DIJEMSS: Dinasti International Journal of Education Management and Social Science
E-ISSN: 2686-6331
P-ISSN: 2686-6358
https://dinastipub.org/DIJEMSS dinasti.info@gmail.com +62 811 7404 455
DINASTI PUBLISHER

DOI: https://doi.org/10.38035/dijemss.v6i4

Enhancing Organizational Performance Through KPI-Based Employee Prediction Using the C4.5 and Random Forest Methods

Grasiella Yustika Rezka Talita Kusumaningtyas[1], Agung Triayudi[2]
[1]Magister Teknologi Informasi, Universitas Nasional, Jakarta, Indonesia, 2018grasiella@student.unas.ac.id
[2]Magister Teknologi Informasi, Universitas Nasional, Jakarta, Indonesia, Agungtriayudi@civitas.unas.ac.id

Corresponding Author: 2018grasiella@student.unas.ac.id[1]

Abstract: Employee performance evaluation is a crucial process for organizations in achieving strategic goals. PT Kreatif Dinamika Integrasi currently conducts employee assessments manually using Microsoft Excel, leading to inefficiencies, potential errors, and subjectivity. This study aims to develop a decision support system (DSS) using the C4.5 and Random Forest algorithms to improve accuracy and fairness in performance evaluation. The research adopts a quantitative approach, encompassing problem analysis, literature review, data collection, machine learning model implementation, and result evaluation. The key performance indicators (KPIs) used include discipline, punctuality, skills, appearance, and education, which serve as attributes in the classification process. The C4.5 algorithm constructs a decision tree to identify patterns, while Random Forest, as an ensemble method, enhances classification accuracy and reduces the risk of overfitting. The results indicate that using Random Forest improves evaluation accuracy from 85% to 87.5%. The implementation of this DSS provides a more reliable framework for salary and bonus prediction, minimizes bias, and enhances decision-making quality. Overall, integrating machine learning models into employee performance evaluation significantly improves efficiency and transparency.

Keyword: C4.5, Performance, Organization, Random Forest

INTRODUCTION

Employee Performance Assessment: A Strategic Component in Organizations

Employee performance assessment is a crucial component of a company, serving as both an evaluation process and a strategy to achieve corporate goals. According to experts, employee performance assessment is a system used to evaluate their performance based on predefined standards to achieve company objectives (Iman Sari, 2022). The benefits of this assessment include helping employees focus on achieving company goals and providing reports on the company's condition to stakeholders (Iman Sari, 2022). The Role of Information Technology in Performance Assessment

Business growth today is closely linked to Information Technology (IT). One of the fastest-growing professions is IT consulting, which requires strong communication skills with clients and a deep understanding of the IT products being implemented. PT Kreatif Dinamika Integrasi, an IT consulting firm collaborating with Microsoft, offers various industry solutions,

including ERP, CRM Dynamics 365, Human Resource, Power Platform, and more. These technologies also play a role in improving the quality of employee performance assessments, reducing manual errors, and accelerating decision-making processes (Kusumaningtyas & Wahyuddin, 2022).

Understanding Performance and Its Indicators

Performance is the outcome achieved by an organization, whether profit-oriented or non-profit, within a specific period. It encompasses both work results and work behaviors demonstrated by employees in fulfilling their duties and responsibilities. Performance is also the result of individual efforts in carrying out their tasks. To measure employee performance, several key indicators are used, such as quality, quantity, time (duration), cost control, supervision, and interpersonal relationships (Dwi Egie P et al., 2019).

According to Mujiastuti (2019), employee performance assessment not only measures work outcomes but also evaluates how well organizational goals are achieved through employee contributions. Performance assessments are guided by specific criteria that serve as benchmarks for success. These criteria may be determined through company discussions or by following established employee performance evaluation methods (Mujiastuti et al., 2019).

The Need for a Decision Support System in Performance Evaluation

PT Kreatif Dinamika Integrasi regularly provides salary increases for employees who have worked for at least one year. However, the salary increment and bonus determination process is still conducted manually using Microsoft Excel, without utilizing a decision support system. This manual approach has several limitations, such as data inaccuracy—where manual evaluations are prone to human errors, including incorrect data entry or bias in assessments. The manual process is also inefficient as it requires more time and effort compared to an automated system. Additionally, manual assessments may not always be objective, as they can be influenced by subjective factors, such as the relationship between the evaluator and the employee. Therefore, a more accurate and efficient decision support system is needed to enhance this process.

The use of Random Forest as a decision-making foundation greatly supports the development of a Decision Support System (DSS), as this method enables more complex and data-driven analysis to determine more objective, transparent, and measurable decisions. With highly positive evaluation results, the consistently high ROC-AUC Score of 0.96 further reinforces the conclusion that Random Forest provides stability in decision-making, which is crucial in human resource management. Alongside this, the development of this DSS aims to facilitate the employee evaluation process and reduce subjectivity in performance assessment. By utilizing decision tree-generated rules as the basis for decision-making, the system delivers fairer and data-driven decisions. Furthermore, leveraging the latest technology, this DSS can accelerate data processing and enhance efficiency in supporting employee performance management decisions.

This DSS is built by integrating the React JS framework for an interactive and responsive user interface, ASP.Net Core for business logic management on the backend, and Microsoft SQL Server as the database to store structured performance evaluation data. The research on Employee Performance Evaluation Using Key Performance Indicators (KPI) at PT. Hilfic (Happy Lailatul Fitriana et al., 2024) employed the KPI method to measure employee performance. The findings indicate a positive and significant impact on employee performance, particularly in terms of quality and quantity, after the company adopted KPI as an assessment tool. Similarly, a study by Putri Iglina Lubis and Ismu Kusumanto on Employee Performance Assessment Using the Key Performance Indicators (KPI) Method found that implementing KPI improved employee performance at CV. Bunda Bakery, with a performance improvement strategy that included monetary rewards and fulfillment of employee rights to achieve better production targets (Iglina Lubis & Kusumanto, 2018).

The study on Applying the C4.5 Algorithm to Determine the Best Employee at Klinik Hosana Jakarta Barat (Untung Surapati & Arif Prasetyo, 2022) demonstrated that the C4.5 algorithm successfully classified employees with an accuracy of 90%, aiding in the selection of the best employees based on criteria such as discipline, accuracy, tidiness, speed, and politeness. Additionally, the research Comparison of NBC, SVM, and C4.5 Models in Measuring Outstanding Employee Performance Post-COVID-19 (Galih & Mindit Eriyadi, 2022) found that SVM achieved the highest accuracy at 95%, followed by C4.5 with 90% accuracy and NBC with 85% accuracy, highlighting SVM's superiority in handling complex data. With the growing need to evaluate organizational performance through employee prediction, this research was conducted to analyze and design a web-based Decision Support System (DSS) using the C4.5 and Random Forest methods to predict employee salary increases at PT Kreatif Dinamika Integrasi. This aims to enhance accuracy and fairness in salary increment assessments through the implementation of a transparent and objective system.

METHOD

The research design is a systematic framework used to guide the entire research process. This study employs a quantitative approach, consisting of several stages: problem analysis, literature review, data collection, method implementation, and result evaluation. To clarify the stages of this research, the research flow design is presented as shown in Figure 1.

Figure 1. Research Design

The data used in this study is primary data obtained directly from PT Kreatif Dinamika Integrasi. This data includes employee performance information over a specific period, collected through documentation methods (based on annual employee performance reports), interviews with HR managers, and direct observations of employee work activities within the company environment. Initial data processing is conducted using Microsoft Excel to facilitate data cleaning and dataset preparation. Subsequently, the implementation of algorithms is carried out using the Python programming language on the Google Colaboratory (Google Colab) platform, employing two classification methods: C4.5 and Random Forest.

The C4.5 method is used to construct a decision tree model based on attributes influencing employee performance. This method involves calculating total entropy, entropy for each attribute, information gain, and gain ratio. The attribute with the highest gain ratio is selected as the root node of the decision tree. The modeling process for the C4.5 method is illustrated in Figure 2.

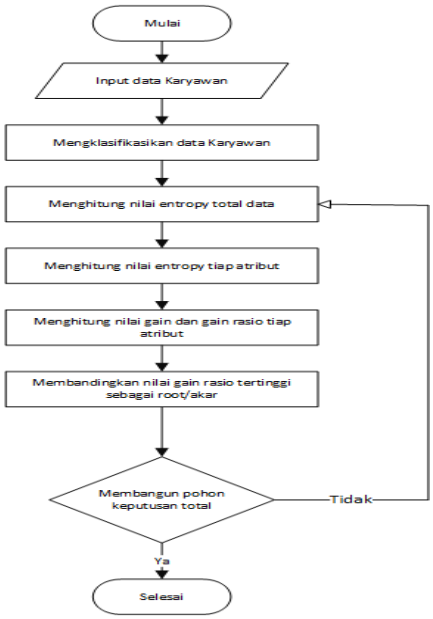

Figure 2. C4.5 Method Modeling

Meanwhile, the Random Forest method is used as an ensemble learning method by constructing multiple random decision trees to generate more stable and accurate predictions. This algorithm works by creating several decision trees based on random subsets of the training data. The results from each tree are then combined (through voting or averaging) to produce the final decision. The modeling process for the Random Forest method is illustrated in Figure 3.

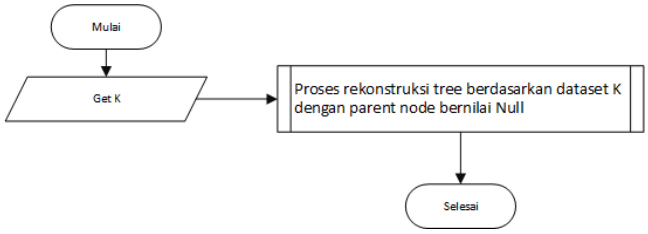

Figure 3. Random Forest Method Modeling

Both methods are then compared based on accuracy, precision, and recall to evaluate the performance of the models in classifying employee performance. The evaluation results will later serve as the basis for decision-making recommendations by the company's managem

RESULTS AND DISCUSSION

Results

The research dataset consists of 200 employee records with five main attributes: Discipline, Punctuality, Skills, Appearance, and Education. This data is used to determine the final decision—whether an employee receives a salary increase only or both a salary increase and a bonus. The dataset was collected through internal company documentation, HR interviews, and direct observations, then processed using Python and implemented into a C#-based system. The C4.5 method was applied to identify the most influential attribute in salary increase decisions. Based on entropy and information gain calculations, the attribute with the highest gain value was Discipline (0.0834), making it the root node in the decision tree model.

The Random Forest model, built based on the results of C4.5, achieved a higher accuracy of 87.5%, compared to 85% for C4.5. Model evaluation was conducted using accuracy, precision, recall, F1-score, Mean Squared Error (MSE), and Mean Absolute Error (MAE)

metrics. To provide a more comprehensive understanding of model performance, three different data scenarios were tested. Each scenario used a different combination of attributes to evaluate how attribute selection impacts model performance in classifying employee performance.

Tabel Entropy dan Information Gain untuk Setiap Atribut:

	Feature	Entropy	Information Gain
0	Kedisiplinan	1.523764	0.821544
1	Ketepatan Waktu	1.523764	0.519787
2	Keterampilan	1.523764	0.685537
3	Penampilan	1.523764	0.646652
4	Pendidikan	1.523764	0.311223

Figure 4. Entropy and Information Gain for Each Attribute

The analysis results indicate that Discipline has the greatest influence, followed by Skills and Appearance. Meanwhile, Punctuality and Education have lower contributions, with Education being the least significant factor in salary increase decisions.

Table 1. Model Evaluation Results Based on Scenarios

Skenario	Metode	Akurasi	ROC-AUC	MSE	MAE
Skenario 1	C4.5	82.00%	0.91	0.32	0.23
	Random Forest	84.50%	0.93	0.25	0.18
Skenario 2	C4.5	85.00%	0.94	0.30	0.20
	Random Forest	87.50%	0.96	0.20	0.15
Skenario 3	C4.5	83.50%	0.92	0.31	0.21
	Random Forest	86.00%	0.95	0.22	0.17

The evaluation results in Table 1 show that Random Forest consistently produces higher accuracy and ROC-AUC scores, as well as lower MSE and MAE, compared to the C4.5 method. This reinforces the finding that ensemble models are more stable and reliable for KPI-based employee performance prediction. Based on the analysis results and the decision tree visualization, several important rules were formed in the decision-making process. These rules illustrate how attributes such as Discipline, Punctuality, Skills, and Education influence the final decision regarding salary increases and bonuses. Discipline was identified as the most dominant factor, serving as the root node in the decision tree. From there, the tree establishes rules that lead to decisions such as Salary Increase or Salary Increase + Bonus, depending on the values of each evaluated attribute. These final results demonstrate how the model can provide clear, data-driven recommendations for decision-making related to employee performance evaluations.

Discussion

The discussion of this research shows that the Random Forest method is able to improve classification performance compared to C4.5. This is because Random Forest is an ensemble method that combines the strength of multiple decision trees to produce more stable and accurate predictions. In terms of evaluation metrics such as accuracy, precision, recall, and F1-score, the Random Forest method outperforms in all aspects. These findings align with previous studies by Aini et al. (2024) and Sastra Ompusunggu et al. (2023), which also found that Random Forest excels in prediction stability and resilience against overfitting.

Additionally, comparisons with previous studies, such as Untung Surapati & Arif Prasetyo (2022), which used the C4.5 algorithm to determine the best employees, indicate that

using ensemble methods like Random Forest can enhance the objectivity of decision support systems in the human resource management context. The scenario comparison provides strong evidence that feature selection and data preprocessing significantly impact model performance. However, Random Forest consistently demonstrates the best performance across all scenarios. This strengthens the confidence that this model can be effectively and efficiently applied as a decision support system for KPI-based employee performance evaluations.

CONCLUSION

In this study, the C4.5 and Random Forest methods were implemented to classify employee salary increases based on various attributes across three different scenarios. The first scenario, which utilized Discipline as the root node, demonstrated that Random Forest achieved a higher accuracy (84.50%) compared to C4.5 (82.00%), along with a lower MSE (0.25 for Random Forest and 0.32 for C4.5) and lower MAE (0.18 for Random Forest and 0.23 for C4.5). In the second scenario, where Punctuality was used as the root node, Random Forest once again outperformed C4.5, achieving an accuracy of 87.50% with an MSE of 0.20, whereas C4.5 only reached an accuracy of 85.00% with an MSE of 0.30. Similarly, in the third scenario, which employed Skills as the root node, Random Forest continued to demonstrate superior performance, attaining an accuracy of 86.00% and an MSE of 0.22, while C4.5 only achieved 83.50% accuracy with an MSE of 0.31.

Overall, the evaluation results consistently indicate that Random Forest produces higher accuracy and lower MSE and MAE values compared to C4.5 in all three tested scenarios. Furthermore, Random Forest exhibited a higher ROC-AUC score in each scenario, with its highest value reaching 0.96 in the second scenario. These findings reinforce the conclusion that ensemble models like Random Forest provide greater stability and reliability for KPI-based employee performance predictions. While C4.5 remains more interpretable, Random Forest demonstrates superior accuracy and stability due to its ability to reduce bias associated with single models. Taking into account accuracy, error metrics, and the ROC-AUC Score, it can be concluded that Random Forest delivers more precise and stable classifications compared to C4.5, particularly within the three tested scenarios in this study.

REFERENCE

Aini, N., Arif, M., Agustin, I. T., & Toyibah, Z. B. (2024). Implementasi Algoritma Random Forest untuk Klasifikasi Bidang MSIB di Prodi Pendidikan Informatika. *Jurnal Informatika, 11*(1), 11–16. https://doi.org/10.31294/inf.v11i1.20637

Anwar Prabu Mangkunegara, A. A. (2010). *Manajemen sumber daya manusia perusahaan.* PT Remaja Rosdakarya.

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

Deep Prakash, C., & Verma, S. (2022). A new in-form and role-based deep player performance index for player evaluation in T20 cricket. *Decision Analytics Journal, 2*. https://doi.org/10.1016/j.dajour.2022.100025

Devia, E. (2023). *Penerapan decision tree dengan algoritma C4.5 untuk menentukan rekomendasi kenaikan jabatan karyawan.*

Dinisari, M. C. (2019). Microsoft ganjar penghargaan untuk Kreatif Dinamika Integrasi. *Bisnis.com.* https://teknologi.bisnis.com/read/20191127/84/1175148/microsoft-ganjar-penghargaan-untuk-kreatif-dinamika-integrasi

Dwi Egie, P., Sumardi, H. R., & Nunung, N. (2019). Pengaruh kompensasi dan gaya kepemimpinan terhadap kinerja karyawan di Sekretariat DPRD Kabupaten Indramayu.

Es-Sabery, F., & Hair, A. (2020). A MapReduce C4.5 decision tree algorithm based on fuzzy rule-based system. *Fuzzy Information and Engineering, 11*(4), 446–473. https://doi.org/10.1080/16168658.2020.1756099

Fitriana, H. L., Feriantono, I., Laily, S. I., & Jakaria, R. B. (2024). Penilaian kinerja karyawan menggunakan Key Performance Indicator (KPI) pada PT. Hilfic. *Venus: Jurnal Publikasi Rumpun Ilmu Teknik, 2*(4), 333–341. https://doi.org/10.61132/venus.v2i4.448

Galih, & Eriyadi, M. (2022). Perbandingan model NBC, SVM, dan C4.5 dalam mengukur kinerja karyawan berprestasi pasca pandemi Covid-19. *Jurnal Informatika, 9*(2). http://ejournal.bsi.ac.id/ejurnal/index.php/ji

Gibson, J. L., Ivancevich, J. M., & Donnelly, J. H. (1996). *Organization: Perilaku, struktur, proses jilid 1-8/E.* (N. Adiarni, Trans.; L. Saputra, Ed.). Binarupa Aksara.

Ignatius, W. (2005). *Dasar-dasar ilmu organisasi.* Andi. https://elibrary.bsi.ac.id/readbook/205321/dasar-dasar-ilmu-organisasi

Kathirisetty, N., Jadeja, R., Garg, D., & Thakkar, H. K. (2022). On the design of student assessment model based on intelligence quotient using machine learning. *IEEE Access, 10*, 48733–48746. https://doi.org/10.1109/ACCESS.2022.3171807

Kompascom. (2024). PT Kreatif Dinamika Integrasi raih Microsoft Partner of the Year Indonesia 2024. https://money.kompas.com/read/2024/07/04/205947026/pt-kreatif-dinamika-integrasi-raih-microsoft-partner-of-the-year-indonesia

Kusumaningtyas, G. Y. R. T., & Wahyuddin, M. I. (2022). Implementasi algoritma C4.5 dan simple additive weight untuk menentukan KPI karyawan. *Building of Informatics, Technology and Science (BITS, 3)*(4), 519–527. https://doi.org/10.47065/bits.v3i4.1356

López-Santiago, J., Md Som, A., Ruiz-Garcia, L., Zubelzu Mínguez, S., & Gómez Villarino, M. T. (2024). Assessment of environmental management performance in wineries: A survey-based analysis to create key performance indicators. *Environments - MDPI, 11*(7). https://doi.org/10.3390/environments11070139

Lubis, P. I., & Kusumanto, I. (2018). Penilaian kinerja karyawan menggunakan metode Key Performance Indicators (KPI) (Studi kasus: CV. Bunda Bakery Pekanbaru). *Jurnal Sains, Teknologi dan Industri, 15*(2), 37–45.

Naufal, N. R. A., & Kholik, M. A. (2024). Sistem pendukung keputusan penilaian kinerja karyawan menggunakan algoritma decision tree. https://mail.jurnal.stmik-yadika.ac.id/index.php/spirit/article/view/346/324

Novianti, R. A. (2016). Pengaruh stres kerja terhadap motivasi kerja dan dampaknya terhadap kinerja karyawan di bagian funding officer dan accounting officer PT. Bank Rakyat Indonesia (Persero), Tbk. Cabang Bangkalan, Madura. https://ejournal.unesa.ac.id/index.php/jim/article/view/17872

Ompusunggu, E. S., Nainggolan, A., Sihombing, M. K., Sains, F., & Teknologi, D. (2023). Penentuan kelayakan promosi pegawai menggunakan algoritma random forest classifier dan XGBoost classifier. *Jurnal TEKINKOM, 6*(2). https://doi.org/10.37600/tekinkom.v6i2.949

Pahlevi, O., & Handrianto, Y. (2023). Implementasi algoritma klasifikasi random forest untuk penilaian kelayakan kredit. *Jurnal, 5*(1). http://ejournal.bsi.ac.id/ejurnal/index.php/infortech

Rahmayani, S. (2025). SPK seleksi calon karyawan PT Jasa Raharja menggunakan metode kombinasi ARAS dan weighted product. 206–222. https://doi.org/10.59841/saber.v3i1.2228

Robbins, S. P., & Judge, T. A. (2013). *Organizational behavior (15th ed.).* Pearson Prentice Hall. https://digilib.stekom.ac.id/ebook/view/ORGANIZATIONAL-BEHAVIOR-15TH-EDITION

Rochmat Asyhari, M., Alijoyo, A., & Vernanda, D. (2024). Penerapan algoritma C4.5 dan metode 360° pada sistem informasi penilaian karyawan. *Jurnal Teknologi Komputer, 18*(1). https://doi.org/10.33365/jtk.v18i1.3215

Surapati, U., & Prasetyo, A. (2022). Penerapan algoritma C4.5 dalam menentukan karyawan terbaik pada Klinik Hosana Jakarta Barat. *Jurnal Pendidikan dan Konseling, 4*. https://journal.universitaspahlawan.ac.id/index.php/jpdk/article/download/6737/5086/22159

Wibowo, P. D. (2014). *Perilaku dalam organisasi.* Rajawali Pers.

Witasari, D., & Jumaryadi, Y. (2020). Aplikasi pemilihan karyawan terbaik dengan metode simple additive weighting (Studi kasus Citra Widya Teknik). *Teknologi Informatika dan Komputer*. https://doi.org/10.24853/justit.10.2.115-122

Yuliana, H., Hendrawan, Iskandar, & Musashi, Y. (2024). Estimating base station traffic and throughput using machine learning based on hourly key performance indicator (KPI) network analysis. *IEEE Access, 12*, 116285–116301. https://doi.org/10.1109/ACCESS.2024.3447098

Yurtay, Y., et al. (2023). Improvement and implementation of sustainable key performance indicators in supply chain management. *IEEE Access, 11*, 41913–41927. https://doi.org/10.1109/ACCESS.2023.3271138