International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

A Functional Data-Driven Approach for Accurate Employee Presence Forecasting

M Tejasri[1], C Venkatesh[2], E Pavithra[3]

[1]P.G Scholar, Department of MCA, Sri Venkatesa Perumal College of Engineering & Technology, Puttur, E-mail: tejasrimelam8@gmail.com , ORC-ID: https://orcid.org/0009-0006-8910-2688

[2] Assistant Professor, Department of CSE(AI & ML), Sri Venkatesa Perumal College of Engineering & Technology, Puttur, E-mail: chevireddyvenkatesh22@gmail.com, ORC-ID: https://orcid.org/0009-0007-1038-5978

[3]Assistant Professor, Department of MCA, Sri Venkatesa Perumal College of Engineering & Technology, Puttur, E-mail: epavithra526@gmail.com, ORC-ID: https://orcid.org/0009-0006-8871-4551

Abstract: In smart workplaces, figuring out when employees will be there is important for making the best use of room and resources. This research shows a smart way to make predictions using a dataset of employee presences that includes information about time and place. After data preprocessing gets rid of null values and duplicates, exploratory analysis and association visualization are used to see patterns in how people behave. Several prediction methods are used, such as AdaBoost, SARIMAX, FTSA, LSTM, DFMM, RobFTS, CNN, and an ensemble Voting Regressor that combines XGBoost, Gradient Boosting, and K-Nearest Neighbors. Functional Data Analysis is used to describe how things depend on time and how occupancy patterns change over time. The DFMM model gets 4.68% MAPE and a R^2 value of 91.3%, which means it is very accurate. On the other hand, the Voting Regressor gets better generalization with a R^2 value of 93.1%. Explainable AI methods, like LIME and SHAP, are used to figure out what forecasts mean and find features that are important. For real-world use, the system is built on the Flask framework, which provides a simple web interface with safe user signup and signin through SQLite, file-based user input, processing on the back end, and interactive visualization. Users upload test files and choose the time frames for the predictions. The system then shows weekly and daily employee presence forecasts, including predictions for the next four or eight weeks. This lets managers make smart choices about how to plan the workplace and use resources most efficiently.

***"Index Terms -** Functional Data Analysis, DFMM, Voting Regressor, Employee Presence Prediction, Explainable AI, LIME, SHAP, Smart Workplace"*

1. INTRODUCTION

Information and Communication Technologies (ICT) and smart technology are having a big impact on modern workplaces, which is changing how things are done and how productive employees are [1, 7]. The idea of intelligent office environments came about because of rising property prices, changing occupancy rates, and the need for long-term control of workspaces [6]. These places use data analytics, sensor networks, and automation systems to keep an eye on the environment and the distribution of workers. This lets companies make choices based on data that improve comfort, resource use, and sustainability [2]. By using predictive analytics and artificial intelligence to look at patterns in employee behavior and involvement, data-driven strategies in Human Resource Management (HRM) have made it even easier for companies to make decisions [3, 4].

Traditional office layouts have changed because of the move toward mixed and flexible working models, especially after the COVID-19 pandemic. This has led to the creation of Activity-Based Workplaces (ABW) and agile work arrangements [8]. These flexible systems give workers freedom, but they make it hard to predict occupancy and run shared spaces efficiently. More and more, machine learning and predictive models are being used in HR analytics to deal with complex workforce dynamics, predict employee turnover, and make engagement strategies work better [5]. Artificial intelligence has also changed important HR tasks like hiring, reviewing performance, and job advancement. It makes sure that decisions made by organizations are fair and reduces bias [4]. Also, workplace analytics now focus on finding behavioral indicators that are linked to job satisfaction, stress levels, and performance outcomes. This makes it easier to control the health and happiness of employees [9].

The main goal is to create and use an advanced machine learning and forecasting framework that is data-driven and smart to look at and predict patterns of employee presence and involvement. Through real-time data insights, this method aims to make the best use of space, improve the accuracy of decisions, and support long-term, flexible, and effective management of the workplace [10].

2. RELATED WORK

Modern workplaces have changed over time thanks to new ways of working and technology advances that make them more flexible, productive, and engaging for employees. New Ways of Working (NWOW) have had a big impact on company structures by encouraging employees to be self-sufficient, share their knowledge, and be flexible in hybrid work settings [11]. Transformational leadership and social contact have been found to be important in supporting intrapreneurial behavior and innovation within NWOW frameworks. This shows how leadership dynamics can make flexible workplace models possible. Changes in this area have also been linked to changes in how motivated and engaged workers are, which means that companies need to use digital tools to keep an eye on and support flexible work schedules [12].

The COVID-19 pandemic sped up the acceptance of flexible and remote work models, which completely changed how companies handle human resources and how they set up their workspaces. Studies have shown that during the pandemic, there were big changes in how engaged, tired, and productive employees felt. For example, public employees reported different levels of job satisfaction based on the availability of digital tools and the level of support they received from managers [12]. Activity-Based Workplaces (ABW) became the most popular type of workspace after the pandemic. They stress flexibility and making good use of space in ways that go beyond standard office layouts [13]. The idea behind ABW makes flexible work arrangements possible, so workers can pick the place where they work based on their personal preferences and the needs of their jobs. But it also makes it hard to control occupancy and guess how many employees will be there, especially in workplaces that change often and have fluctuating attendance rates [14].

As digital office design has changed, people have had to rethink how they organize their time and space at work. Large-scale benchmarking studies, like the ones done by IDET, look at performance metrics and occupancy trends across many companies [15]. These studies look at how data analytics and automation can be used together in workspace management. These studies show how important it is to use data-driven methods to manage the efficiency of workspaces, the amount of energy used, and the health and happiness of employees. The use of advanced data analytics techniques has made it possible to predict occupancy trends and adjust environmental controls like lighting, ventilation, and temperature to match what people are doing in real time [16]. The first time functional autoregressive models were used was for hydrological forecasting. They have since shown promise in modeling complex temporal dependencies in occupancy data, especially when affected by outside environmental factors [16].

Systematic reviews and empirical study have helped build on the progress made in building occupancy prediction and workplace analytics over the past few years. Li et al. [18] did a thorough review that stressed how important it is to combine statistical, functional, and machine learning models for correct occupancy prediction. The study put current models into groups based on the sources of their input data, the level of detail in the time data, and the learning algorithms they used. It showed how hybrid frameworks that combine deep learning with functional data analysis are becoming more useful. By combining these two things, models can show how occupancy behavior and external factors are related in ways that aren't straight lines. Similarly, Motuziene et al. [19] looked at how office occupancy changed during and after the COVID-19 pandemic. They showed how policies for online work and social distance changed occupancy rates and affected how future office buildings were designed. Their results showed how important it is to have adaptive forecast systems that can learn from irregular and sparse datasets that show hybrid attendance trends.

Smart building management has come a long way thanks to machine learning and sensor systems that are built on the Internet of Things (IoT). Pesic et al. [20] created BLEMAT, a system for smart occupancy detection and prediction based on data analytics and machine learning. Using Bluetooth Low Energy (BLE) sensors and machine learning algorithms, the system processes occupancy data in real time, making predictions much more accurate and improving the speed of operations in smart building systems. BLEMAT showed that combining sensor-based inputs with advanced analytics can make building automation systems work better, make people more comfortable, and waste less energy.

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
SN: 3068-272X www.ijdim.com Original Research Paper

3. MATERIALS AND METHODS

A smart predictive framework for estimating employee presence in smart workplaces is introduced by the suggested system. This framework combines temporal, functional, and ensemble learning methods. Several algorithms are used in this method, such as AdaBoost for boosting-based prediction, SARIMAX for time series forecasting with exogenous factors [23], LSTM for sequential dependency modeling, FTSA for functional time series analysis [25], DFMM for dynamic functional mixture modeling, and RobFTS for robust noise-resistant forecasting. Deep learning and ensemble extensions like Convolutional Neural Networks and a Voting Regressor that combines XGBoost, Gradient Boosting, and KNN are used to make the generalization better. The steps in the workflow are cleaning, normalization, feature analysis, training the model, and evaluation with MAE, MAPE, RMSE, R^2, AIC, and BIC. The trained models are put into use through a Flask-based web service that lets users securely sign up and log in, handle input from files, do back-end processing, and see predictions in real time. This unified structure makes things more accurate, reliable, and easy to understand, which lets better systems predict when employees will be present [22].

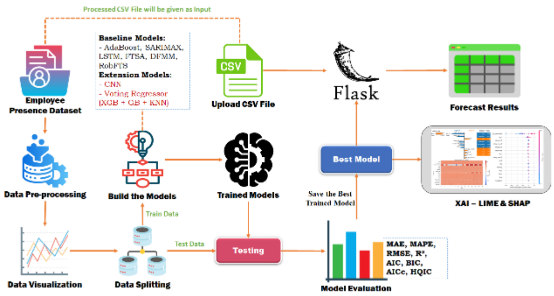


Fig.1 Proposed Architecture

Figure 1 shows a machine learning process for predicting when employees will be present. The process starts with the Employee Presence Dataset and pre-processing the data. Next, the data is visualized and split into training and test sets. The next step is to build and train Baseline and Extension Models. After being trained, the models are tested and evaluated using measures such as MAE and RMSE. The Best Model is saved, added using Flask, and can take a processed CSV file as input to make Forecast Results. These results are then described with XAI—LIME and SHAP.

i) Dataset Collection:

The dataset used in this study is called Employee-Presence-dataset, and it is made up of sensor and environmental data from real workplaces that were gathered from September 2022 to September 2023. It has 250 daily records with information like date, number of people, weather, amount of rain, and total reservations. These records show how occupancy changes when environmental and operational conditions change. Dealing with null values, getting rid of duplicates, and normalizing the data were all parts of data preparation that made sure it was consistent and reliable. Figure 2 shows the structure of the dataset and how key factors that affect employee presence are related to each other. These variables help with time series and functional analysis for forecasting models like FTSA and DFMM, which are good at showing how workplace occupancy rates change over time [24]. This large collection makes it possible to build strong predictive models for smart workplace analytics.

	Date	GLOBAL	D1	D2	D3	D4	jour_feriÃ.	pont.congÃ.	holiday	jour_semaine	Semaine	Annee	Annee_et_Semaine	Temp	pluie
0	05/09/2022	240	0	172	55	13	0	0	0	lundi	36	2022	2022 Sem 36	20	3.1
1	06/09/2022	542	4	336	169	33	0	0	0	mardi	36	2022	2022 Sem 36	17	0.4
2	07/09/2022	657	1	399	217	40	0	0	0	mercredi	36	2022	2022 Sem 36	17	0.5
3	08/09/2022	607	2	377	180	48	0	0	0	jeudi	36	2022	2022 Sem 36	15	0.4
4	09/09/2022	461	5	300	120	36	0	0	0	vendredi	36	2022	2022 Sem 36	15	0.4
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
245	25/09/2023	624	6	325	229	64	0	0	0	lundi	39	2023	2023 Sem 39	14	0.0
246	26/09/2023	713	6	374	259	74	0	0	0	mardi	39	2023	2023 Sem 39	16	0.0
247	27/09/2023	623	6	335	231	51	0	0	0	mercredi	39	2023	2023 Sem 39	17	0.0
248	28/09/2023	659	6	354	232	67	0	0	0	jeudi	39	2023	2023 Sem 39	16	0.0
249	29/09/2023	408	6	239	126	37	0	0	0	vendredi	39	2023	2023 Sem 39	19	0.1

250 rows × 20 columns

Fig.2 Dataset Collection

ii) Pre-processing:

Before training a model, data preprocessing is necessary to make sure that the information is reliable and of good quality. It includes removing missing values, duplicates, and inconsistencies from data and cleaning, transforming, and organizing it so that accurate feature extraction and fast performance of predictive algorithms are possible.

***a) Data Preprocessing*:** Systematic preprocessing is done on the information to make it more accurate and consistent so that it can be used for predictive analysis. Null value removal gets rid of records that are missing information that could mess up model learning. Duplicate removal makes sure that each record is unique and stops data from being duplicated. All the numbers are reset to make sure the right order is kept, and categorical variables are found to be encoded or transformed. Data normalization brings all the features of a set of data to the same scale, which improves the speed and convergence of an algorithm. This process makes sure that the training dataset is clean, balanced, and ready for correlation analysis and display. This sets

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
SN: 3068-272X www.ijdim.com Original Research Paper

the stage for building strong time series and functional forecasting models.

***b) Data Visualization*:** Data visualization helps us understand trends in employee participation and how different factors in the workplace are connected. Visualizations of the number of people who come every day on a weekly basis show changes in attendance and trends of occupancy over time, which helps us understand how behavior changes over time. It is possible to find dependencies between variables like temperature, rainfall, and total reservations by plotting a correlation matrix. This shows how these factors affect workplace presence. These visual details help choose the features and improve the model, which makes the predicted results easier to understand. Visualization helps with experimental analysis and making smart decisions for intelligent workplace management systems by connecting the steps of prepping data and building models.

iii) Train & Test:

The dataset is split into training and testing groups in an 80:20 split to make sure the model learns well and is evaluated fairly. Eighty percent of the data is used as the training set to find patterns in employee presence linked to time and behavior. The other twenty percent is used as the testing set to check for generalization and predictive reliability. This split helps make sure that seasonal, environmental, and operational changes are shown fairly in both groups. By keeping the flow of time, the method reduces data leaks and makes sure that the predictive framework accurately models real-life situations in smart workplaces that are always changing.

iv) Algorithms:

AdaBoost: AdaBoost is an ensemble learning method that builds a strong predictor by combining several weak learners, which are usually decision trees, over and over again. By giving instances that were wrongly classified more weight, it lets later models focus on tough samples, which improves accuracy and stability. AdaBoost does a good job of capturing changes in behavior and improving the accuracy of occupancy forecasting when it comes to predicting staff presence. Because it is adaptive, it can handle relatively noisy data while still being easy to understand. The algorithm's boosting process makes predictions more accurate by using the group performance of weak learners. This creates a more accurate and fair model for prediction tasks in the real world [21].

The equation below defines AdaBoost's final strong classification model.

$$H(x) = sign\left(\sum_{t=1}^{T} \propto_t h_t(x)\right) \quad (1)$$

SARIMAX: Seasonal AutoRegressive Integrated Moving Average with eXogenous factors (SARIMAX) builds on ARIMA by adding external variables to time series predictions and modeling seasonal effects. It does a good job of capturing seasonal, trend, and cyclical patterns while also taking into account outside factors like temperature or daily effects. In predicting an employee's present, SARIMAX finds temporal dependencies and periodic behaviors in attendance data. This gives a strong statistical base for comparing models in the future. This method makes things easier to understand by taking trends and outside factors into account. It gives us useful information about how employees change over time. SARIMAX works best with organized, time-sensitive data from the workplace that changes regularly.

FTSA: Functional Time Series Analysis (FTSA) looks at time series data as smooth functional observations instead of discrete points. This lets you analyze underlying trends and temporal structures. Taking into account smooth changes and connections over time, it gives a full picture of how dynamic processes work. In employee presence prediction, FTSA models habits like continuous attendance and accurately predicts how people will be occupied in the future. It makes the data easier to understand and finds small trends that are missed in discrete models by showing the data as functions. FTSA gives us a mathematically rich way to deal with structured data that changes over time. This lets us make accurate predictions about the future even in complicated work settings.

LSTM: A recurrent neural network called Long Short-Term Memory (LSTM) is made to find long-range relationships in sequential data. It successfully deals with vanishing gradient problems through its input, forget, and output gates, which allow learning over long temporal sequences. LSTM models temporal dependencies and adapts to changing workplace behavior trends to predict when an employee will be present. It can deal with variable and dynamic changes in attendance, making it very accurate and flexible. Because LSTM can remember context over time, it makes predictions more accurate and is an important part of deep learning for looking at sequential workplace data [26].

The equation below represents the hidden state update in LSTM.

$$h_t = \sigma(w_o \cdot [h_{t-1}, x_t] + b_o) \cdot \tanh(C_t) \quad (2)$$

DFMM: The Dynamic Factor Mixture Model (DFMM) takes into account both factor analysis and mixture modeling to find hidden structures and temporal relationships in time series data. It lowers the number of dimensions while dynamically finding hidden factors that affect results. In employee present prediction, DFMM finds the underlying behavioral and environmental factors that cause differences in attendance. This mixed statistical method handles doubt well and can adapt to changing conditions over time. DFMM makes generalization better by adding hidden dynamic features. This lets models of complex temporal behaviors be more flexible. It gives us a strong probabilistic way to predict how workplace usage will change over time [27].

RobFTS: Robust Functional Time Series (RobFTS) builds on standard FTSA by making it more resistant to data outliers and other problems. It keeps working well even when there is noise and nonstationarity, which makes it very useful for making predictions in the real world. When predicting an employee's appearance, RobFTS deals with changes that are hard to predict because of outside or behavioral issues. It gives smooth and understandable predictions of time even when the data is noisy. The method uses both functional regression and strong statistical estimation to make sure that the results are consistent and reliable. RobFTS improves the stability and accuracy of predictions, which helps make forecasting models in workplace analytics more reliable and easy to understand [28].

CNN: A deep learning model called Convolutional Neural Network (CNN) can easily pull out hierarchical patterns from structured data. CNN learns complex temporal and spatial relationships without having to do feature engineering by hand. It does this through convolutional and pooling layers. CNN finds complex, nonlinear links between external factors and occupancy trends that help them guess how many employees will be present. Its hierarchical feature extraction makes model adaptation better, which lets it make accurate predictions even with datasets that aren't all the same. CNN improves the accuracy of predictions by recognizing both local and global dependencies in the input sequence. This provides a scalable method for smart modeling of office occupancy [29].

The following equation represents the core operation in CNNs.

$$S(i,j) = \sum_m \sum_n I(i+m, j+n) \cdot K(m,n) \quad (3)$$

Voting Regressor: The Voting Regressor is an ensemble method that takes results from several base models, like XGBoost, Gradient Boosting, and K-Nearest Neighbors, and adds them all together to make a single output. It works by weighting or averaging model results to make them more accurate and less variable. The Voting Regressor uses the strengths of different models to predict employee attendance. It does this by balancing the trade-offs between bias and variance to make the results more reliable. This group method makes sure that the results are the same and stable across different types of data. This makes it a good choice for workplace prediction systems that need to work in a changing environment and need to be easy to understand [30].

The equation below represents the majority voting process in classifiers.

$$\hat{y} = argmax_c \left(\sum_{i=1}^{n} II\ (\hat{y}_i = c) \right) \quad (4)$$

v) Integration of XAI & Flask:

Explainable Artificial Intelligence (XAI) makes models more clear by showing people how predictions are made, which helps them understand what each feature does. Techniques like LIME and SHAP help understand model choices by pointing out important factors that affect outcomes. This makes predictive systems more trustworthy, accountable, and fair. In employee presence prediction, XAI lets stakeholders see and understand how models behave, which helps managers make choices based on facts and data.

The Flask framework lets users connect with predictive models in real time. It is a lightweight and efficient web application platform. It lets machine learning backends work with simple user interfaces that let people enter data, make model predictions, and see results right away. Flask makes sure that clever workplace analytics systems can be used by many people and can be scaled up as needed.

4. RESULTS & DISCUSSIONS

MAE: Absolute Error is the amount of mistake when you measure something. It's the difference between what was recorded and what was "true." To

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
SN: 3068-272X www.ijdim.com Original Research Paper

give you an example, if the scale says you weigh 90 pounds but you know you really weigh 89 pounds, the scale is off by 1 pound.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \quad (5)$$

MAPE: A performance indicator called Mean Absolute Percentage Error (MAPE) is used to find out how accurate a model is at making predictions or forecasts. It gives prediction errors as a percentage, which makes it easy to understand the mean difference between what was expected and what actually happened. A model is more accurate if its MAPE number is lower.

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{A_i - P_i}{A_i}\right| \times 100 \quad (6)$$

RMSE: Root mean square error (RMSE) is a way to find the average difference between what a statistical model said would happen and what actually happened. It is the standard deviation of the residuals in math terms. The residuals show how far away the regression line is from the data points.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}\left||y(i) - \hat{y}(i)|\right|^2}{N}} \quad (7)$$

R2 Score: To find the sum squared regression, add up all the residuals. To find the total sum of squares, add up all the squares that show how far the data is from the mean.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2} \quad (8)$$

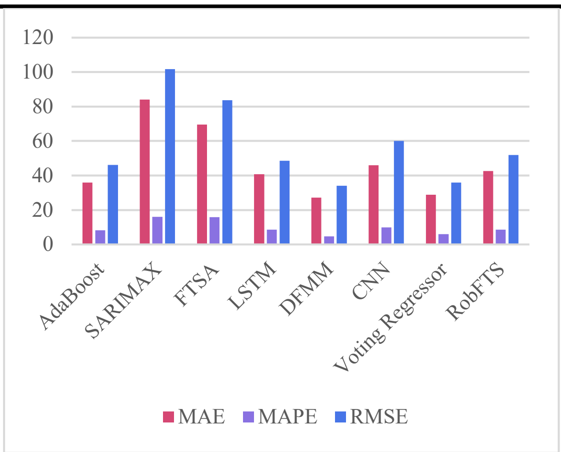
Table.1 Performance Evaluation

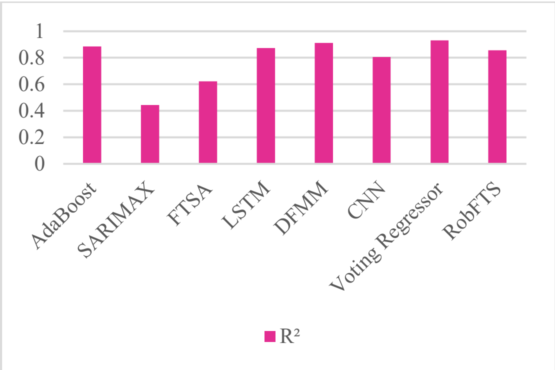
Model	**MAE**	**MAPE**	**RMSE**	**R²**	**AIC**	**BIC**	**AICc**	**HQIC**
AdaBoost	36.005000	8.271	46.131	0.886	546.541	576.810	563.541	558.025
SARIMAX	84.021000	15.955	101.755	0.444	634.067	673.796	668.290	649.140
FTSA	69.583000	15.764	83.743	0.623	614.976	654.704	649.198	630.049
LSTM	40.731000	8.538	48.484	0.874	521.417	523.308	521.502	522.134
DFMM	27.129000	4.683	33.983	0.913	1082905.240	1742718.072	NaN	1265909.278
CNN	45.941002	9.804	60.034	0.806	276254.356	537055.027	NaN	375201.767
Voting Regressor	**28.891000**	**6.056**	**35.924**	**0.931**	**492.034**	**493.925**	**492.119**	**492.751**
RobFTS	42.540000	8.671	51.995	0.857	559.528	600.695	600.008	575.085

Model performance measures are shown in Table 1. The Voting Regressor had the best accuracy with the lowest MAE, MAPE, and RMSE.

Fig.3 Comparison Graph

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
SN: 3068-272X www.ijdim.com Original Research Paper





There are four model performance measures shown in Figure 3. They are MAE (red), MAPE (purple),RMSE (blue), and R2 (pink).

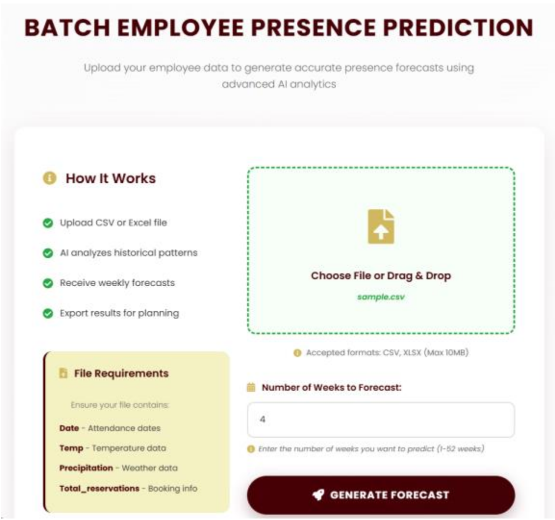
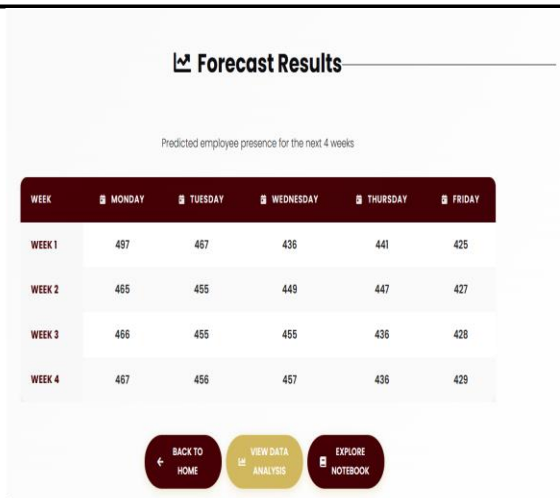


Fig.4 Upload Input Data

The Batch Employee Presence Prediction interface is shown in Figure 4. This is where users add a data file and choose how many weeks they want the forecast to cover.



WEEK	MONDAY	TUESDAY	WEDNESDAY	THURSDAY	FRIDAY
WEEK 1	497	467	436	441	425
WEEK 2	465	455	449	447	427
WEEK 3	466	455	455	436	428
WEEK 4	467	456	457	436	429

Fig.5 Predict Result

Figure 5 shows the Forecast Results table, which shows how many employees are expected to be present every day for the next four weeks.

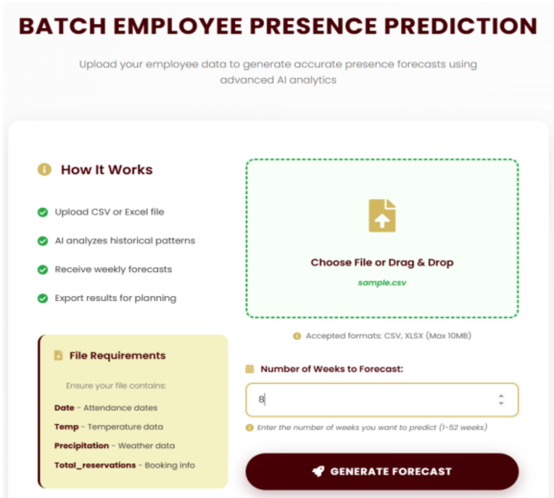
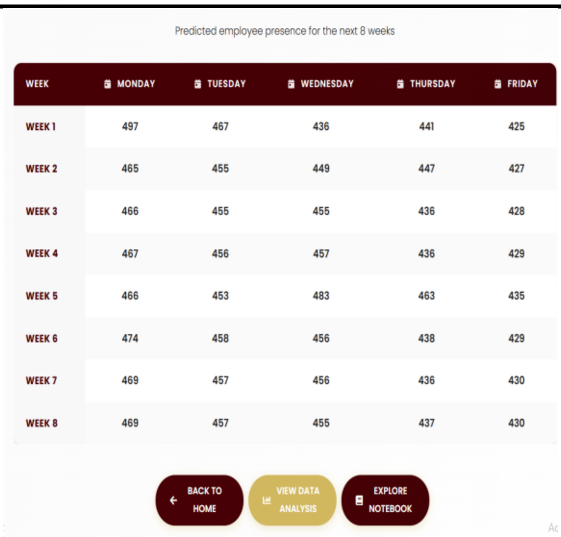


Fig.6 Upload another Input

Figure 6 shows the Batch Employee Presence Prediction interface, with the ability to add files and choose the length of the forecast.

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
SN: 3068-272X www.ijdim.com Original Research Paper


Predicted employee presence for the next 8 weeks

WEEK	MONDAY	TUESDAY	WEDNESDAY	THURSDAY	FRIDAY
WEEK 1	497	467	436	441	425
WEEK 2	465	455	449	447	427
WEEK 3	466	455	455	436	428
WEEK 4	467	456	457	436	429
WEEK 5	466	453	483	463	435
WEEK 6	474	458	456	438	429
WEEK 7	469	457	456	436	430
WEEK 8	469	457	455	437	430

Fig.7 Final Outcome

In Fig. 7, the expected daily appearance of employees is shown as a table for eight weeks.

5. CONCLUSION

The study successfully shows an advanced predictive method for figuring out how many employees are in smart workplaces by using functional data analysis and a mix of machine learning and functional data analysis. The modeling of temporal and behavioral occupancy dynamics was made possible by thorough preprocessing, exploratory analysis, and feature assessment. The DFMM model was the most accurate of the ones that were tested. It had a 4.68% MAPE and a R^2 value of 91.3%, which showed that it could reliably predict complex time-dependent trends. Using XGBoost, Gradient Boosting, and K-Nearest Neighbors together in the ensemble Voting Regressor made it even more stable, with a R^2 score of 93.1% showing better generalization in a variety of situations. Explainable AI techniques, like LIME and SHAP, improved openness by drawing attention to important factors that affect the results of predictions. The Flask platform was used to make the framework usable in the real world. It provides a simple web interface with safe signup and signin for users, file-based input handling, backend preprocessing, and interactive visualization. Users share test files and choose the time frames for the predictions. The system then shows daily details of employee presence predictions for four or eight weeks. Overall, the answer helps with smart decision-making, data-driven space planning, and operational efficiency in today's smart workplaces around the world.

To make predictions more accurate, future study can look into adding more contextual and behavioral factors, like employee schedules, department roles, and environmental conditions. Using real-time data from IoT sensors and spatial analytics could make it possible to model occupancy trends in a way that is always changing. Adding more deep temporal models, such as Transformers or Graph Neural Networks, to the framework might help it understand more complicated relationships between time and place. Creating ways for flexible learning to keep adding new data to model parameters will make long-term robustness better. Adding causal inference methods to Explainable AI could also help us understand how features affect each other better, giving us clearer and more useful information for smart office management and eco-friendly space use.

REFERENCES

[1] Yuan, S., Kroon, B., & Kramer, A. (2024). Building prediction models with grouped data: A case study on the prediction of turnover intention. *Human Resource Management Journal*, *34*(1), 20-38.

[2] Alabi, O. A., Ajayi, F. A., Udeh, C. A., & Efunniyi, C. P. (2024). Data-driven employee engagement: A pathway to superior customer service. *World Journal of Advanced Research and Reviews*, *23*(3), 923-933.

[3] Okon, R. I. C. H. A. R. D., Odionu, C. S., & Bristol-Alagbariya, B. E. R. N. A. D. E. T. T. E. (2024). Integrating data-driven analytics into human resource management to improve decision-making and organizational effectiveness. *IRE Journals*, *8*(6), 574.

[4] Ghedabna, L., Ghedabna, R., Imtiaz, Q., Faheem, M. A., Alkhayyat, A., & Hosen, M. S. (2024). Artificial intelligence in human resource management: revolutionizing recruitment, performance, and employee development. *Nanotechnology Perceptions*, *20*, 52-68.

[5] Adeusi, K. B., Amajuoyi, P., & Benjami, L. B. (2024). Utilizing machine learning to predict employee turnover in high-stress sectors. *International Journal of Management & Entrepreneurship Research*, *6*(5), 1702-1732.

[6] H. H. Nijp, D. G. J. Beckers, K. van de Voorde, S. A. E. Geurts, and M. A. J. Kompier, ''Effects of new ways of working on work hours and work location, health and job-related outcomes,''

Chronobiol. Int., vol. 33, no. 6, pp. 604–618, Jul. 2016.

[7] K. Renard, F. Cornu, Y. Emery, and D. Giauque, ‘‘The impact of new ways of working on organizations and employees: A systematic review of literature,’’ Administ. Sci., vol. 11, no. 2, p. 38, Apr. 2021.

[8] L. L. T. Brummelhuis, A. B. Bakker, J. Hetland, and L. Keulemans, ‘‘Do new ways of working foster work engagement?’’ Psicothema, vol. 24, no. 1, pp. 113–120, Feb. 2012.

[9] D. Knights and C. A. Clarke, ‘‘It's a bitter sweet symphony, this life: Fragile academic selves and insecure identities at work,’’ Org. Stud., vol. 35, no. 3, pp. 335–357, Mar. 2014.

[10] Singh, M., Tiwari, S. K., Swapna, G., Verma, K., Prasad, V., Patidar, V., Sharma, D. & Mewada, H. (2023). A Drug-Target Interaction Prediction Based on Supervised Probabilistic Classification. Journal of Computer Science, 19(10), 1203-1211. https://doi.org/10.3844/jcssp.2023.1203.1211

[11] R. Gerards, S. van Wetten, and C. van Sambeek, ‘‘New ways of working and intrapreneurial behaviour: The mediating role of transformational leadership and social interaction,’’ Rev. Managerial Sci., vol. 15, no. 7, pp. 2075–2110, Oct. 2021.

[12] D. Giauque, K. Renard, F. Cornu, and Y. Emery, ‘‘Engagement, exhaustion, and perceived performance of public employees before and during the COVID-19crisis,’’ Public Personnel Manage., vol. 51, no. 3, pp. 263–290, Sep. 2022.

[13] A. Badungodage, ‘‘New ways of working (NWOW) beyond COVID-19; with special reference to the activity based workplace (ABW),’’ Univ. Canberra, Canberra, ACT, Australia, Tech. Rep., Jun. 2021.

[14] X. Baron, ‘‘Repenser l'espace et le temps du travail intellectuel,’’ L'Expansion Manage. Rev., vol. 142, no. 3, pp. 100–108, Sep. 2011.

[15] (2022). Benchmarks De L'IDET. [Online]. Available: https://www.idet. fr/sinformer/benchmarks/

[16] Gudditti, V., & Krishna, P. V. (2021). Adaptive Light Weight Encryption Algorithm for Securing Multi-Cloud Storage. Turkish Journal of Computer and Mathematics Education (TURCOMAT), 12(9), 545–554.

[17] (2020). PRISMA 2020 Flow Diagram. [Online]. Available: https://www. prisma-statement.org/prisma-2020-flow-diagram

[18] T. Li, X. Liu, G. Li, X. Wang, J. Ma, C. Xu, and Q. Mao, ‘‘A systematic review and comprehensive analysis of building occupancy prediction,’’ Renew. Sustain. Energy Rev., vol. 193, Apr. 2024, Art. no. 114284.

[19] V. Motuziene, J. Bielskus, V. Lapinskiene, G. Rynkun, and J. Bernatavi ciene, ‘‘Office buildings occupancy analysis and prediction associated with the impact of the COVID-19 pandemic,’’ Sustain. Cities Soc., vol. 77, Feb. 2022, Art. no. 103557.

[20] S. Pesic, M. Tosic, O. Ikovi, M. Radovanovic, M. Ivanovic, and D. Boscovic, ‘‘BLEMAT: Data analytics and machine learning for smart building occupancy detection and prediction,’’ Int. J. Artif. Intell. Tools, vol. 28, no. 6, Sep. 2019, Art. no. 1960005.

[21] Dr, K, Pushpa Latha., Mr, M, N, Mallikarjuna Reddy., Dr, B, Rajalingam., Malleswari Akurati., Dr, G, Swapna., Bakkala Santha Kumar., (2026). Blockchain-Enabled Trade Finance Framework for Secure Drug Supply Chain Transactions., International Journal of Drug Delivery Technology, 16(3s), 884-889.

[22] J. O. Ramsay and B. W. Silverman, Functional Data Analysis (Springer Series in Statistics). New York, NY, USA: Springer, 2005.

[23] J. Damon and S. Guillas, ‘‘The inclusion of exogenous variables in functional autoregressive ozone forecasting,’’ Environmetrics, vol. 13, no. 7, pp. 759–774, Nov. 2002.

[24] J. O. Ramsay and C. J. Dalzell, ‘‘Some tools for functional data analysis,’’ J. Roy. Stat. Soc. Ser. B, Stat. Methodol., vol. 53, no. 3, pp. 539–561, Jul. 1991.

[25] Kumar, C. S., Sirisati, R. S., Gudditti, V., Rao, K. S., & Challa, R. K. (2022, December). A smart recommendation system for medicine using intelligent NLP techniques. In 2022 International Conference on Automation, Computing and Renewable Systems (ICACRS) (pp. 1081-1084). IEEE. https://doi.org/10.1109/ICACRS55517.2022.10029078

[26] H. Hullait, D. S. Leslie, N. G. Pavlidis, and S. King, ‘‘Robust function-on function regression,’’ Technometrics, vol. 63, no. 3, pp. 396–409, Jul. 2021.

[27] H. Shi and J. Cao, ‘‘Robust functional principal component analysis based on a new regression framework,’’ J. Agricult., Biol. Environ. Statist., vol. 27, no. 3, pp. 523–543, Sep. 2022.

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
SN: 3068-272X www.ijdim.com Original Research Paper

[28] G. Boente and M. Salibián-Barrera, ‘‘Robust functional principal compo nents for sparse longitudinal data,’’ METRON, vol. 79, no. 2, pp. 159–188, Aug. 2021.
[29] D. Rivera-García, L. A. García-Escudero, A. Mayo-Iscar, and J. Ortega, ‘‘Robust clustering for functional data based on trimming and constraints,’’ Adv. Data Anal. Classification, vol. 13, no. 1, pp. 201–225, Mar. 2019.
[30] J. L. Wang, J. M. Chiou, and H. G. M´ ’uller, ‘‘Functional data analysis,’’ Annu. Rev. Statist., vol. 3, no. 3, pp. 257–295, 2016.