GenWorkBalanceNet: A Hybrid Generative AI-Deep Learning Model for Work-Life Balance-Driven Productivity Prediction in IT Workforces

Chaya J. Swamy
Department of Management Studies, JSS Academy of Technical Education, Bengaluru, India
chaya.s198402@gmail.com (corresponding author)

Aruna Adarsh
JSS Center for Management Studies, JSS Science and Technology University, Mysuru, India
aruna900@gmail.com

Received: 25 July 2025 | Revised: 15 August 2025, 29 August 2025, and 7 September 2025 | Accepted: 9 September 2025

ABSTRACT

This paper presents GenWorkBalanceNet, a hybrid Generative Artificial Intelligence (AI)-Deep Learning (DL) framework designed for predictive modeling of workforce productivity under diverse Work-Life Balance (WLB) policies. The framework integrates three core components: (i) a transformer-based Generative AI module for synthetic scenario generation, (ii) an attention-enhanced Long Short-Term Memory (LSTM) network for forecasting temporal productivity trends, and (iii) a Shapley Additive Explanations (SHAP)-based interpretability layer for interpretable analysis of key productivity determinants. By augmenting historical Human Resource (HR) records and productivity logs with synthetic policy-driven data, GenWorkBalanceNet addresses critical challenges such as data sparsity, limited policy simulation capacity, and the lack of transparency in conventional analytics approaches. Experimental evaluation on an Information Technology (IT) workforce dataset demonstrates that the proposed model outperforms established baselines, such as Random Forest, Gradient Boosting, and standalone LSTM, achieving a 19-28% reduction in Root Mean Square Error (RMSE), a 27-36% reduction in Mean Absolute Error (MAE), and a 2-7% improvement in R^2 score. Scenario-based simulations further reveal that flexible work-hour policies can enhance overall productivity by up to 6-8%, emphasizing the framework's potential as a decision-support tool for HR managers. Overall, GenWorkBalanceNet offers a scalable, interpretable, and data-driven solution for adaptive workforce planning.

Keywords-generative Artificial Intelligence (AI); Work-Life Balance (WLB); Large Language Models (LLMs); Information Technology (IT) workforce; employee productivity; burnout reduction; retention prediction; Human Resource (HR) policy simulation

I. INTRODUCTION

Deep Learning (DL) has become an essential tool for predictive analytics, offering state-of-the-art capabilities in handling high-dimensional and complex datasets. Its ability to automatically extract hierarchical representations from raw data without extensive feature engineering has enabled its successful application in predicting workforce productivity based on several factors, including absenteeism, project delivery performance, workload distribution, and employee wellness factors. Architectures such as Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks have proven particularly effective, as CNNs excel in spatial feature extraction, while LSTMs capture temporal dependencies in sequential organizational datasets, making them highly suitable for analyzing work environment dynamics [1, 2].

However, standalone DL models present inherent limitations when applied to Work-Life Balance (WLB)-driven productivity modeling, which involves multiple heterogeneous and often sparse data sources. For instance, workforce datasets are frequently imbalanced, particularly when capturing scenarios involving newly introduced policies such as hybrid work schedules, flexible shift programs, or employee wellness incentives. This imbalance negatively impacts model generalization, leading to reduced predictive accuracy for underrepresented policy configurations [3]. In addition, conventional DL models rely solely on historical data and cannot simulate hypothetical "what-if" scenarios, such as the potential productivity effect of compressed workweeks or wellness-based incentives, thus limiting their role in proactive Human Resource (HR) policy design [4]. Furthermore, most DL models act as black boxes, meaning there is no justification in the model's final prediction, making it difficult for HR managers to interpret how individual factors (e.g., workload intensity or remote work participation) affect overall employees' productivity, reducing trust in AI-driven decision support [5].

To address these challenges, hybrid modeling approaches integrating Generative AI with predictive DL architectures have gained momentum. Generative AI models, including transformer-based architectures and Variational Autoencoders (VAEs), can generate synthetic yet realistic datasets that are then used for improving training robustness and enabling counterfactual policy simulations [6, 7]. Such simulations allow decision-makers to evaluate the potential productivity scores of new or modified WLB policies before implementation. Furthermore, with the integration of explainability mechanisms, such as Shapley Additive Explanations (SHAP), decision-makers can also assess the contribution of each feature to the final productivity score [8].

In this context, this paper introduces GenWorkBalanceNet, a hybrid architecture to predict productivity scores influenced by WLB strategies. GenWorkBalanceNet advances existing research along three main axes. First, it couples conditional generative modeling with a deep predictive network to improve sample efficiency and robustness under policy heterogeneity and class imbalance (e.g., rare burnout events). Second, it unifies tabular HR Information System (HRIS) data, longitudinal activity traces, and optional textual feedback within a hybrid architecture, comprising a sequence encoder, transformer text encoder, and fusion head, to jointly capture temporal and semantic determinants of productivity. Third, it embeds governance-by-design principles through policy-aware conditioning, subgroup stability testing, and explainability tooling, translating model outputs into actionable and auditable WLB interventions. To achieve these, the architecture incorporates an attention-based LSTM module for temporal workforce analysis and employs SHAP analysis to improve interpretability and transparency. This synthesis directly addresses persistent limitations in prior research, including limited data realism, poor generalization across teams or policies, and insufficient support for counterfactual, decision-oriented analytics in enterprise IT settings [9, 10].

II. METHODOLOGY

GenWorkBalanceNet is a novel hybrid Generative AI-DL framework designed to predict workforce productivity under varying WLB conditions. Figure 1 illustrates the proposed architecture, which mainly consists of three integrated components: (i) a generative policy simulation layer; (ii) an attention-enhanced LSTM predictor; and (iii) a SHAP-based explainability layer. Additionally, Figure 2 illustrates the overall workflow of the proposed GenWorkBalanceNet framework.

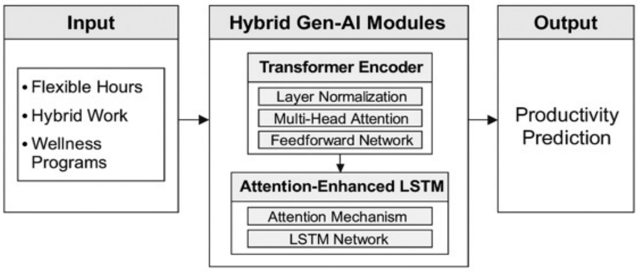

Fig. 1. GenWorkBalanceNet architecture.

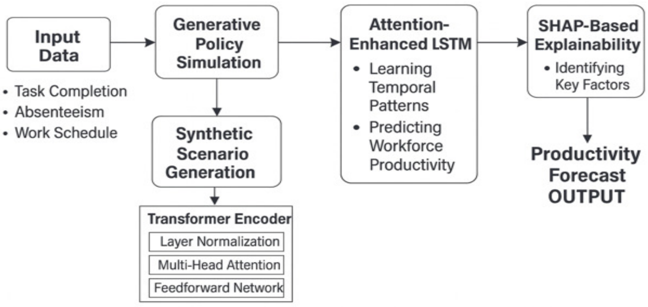

Fig. 2. GenWorkBalanceNet workflow.

A. *Generative Policy Simulation Layer*

The first stage of the proposed model uses a transformer-based encoder-decoder structure to create synthetic workforce data representing policy-driven scenarios. Given an input dataset:

$$X = \{x_1, x_2, \ldots, x_n\} \tag{1}$$

where each element x_i denotes workforce-specific features (e.g., task completion rate, absenteeism, wellness participation, and scheduling details), and the encoder produces contextual representations:

$$H = Encoder(X) \tag{2}$$

while the decoder reconstructs synthetic samples:

$$\hat{X} = Decoder(H) \tag{3}$$

The generated dataset $\hat{X}$ augments the original data, ensuring balanced representation of policy scenarios such as flexible schedules, wellness-driven incentives, or hybrid working models. This approach allows the downstream prediction module to learn from both historical and hypothetical policy interventions, enhancing generalization and adaptability.

B. *Productivity Prediction via Attention-Enhanced LSTM*

The prediction component is built on an LSTM network [11], which is well-suited for capturing temporal dependencies present in workforce productivity data. For each time step t, the model receives a feature input z_t, composed of both original and synthetically generated attributes:

$$h_t, c_t = LSTM\left(z_t, h_{\{t-1\}}, c_{\{t-1\}}\right) \tag{4}$$

where h_t is the hidden state vector and c_t is the memory cell state. To dynamically emphasize relevant features, an attention mechanism is incorporated, with attention weights α_t computed as:

$$\alpha_t = \{exp(W_a h_t)\} / \{\sum_k (W_a h_k)\} \tag{5}$$

and the resulting context vector is given by:

$$c_t^{\{*\}} = \sum_{\{t\}} \alpha_t h_t \tag{6}$$

The final productivity estimate Y_t is computed as:

$$Y_t = W_o c_t^{\{*\}} + b_o \tag{7}$$

where W_o and b_o represent learnable parameters of the output layer.

C. SHAP-Based Explainability Layer

To improve transparency, interpretability, and enhance trust in the model's outputs, the proposed approach integrates SHAP for feature attribution. The contribution of a given feature i is quantified as the contribution score ϕ_i defined as:

$$\phi_i = \sum_{\{S \subseteq F \setminus \{i\}\}} a \frac{|S|!(|F|-|S|-1)!}{|F|!} [f(S \cup \{i\}) - f(S)] \tag{8}$$

where F represents the full feature set, F is a subset excluding feature i, and $f\ (\cdot)$ denotes the trained prediction model.

D. Hybrid Integration of the Generative AI Module and the DL Predictive Model

Figure 3 focuses specifically on the hybrid integration pipeline that connects the Generative AI module and the DL predictive model within the proposed GenWorkBalanceNet framework. The diagram illustrates how synthetic workforce scenarios generated through the transformer-based generative module are fused with attention-enhanced LSTM processing to enable accurate and interpretable productivity predictions.

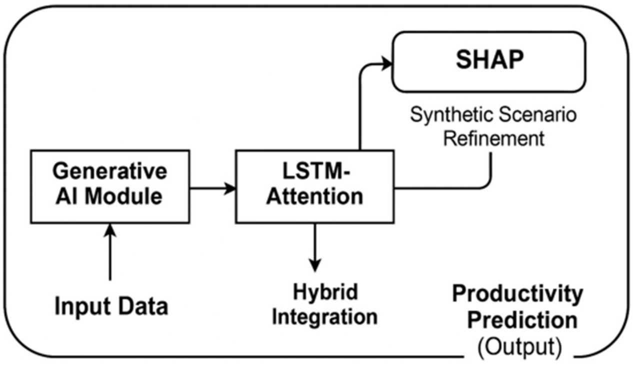

Fig. 3. Hybrid integration pipeline of GenWorkBalanceNet.

The pipeline begins with historical HR datasets containing task completion records, absenteeism trends, wellness program participation, and scheduling information. The Generative AI module produces synthetic workforce scenarios reflecting alternative HR policies such as flexible scheduling, remote work adoption, and wellness incentives. These synthetic data are then merged with historical records and fed into an attention-enhanced LSTM network that models temporal dependencies and dynamically assigns importance to critical workforce attributes. The resulting predictions undergo a SHAP-based interpretability analysis, which identifies key policy and workforce factors that impact productivity outcomes. The final output provides HR managers with actionable productivity forecasts and interpretable insights, enabling proactive workforce planning and policy evaluation [2, 12].

E. Dataset and Preprocessing

The dataset utilized in this study is derived from a publicly available employee performance and productivity dataset [13] hosted on Kaggle. The dataset comprises 100,000 employee records covering various aspects of performance, productivity, and demographics within a corporate environment. Features include job role, department, education, experience, working hours, satisfaction levels, employee engagement scores, overtime hours, job satisfaction ratings, work-from-home frequency, and average project turnaround time, ensuring a holistic representation of productivity-related parameters. All records were preprocessed, with personal identifiers anonymized to ensure ethical and reliable analysis [14]. Other preprocessing steps included:

- Data Cleaning: Removed duplicate records and resolved inconsistent categorical entries.
- Feature Encoding: Converted categorical variables (e.g., department, role, and location) to one-hot encoding.
- Normalization: Scaled continuous variables, such as task completion times and overtime hours, to the [0,1] range using min-max normalization.
- Outlier Treatment: Applied Interquartile Range (IQR) filtering to cap extreme values in absenteeism and overtime features.
- Data Splitting: Partitioned the dataset into training (70%), validation (15%), and testing (15%) sets while maintaining temporal sequence integrity for productivity trends.

The resulting dataset, combining real HR records and synthetic policy-driven data, provides a robust foundation for training and evaluating the proposed GenWorkBalanceNet framework [15].

F. Model Configuration and Parameter Settings

The proposed hybrid framework configuration parameters are presented in Table I. The transformer encoder-decoder and LSTM modules shared the same optimizer (AdamW), learning rate ($5 \cdot 10^{-5}$), batch size (32), and early stopping criterion (patience = 10 epochs).

The experimental evaluation was carried out on a high-performance computing server configured with an NVIDIA RTX A6000 Graphics Processing Unit (GPU) featuring 48 GB of Video Random Access Memory (VRAM), 256 GB of system RAM, and dual Intel Xeon processors to support large-scale DL workloads. The software environment comprised Python 3.11 for overall development, PyTorch 2.3 for DL model implementation, and SHAP 0.41 for explainability analysis.

TABLE I. MODEL CONFIGURATION AND TRAINING SETTINGS

Parameter	Value / Setting	Description
Architecture Type	Transformer encoder-decoder + LSTM-Attention	Hybrid model for policy simulation and productivity prediction.
Encoder-Decoder Layers	4	Number of stacked transformer layers in both encoder and decoder.
Attention Heads	8	Number of self-attention heads per transformer layer.
Hidden Size	512	Dimensionality of hidden representation within each transformer block.
Activation Function	GELU	Non-linear activation for improved learning dynamics.
Positional Encoding	Sinusoidal	Encodes the sequential order of input data.
Normalization	Layer Normalization	Stabilizes training and mitigates internal covariate shift.
Optimizer	AdamW	Optimizer with decoupled weight decay for generalization.
Learning Rate	$5{\cdot}10^{-5}$	Adaptive learning rate for stable convergence.
Batch Size	32	Mini-batch size used per training iteration.
Weight Decay	$1{\cdot}10^{-2}$	Regularization factor to prevent overfitting.
Max Epochs	100	Upper limit of training iterations.
Early Stopping Criterion	Patience = 10 epochs	Training stops if validation loss does not improve for 10 consecutive epochs.
LSTM Layers	2 (Bidirectional)	Number of recurrent layers for temporal modeling.
LSTM Hidden Size	256 (per direction)	Hidden dimension in each LSTM direction.
Hardware Setup	NVIDIA RTX A6000 GPU	Environment used for model training and evaluation.
Framework	PyTorch 2.3	DL framework utilized for implementation.

III. RESULTS AND DISCUSSION

A. *Comparative Analysis of GenWorkBalanceNet vs Baseline Methods*

The models were evaluated using Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and R^2 Score to measure prediction accuracy based on the productivity index. The productivity index represents a composite score derived from normalized indicators, including task completion rate, punctuality, engagement in wellness programs, and remote work efficiency. Each component (task completion rate, attendance consistency, overtime hours, project turnaround time, and wellness program participation) was scaled between 0 and 10, and the final index reflects an aggregated measure of workforce productivity under varying policy scenarios.

Figure 4 presents a comparative analysis of model performance based on the RMSE. The proposed GenWorkBalanceNet achieved the lowest RMSE value of 3.12 and the closest alignment between predicted and actual productivity scores. In contrast, Random Forest (4.35) and Gradient Boosting (4.12) exhibited higher error margins, reflecting lower prediction precision. The standalone LSTM (3.84) performed moderately well but remained less accurate than GenWorkBalanceNet.

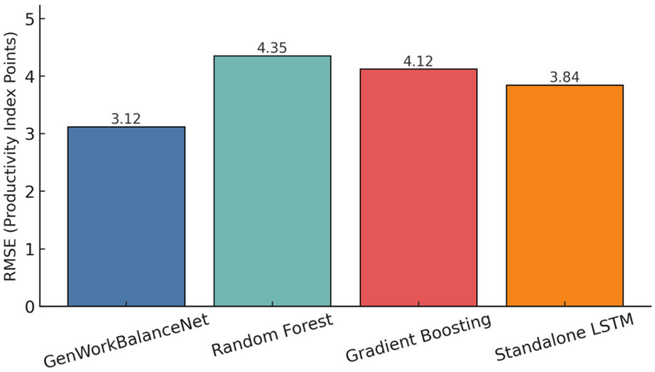

Fig. 4. RMSE comparison across models.

Figure 5 compares the MAE values of the evaluated models. GenWorkBalanceNet achieved the lowest MAE of 1.60, demonstrating higher consistency and reduced deviation between predicted and actual outcomes. In comparison, Random Forest (2.50) and Gradient Boosting (2.40) produced larger errors, while the standalone LSTM (2.20) performed moderately but less effectively than GenWorkBalanceNet.

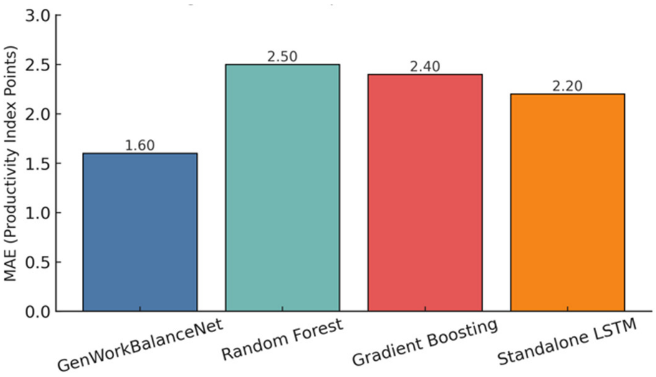

Fig. 5. MAE comparison across models.

Figure 6 depicts the R^2 values, representing each model's ability to explain the variance in productivity predictions. GenWorkBalanceNet attained the highest R^2 value of 0.92, indicating a stronger correlation between predicted and actual productivity compared to other models. In contrast, Random Forest (0.86) and Gradient Boosting (0.88) demonstrated lower explanatory power, while the standalone LSTM (0.90) performed competitively but slightly below GenWorkBalanceNet.

In relative terms, GenWorkBalanceNet reduced RMSE by 19-28%, MAE by 27-36%, and improved R^2 by 2-7% compared with all baselines, validating the effectiveness of the proposed model.

Moreover, Figure 7 illustrates the weekly variation in productivity scores, comparing actual productivity scores from the dataset in [13] against predictions from the proposed model and the second-best performing model (standalone LSTM). As shown, GenWorkBalanceNet closely follows the actual trend, maintaining minimal deviation throughout the observation period, while the standalone LSTM slightly underestimates productivity in certain intervals, showing delayed responses to fluctuations.

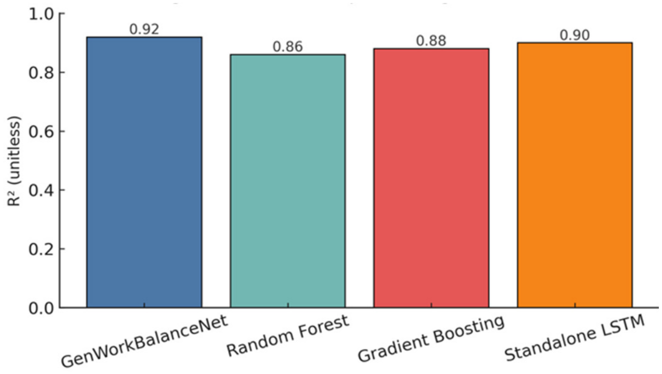

Fig. 6. MAE comparison.

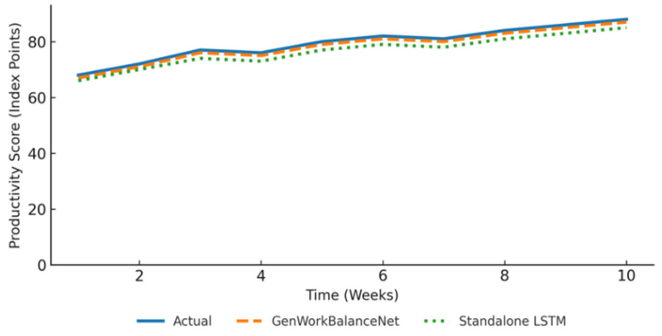

Fig. 7. Productivity trend prediction.

B. *SHAP-Based Feature Importance and Interpretability Analysis*

Table II summarizes the mean absolute SHAP values for the ten most influential features based on our model's predictions. The analysis indicates that task completion rate, wellness participation, and remote work frequency exhibit the strongest positive contributions to productivity, while overtime hours and absenteeism rate show pronounced negative impacts. Collectively, these five factors explain approximately 68% of the total model variance. Intermediate behavioral attributes such as engagement, peer collaboration, and skill development demonstrate moderate positive effects, highlighting the importance of collaborative and learning-oriented environments. In contrast, demographic features, including experience and education level, had a limited influence. Overall, the interpretability results validate that balanced workload management, flexibility, and wellness participation drive an estimated 6-8% improvement in productivity, consistent with the model's quantitative performance gains.

TABLE II. SHAP-BASED FEATURE IMPORTANCE RANKING

Rank	**Feature**	**Mean |SHAP| Value**	**Influence Direction**
1	Task Completion Rate	0.43	Positive
2	Wellness Participation	0.41	Positive
3	Remote Work Frequency	0.37	Positive
4	Overtime Hours	0.39	Negative
5	Absenteeism Rate	0.41	Negative
6	Engagement Score	0.28	Positive
7	Peer Collaboration	0.25	Positive
8	Skill Development	0.24	Positive
9	Experience (Years)	0.12	Neutral
10	Education Level	0.09	Neutral

IV. CONCLUSION

The proposed GenWorkBalanceNet framework effectively integrated Generative Artificial Intelligence (AI) for synthetic policy-driven scenario generation with an attention-enhanced Long Short-Term Memory (LSTM) predictor and Shapley Additive Explanations (SHAP)-based interpretability, addressing key limitations of traditional workforce analytics models such as data sparsity, limited policy simulation, and lack of transparency. By enriching historical Human Resource (HR) datasets with realistic synthetic scenarios, the model improves generalization and enables proactive "what-if" analysis of Work-Life Balance (WLB) strategies.

Experimental results demonstrated that GenWorkBalanceNet outperformed baseline methods such as Random Forest, Gradient Boosting, and standalone LSTM, with a 19-28% reduction in Root Mean Square Error (RMSE), 27-36% reduction in Mean Absolute Error (MAE), and 2-7% improvement in R^2 score. Additionally, the SHAP analysis identified task completion rate, wellness participation, and remote work frequency as positive drivers of productivity, while overtime hours and absenteeism rate exhibited negative impacts, jointly explaining 68% of model variance. Scenario analysis further revealed that flexible work-hour policies could enhance productivity by up to 6-8%, highlighting the model's value as a decision-support tool for HR management. Future research will extend this framework toward real-time workforce analytics, cross-organizational validation, and adaptive generative modeling to support dynamic, data-driven HR decision-making.

REFERENCES

[1] H. Sinha, "Predicting Employee Performance in Business Environments Using Effective Machine Learning Models," *International Journal of Novel Research and Development*, vol. 9, no. 9, pp. a875-881, Sep. 2024, https://doi.org/10.5281/ZENODO.13771036.

[2] S. D. Paigude and S. R. Shikalgar, "Deep Learning Model for Work-Life Balance Prediction for Working Women in IT Industry," in *Proceedings of the 4th International Conference on Information Management & Machine Intelligence*, Jaipur India, Dec. 2022, pp. 1–8, https://doi.org/10.1145/3590837.3590846.

[3] C. Gupta, K. V. S. Rajeswara Rao, and P. Datta, "Support Vector Machine Based Prediction of Work-Life Balance Among Women in Information Technology Organizations," *IEEE Engineering Management Review*, vol. 50, no. 2, pp. 147–155, Jun. 2022, https://doi.org/10.1109/EMR.2022.3152520.

[4] M. Parekh and D. Shah, "Understanding Work-Life Balance: An Analysis of Quiet Quitting and Age Dynamics using Deep Learning."

International Research Journal of Engineering and Technology (IRJET), vol 10, no 6, pp. 1230-1235, Jun. 2023, https://doi.org/10.13140/RG.2.2.21097.47204.

[5] S. Das, S. Chakraborty, G. Sajjan, S. Majumder, N. Dey, and J. M. R. S. Tavares, "Explainable AI for Predictive Analytics on Employee Attrition," in *Soft Computing and Its Engineering Applications*, vol. 1788, K. K. Patel, K. C. Santosh, A. Patel, and A. Ghosh, Eds. Cham: Springer Nature Switzerland, 2023, pp. 147–157.

[6] A. Ardichvili, K. Dirani, S. Jabarkhail, W. El Mansour, and S. Aboulhosn, "Using generative AI in human resource development: an applied research study," *Human Resource Development International*, vol. 27, no. 3, pp. 388–409, May 2024, https://doi.org/10.1080/13678868.2024.2337964.

[7] P. Phattharajiranan and V. Muangsin, "Generating Synthetic Population Using Transformer Based Networks," in *2023 27th International Computer Science and Engineering Conference (ICSEC)*, Samui Island, Thailand, Sep. 2023, pp. 193–200, https://doi.org/10.1109/ICSEC59635.2023.10329722.

[8] W. Narkbunnum and K. Hinthaw, "Interpretable Gradient Boosted Modeling of Employee Attrition: A SHAP-Based Framework for HR Analytics," *International Journal of Analysis and Applications*, vol. 23, Oct. 2025, Art. no. 236, https://doi.org/10.28924/2291-8639-23-2025-236.

[9] C. Fan, X. Liao, and X. Yang, "Artificial intelligence and enterprise total factor productivity: A human capital requirement perspective," *International Review of Economics & Finance*, vol. 104, Dec. 2025, Art. no. 104661, https://doi.org/10.1016/j.iref.2025.104661.

[10] J.-L. He, J.-H. Wang, C.-M. Lo, and Z. Jiang, "Human Activity Recognition via Attention-Augmented TCN-BiGRU Fusion," *Sensors*, vol. 25, no. 18, Sep. 2025, Art. no. 5765, https://doi.org/10.3390/s25185765.

[11] A. Tak, "The Data Mining Techniques for Analyzing Employee Performance and Productivity," *International Journal of Science and Research (IJSR)*, vol. 10, no. 10, pp. 1575–1578, Oct. 2021, https://doi.org/10.21275/SR231208202957.

[12] M. U. Tariq, "AI and Work-Life Balance: Transforming Employee Wellbeing in the Modern Workplace," in *Advances in Computational Intelligence and Robotics*, E. Ahmed, A. R. Babar, A. Samad, R. I. Ahmed, and G. Beydoun, Eds. IGI Global, 2025, pp. 85–110.

[13] *Employee Performance and Productivity Data.* (2024), M. Finsterwald. [Online]. Available: https://www.kaggle.com/datasets/mexwell/employee-performance-and-productivity-data.

[14] Y.-T. Chuang, H.-L. Chiang, and A.-P. Lin, "Insights from the Job Demands–Resources Model: AI's dual impact on employees' work and life well-being," *International Journal of Information Management*, vol. 83, Aug. 2025, Art. no. 102887, https://doi.org/10.1016/j.ijinfomgt.2025.102887.

[15] S. Garg, S. Sinha, A. K. Kar, and M. Mani, "A review of machine learning applications in human resource management," *International Journal of Productivity and Performance Management*, vol. 71, no. 5, pp. 1590–1610, May 2022, https://doi.org/10.1108/IJPPM-08-2020-0427.