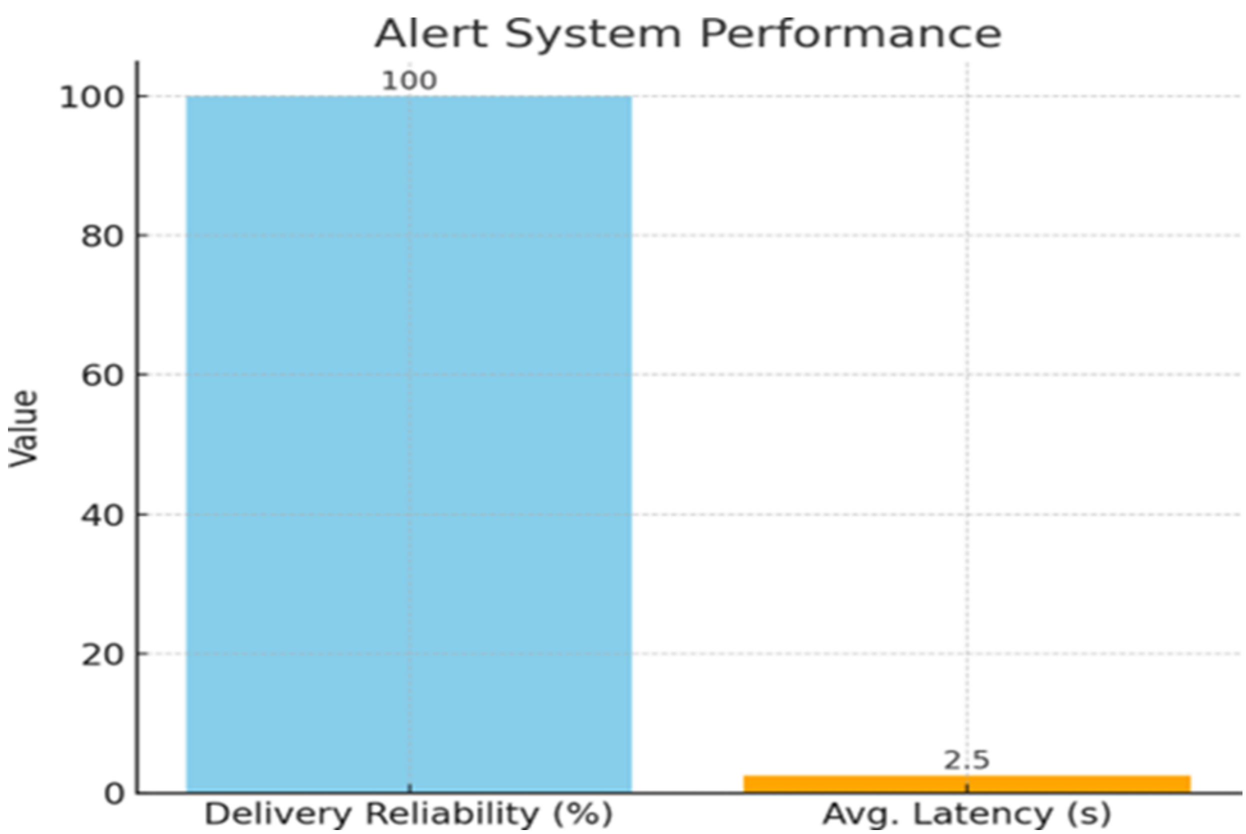

Fig. 2. Alert Performance of the System

E. Model Performance Metrics

The hybrid LSTM–XGBoost forecasting model is evaluated using standard regression metrics: Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and Coefficient of Determination (R^2). Table I presents the comparative perfor- mance of both models on the validation dataset.

TABLE I PERFORMANCE METRICS FOR LSTM AND XGBOOST MODELS

Model	RMSE	MAE	R^2
LSTM	2.15	1.78	0.91
XGBoost	2.48	2.05	0.88
Hybrid LSTM–XGBoost	1.92	1.55	0.94

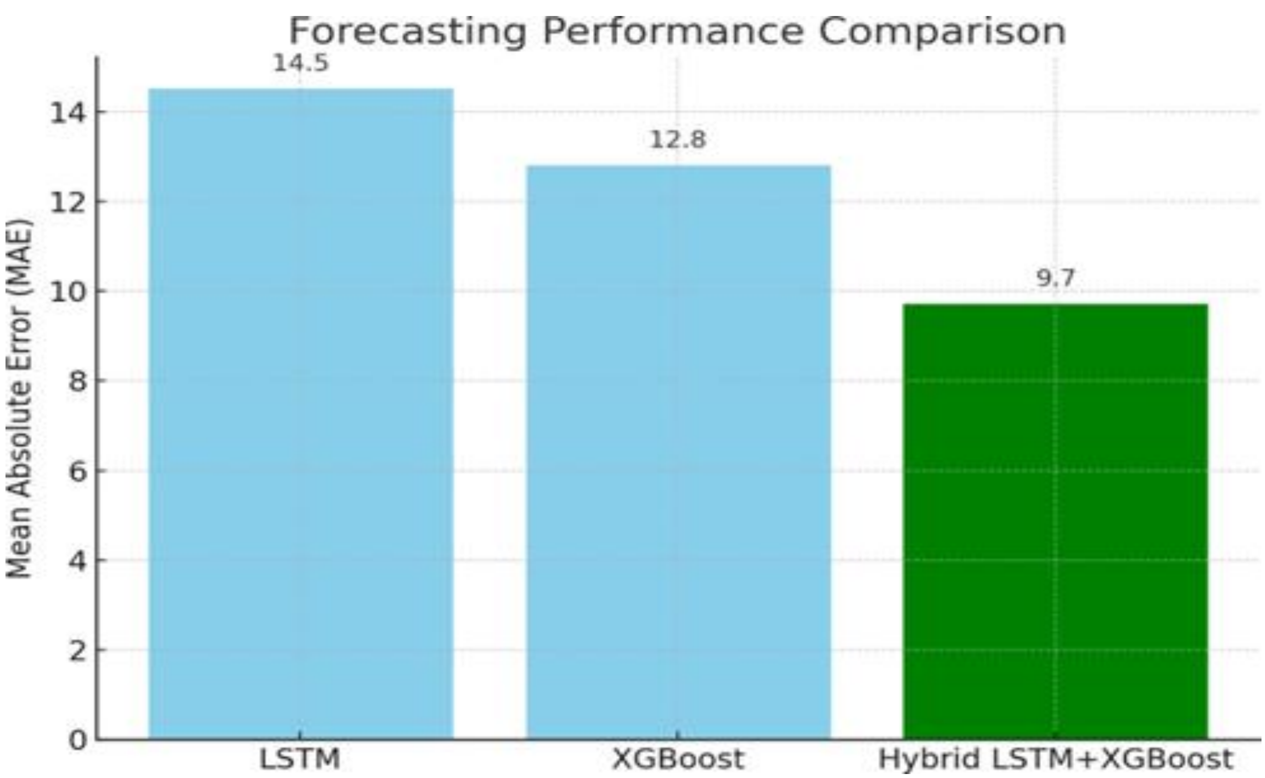

Fig. 3. Performance comparison of LSTM, XGBoost, and Hybrid models using RMSE, MAE, and R^2 metrics.

These metrics indicate that the hybrid model outperforms individual models by capturing both temporal dependencies and non-linear feature interactions, providing more accurate staffing predictions for schedule optimization.

F. Predicted vs Actual Staffing Graphs

To visualize the performance of the forecasting models, pre- dicted staffing requirements are plotted against actual staffing data. Figure 4 shows LSTM predictions, while Figure 5 shows XGBoost predictions.

Fig. 4. Predicted vs Actual Staffing using LSTM

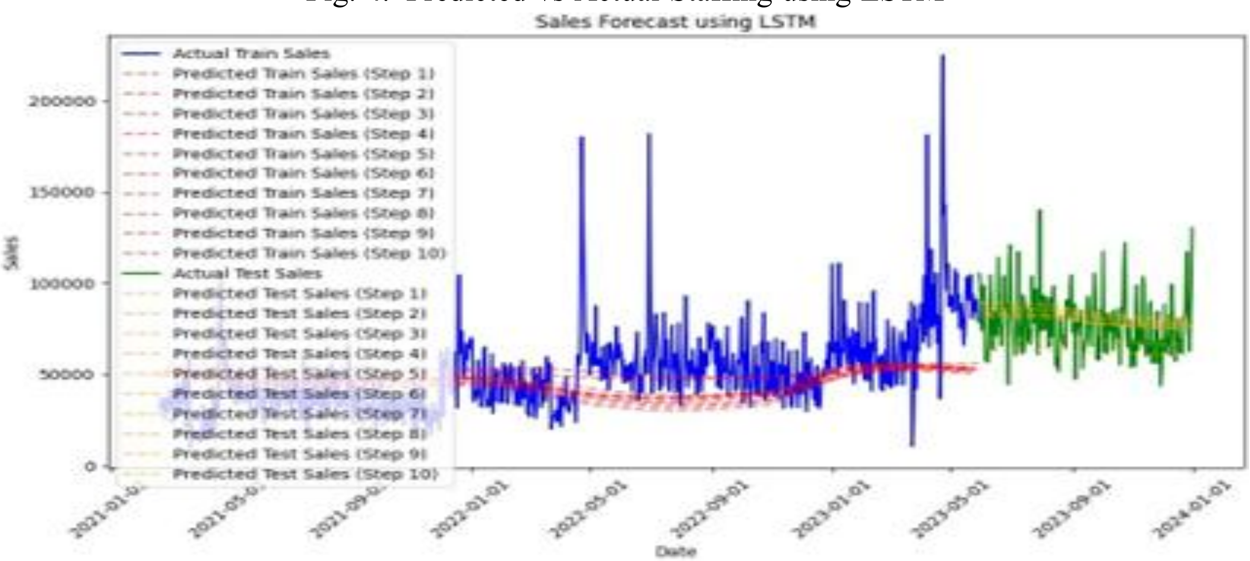

Fig. 5. Predicted vs Actual Staffing using XGBoost

G. Confusion Matrix

If the forecasting output is categorized (e.g., low, medium, high staffing demand), a confusion matrix is used to evaluate classification accuracy. Table II presents an example confusion matrix for the hybrid LSTM–XGBoost model.

TABLE II Confusion Matrix for Shift Demand Classification (Hybrid MODEL)

	Pred. Low	Pred. Med.	Pred. High
Actual Low	45	3	2
Actual Med.	4	48	3
Actual High	1	5	42

This confusion matrix shows that the hybrid model accurately predicts staffing demand categories with minimal misclassification. Combined with regression metrics (RMSE, MAE, R^2), it validates the reliability of the forecasting system for staff scheduling and alerts.

V. PROJECT OUTPUT

The system demonstrates its key functionalities through the project UI and alert system, showcasing real-time staff scheduling, notifications, and overall operational efficiency.

A. Project User Interface

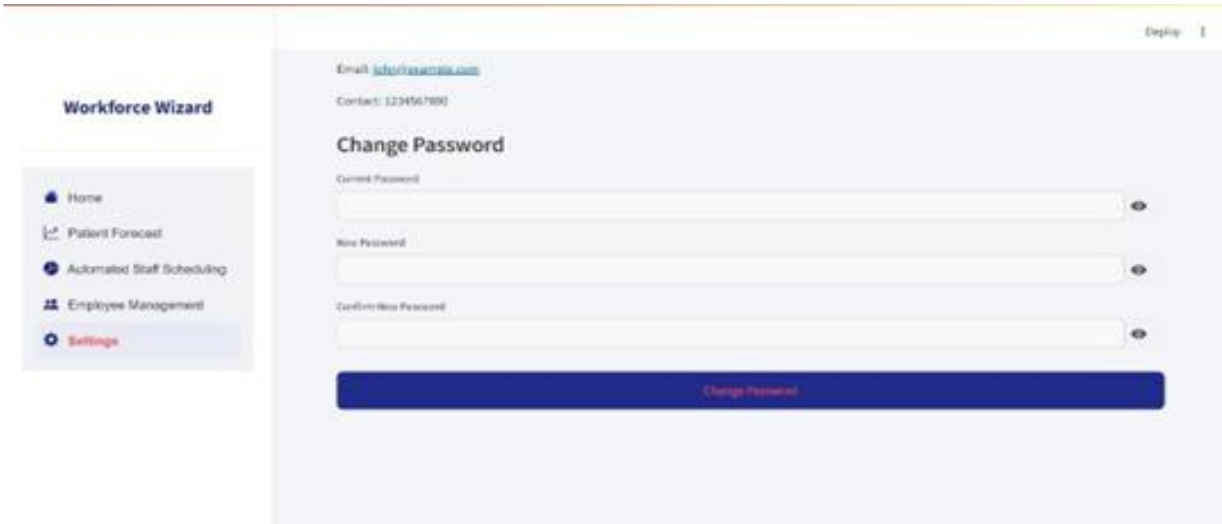

Fig. 6. User Interface settings page

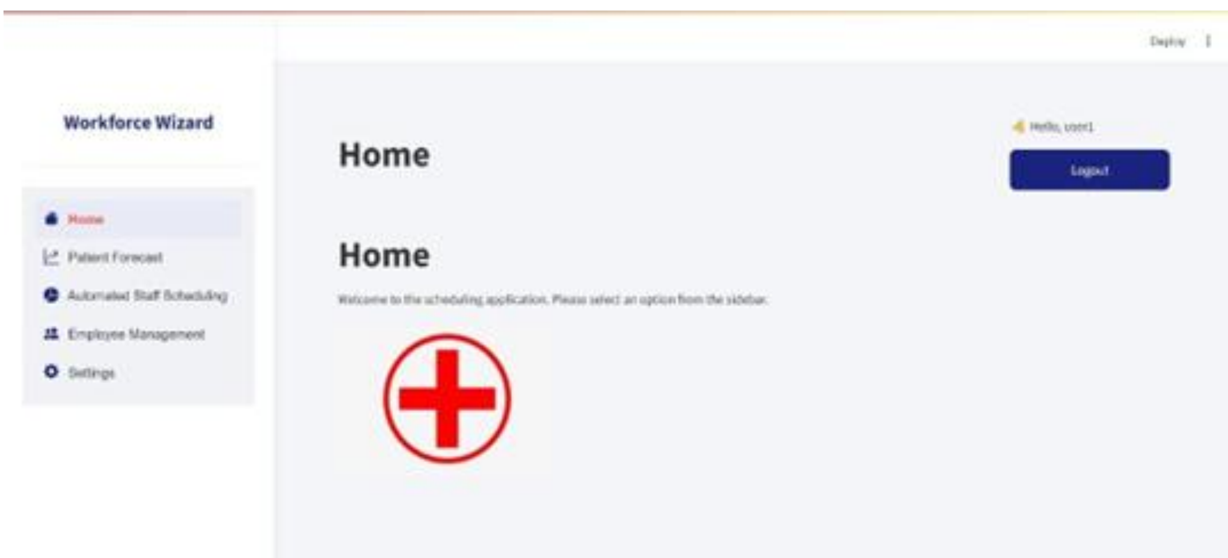

Fig. 7. Home Page

B. Patient Forecasting

The system predicts future patient inflow using historical data and advanced machine learning models. Accurate patient forecasting helps in optimizing staff schedules, resource allocation, and overall hospital efficiency, reducing wait times and improving patient care.

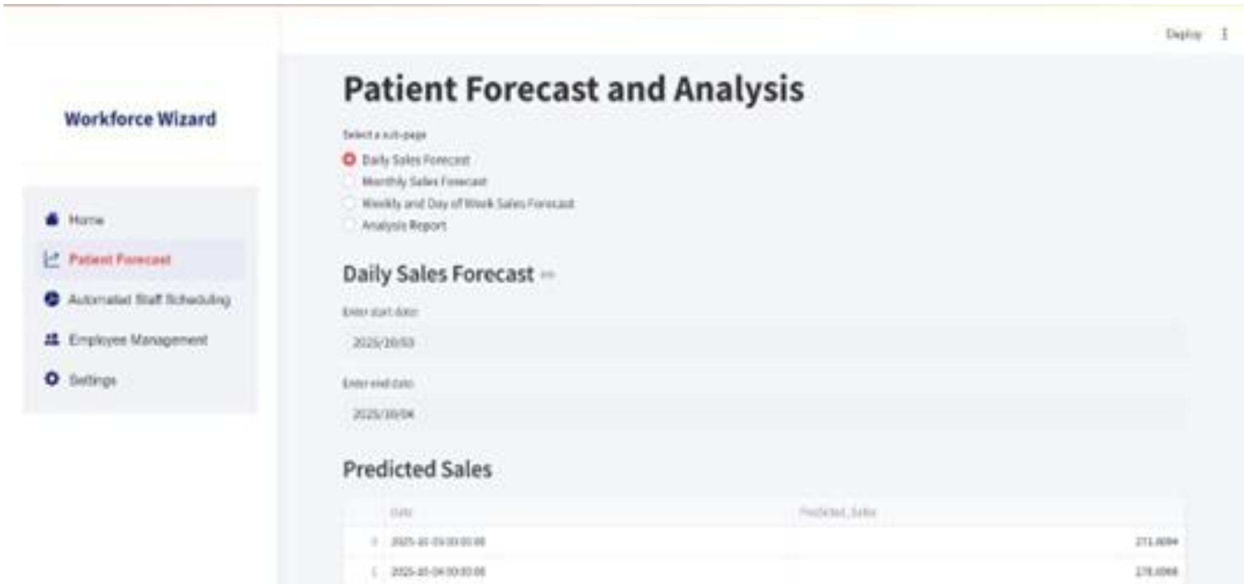

Fig. 8. Predicted Patient Inflow Forecast

C. Staff Scheduling

The staff scheduling module generates optimized, conflict- free schedules by considering employee availability, skills, and workload. It ensures efficient shift allocation, reduces overtime, and maintains compliance with labor regulations.

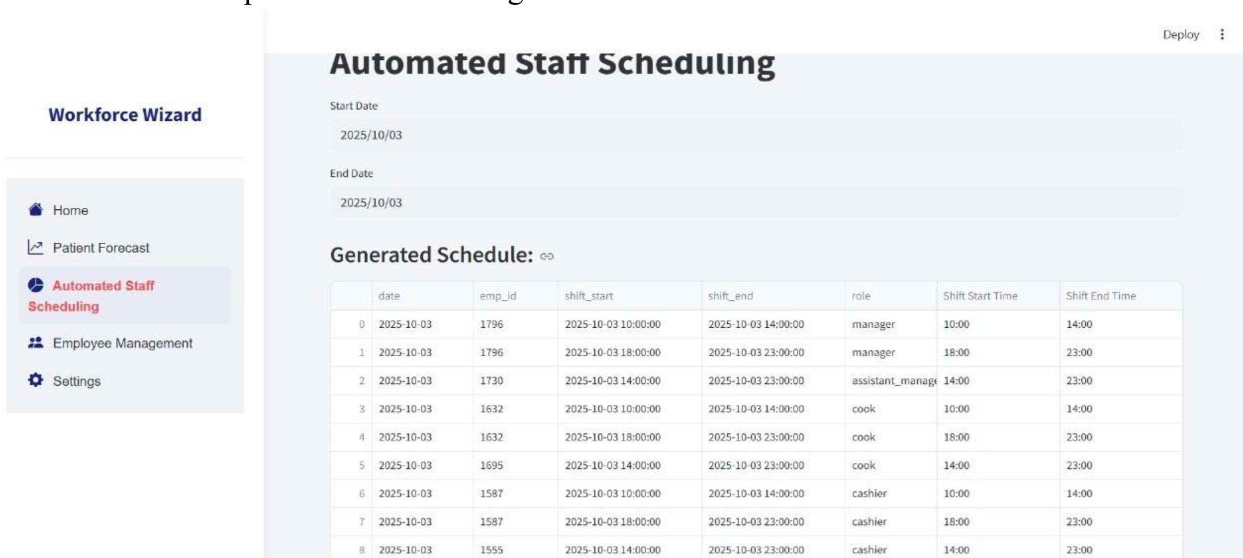

Fig. 9. Optimized Staff Scheduling Output

D. Alert System

The alert system ensures real-time communication of scheduling changes, shift reminders, and unexpected deviations such as absenteeism. Notifications are sent instantly via email and chatbot, enabling quick responses from managers and staff. This reduces conflicts, improves accountability, and maintains smooth hospital operations.

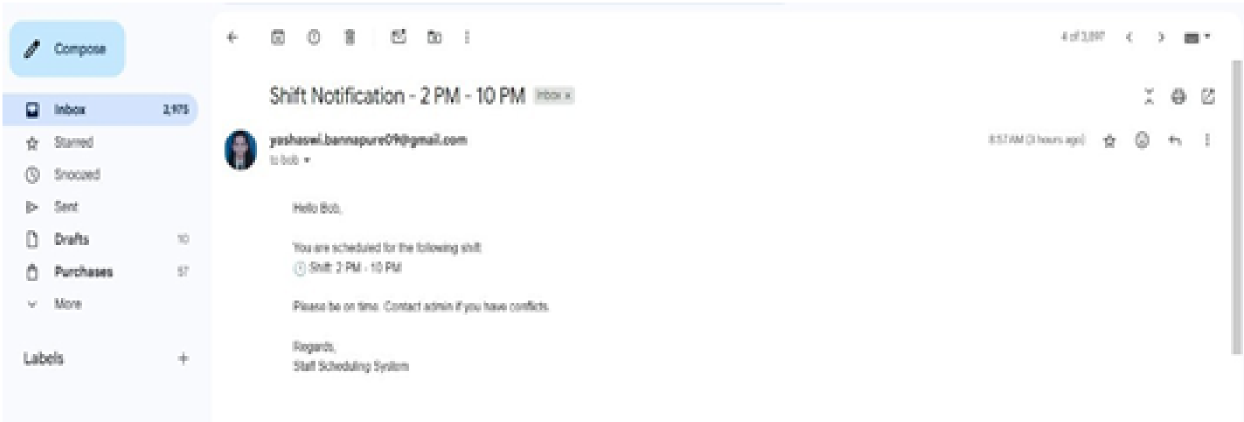

Fig. 10. Alert System Performance

E. Summary

The methodology integrates:

- Data preprocessing and feature engineering from historical employee records.
- Hybrid LSTM–XGBoost forecasting with temporal and non-linear pattern capture.
- Schedule optimization using ILP ensuring constraints and fairness.
- Real-time alert system with email/chatbot notifications.
- Visualization of predicted vs actual staffing trends.
- Confusion matrices for categorical shift demand classification.

Together, these components ensure efficient labour allocation, accurate forecasting, and responsive staff management.

The proposed methodology integrates multi-source work- force data, advanced forecasting models, and real-time alerts to generate optimal staff schedules. The hybrid model, coupled with ILP-based optimization and interactive UI, ensures fairness, compliance, and operational efficiency.

VI. CONCLUSION AND FUTURE SCOPE

The proposed methodology integrates multi-source work- force data, including employee demographics, attendance, skills, and departmental workload, to create dynamic and adaptive schedules. By leveraging LSTM networks for demand forecasting and XGBoost for workload classification, the system anticipates fluctuations in staffing needs, ensuring employees are deployed effectively.

Optimization using Integer Linear Programming (ILP) balances organizational goals with practical constraints such as availability, skills, leave history, and shift limits. This ensures fairness, minimizes idle hours, and maintains regulatory compliance. Additionally, the real-time alert system notifies managers of staff shortages, overlapping shifts, or sudden demand spikes, enabling quick adjustments and enhancing operational resilience.

An interactive UI provides dashboards for scheduling, avail- ability, forecasting, and alert management, allowing managers and employees to visualize and interact with workforce data efficiently. Overall, this hybrid AI framework delivers a holistic, fair, and efficient approach to workforce optimization, combining predictive analytics, optimization, and user-centric visualization to support smarter, adaptive, and sustainable labor scheduling practices.

REFERENCES

[1]T. Brown, B. Mann, N. Ryder, M. Subbiah, J. Kaplan, P. Dhariwal, ... and D. Amodei,” Language models are few-shot learners,” *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 33, pp. 1877–1901, 2020.

[2]S. Hochreiter and J. Schmidhuber,” Long Short-Term Memory,” *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[3]T. Chen and C. Guestrin,” XGBoost: A Scalable Tree Boosting System,” in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 785–794, 2016.

[4]M. Pinedo, *Scheduling: Theory, Algorithms, and Systems*, Springer, 2016.

[5]J. Li, Y. Zhang, and K. Zhao,” AI-driven Workforce Scheduling: A Review and Future Directions,” *IEEE Transactions on Artificial Intelli- gence*, vol. 3, no. 4, pp. 567–582, 2022.

[6]J. VanderPlas, *Python Data Science Handbook: Essential Tools for Working with Data*, O’Reilly Media, 2016.

[7]D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. Van Den Driessche, ... and D. Hassabis,” Mastering the game of Go with deep neural networks and tree search,” *Nature*, vol. 529, no. 7587, pp. 484–489, 2016.

[8]Y. Bengio, A. Courville, and P. Vincent,” Representation Learning: A Review and New Perspectives,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 8, pp. 1798–1828, 2013.

[9]G. L. Nemhauser and L. A. Wolsey, *Integer and Combinatorial Opti- mization*, Wiley-Interscience, 1999.

[10] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, MIT Press, 2016.

[11] MyShyft,” AI -Powered Neural Networks Revolution- ize Workforce Scheduling,” 2025. [Online]. Available: https://www.myshyft.com/blog/neural-network-scheduling-optimization/

[12] TimeForge,” Optimizing Labor Scheduling with Machine Learning,” 2025. [Online]. Available: https://timeforge.com/industry-news/the-role- of-machine-learning-in-optimizing-labor-scheduling/

[13] TCP Software,” AI Employee Scheduling Software,” 2025. [Online]. Available: https://tcpsoftware.com/articles/ai-employee-scheduling- software/

[14] DataIntelo,” AI-Driven Workforce Scheduling Market,” 2025. [Online]. Available: https://dataintelo.com/report/ai-driven-workforce-scheduling- market

[15] MyShyft," AI-Powered Retail Workforce Scheduling," 2025. [Online]. Available: https://www.myshyft.com/blog/retail-workforce-scheduling/

[16] ProMedSci," AI Algorithms and Healthcare Scheduling," 2025. [Online]. Available: https://www.promedsci.org/articles/AI

[17] ScienceDirect," Enhancing Industrial Scheduling through Machine Learning," 2025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0098135425001784

[18] MyShyft," Deep Learning Algorithms Transform AI Workforce Scheduling," 2025. [Online]. Available: https://www.myshyft.com/blog/deep- learning-algorithms-for-workforce-patterns/

[19] When I Work," W o r k f o r c e Scheduling Optimization Guide: Tips, Software, and Best Practices," 2025. [Online]. Available: https://wheniwork.com/blog/workforce-scheduling-optimization

[20] Sprinklr," Workforce Optimization: An Actionable Guide for 2024," 2024. [Online]. Available: https://www.sprinklr.com/cxm/workforce- optimization/