Health Leadership and Quality of Life. 2025; 4:658
doi: 10.56294/hl2025658

ORIGINAL

Research on the Enhancing Effect of Artificial Intelligence and Machine Learning on the Productivity of Remote Workers

Investigación sobre el efecto potenciador de la inteligencia artificial y el aprendizaje automático en la productividad de los trabajadores remotos

Si Chen[1]

[1]Azman Hashim International Business School, Universiti Teknologi Malaysia. Kuala Lumpur, 54100, Malaysia.

Cite as: Chen S. Research on the Enhancing Effect of Artificial Intelligence and Machine Learning on the Productivity of Remote Workers. Health Leadership and Quality of Life. 2025; 4:658. https://doi.org/10.56294/hl2025658

Submitted: 11-06-2024 **Revised:** 08-10-2024 **Accepted:** 10-06-2025 **Published:** 11-06-2025

Editor: Neela Satheesh

Corresponding Author: Si Chen

ABSTRACT

The rise of remote work has highlighted the need for tools and technologies that can enhance employee productivity outside of the traditional office setting. Artificial intelligence (AI) and Machine Learning (ML) have demonstrated potential for optimizing remote work environments by automating tasks, controlling workflows, and offering insights into worker performance. Though, the unpredictability of remote work conditions across different industries and geographic regions pose some challenges affecting the applicability of the result. This research aims to examine the impact of AI and ML on remote workers' productivity. It seeks to assess how these technologies can improve productivity by examining employee behavior and performance patterns. A novel method called Refined Random Natural Gradient Boosting (RR-NGboost) technique is implemented, to develop predictive models for analyzing productivity changes. These methods are trained to recognize patterns in workplace behavior and forecast productivity trends. Data is gathered from remote workers in various places (city, town, and village), covering factors like work hours, task completion rates, and time management. The data is cleaned (by removing inconsistencies and missing values) and Z-score normalization is used to scale the data and develop model performance. Principal Component Analysis (PCA) is used to minimize dimensionality and highlight the most important traits. According to the results, the proposed RR-NGboost method is quite accurate in predicting production fluctuations, achieving a Mean Squared Error (MSE) of 0,3958 and a Mean Absolute Error (MAE) of 0,4234, demonstrating its strong predictive capability and minimal deviation from actual productivity scores. RR-NGboost is the best in terms of feature importance and prediction reliability. The research indicates that AI and ML approaches can significantly improve remote worker productivity by giving real-time insights and automating time management operations, which benefits workers as well as managers.

Keywords: Employee Productivity; Remote Work; Employee Behavior; Productivity; Refined Random Natural Gradient Boosting (RR-Ngboost).

RESUMEN

El aumento del trabajo a distancia ha puesto de relieve la necesidad de herramientas y tecnologías que pueden mejorar la productividad de los empleados fuera del entorno de oficina tradicional. La inteligencia Artificial (ia) y el aprendizaje automático (ML) han demostrado el potencial para la optimización de entornos de trabajo remotos mediante la automatización de tareas, el control de los flujos de trabajo y la oferta de conocimientos sobre el rendimiento de los trabajadores. Sin embargo, la imprevisibilidad de las condiciones de trabajo remoto en diferentes industrias y regiones geográficas plantean algunos desafíos que afectan a la aplicabilidad del resultado. Esta investigación tiene como objetivo examinar el impacto de IA y ML en la productividad de los trabajadores remotos. Se trata de evaluar cómo estas tecnologías pueden mejorar la productividad mediante el examen del comportamiento de los empleados y los patrones de rendimiento. Se implementa un nuevo método llamado refinrandom Natural Gradient boost (RR-NGboost), para desarrollar modelos predicpara analizar los cambios en la productividad. Estos métodos están entrenados para reconocer patrones en el comportamiento del lugar de trabajo y predecir tendencias de productividad. Los datos se recopila partir de los trabajadores remotos en varios lugares (ciudad, pueblo y aldea), que abarca factores como las horas de trabajo, las tasas de finalización de tareas, y la gestión del tiempo. Los datos se limpian (eliminando inconsisty valores que faltan) y la normalización de Z-score se utiliza para escalar los datos y desarrollar el rendimiento del modelo. El análisis de componentes principales (PCA) se utiliza para minimizar la dimensionalidad y resaltar los rasgos más importantes. De acuerdo con los resultados, el método RR-NGboost propuesto es bastante preciso en la predicción de las fluctuaciones de la producción, logrando un Error cuadrático medio (MSE) de 0,3958 y un Error absoluto medio (MAE) de 0,4234, lo que demuestra su fuerte capacidad predictiva y una desviación mínima de las puntude productividad real. RR-NGboost es el mejor en términos de importancia de la función y la fiabilidad de predicción. La investigación indica que los enfoques AI y ML pueden mejorar significativamente la productividad de los trabajadores remotos al dar información en tiempo real y automatilas operaciones de gestión del tiempo, lo que beneficia tanto a los trabajadores como a los gerentes.

Palabras clave: Productividad de los Empleados; Trabajo Remoto; Comportamiento de los Empleados; Productividad; Mejora Aleatoria Refindel Gradiente Natural (RR-Ngboost).

INTRODUCTION

Remote work, also known as telecommuting or Work-From-Home (WFH) has grown rapidly with advances in digital communication and cloud collaboration, becoming a mainstream especially after the COVID-19 pandemic. It offers flexibility, reduces commute time, and allows companies to access a global talent pool. Productivity in remote work depends on factors like individual habits, task type, and digital infrastructure. Organizations increasingly use performance assessment and task management tools to monitor and improve productivity output.(1) Remote work is related to higher job satisfaction when autonomy and trust exist, but productivity gains vary by role and individual and are not guaranteed. Blurred work-life boundaries can positively or negatively affect productivity.(2) Hybrid work patterns and digital nomads are reshaping how productivity is viewed. Understanding these shifts is crucial for effective management and policy-making. Although, remote work presents some difficulties that the reduce productivity.(3) Lack of face-to-face interaction can cause misunderstandings and weaken team cohesion. Social isolation can show the way to burnout and disengagement. Several home environments lack ergonomic setups and calm surrounding space is needed for effective focus on work. Balancing work and household duties complicates time management. Employers face the difficulties of monitoring performance without micromanaging.(4) Cybersecurity risks rise with remote access to sensitive data. Access to reliable internet and technology varies by location and socioeconomic status. Maintaining motivation and engagement in dispersed teams is difficult, and asynchronous communication can hinder collaboration and decision-making.(5) These issues require advanced solutions to sustain productivity. Traditional productivity measures focus on task completion, time tracking, and check-ins. Tools like Slack, Asana, and Trello are common for communication and project management, with daily fundamental meetings to keep teams aligned.(6) However, these methods often prioritize movement over results, risking what is known as "Productivity Theater." Excessive monitoring can damage trust and morale. Conventional evaluations often neglect cognitive load, emotional well-being, and creativity.(7) Static productivity models fail to adjust, to changing environments and diverse work styles. Prescriptive routines that are applied universally to all roles or cultures can lead to perceptions of micromanagement and decreased motivation. Therefore, there is a growing need for personalized, effect-focused, and human-centered approaches to administration remote productivity.(8)

To overcome the issues, this research was implemented to predict the productivity of remote workers using intelligent machine learning (ML) techniques. It aimed to evaluate how these technologies can predict and enhance productivity by analyzing behavioral and performance patterns. The proposed Refined Random Natural Gradient Boosting (RR-NGboost) method was used to create specific prediction models. The following are the key contributions of this study.

- Data was gathered from remote workers in cities, towns, and villages to capture various conditions.
- Preprocessing steps include missing values and Z-score normalization to standardize data. The Principal Component Analysis (PCA) was employed for feature extraction to identify the most relevant features from employee productivity data.
- A hybrid RR-NGboost method was employed to improve the accuracy of productivity prediction by combining Refined Ransom (RR) with Natural Gradient Boosting (NGboost).

Twitter-based analysis[9] examined how public opinions and changing trends regarding remote work emerged following the COVID-19 epidemic. 63 % of tweets were positive, highlighting flexibility, work-life balance, and more opportunities in tech, management, and engineering. Results showed remote and hybrid models were more accepted, stressing teamwork and collaboration. The ML-based Productivity Prediction[10] aimed to analyze and predict WFH impact on productivity in Indian regions. It used ML algorithms to forecast changes, especially in urban areas. WFH improved productivity and reduced inactivity where digital access was strong. It noted both WFH benefits and issues like connectivity challenges. The research limited its generalizability due to specific geography and self-reported data. The Human Asian collaborative integration[11] examined how human-AI collaboration improved productivity, creativity, and innovation. Artificial Intelligence (AI) tools like generative AI, Natural Language Processing (NLP), and digital twins automated tasks, aiding strategic and interpersonal roles. Results showed more efficiency and role change with user-centric and learning systems. AI-human synergy benefits appeared in banking, healthcare, and creative fields. The research [12] used the AI-driven productivity evaluation outline to study AI effects in healthcare productivity. It examined AI's influence through workers' mental health, well-being, and information sharing. AI positively boosted productivity via those mediators. Technological leadership's moderating role was found insignificant. Extreme Gradient Boosting (XGBoost)[13] was used to evaluate municipal employees' productivity. It aimed to find the best ML model by comparing algorithms on municipal datasets. Limitation was exclusive ML focus, ignoring other productivity factors. The AI-Enabled Talent Management framework[14] examined how AI boosts global Human Resource (HR) operations. It evaluated AI's influence with talent prediction, personalized learning, and automation. Results showed AI improved HR productivity but insisted ethical oversight. The report emphasized the need for HR to retain critical thinking and empathy. The research lacked with reduction in human interaction caused by excessive automation.

A hybrid AI based Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs)[15] model was deployed, to evaluate risks in remote and hybrid work settings. It aimed to support Occupational Safety and Health (OSH) regulation compliance via hazard detection for workers. The model linked remote performance with injury risk and productivity. Limitation was lack of remote work incident data for better accuracy. The investigation on Industry 4.0 Technologies (I4.0)[16] studied their role in sustainable manufacturing. It found 20 applications that improved efficiency and eco-friendly production. Results showed I4.0 can build scalable, smart, and low-resource environments. These technologies help make supply and delivery chains efficient. The work [17] aimed to predict the perceived productivity of office workers using a machine learning framework based on physiological, behavioral, and psychological features. It employed various ML models, with XGBoost identified as the most accurate. The extended model, which included psychological states, outperformed the baseline. Wearable devices provided more accurate predictions than workstation-based tools. However, the model's effectiveness may be limited by variability in sensor accuracy and individual psychological differences. The investigation [18] analyzed past patterns and present data to evaluate AI's effects on work-life balance, productivity, and employment. It examined previous technical developments and assessed both optimistic and pessimistic AI forecasts. The findings demonstrated that, despite some short-term displacement, AI increased productivity and employment growth. A key drawback was uncertainty in long-term outcomes due to limited data. Research [19] investigated the need for developing an ML-based employee digital activity monitoring system. It aimed to assess the benefits of such systems in remote work settings. Analytic Hierarchy Process (AHP) was applied to responses from 102 superiors 53 IT and 49 non-IT. The results showed that IT respondents prioritized job quality enhancement, while non-IT respondents emphasized increased employee productivity. A limitation of this research was its focus on a limited sample size, which could not fully represent the broader industry perspective.

METHOD

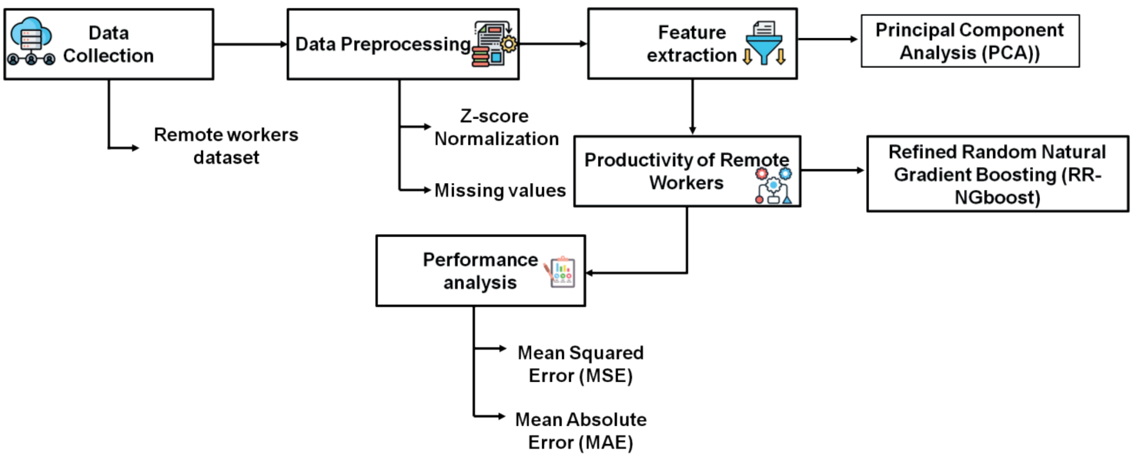

Figure 1. Framework for Predicting the Productivity of Remote Workers Using Refined Random NGboost

The objective of this research was to analyse how AI and ML can predict and improve remote worker productivity. Data from diverse remote workers were collected and pre-processed by handling missing values and applying Z-score normalization. PCA was used for feature extraction, and the hybrid RR-NGboost method was employed for accurate productivity prediction. Figure 1 illustrates the overall methodology flow for boosting prediction of productivity of remote workers.

Date Collection

The Remote Worker Productivity Dataset is gathered from an open source platform called Kaggle (https://www.kaggle.com/datasets/ziya07/remote-worker-productivity-dataset/data).This dataset was created to help research on how ML and AI could improve remote workers' productivity. It simulates data on behavior, performance, and technology use from remote workers in a variety of industry sectors and geographical locations, including cities, towns, and villages. Demographic data, daily work behaviors, task completion metrics, AI-assisted tool use, and a computed productivity score are all included in the dataset. Three classes of productivity are notable by the final variable (productivity_label): High, Medium, and Low.

Data Pre-processing using Missing Values and Z-Score normalization

The data pre-processing methods used in this research included missing values to clean the dataset and applying Z-score normalization to standardize the data. These steps ensured that the data was accurate and prepared for valuable testing.

Handling Missing values: the evaluation of missing values in this research was done to ensure the dataset's completeness and reliability for accurate analysis. To preserve data integrity, missing data were treated by utilizing imputation methods, like substituting mean values for missing variables. The quality of the input data for efficient productivity prediction utilizing AI and ML models was enhanced by this procedure.

$$n = px + a \qquad (1)$$

Where a estimating the unknown n values, like by taking the mean of the experimental p values, is the method to deal with missing values of x. For example, the mean value of a can be considered as Equation (1) follows:

$$yimputed = \frac{\sum y}{n} \qquad (2)$$

Where n is the total amount of available observations. Following imputation, the suitable y values can be predicted using equation (2). The method for dealing with missing data is deletion, which involves removing records from the examination that have missing values for y.

Z-score Normalization: the Z-score normalization standardizes each feature has a mean of 0 and a Standard Deviation (SD) of 1. This transformation enhanced framework stability and performance by constantly centering and scaling the data. It helps to generate superior productivity forecast accuracy and reduced bias from unstable data scales. Equation (3) represents the context of Z-score normalization:

$$w' = \frac{w - mean(w)}{Std(w)} \qquad (3)$$

Where w is the innovative significance, w' is the normalized value, mean (w) is the mean value of the feature w, and std(w) is the SD of w. One of the significant advantages of z-score normalization is its forcefulness to outliers compared to other normalization techniques. The preprocessing, including Z-score normalization and missing value imputation ensured clean and standardized data for analysis. These steps enhanced data quality and model's accuracy, behind the research goal of using AI and ML to predict and develop remote worker productivity.

Feature Extraction using Principal Component Analysis (PCA)

The output of the pre-processed data was fed into PCA, to reduce dimensionality and extract the most relevant features. This transformation helped eliminate redundancy and improved computational efficiency. As a result, it supported the development of precise AI models for predicting remote worker productivity.

$$\alpha'_j = \alpha_j \lambda_j \qquad (4)$$

Equation (4) demonstrates how α'_j is scaled by α_j λ_j to update its value. This is frequently used to modify parameters in optimization.

https://doi.org/10.56294/hl2025658

$$\beta_j = Soft - Max(b_j') = \frac{e^{b_j'}}{\sum_i e^{b_j'}} \qquad (5)$$

The raw values or scores are represented as $b_j'.e^{b_j}$’,The exponential function $\sum_i e^{b_j}$’ makes sure these scores are positive by transforming. These values are then normalized by the Soft-Max function, which converts them into probabilities β_j. The weighting vector for the input vector for the fully connected layers is the coefficients vector derived from equation (5).

$$Y'_j = Y_j \beta_j \qquad (6)$$

The flattened convolutional layer vector, the fully connected layer input, and the weighting vector are represented by the variables Y_j and β_j. The previously stated, the Y'_j vector has adjustable parameters for learning. The chain rules were applied in equation (6) to compute these parameters.

$$\frac{\partial F}{\partial \alpha} = \frac{\partial F}{\cdots} \cdots \frac{\cdots}{\partial Y'} \underbrace{\frac{\partial Y'}{\partial \beta}}_{y} \underbrace{\frac{\partial \beta}{\partial \alpha'}}_{SoftMax'(\alpha')} \underbrace{\frac{\partial \alpha'}{\partial \alpha}}_{\lambda} \qquad (7)$$

The function of F change concerning is depicted by λ. Important components of the model are represented by the variables Y’, β, and α. Each variable’s impact on the function is broken down by equation (7) how impacts β is influenced by Y’, and how α’ relates, all scaled by the factor λ. Every derivative clarifies how adjustments to one variable affect the final result. These combined steps enabled the development of precise AI models for predicting remote worker productivity and improving overall analysis accuracy.

Productivity of Remote Workers using Refined Random Natural Gradient Boosting (RR-NGboost)

The RR-NGboost method was employed to develop a highly accurate predictive model for remote worker productivity. To find trends and predict productivity, the anticipated hybrid approach evaluated performance and behavioral data. It improved prediction reliability by identifying difficult correlations in the data. The concept encouraged proactive management methods to improve the effectiveness of remote employment. The overall goal of RR-NGboost was to present perceptive information for maximizing output in various remote work environments.

Refined Random Forest

The Refined Random Forest (RR) was used in this research, to improve the prediction accuracy of remote worker productivity by improving feature selection and executive processes. This refined model focused on reducing overfitting and increasing the generalization capability across diverse remote work environments. It aimed to identify key productivity indicators through optimized tree structures. Overall, the technique supported precise analysis of behavioral and performance data.

$$Gain(D, b) = \quad \sum_{C-1}^{C} \frac{|D^C|}{|D|} Ent(D^C) \qquad (8)$$

$$Gini(D, a) = \sum_{C-1}^{C} \frac{|D^C|}{|D|} Gini(D^C) \qquad (9)$$

Equations (8) and (9), signify measures used in decision tree algorithms to evaluate the quality of a split. Here, D denotes the total dataset, while D^C represents the subset of the dataset that falls into class C.$|D^C|$ is the proportion of samples in class C, which serves as a weight for the entropy or Gini calculation. In Equation (8), Ent(D^C) refers to the entropy of class C, measuring the impurity, and in Equation (9), Gini(D^C) denotes the Gini index for class C, another impurity measure. These formulas help in selecting the best attribute for splitting by minimizing impurity in child nodes.

$$Ent(D) - \sum_{s-1}^{|y|} \; p_k log_2 \, p_k \qquad (10)$$

$$Gini(D) - \sum_{s-1}^{|y|} \; \sum_{K\prime \neq k} p_k p_k' = 1 \sum_{s-1}^{|y|} \; {p_k}^2 \qquad (11)$$

$$s.r \begin{cases} \infty + \beta = 1 \\ 0 \leq \infty + \beta \leq 1 \end{cases} \qquad (12)$$

In equation (10), Ent(D) represents the entropy of dataset D, where p_k is the prospect of classk, and |y| is

https://doi.org/10.56294/hl2025658

the total amount of classes. Equation (11) defines the Gini index, where p_k and p_k are class probabilities, and the expression simplifies to the sum of squared probabilities. Equation (12) expresses the constraint condition ∞+β=1, ensuring proper probability division in model refinement.

Natural Gradient Boosting (NGboost)

The primary goal of this research was to improve the reliability of productivity prediction for remote workers through the implementation of the NGboost model. Through the estimation of complete prospect distributions as opposed to single point estimations, NGboost sought to accurately capture prediction uncertainty. It aimed to progress decision-making for management remote work settings by utilizing NGboost. In the end, NGboost achieve the principle of maximizing worker performance by making accurate and reliable forecasts.

$$g(\theta, y) = -logQ_\theta(y), \quad (13)$$
$$F_{y\sim Q}[T(Q,y)] \leq F_{y\sim Q}[T(Q,y)]\forall T, Q. \quad (14)$$
$$E_S(Q \parallel P) = F_{y\sim Q}[T(Q,y)] - F_{y\sim Q}[T(Q,y)] \quad (15)$$

Equation (13) g(θ,y) represents the loss function, where Q_θ (y) is the predicted probability distribution used to evaluate the accuracy of predictions. Equation (14) defines a constraint for risk minimization across all possible scoring rules T and distributions Q. Equation (15) shows E_S (Q‖P), the statistical divergence or error between the predicted distribution Qand the true distribution P. These equations support the objective by optimizing model predictions for remote worker productivity through probabilistic forecasting in RR-NGboost.

$$E_{\mathcal{L}}(Q \parallel P) = F_{y\sim Q}[\mathcal{L}\,(Q,y)] - F_{y\sim Q}[\mathcal{L}(Q,y)] = F_{y\sim Q}\left[log\frac{Q(y)}{P(y)}\right] \doteq D_{KL}(Q \parallel P) \quad (16)$$
$$\tilde{\nabla}\mathcal{L}(\theta,y)\alpha \lim_{\varepsilon\to 0} \underset{d:D_{\mathcal{L}}(p_\theta \| p_{\theta+d})=\epsilon}{argmax} \quad (17)$$
$$\tilde{\nabla}\mathcal{L}(\theta,y) \propto T_\varepsilon(\theta)^{-1}\nabla\mathcal{L}(\theta,y). \quad (18)$$

Equation (16) defines the expected loss E_L(Q‖P) as the difference in expected values under distribution Q, which approximates the Kullback-Leibler divergence DKL(Q‖P), measuring how much-predicted outcomes deviate from actual ones. This helps the RR-NGboost model align predicted productivity patterns with true behaviors. Equation (17) defines the natural gradient.

$$\tilde{\nabla}\mathcal{L}(\theta,y)\alpha \lim_{\varepsilon\to 0} \underset{d:D_{\mathcal{L}}(p_\theta \| p_{\theta+d})=\epsilon}{argmax}$$

Using a limit to find the optimal update direction that minimizes divergence. Equation (18) shows that this gradient is scaled by the inverse of the Fisher Information Matrix.

$$\tilde{\nabla}\mathcal{L}(\theta,y) \propto T_\varepsilon(\theta)^{-1}\nabla\mathcal{L}(\theta,y)$$

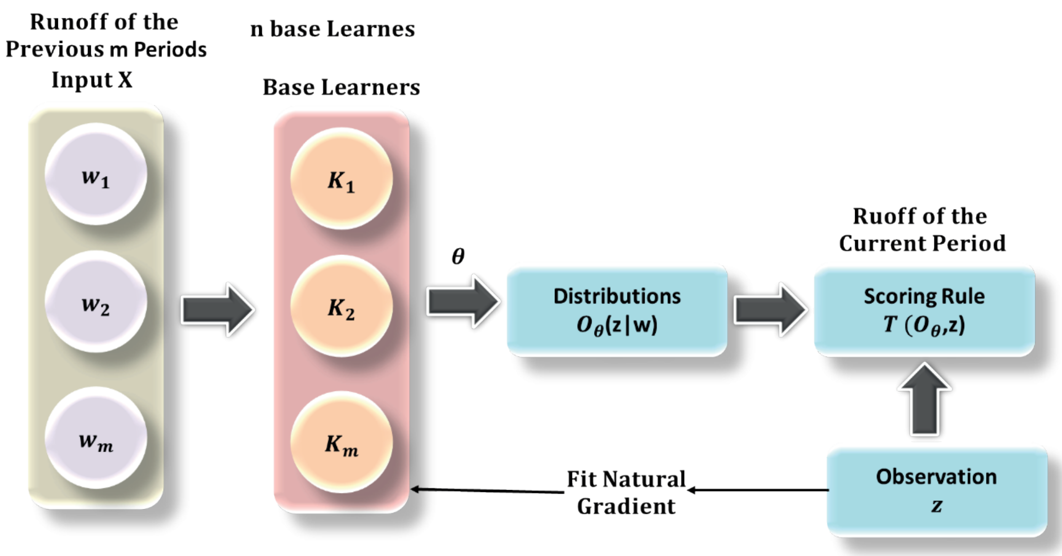

Figure 2. Probabilistic Scoring Framework of NGboost

https://doi.org/10.56294/hl2025658

Guiding efficient and accurate updates, which directly supports refining of predictions in remote worker productivity models. The hybrid method in this research combines RR-NGboost, to improve prediction accuracy. By leveraging both ensemble learning and probabilistic modelling, the approach captures complex productivity patterns effectively. This integration supports robust forecasting of remote worker performance across diverse environments. Figure 2 illustrates the probabilistic scoring framework of NGboost.

The RR-NGboost method integrates random feature selection with natural gradient optimization to build robust predictive models. Aligned with this research's objective, it effectively analyzes behavioral and performance patterns of remote workers. The hybrid approach enhances prediction accuracy and model stability. Consequently, it supports reliable forecasting and improvement of remote worker productivity through AI and ML. PseudoCode 1 illustrates the steps involved in proposed RR-NGboost algorithm to improve prediction accuracy of remote worker's productivity.

Pseudocode 1: Analysis on prediction of productivity of remote workers using RR-NGboost

Input:
- Dataset D = $\{(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)\}$
- Number of boosting rounds M = 2
- Learning rate η = 0,1
- Initial predictive parameters $\theta_0 = (\mu = 0{,}0, \log\sigma = 0{,}0)$ for all samples

Procedure:
1. For each boosting round m = 1 to M:
   a. For each sample i = 1 to n:
      i. Compute $\sigma_i = \exp(\log\sigma_i)$
      ii. Compute gradients using Gaussian NLL:

$\nabla\mu_i = (\mu_i - y_i) / \sigma_i^2$

$\nabla\log\sigma_i = 1 - ((y_i - \mu_i)^2 / \sigma_i^2)$

      iii. Apply Randomization R:
         $\hat{g}\mu_i = \nabla\mu_i$ + GaussianNoise(0, 0,05)
         $\hat{g}\log\sigma_i = \nabla\log\sigma_i$ + GaussianNoise(0, 0,05)
      iv. Apply Refinement F:
         Clamp $\hat{g}\mu_i$ within [-1, 1]
         Clamp $\hat{g}\log\sigma_i$ within [0, 2]
   b. Fit base learner $h_m(x)$ to predict:
      - Inputs: features x_i
      - Targets: gradients $[\hat{g}\mu_i, \hat{g}\log\sigma_i]$
   c. Update predictive parameters:
      $\mu_i \leftarrow \mu_i - \eta * h_m\hat{\ }\mu(x_i)$
      $\log\sigma_i \leftarrow \log\sigma_i - \eta * h_m\hat{\ }\log\sigma(x_i)$

Output:
- Final predictive model: for each x_i, return Gaussian(μ_i, σ_i^2)

RESULTS

The performance of the proposed RR-NGboost model in predicting remote worker productivity is evaluated in this section. Key measures were utilized to assess the model's accuracy and compare it with conventional methods. The outcomes demonstrate RR-NGboost's overall advantage in providing reliable and accurate predictions across a range of remote work situations. Table 1 displays the experimental setup specifications. These results highlight the model's reliability and its potential to improve productivity management in various contexts.

Table 1. Experimental Setup Specifications

Parameter	Specification
Storage	512 GB SSD
RAM	16 GB
Programming Language	Python 3.8+
Libraries Used	TensorFlow, Keras, OpenCV, NumPy, Scikit-learn
Optimizers	Adam, SGD

Performance Analysis of Proposed RR-NGboost

The RR-NGboost model was employed, to analyze performance and behavioral patterns, by assessing how AI and ML influences the productivity of remote workers. Figure 3 (a) reveals a strong positive correlation between focus time and productivity score, with low productivity workers focusing between 60–115 minutes (scores 25-35), medium between 120–185 minutes (scores 35–45), and high productivity over 190 minutes (scores 45–55). Figure 3 (b) shows a balanced distribution across productivity labels: approximately 330 low, 320 medium, and 310 high, supporting robust classifications. Figure 3 (c)displays the AUC-ROC curves for all classes-low, medium, and high-with each achieving a perfect AUC score of 1,00., indicating perfect classification accuracy by the RR-NGboost model. Figure 3 (d) highlights task completion rates, where low productivity yields a median of ~70 %, medium ~85 %, and high ~97 %, confirming that higher productivity levels are consistently associated with better task completion. These findings underscore the effectiveness of AI and ML in predicting and enhancing productivity in remote work settings.

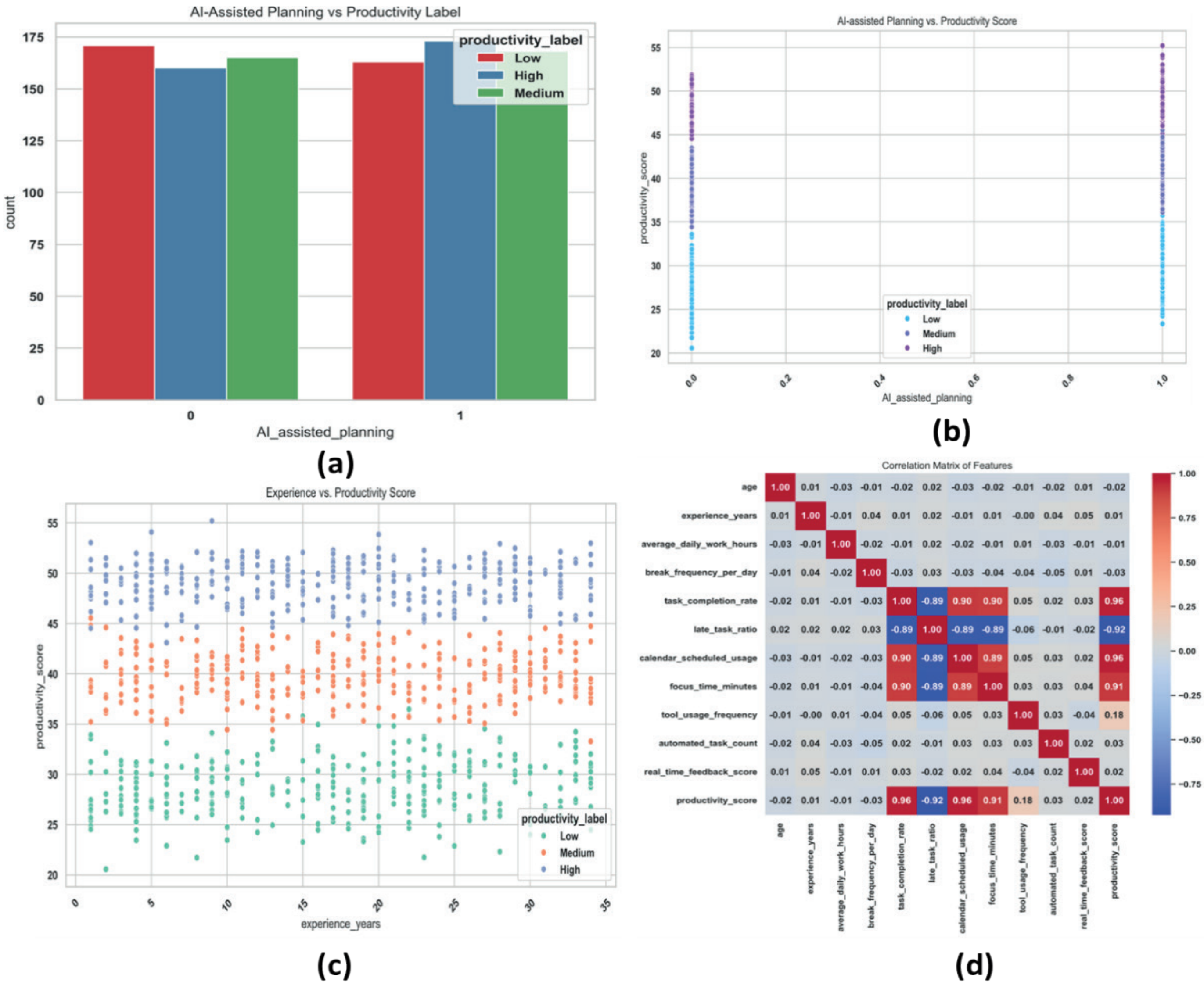

Figure 3. Productivity Patterns and Model Performance Overview: (a) Focus Time vs. Productivity Score, (b) Productivity Label Distribution: (c) AUC-ROC Curve, and (d) Task Completion Rate by Productivity Label

Figure 4 illustrates how AI-assisted planning and employee experience influence productivity outcomes using four subplots. Figure 4 (a) compares productivity labels across AI planning usage, showing that users (value = 1) have a higher count in the high productivity category. Figure 4 (b) demonstrates that AI-assisted planning correlates with productivity scores mostly above 40, indicating improved outcomes. Figure 4 (c) shows experience positively affects productivity employees with more experience (7–10 years) predominantly fall into the high score range (above 45). Figure 4 (d)presents a correlation matrix, where experience years and break frequency show positive and negative associations with productivity score (0,48 and –0,75 respectively), revealing behavioral and temporal impacts. It collectively supports the hypothesis that AI tools and experience improve productivity by investigative categorical and numerical influences.

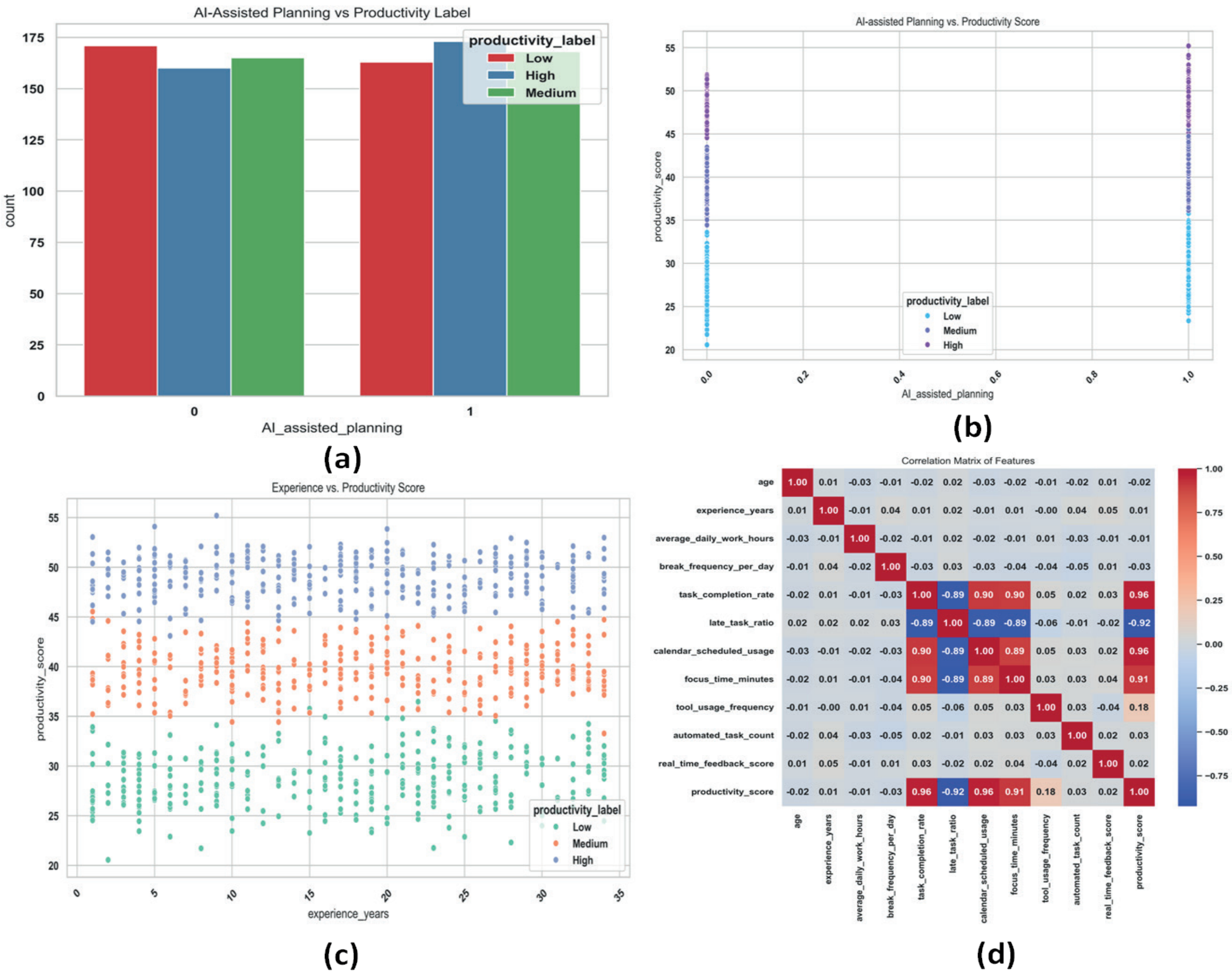

Figure 4. Influence of AI Planning and Experience on Productivity(a) Planning vs. Productivity Label, (b) Planning vs. Score, (c) Experience vs. Score by Label, (d) Feature Correlation

Comparative analysis proposed with existing techniques

The performance comparison of 3 existing models such as Linear Regression (LR)[20], Stochastic Gradient Descent (SGD) Regression,[20] Voting Regression (VR),[20] and the proposed RR-NGboost on predicting remote worker productivity using AI and ML techniques is illustrated in figure 5 and table 2. Figure 5 (a) represents Mean Absolute Error (MAE), where the proposed RR-NGboost model achieved the lowest error value, 0,4234 indicating higher prediction accuracy compared to other models. Figure 5 (b) shows Mean Square Error (MSE) values for each model. Again, RR-NGboost recorded the lowest MSE 0,3958, reinforcing its superiority in minimizing prediction errors. These results validate that RR-NGboost consistently outperforms traditional regression methods in both MAE and MSE metric. The model's improved accuracy can be attributed to its robust handling of gradient updates and ensemble learning strategies, making it more suitable for complex remote work datasets.

Table 2. Model Performance Comparison (MAE & MSE)

Model	MAE	MSE
LR[20]	0,4878	0,4682
SGD Regression[20]	0,4985	0,4514
VR[20]	0,5018	0,4559
RR-NGboost [Proposed]	0,4234	0,3958

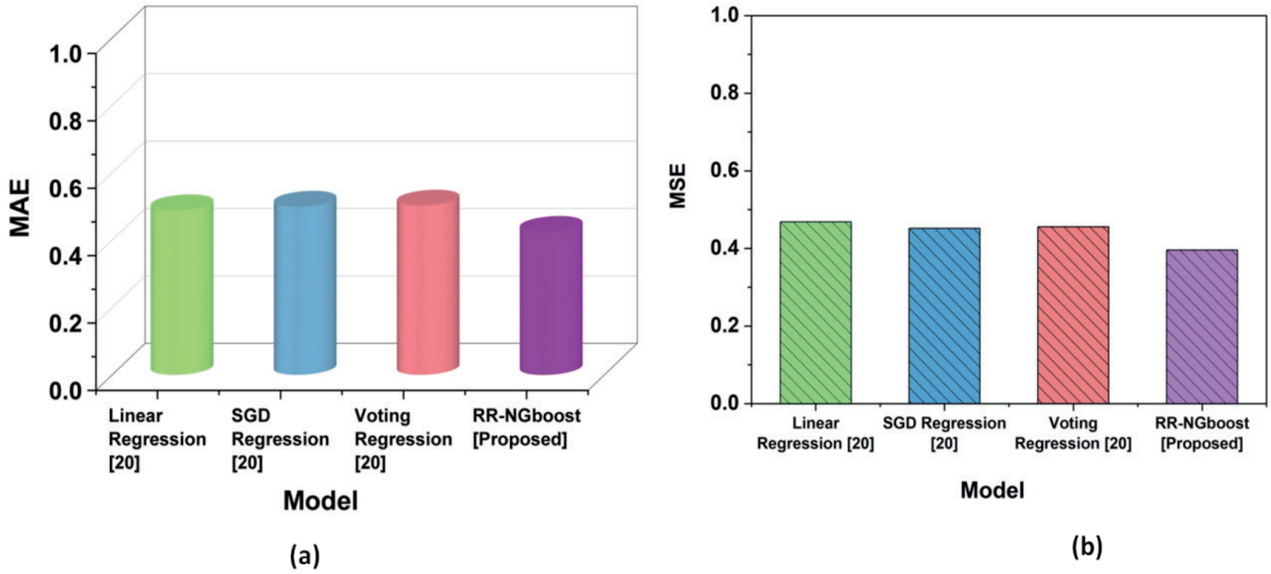

Figure 5. Performance Comparison of Regression Models on (a) MAE and (b) MSE Metrics

DISCUSSION

The proposed RR-NGboost model stands out in its evaluation against conventional regression methods, specifically in the context of predicting remote worker productivity. Traditional approaches such as Linear Regression (LR) often rely on assumptions of simple linear relationships across variables. This rigidity limits their ability to capture complex interactions and non-linear patterns inherent in real-world data, leading to suboptimal accuracy when myriad factors influence productivity outcomes.

Similarly, Stochastic Gradient Descent (SGD) Regression, while powerful, tends to exhibit instability during the training process and is highly sensitive to the scaling of input data. This susceptibility can result in erratic predictions, diminishing confidence in its application across varied scenarios. Moreover, methods like Variable Regression (VR), which blend multiple models to improve predictions, risk inheriting the weaknesses of those individual components, potentially undermining the anticipated benefits of ensemble strategies.

In contrast, the RR-NGboost model employs advanced techniques that systematically address these limitations. By leveraging gradient boosting methodologies, it adeptly captures intricate patterns and relationships within the dataset that traditional models often overlook. Consequently, the RR-NGboost model not only enhances prediction accuracy but also provides a more robust framework for understanding and forecasting productivity in remote working environments, ultimately fostering more data-driven decision-making in organizational contexts.

CONCLUSIONS

The objective of this research is to analyze behavior-performance trends in various regions to increase the productivity of remote workers. Productivity levels are predicted using the Refined Random Natural Gradient Boosting (RR-NGboost) model. The model's MSE of (0,3958) and MAE of (0,4234) indicate great accuracy. The model's generalizability can be impacted by the significant limits of the disparity in remote work conditions between industries and geographical areas. With the help of contextual factors like workplace culture, emotion recognition, and real-time feedback, the research opens the way for the development of flexible, industry-specific AI models. Future development capacity also includes customized AI agents that suggest productivity interventions and optimize daily routines, further assisting staff members and organizational objectives.

BIBLIOGRAPHIC REFERENCES

1. George, T.J., Atwater, L.E., Maneethai, D. and Madera, J.M., 2022. Supporting the productivity and wellbeing of remote workers: Lessons from COVID-19. Organizational Dynamics, 51(2), p.100869. https://doi.org/10.1016/j.orgdyn.2021.100869

2. Kurdy, D.M., Al-Malkawi, H.A.N. and Rizwan, S., 2023. The impact of remote working on employee productivity during COVID-19 in the UAE: the moderating role of job level. Journal of Business and Socio-economic Development, 3(4), pp.339-352. https://doi.org/10.1108/JBSED-09-2022-0104

3. Tleuken, A., Turkyilmaz, A., Sovetbek, M., Durdyev, S., Guney, M., Tokazhanov, G., Wiechetek, L., Pastuszak, Z., Draghici, A., Boatca, M.E. and Dermol, V., 2022. Effects of the residential built environment on

remote work productivity and satisfaction during COVID-19 lockdowns: An analysis of workers' perceptions. Building and Environment, 219, p.109234. https://doi.org/10.1016/j.buildenv.2022.109234

4. Howe, L.C. and Menges, J.I., 2022. Remote work mindsets predict emotions and productivity in the home office: A longitudinal study of knowledge workers during the Covid-19 pandemic. Human-Computer Interaction, 37(6), pp.481-507. https://doi.org/10.1080/07370024.2021.1987238

5. Rañeses, M.S., Bacason, E.S. and Martir, S., 2022. Investigating the Impact of Remote Working on Employee Productivity and Work-life Balance: A Study on the Business Consultancy Industry in Dubai, UAE. International Journal of Business & Administrative Studies, 8(2). https://dx.doi.org/10.20469/ijbas.8.10002-2

6. Farooq, R. and Sultana, A., 2022. The potential impact of the COVID-19 pandemic on work-from-home and employee productivity. Measuring Business Excellence, 26(3), pp.308-325. http://dx.doi.org/10.1108/MBE-12-2020-0173

7. Demerouti, E., 2023. Effective employee strategies for remote working: An online self-training intervention. Journal of Vocational Behavior, 142, p.103857. https://doi.org/10.1016/j.jvb.2023.103857

8. Kowalski, G. and Ślebarska, K., 2022. Remote working and work effectiveness: a leader perspective. International Journal of Environmental Research and Public Health, 19(22), p.15326. https://doi.org/10.3390/ijerph192215326

9. Hegde, N.P., Vikkurty, S., Kandukuri, G., Musunuru, S. and Hegde, G.P., 2022. Employee sentiment analysis towards remote work during COVID-19 using Twitter data. International Journal of Intelligent Engineering and Systems, 15(1), pp.75-84. DOI: 10.22266/ijies2022.0228.08

10. Sungheetha, A. and Sharma, R., 2020. A comparative machine learning study on IT sector edge nearer to working from home (WFH) contract category for improving productivity. Journal of Artificial Intelligence, 2(04), pp.217-225. https://doi.org/10.36548/jaicn.2020.4.004

11. Patil, D., 2024. Human-Artificial Intelligence Collaboration In The Modern Workplace: Maximizing Productivity And Transforming Job Roles. Available at SSRN 5057414. https://dx.doi.org/10.2139/ssrn.5057414

12. Shaikh, F., Afshan, G., Anwar, R.S., Abbas, Z. and Chana, K.A., 2023. Analyzing the impact of artificial intelligence on employee productivity: the mediating effect of knowledge sharing and well-being. Asia Pacific Journal of Human Resources, 61(4), pp.794-820. https://doi.org/10.1111/1744-7941.12385

13. Bijalwan, P., Gupta, A., Mendiratta, A., Johri, A. and Asif, M., 2024. Predicting the productivity of municipality workers: A comparison of six machine learning algorithms. Economies, 12(1), p.16. https://doi.org/10.3390/economies12010016

14. Pryiatelchuk, O. and Aizenberh, T., 2024. Implementation of Ai Systems For Analysis of Productivity And Development Of Talent In Global Teams. Actual Problems of International Relations, 1(161), pp.129-135. https://doi.org/10.17721/apmv.2024.161.1.129-135

15. Simoncelli, G., Bernardi, M.L., De Angelis, L., Anastasi, S., Bonafede, M., Artenio, E. and Pecori, R., 2024. Predictive functions of artificial intelligence for risk assessment in remote hybrid work. Artificial Intelligence and Social Computing, 122(122). https://doi.org/10.54941/ahfe1004639

16. Javaid, M., Haleem, A., Singh, R.P., Suman, R. and Gonzalez, E.S., 2022. Understanding the adoption of Industry 4.0 technologies in improving environmental sustainability. Sustainable operations and computers, 3, pp.203-217. https://doi.org/10.1016/j.susoc.2022.01.008

17. Awada, M., Becerik-Gerber, B., Lucas, G. and Roll, S.C., 2023. Predicting office workers' productivity: A machine learning approach integrating physiological, behavioral, and psychological indicators. Sensors, 23(21), p.8694. https://doi.org/10.3390/s23218694

18. George, A.S., 2024. Automated Futures: Examining the Promise and Peril of AI on Jobs, Productivity, and Work-Life Balance. Partners Universal Innovative Research Publication, 2(6), pp.1-17. DOI:10.5281/zenodo.14544519

19. Awan, W.N., Paasivaara, M., Gloor, P.A. and Salman, I., 2023. Creating Happier and More Productive Software Engineering Teams through AI and Machine Learning. In ICSOB Companion.

20. Razali, M.N., Ibrahim, N., Hanapi, R., Zamri, N.M. and Manaf, S.A., 2023. Exploring Employee Working Productivity: Initial Insights from Machine Learning Predictive Analytics and Visualization. Journal of Computing Research and Innovation, 8(2), pp.235-245. https://dx.doi.org/10.24191/jcrinn.v8i2.362

FINANCING
No financing.

CONFLICT OF INTEREST
The authors declare that there is no conflict of interest.

AUTHORSHIP CONTRIBUTION
Data curation: Si Chen.
Formal analysis: Si Chen.
Project management: Si Chen.
Supervision: Si Chen.
Display: Si Chen.
Drafting - original draft: Si Chen.
Writing - proofreading and editing: Si Chen.