NUANSA INFORMATIKA ***p-ISSN :1858-3911*** **,** ***e-ISSN : 2614-5405***
Volume 20 Number 2 July 2026 https://journal.fkom.uniku.ac.id/ilkom

Hybrid PSO-Adaptive Boosting Regression for Employee Salary Prediction and Recommendation

Muhammad Amran Hakim Siregar*[1], Bachtiar Ramadhan[2], Syafrial Fachri Pane[3]
[1] Faculty of Information Technology & Actuarial Science, Computer Science,
Mitra Bangsa University, Indonesia
[2,3] School of Information Technology, Bachelor of Applied Informatics,
University of Logistics and International Business, Indonesia
E-mail: *[1]amranhakimsiregar@umiba.ac.id

Abstract

Recommending appropriate employee salaries is important for supporting employee performance and data-driven managerial decisions. This study develops a hybrid machine learning model to recommend employee salaries and identify influential factors affecting monthly income. The dataset was obtained from Kaggle and consisted of 1,029 employee records with 34 variables covering company, personal, and demographic characteristics. Data preprocessing included categorical encoding, missing-value handling, duplicate checking, and outlier removal using the Interquartile Range method. The proposed approach combines Particle Swarm Optimization for variable optimization with an AdaBoost Regressor selected through TPOT Regression. Model performance was evaluated using R-Square and Mean Absolute Percentage Error. The PSO-AdaBoost Regressor achieved an R-Square value of 0.88 and a MAPE value of 0.22. Feature importance analysis identified Job Level as the most influential feature, with a score of 0.97156. The results were implemented in a Django-based web application.

Keywords*—Salary Recommendations, AdaBoost Regression, Particle Swarm Optimization, TPOT Regression, Django Framework*

Submission: February 15, 2026 Accepted: June 18, 2026 Published: July 15, 2026
DOI Article: https://doi.org/*10.25134/ilkom.v20i2.561*

1. INTRODUCTION

The development of digital technology in the Industry 4.0 era has encouraged organizations to adopt data-driven decision-making in various managerial activities, including human resource management [1], [2], [3], [4], [5]. One important aspect of human resource management is employee compensation. Salary is closely related to employee welfare, motivation, productivity, and organizational performance [6], [7]. Therefore, determining an appropriate salary is an important managerial decision. However, salary recommendation is not a simple process because it involves various factors, such as employee age, job position, work experience, job level, education, and other organizational attributes.

In practice, many companies still face challenges in determining salary recommendations objectively. Salary decisions that are not supported by data-driven analysis may lead to inconsistency between employee characteristics and compensation levels. This condition can affect employee satisfaction and performance. Therefore, a predictive model is needed to assist companies in analyzing employee-related factors and generating salary recommendations based on available data.

Previous research has modeled employee salary prediction using several factors, including age, employee position, length of service, work experience, and salary. These factors were examined using the Spearman correlation approach, and the prediction model was developed using multivariate linear regression [8]. The prediction results were also visualized in real time using a Django-based web application. However, the previous model still had limitations. The dataset contained outliers, which reduced model performance and affected prediction accuracy [9], [10]. In addition, linear regression has limitations in capturing complex and non-linear relationships among employee attributes. These limitations indicate the need for a more robust prediction approach.

To address these problems, this study proposes a hybrid approach that combines Particle Swarm Optimization (PSO) and AdaBoost Regression. A hybrid approach integrates different computational methods to improve model performance and optimization

p-ISSN :1858-3911 **,** ***e-ISSN : 2614-5405***
https://journal.fkom.uniku.ac.id/ilkom

capability [11], [12]. PSO is used to optimize influential variables because it is an adaptive stochastic metaheuristic optimization algorithm that can search for optimal solutions in complex data spaces [13], [14], [15]. Meanwhile, AdaBoost Regression is used as the prediction model because it can improve predictive performance by combining multiple weak learners into a stronger model. In this study, TPOT Regression is also used to select the most suitable regression pipeline for the dataset.

The proposed PSO–AdaBoost Regression model is expected to improve employee salary prediction by optimizing relevant variables and producing more reliable prediction results. The model performance is evaluated using R-Square and Mean Absolute Percentage Error (MAPE). A good regression model is indicated by an R-Square value close to 1, while MAPE values should be close to 0 [16], [17], [18]. In addition, feature importance analysis is used to identify the most influential factors affecting monthly income.

The contributions of this study are as follows. First, this study proposes a hybrid PSO–AdaBoost Regression approach for employee salary prediction and recommendation. Second, this study evaluates the proposed model using standard regression evaluation metrics, namely R-Square and MAPE. Third, this study identifies influential salary-related factors using feature importance analysis. Fourth, this study implements the prediction results into a web-based application using the Django framework to support salary recommendation and data visualization.

2. METHOD

Figure 1 shows the research methodology used in this study.

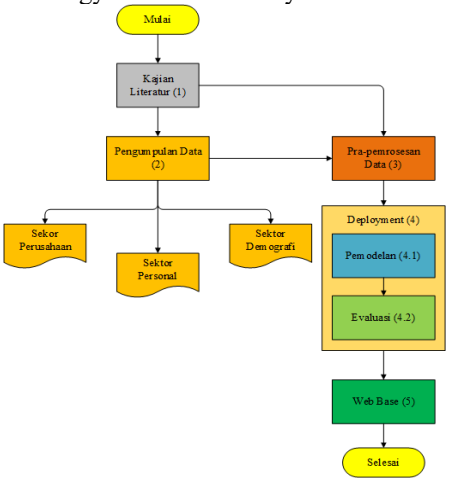


Figure 1. Research Method

To build a hybrid machine learning model for predicting employee salaries, the following stages were conducted.

1) Literature review: relevant studies related to salary prediction, machine learning regression, feature optimization, and hybrid machine learning approaches were collected and reviewed.
2) Data collection, using an employee dataset obtained from Kaggle. The dataset consisted of 1,029 records and 34 variables grouped into corporate, personal, and demographic attributes. The target variable in this study is Monthly Income.
3) Data preprocessing, including categorical data encoding, missing value handling, duplicate checking, and outlier removal using the Interquartile Range (IQR) method. After preprocessing, the dataset was divided into training and testing sets using an 80:20 train-test split. To avoid data leakage, feature optimization and model selection were conducted only on the training data, while the testing data was used only for final evaluation.
4) Modeling and evaluation consisted of feature optimization, model selection, and model performance evaluation.
   a. Feature optimization and model selection were conducted using PSO and TPOT Regression. PSO was used to identify the most relevant variables for predicting Monthly Income, with the fitness function based on prediction error, where a lower error value indicated a better feature subset. After the optimization process, TPOT Regression was applied to select the best regression pipeline using 5 generations and a population size of 20. The selected model was AdaBoost Regressor with n_estimators = 100, learning_rate = 0.001, and loss = exponential.
   b. Model evaluation using R-Square and Mean Absolute Percentage Error (MAPE) to measure the predictive performance of the proposed PSO–AdaBoost Regression model.
5) Web-based implementation, implementing the hybrid method into a web-based application using the Django framework to generate salary recommendations and visualize the prediction results

p-ISSN :1858-3911* , *e-ISSN : 2614-5405
https://journal.fkom.uniku.ac.id/ilkom

To complete the research method in accordance with the steps above, several methods are needed, as described below.

2.1. Interquartile Range (IQR)

The IQR method provides a visualization of outlier distribution using box plots. This makes it easier to see data that falls within the quartile range [19]. Mathematically, IQR can be written as equation 1.

$$IQR = Q3 - Q1 \quad (1)$$

To determine outliers, it is necessary to define the range or limits of the data. Data that falls outside these limits is referred to as an outlier. For the IQR method, this consists of the upper limit (Q3) and lower limit (Q1). Data that is smaller than Q1 or larger than Q3 is considered an outlier.

2.2. Tree-Based Pipeline Optimizaation Tool (TPOT)

Tree-Based Pipeline Optimization Tool (TPOT) is an automated machine learning tool that uses genetic programming to identify an appropriate machine learning pipeline for a given dataset [20]. TPOT explores several combinations of preprocessing techniques, feature selection methods, and regression models to obtain a pipeline that provides good predictive performance while maintaining reasonable model complexity [21]. In this study, TPOT Regression was used to support model selection for employee salary prediction. This is important because salary prediction involves various employee attributes that may have different relationships with monthly income. By using TPOT, the model selection process is not determined manually, but is guided by data-driven pipeline optimization. Based on the TPOT Regression process, AdaBoost Regressor was selected as the most suitable regression model for the dataset used in this study.

2.3. Adaptive Boosting Regression

Adaptive Boosting, or AdaBoost, is an ensemble learning method that combines several weak learners to form a stronger predictive model [20]. In the regression context, AdaBoost Regressor works by sequentially training weak regression models and giving higher attention to observations with larger prediction errors. The final prediction is obtained by combining the results of multiple weak regressors through a weighted aggregation mechanism. Therefore, AdaBoost Regression is different from AdaBoost classification because the main objective is to predict continuous numerical values rather than class labels.

Figure 2 shows the AdaBoost process.

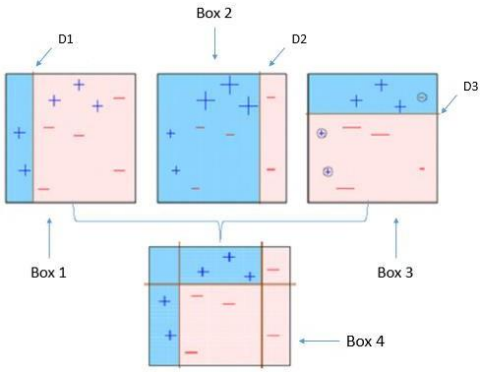


Figure 2. Adaboost Regression Process

AdaBoost Regression has several advantages, including its ability to improve prediction performance by combining multiple weak learners, its relative robustness to overfitting, and its interpretability through feature importance analysis [22]. AdaBoost Regressor was selected in this study because the target variable, Monthly Income, is a continuous variable. Salary prediction may involve complex and non-linear relationships among factors such as job level, work experience, job role, education, and other employee attributes. Compared with simple linear regression, AdaBoost Regressor can capture more flexible relationships in the data by combining multiple weak prediction models. In addition, AdaBoost provides feature importance information, which is useful for identifying the most influential factors in salary recommendation.

2.4. Particle Swarm Optimization (PSO)

Particle Swarm Optimization (PSO) is a population-based metaheuristic optimization method inspired by the social behavior of bird flocks or fish schools [23]. In PSO, each particle represents a possible solution in the search space. Each particle updates its position based on its own best experience and the best solution found by the swarm. The velocity and position of each particle are updated using the following equations [24] :

$$v_{ij}^{k+1} = v^k + c1\left(x_{ij}^{pb} - x_{ij}^k\right) + c2(x_{ij}^{gb} - x_{ij}^k) \quad (2)$$

p-ISSN :1858-3911 **,** ***e-ISSN : 2614-5405***
https://journal.fkom.uniku.ac.id/ilkom

$$x_{ij}^{k+1} = x_{ij}^{k} + v_{ij}^{k+1} \quad (3)$$

where v_{ij} represents the velocity of particle i in dimension j, x_{ij} represents the position of particle i in dimension j, w is the inertia weight, c_1 and c_2 are acceleration coefficients, r_1 and r_2 are random values between 0 and 1, $pbest_{ij}$ is the best position found by the particle, and $gbest_j$ is the best position found by the swarm. In this study, PSO was selected because employee salary prediction involves many independent variables, and not all variables contribute equally to Monthly Income. PSO is used to optimize the search for influential variables so that the regression model can focus on more relevant factors. By combining PSO with AdaBoost Regressor, this study aims to improve prediction performance and obtain more reliable salary recommendations.

2.5. Mean Absolute Percentage Error (MAPE)

Mean Absolute Percentage Error (MAPE) is an evaluation metric used to measure prediction error in percentage form. MAPE shows the average percentage difference between actual values and predicted values. The formula of MAPE is written as follows [25] :

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{|y'_t - y_t|}{y'_t} \times 100\% \quad (4)$$

where y_i is the actual value, $\hat{y}_i$ is the predicted value, and n is the number of observations. In this study, MAPE was used because it provides an interpretable error measurement for salary prediction. A lower MAPE value indicates that the predicted salary is closer to the actual salary value.

2.6. R-Square (R2)

R-Square, or the coefficient of determination, is used to measure how well the independent variables explain the variation in the dependent variable. The formula of R-Square is written as follows [26] :

$$R^2 = 1 - \frac{\sum(y_i - p_i)^2}{\sum(y_i - a)^2} \quad (5)$$

where y_i is the actual value, $\hat{y}_i$ is the predicted value, and $\bar{y}$ is the mean of the actual values. An R-Square value closer to 1 indicates that the model explains a greater proportion of the variation in Monthly Income.

2.7. Django Framework

Django is a Python-based web framework that supports rapid web application development using the Model-View-Template architecture [27]. In this study, Django was used as the implementation framework because it is compatible with Python-based machine learning workflows. The use of Django allows the prediction model to be integrated into a web-based application, making the salary recommendation results easier to access and visualize. Therefore, Django is not used as a prediction method, but as a deployment and visualization framework to present the output of the PSO–AdaBoost Regression model in a practical system.

3. RESULTS

3.1. Dataset and Data Preprocessing

In this study, the dataset used for hybrid modeling was an employee dataset from Kaggle consisting of 1,029 records and 34 variables with three categories of variable, namely company data, personal data, and demographic data. Before the modeling process, data preprocessing was conducted by encoding categorical variables into numerical values, handling missing values, checking duplicate data, and removing outliers using the Interquartile Range (IQR) method.

The categorical variables, such as Department, Education Field, Gender, Job Role, Marital Status, Over 18, and Over Time, were converted into numerical values. Based on the initial data inspection, no duplicate records were found, while missing values were found in 0.8% of the data. These missing values were handled using mean imputation. Furthermore, 114 outlier records were detected using the IQR method. After removing these outlier records, the final dataset used for modeling consisted of 915 records.

Figure 3 shows the boxplot of Monthly Income before outlier removal. The boxplot was used to identify extreme values outside the lower and upper quartile boundaries. Removing outliers reduced extreme data variation and helped improve the stability of the modeling process, especially because salary prediction is sensitive to unusually high or low values. Therefore, the final dataset was considered cleaner and more suitable for PSO-based feature optimization and AdaBoost Regression modeling.

p-ISSN :1858-3911* , *e-ISSN : 2614-5405
https://journal.fkom.uniku.ac.id/ilkom

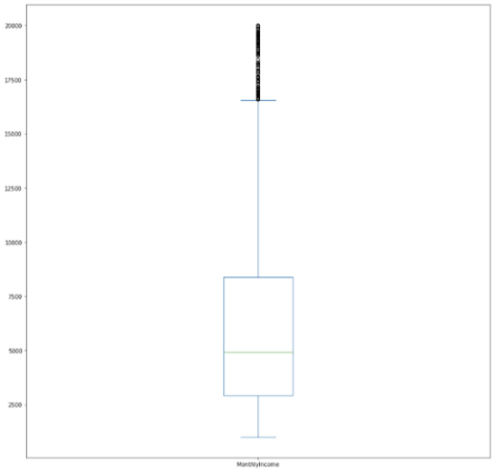


Figure 3. Boxplot of Monthly Income before Outlier Removal

Data that falls outside of the first and third quartiles is referred to as outlier data. In order to exclude the outlier data from the dataset, data that falls between the first and third quartiles is then collected.

3.2. Modelling

After the preprocessing stage, the cleaned dataset was used to develop the proposed hybrid model. The model development process consisted of three sequential stages: feature optimization using Particle Swarm Optimization (PSO), model selection using TPOT Regression, and final modeling using the regression model selected by TPOT. First, PSO was applied to optimize the feature selection process for predicting Monthly Income. Each particle represented a candidate feature subset, and the fitness function was based on prediction error, where a lower error value indicated a better feature subset.

In this study, the PSO parameters were set as follows: number of particles = 10, number of iterations = 3500, inertia weight = 0.3647, cognitive coefficient = 0.3365, and social coefficient = 0.0772. Based on the PSO process, the best cost value obtained was 188.87. The selected feature obtained from the PSO optimization process was Job Level. The optimized feature result was then used in the next stage for model selection and regression modeling.

3.2.1. Variable Optimization Using Particle Swarm Optimization

Particle Swarm Optimization (PSO) was used to optimize the selection of relevant features for employee salary prediction. This process was conducted before the final regression modeling stage to identify variables that contributed to predicting Monthly Income. Based on the PSO calculation, the best cost value obtained was 188.87, as shown in Figure 4.

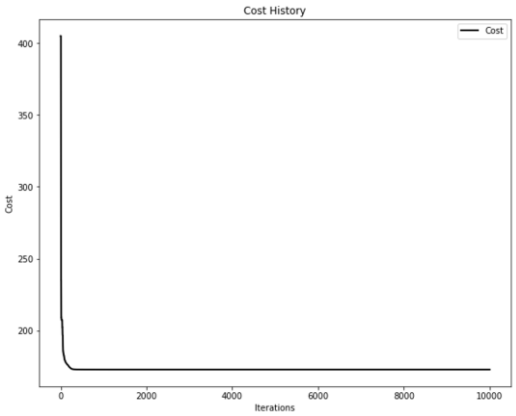


Figure 4. Cost History Graph

3.2.2. Model Selection using TPOT Regression

After the feature optimization process, TPOT Regression was used to select the best regression pipeline. The TPOT configuration used 5 generations and a population size of 20. Since this study is a regression problem, the model selection process was evaluated using regression-based scoring. The result showed that AdaBoost Regressor was selected as the best model for employee salary prediction. The selected AdaBoost Regressor parameters were n_estimators = 100, learning_rate = 0.001, and loss = exponential.

3.2.3. Modelling using AdaBoost Regression

The final modeling stage was performed using the AdaBoost Regressor selected by TPOT Regression. The model was trained using n_estimators = 100, learning_rate = 0.001, and loss = exponential. AdaBoost Regressor was used because the target variable, Monthly Income, is a continuous variable. After the model was trained, feature importance analysis was conducted to identify the relative contribution of each variable to the target variable. The feature importance scores are presented in Table 1.

p-ISSN :1858-3911* , *e-ISSN : 2614-5405
https://journal.fkom.uniku.ac.id/ilkom

No	Feature	Score
1	Age	0,00011
2	Daily Rate	0,00042
3	Department	0,00027
4	Distance From Home	0,00026
5	Education	0,00004
6	Education Field	0,00017
7	Employee Count	0,00000
8	Employee Number	0,00122
9	Environment Satisfaction	0,00000
10	Gender	0,00000
11	Hourly Rate	0,00022
12	Job Involvement	0,00000
13	Job Level	0,97156
14	Job Role	0,00250
15	Job Satisfaction	0,00000
16	Marital Status	0,00005
17	Monthly Rate	0,00117
18	Num Companies Worked	0,00040
19	Over 18	0,00000
20	Over Time	0,00002
21	Percent Salary Hike	0,00122
22	Performance Rating	0,00003
23	Relationship Satisfaction	0,00000
24	Standard Hours	0,00000
25	Stock Option Level	0,00005
26	Total Working Years	0,01769
27	Training Times Last Year	0,00053
28	Work Life Balance	0,00017
29	Years At Company	0,00020
30	Years In Current Role	0,00110
31	Years Since Last Promotion	0,00009
32	Years With Curr Manager	0,00046

Table I : Feature Importance Scores of AdaBoost Regressor

From Table I, Job Level had the highest feature importance score, with a value of 0.97156. This result indicates that Job Level was the most influential feature in predicting Monthly Income. Therefore, Job Level can be considered the most important factor in the employee salary recommendation model. However, because the feature importance score of Job Level was dominant, this result also indicates that the model strongly depends on this variable. This finding is reasonable because employees with higher job levels generally receive higher monthly income. Therefore, the relationship between Job Level and Monthly Income should be interpreted carefully. After the feature importance analysis, model evaluation was conducted to determine the predictive performance of the proposed PSO–TPOT–AdaBoost Regression model.

3.3. Model Evaluation

The hybrid model was evaluated using R-Square and Mean Absolute Percentage Error (MAPE). These metrics were used to provide a complete evaluation of the regression model. R-Square was used to measure how well the independent variables explained the variation in Monthly Income. MAPE was used to measure the prediction error in percentage form.

The proposed PSO–TPOT–AdaBoost Regressor model obtained an R-Square value of 0.88, indicating that 88% of the variation in Monthly Income could be explained by the independent variables used in the model. The remaining 12% may be influenced by other factors not included in the dataset. The model also obtained a MAPE value of 0.22, equivalent to 22%. To maintain consistency, all MAPE values in this study are reported in the same format. Based on these evaluation results, the proposed hybrid model showed good performance for employee salary prediction.

The comparison results between the proposed model and baseline models are presented in Table 2. All models were evaluated using the same evaluation metrics, namely R-Square and MAPE, to ensure a fair comparison.

Model	R2	MAPE
PSO-Linear Regression	0.80	1.04
PSO-Decision Tree Regression	0.72	2.01
PSO-Random Forest Regression	0.85	1.18
PSO-AdaBoost Regression	**0.88**	**0.22**

Table II : Comparison Regression Model Performance

Based on Table II, the proposed PSO–AdaBoost Regressor model produced competitive performance compared with the baseline models. The use of PSO contributed to feature optimization, TPOT Regression contributed to regression model selection, and AdaBoost Regressor was used as the final model selected by TPOT. Therefore, the proposed hybrid approach can be considered suitable for employee salary prediction.

p-ISSN :1858-3911 **,** ***e-ISSN : 2614-5405***
https://journal.fkom.uniku.ac.id/ilkom

3.4. Django Framework Implementation

The final PSO–AdaBoost Regressor model was implemented into a web-based application using the Django framework. Django was used as the deployment framework because it supports Python-based machine learning integration and allows the prediction model to be accessed through a web interface. The system was designed to receive employee-related input variables, process the input through the trained prediction model, and display the predicted salary recommendation as output. The workflow of the web-based system consists of three main stages. First, the user inputs employee data through the web form, such as Job Level and other required employee attributes. Second, the Django backend processes the input data and sends it to the trained prediction model. Third, the system displays the salary prediction result and related visualization to help users understand the recommendation output.

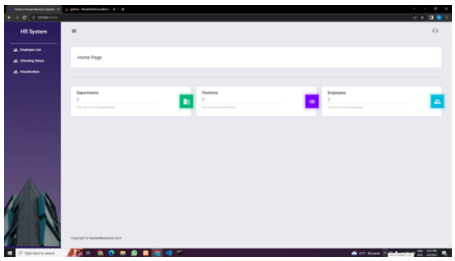


Figure 5. Salary Prediction Result Page

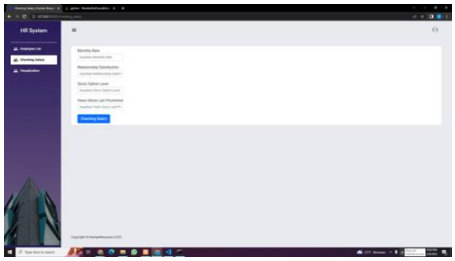


Figure 6. Employee Data Input Page

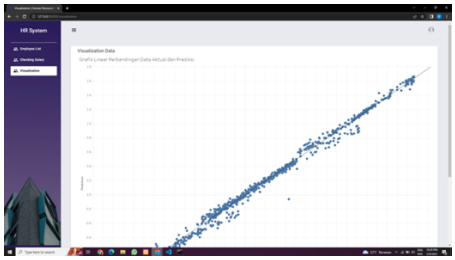


Figure 7. Salary Recommendation Visualization Page

The implementation interface is shown in Figures 5, 6, and 7. Figure 5 shows the input page for entering employee data. Figure 6 shows the prediction result page generated by the model. Figure 7 shows the visualization page used to present salary recommendation results. Therefore, the Django implementation functions as a practical interface for integrating the machine learning model into a salary recommendation system.

4. CONCLUSIONS

This study proposed a hybrid machine learning approach using Particle Swarm Optimization (PSO), TPOT Regression, and AdaBoost Regressor to support employee salary recommendation. Based on the modeling results, TPOT Regression selected AdaBoost Regressor as the best regression model, while PSO was used to optimize the feature selection process. The model evaluation showed that the proposed PSO–AdaBoost Regressor achieved an R-Square value of 0.88 and a MAPE value of 0.22, indicating that the model had good predictive performance in estimating Monthly Income.

The feature importance analysis showed that Job Level was the most influential variable, with the highest feature importance score of 0.97156. This result indicates that Job Level has a strong relationship with salary prediction because higher job levels are generally associated with higher responsibilities and compensation. However, this also shows that the model may be highly dependent on this variable, so the prediction results should be interpreted carefully.

The Django-based implementation shows that the proposed model can be integrated into a web-based salary recommendation system. However, the system should be used as a decision support tool, not as the sole basis for salary decisions. Final salary decisions should still consider other factors that may not be included in the dataset, such as employee performance, competency, workload, company policy, fairness, and managerial judgment. Future research is recommended to use larger and more diverse datasets, include additional human resource variables, and compare the proposed approach with other optimization or ensemble learning methods.

p-ISSN :1858-3911* , *e-ISSN : 2614-5405
https://journal.fkom.uniku.ac.id/ilkom

5. SUGGESTIONS

The following are some recommendations for additional research based on the findings of this study :

1. In order for the salary recommendation model to more accurately reflect actual conditions, future studies are recommended to include other non-technical factors including competence, certification, job risk level, and individual performance indicators (KPIs).
2. To examine the potential for increasing the model's accuracy and stability, PSO may be compared with other optimization techniques like Genetic Algorithm (GA), Ant Colony Optimization (ACO), or Hybrid PSO-GA.
3. The model can be tested on cross-industry and cross-country datasets to measure the generalization of the model in various organizational structures and wage systems.
4. Further research can also develop deep learning or multi-model ensemble approaches as comparators to the PSO–AdaBoost Regression model in big data scenarios.
5. The implementation of a web-based system can be expanded with explainable AI (XAI) features so that management can understand the reasoning behind the salary recommendations generated by the model.

REFERENCES

[1] B. Prasetyo and U. Trisyanti, "Revolusi Industri 4.0 dan Tantangan Perubahan Sosial," *Prosiding SEMATEKSOS 3*, pp. 22–27, 2018.

[2] H. Prasetyo and W. Sutopo, "Perkembangan Keilmuan Teknik Industri Menuju Era Industri 4.0," Surakarta, May 2017.

[3] O. C. Pangaribuan and I. Irwansyah, "Media Cetak Indonesia di Era Revolusi Industri 4.0," *Jurnal Pewarta Indonesia*, vol. 1, no. 2, pp. 134–145, Oct. 2019, doi: 10.25008/jpi.v1i2.11.

[4] A. A. Shahroom and N. Hussin, "Industrial Revolution 4.0 and Education," *International Journal of Academic Research in Business and Social Sciences*, vol. 8, no. 9, Oct. 2018, doi: 10.6007/ijarbss/v8-i9/4593.

[5] S. Kergroach, "Industry 4.0: New challenges and opportunities for the labour market," *Foresight and STI Governance*, vol. 11, no. 4, pp. 6–8, 2017, doi: 10.17323/2500-2597.2017.4.6.8.

[6] Y. Adrianova Eka Tuah, P. Studi Pendidikan Komputer, and S. Persada Khatulistiwa Sintang, "Implementasi Model Regresi Linear Sederhana Untuk Prediksi Gaji Berdasarkan Pengalaman Laman Bekerja," 2020.

[7] M. I. Manda and S. Ben Dhaou, "Responding to the challenges and opportunities in the 4th industrial revolution in developing countries," in *ACM International Conference Proceeding Series*, Association for Computing Machinery, 2019, pp. 244–253. doi: 10.1145/3326365.3326398.

[8] A. N. Puteri, A. Arizal, and A. D. Achmad, "Feature Selection Correlation-Based pada Prediksi Nasabah Bank Telemarketing untuk Deposito," *Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, vol. 20, no. 2, pp. 335–342, May 2021, doi: 10.30812/matrik.v20i2.1183.

[9] P. N. Dacipta and R. E. Putra, "Sistem Klasifikasi Limbah Menggunakan Metode Convolutional Neural Network (CNN) Pada Web Service Berbasis Framework Flask," *Journal of Informatics and Computer Science*, vol. 03, 2022.

[10] S. P. Utami and G. Putra, "Model ARIMA Musiman dengan Outlier untuk Peramalan Penumpang Pesawat Terbang Domestik Bandara Soekarno Hatta di Masa Pandemi Covid-19," *Jurnal Perhubungan Udara*, pp. 17–26, 2021, doi: 10.25104/wa.v.

[11] B. Wang, W. Li, A. Bradlow, E. Bazuaye, and A. T. Y. Chan, "Improving triaging from primary care into secondary care using heterogeneous data-driven hybrid machine learning," *Decis. Support Syst.*, p. 113899, Nov. 2022, doi: 10.1016/j.dss.2022.113899.

[12] S. Fachri Pane, M. Dwi Sulistiyo, and A. Akbar Gozali, "Multi-Temporal Factors to Analyze Indonesian Government

p-ISSN :1858-3911* , *e-ISSN : 2614-5405
https://journal.fkom.uniku.ac.id/ilkom

Policies regarding Restrictions on Community Activities during COVID-19 Pandemic," *International Journal on Informatics Visualization*, pp. 2263–2269, Dec. 2023, doi: https://doi.org/10.47738/jads.v6i4.981.

[13] A. Khare and S. Rangnekar, "A review of particle swarm optimization and its applications in Solar Photovoltaic system," 2013, *Elsevier Ltd*. doi: 10.1016/j.asoc.2012.11.033.

[14] S. F. Pane, M. D. Sulistiyo, A. A. Gozali, and Adiwijaya, "Hybrid Multi-Objective Metaheuristic Machine Learning for Optimizing Pandemic Growth Prediction," *Journal of Applied Data Sciences*, vol. 6, no. 4, pp. 2938–2958, Dec. 2025, doi: 10.47738/jads.v6i4.981.

[15] S. Sarkar, S. Vinay, R. Raj, J. Maiti, and P. Mitra, "Application of optimized machine learning techniques for prediction of occupational accidents," *Comput. Oper. Res.*, vol. 106, pp. 210–224, Jun. 2019, doi: 10.1016/j.cor.2018.02.021.

[16] F. Fangestu, Y. Ade Putra, L. Liawardi, and H. Adam Afiezan, "Pengaruh Kinerja Keuangan, Tata Kelola Perusahaan Yang Baik (GCG) dan Tanggung Jawab Sosial Perusahaan (CSR) Terhadap Nilai Perusahaan Pertambangan Yang Ada di Bursa Efek Indonesia Pada Periode 2014-2018," *Jurnal Ilmiah MEA (Manajemen, Ekonomi, dan Akuntansi)*, vol. 4, no. 3, pp. 777–793, 2020.

[17] A. Bode, "K-Nearest Neighbor Dengan Feature Selection Menggunakan Backward Elimination Untuk Prediksi Harga Komoditi Kopi Arabika," *ILKOM Jurnal Ilmiah*, vol. 9, pp. 188–195, Aug. 2017.

[18] O. Somantri, "Prediksi Kekuatan Gempa Bumi Indonesia Berdasarkan Nilai Magnitudo Menggunakan Neural Network," *Seminar Nasional Informatika Bela Negara (SANTIKA)* , vol. 2, pp. 203–207, 2021.

[19] R. Siringoringo, R. Perangin Angin, and B. Rumahorbo, "Model Klasifikasi Genetic-XGBoost Dengan T-Distributed Stochastic Neighbor Embedding Pada Peramalan Pasar," *Jurnal Times*, vol. XI, no. 1, 2022, [Online]. Available: https://archive.ics.uci.edu/ml/datasets/online+retail

[20] R. S. Olson, N. Bartley, R. J. Urbanowicz, and J. H. Moore, "Evaluation of a tree-based pipeline optimization tool for automating data science," in *GECCO 2016 - Proceedings of the 2016 Genetic and Evolutionary Computation Conference*, Association for Computing Machinery, Inc, Jul. 2016, pp. 485–492. doi: 10.1145/2908812.2908918.

[21] J. D. Romano, T. T. Le, W. Fu, and J. H. Moore, "TPOT-NN: augmenting tree-based automated machine learning with neural network estimators," *Genet. Program. Evolvable Mach.*, vol. 22, no. 2, pp. 207–227, Jun. 2021, doi: 10.1007/s10710-021-09401-z.

[22] A. M. Majid and M. N. Dwi Miharja, "Penerapan Metode Discretization dan Adaboost untuk Meningkatkan Akurasi Algoritma Klasifikasi dalam Memprediksi Penyakit Jantung," *Indonesian Journal of Business Intelligence (IJUBI)*, vol. 5, no. 2, p. 70, Dec. 2022, doi: 10.21927/ijubi.v5i2.2689.

[23] J. C. Bansal, "Particle swarm optimization," in *Studies in Computational Intelligence*, vol. 779, Springer Verlag, 2019, pp. 11–23. doi: 10.1007/978-3-319-91341-4_2.

[24] Y. Liu *et al.*, "Optimization of five-parameter BRDF model based on hybrid GA-PSO algorithm," *Optik (Stuttg).*, vol. 219, Oct. 2020, doi: 10.1016/j.ijleo.2020.164978.

[25] K. Warneke, S. D. Siegel, J. Afonso, and S. Wallot, "What the mean absolute percentage error (MAPE) should adopt from Bland–Altman analyses," *German Journal of Exercise and Sport Research*, Dec. 2025, doi: 10.1007/s12662-025-01084-3.

[26] T. Raykov and C. DiStefano, "Evaluating Change in Adjusted R-Square and R-Square Indices: A Latent Variable Method Application," *Educ. Psychol. Meas.*, vol. 86, no. 3, pp. 627–

p-ISSN :1858-3911 **,** ***e-ISSN : 2614-5405***
https://journal.fkom.uniku.ac.id/ilkom

641, Jun. 2026, doi: 10.1177/00131644251329178.

[27] B. Putra Aji, A. Hernawan, and A. Nurjihadi, "Sistem Informasi Surat Elektronik Untuk Akademik Mataram (Dengan Python Django Framework)," *JBegaTI*, vol. 3, no. 2, pp. 252–262, 2022, [Online]. Available: http://begawe.unram.ac.id/index.php/JBTI/