



Klasifikasi Kelayakan Gaji Guru Menggunakan Algoritma *Decision Tree* (Studi Kasus SMA YPKPP)

Hendi Suhendi¹, Femmy Novica Ramadani²

¹⁻² Program studi Teknik Informatika, Fakultas Teknologi Informasi, Universitas Adhirajasa Reswara Sanjaya, Indonesia

hendi2708@ars.ac.id¹, femmynovica097@gmail.com²

*Penulis Korespondensi: hendi2708@ars.ac.id

Abstract. *This study aimed to examine the eligibility of teacher honorarium by implementing a classification method using the Decision Tree algorithm. The primary issue addressed in this research is the absence of a fair and data-driven salary system at SMA YPKPP. A classification approach was employed to categorize teachers into "Eligible" and "Not Eligible" groups based on attributes such as teaching hours, hourly wage, education level, job position, certification allowance, and school status. The classification model was developed using RapidMiner software. The dataset was divided into training and testing sets using a split data technique. The model was evaluated using metrics such as accuracy, precision, recall, and confusion matrix. The results indicated that the Decision Tree model achieved an accuracy of 93.75% in classifying teacher honorarium eligibility. Teaching hours and hourly wage were identified as the two most influential variables in the classification process. As a form of validation, additional statistical analysis was conducted using SPSS. The Pearson correlation test showed a significant relationship between teaching hours and hourly wage with the total honorarium received. Multiple Linear regression analysis resulted in an R Square value of 0.860, indicating that 86% of the variation in honorarium can be explained by the two variables. This study is expected to serve as a foundation for more objective and data-driven decision-making in the teacher compensation system. The findings demonstrate that a combination of data mining and statistical analysis approaches can be used to develop a transparent, fair, and efficient salary system.*

Keywords: *Classification; Decision Tree; Honorarium Teacher; RapidMine; SPSS.*

Abstrak. Penelitian ini bertujuan untuk mengkaji kelayakan honorarium guru dengan menerapkan metode klasifikasi menggunakan algoritma Decision Tree. Permasalahan utama yang diangkat dalam penelitian ini adalah belum adanya sistem penggajian yang adil dan berbasis data objektif di lingkungan SMA YPKPP. Untuk itu, pendekatan klasifikasi digunakan guna membedakan guru dalam kategori "Layak" dan "Tidak Layak" berdasarkan atribut seperti jumlah jam mengajar, honor per jam, tingkat pendidikan, jabatan, tunjangan sertifikasi, dan status sekolah. Model klasifikasi dibangun dengan memanfaatkan perangkat lunak RapidMiner. Data dibagi menjadi data latih dan data uji menggunakan teknik split data. Evaluasi model dilakukan menggunakan metrik akurasi, presisi, recall, dan confusion matrix. Hasil penelitian menunjukkan bahwa model Decision Tree memiliki akurasi sebesar 93,75% dalam mengklasifikasikan kelayakan honor guru. Jumlah jam mengajar dan honor per jam merupakan dua variabel yang paling dominan dalam menentukan hasil klasifikasi. Sebagai bentuk validasi, dilakukan analisis statistik tambahan menggunakan SPSS. Hasil uji korelasi menunjukkan hubungan yang signifikan antara jumlah jam dan honor per jam terhadap total honor yang diterima. Regresi linear berganda menghasilkan nilai R Square sebesar 0,860 yang berarti 86% variasi honor dapat dijelaskan oleh dua variabel tersebut. Penelitian ini diharapkan dapat menjadi dasar bagi pengambilan keputusan yang lebih objektif dan berbasis data dalam sistem penggajian guru. Hasilnya menunjukkan bahwa kombinasi pendekatan data mining dan analisis statistik dapat digunakan untuk membangun sistem penggajian yang transparan, adil, dan efisien.

Kata Kunci: *Decision Tree; Guru Honorer; Klasifikasi; Rapid Miner; SPSS.*

1. LATAR BELAKANG

Kesejahteraan guru merupakan salah satu faktor penting yang menentukan mutu pendidikan. Dalam praktiknya, sistem penggajian guru di sekolah swasta seperti SMA YPKPP sering kali belum berbasis data yang objektif, sehingga terjadi ketimpangan honor antara guru dengan beban kerja dan kualifikasi yang sepadan. Ketidakjelasan indikator penentu honorarium

serta keputusan yang masih bersifat subjektif menjadi persoalan utama dalam pengelolaan sumber daya manusia di sektor pendidikan.

Berdasarkan studi terdahulu oleh Wulandari (2018) dan Rahmawati & Susanti (2020), ditemukan bahwa tingkat pendidikan, sertifikasi, serta beban kerja memiliki pengaruh signifikan terhadap besaran honor guru. Namun, pendekatan dalam penelitian tersebut umumnya masih menggunakan metode statistik deskriptif dan regresi sederhana. Dalam penelitian ini, penulis mengusulkan pendekatan yang lebih komprehensif dan modern melalui kombinasi metode data mining berbasis algoritma Decision Tree dan validasi statistik menggunakan SPSS.

Algoritma Decision Tree dipilih karena mampu mengidentifikasi pola klasifikasi secara visual dan mudah dipahami, sementara SPSS digunakan untuk menguji hubungan antar variabel dan mengukur signifikansi secara statistik. Kombinasi ini memberikan dua perspektif penting: kemampuan prediktif dan kekuatan analisis inferensial, sehingga mendukung pengambilan keputusan yang lebih akurat dan adil dalam menentukan kelayakan honor guru.

Penelitian ini diharapkan memberikan kontribusi inovatif dalam penyusunan sistem penggajian guru berbasis data, dengan menekankan pentingnya penggunaan teknologi analitik untuk transparansi dan efisiensi dalam manajemen pendidikan.

2. KAJIAN TEORITIS

Data mining merupakan sebuah teknik analisis data yang dirancang untuk mengolah dan mengekstraksi informasi dari kumpulan data dalam jumlah besar. Tujuannya adalah menemukan pola-pola tersembunyi, hubungan, atau tren yang tidak terlihat secara langsung namun memiliki nilai penting sebagai dasar pengambilan keputusan yang lebih efektif dan efisien. Dalam konteks dunia pendidikan, khususnya pada pengelolaan honorarium guru, data mining berperan penting dalam melakukan klasifikasi kelayakan honor berdasarkan berbagai atribut yang berkaitan dengan guru. Atribut-atribut tersebut meliputi masa kerja, jenjang pendidikan, jumlah jam mengajar per minggu, sertifikasi profesi, serta faktor-faktor lain yang memengaruhi besaran honor yang diterima oleh tenaga pendidik.

Han et al. (2020) menguraikan bahwa data mining terdiri dari beberapa teknik utama, yaitu klasifikasi, klastering, regresi, dan asosiasi. Dari teknik-teknik ini, klasifikasi menjadi salah satu metode yang paling tepat digunakan untuk mengelompokkan data ke dalam kategori tertentu, sehingga sangat relevan apabila diaplikasikan untuk menentukan kelayakan honorarium guru secara sistematis dan berbasis data. Dengan klasifikasi, institusi pendidikan dapat

mengkategorikan guru sesuai dengan tingkat kelayakan honor yang layak diterima berdasarkan parameter objektif dan terukur.

Pendapat ini diperkuat oleh Fayyad et al. (1996) yang menyatakan bahwa data mining merupakan bagian integral dari proses *Knowledge Discovery in Databases* (KDD).

KDD adalah tahap dalam ekstraksi pengetahuan yang berguna dari kumpulan data besar dengan cara mengidentifikasi pola, korelasi, atau hubungan signifikan yang dapat memberikan wawasan baru. Dalam proses ini, data mining berperan sebagai inti dari eksplorasi data untuk menemukan pola tersembunyi yang tidak mudah terlihat oleh analisis konvensional.

Lebih lanjut, Witten & Frank (2019) menjelaskan bahwa data mining merupakan sebuah proses otomatis atau semi-otomatis yang bertujuan mengungkap pola-pola tersembunyi dalam data yang kompleks. Pola-pola tersebut selanjutnya dapat dimanfaatkan untuk keperluan prediksi dan klasifikasi, yang sangat membantu dalam pengambilan keputusan yang lebih akurat dan objektif. Pendekatan ini sangat bermanfaat dalam konteks pengelolaan honorarium guru, dimana keputusan terkait besaran honor seringkali membutuhkan analisis yang mendalam dan komprehensif atas berbagai data atribut guru.

Tan et al. (2022) juga menegaskan bahwa data mining adalah sebuah pendekatan sistematis untuk mengeksplorasi informasi tersembunyi dalam kumpulan data yang sangat besar. Mereka menekankan bahwa penerapan teknik data mining telah meluas ke berbagai sektor, termasuk pendidikan, bisnis, dan kedokteran. Di sektor pendidikan, aplikasi data mining tidak hanya digunakan untuk mengukur performa akademik siswa, tetapi juga untuk mengevaluasi sistem penggajian guru sehingga lebih transparan, adil, dan berbasis data nyata.

Secara khusus, dalam pengelolaan honorarium guru, data mining memungkinkan pihak manajemen sekolah untuk mengidentifikasi pola kelayakan honor berdasarkan data historis yang dimiliki oleh setiap guru.

Dengan demikian, proses pengambilan keputusan terkait penggajian dapat dilakukan dengan analisis yang objektif, menghindari keputusan subjektif yang berpotensi menimbulkan ketidakadilan. Pendekatan ini juga memungkinkan identifikasi faktor-faktor dominan yang mempengaruhi kelayakan honor sehingga kebijakan yang diambil dapat lebih tepat sasaran dan efisien. Pendapat serupa juga disampaikan oleh Larose (2005) yang menjelaskan bahwa data mining merupakan serangkaian langkah analisis statistik dan algoritmik untuk menemukan pola yang bermakna dari data, yang biasanya tersimpan dalam data *warehouse* atau *respository* besar.

Dengan demikian, dapat disimpulkan bahwa data mining merupakan gabungan dari teknik statistik, pembelajaran mesin dan basis data yang digunakan untuk mengekstraksi pengetahuan baru dari data berukuran besar. Dalam konteks pendidikan, data mining sering diterapkan untuk menganalisis kinerja siswa, efektivitas pengajaran, maupun penilaian kompensasi tenaga pengajar secara lebih objektif dan akurat.

3. METODE PENELITIAN

Penelitian ini merupakan jenis penelitian kuantitatif deskriptif melalui pendekatan analitik data. Objek dari penelitian ini adalah sistem penggajian guru di SMA YPKPP, dengan fokus pada faktor-faktor yang mempengaruhi besaran gaji guru. Data yang digunakan mencakup latar belakang guru seperti tingkat pendidikan terakhir, jumlah jam mengajar per minggu, status kepegawaian (tetap atau kontrak), lama pengalaman kerja, kepemilikan sertifikasi pendidik, serta data nominal gaji bulanan. Populasi dalam penelitian ini adalah seluruh guru yang aktif mengajar di SMA YPKPP dalam tiga tahun terakhir. Populasi ini mencakup guru tetap maupun kontrak yang memiliki data administrasi lengkap dan terdokumentasi. Untuk pemilihan sampel, digunakan teknik purposive sampling, yaitu pemilihan data guru yang memenuhi kriteria tertentu seperti kelengkapan data atribut yang relevan (misalnya informasi pendidikan, masa kerja, dan sertifikasi).

Serta keterwakilan dalam variasi status kepegawaian. Sampel inilah yang kemudian digunakan dalam proses klasifikasi dan analisis menggunakan metode data mining.

Algoritma yang digunakan

Dalam praktik data mining, pemilihan algoritma yang sesuai sangat penting agar proses pencarian pola dan informasi tersembunyi dari data dapat dilakukan secara akurat dan efisien. Algoritma Decision Tree seperti ID3, C4.5, dan CART digunakan untuk mengklasifikasikan data ke dalam kategori tertentu. Algoritma ini membagi data berdasarkan atribut yang memberikan kontribusi terbesar dalam pengurangan ketidakpastian. Dasar perhitungan utamanya adalah Entropi dan Information Gain.

Teknik pengumpulan data dilakukan melalui dokumentasi dan observasi terhadap data administrasi guru yang tersedia di SMA YPKPP. Pengolahan data dilakukan melalui tahap pra-pemrosesan seperti:

- a. Data Cleaning (menghapus nilai kosong)
- b. Normalisasi (standarisasi format data)
- c. Seleksi Atribut (memilih variabel yang relevan untuk klasifikasi)
- d. Setelah itu dilakukan penerapan algoritma *Decision tree*.

Teknik Analisis Data

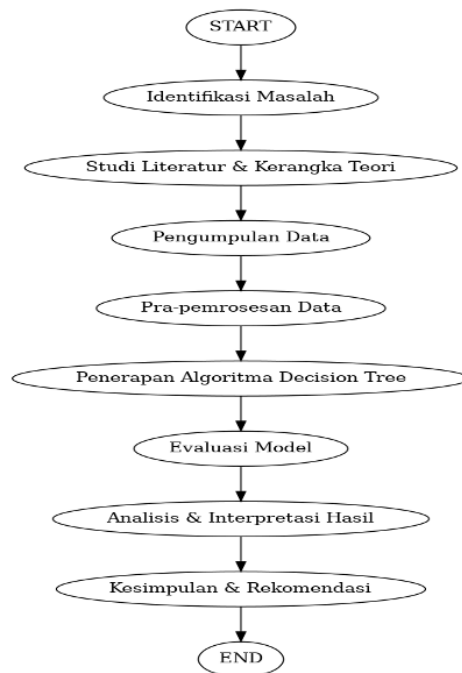
Analisis data dilakukan dengan metode klasifikasi menggunakan algoritma Decision Tree. Langkah-langkah dalam proses analisis meliputi pembuatan model, pelatihan (*training*), pengujian (*testing*), dan evaluasi model.

Model yang dihasilkan kemudian dievaluasi menggunakan metrik evaluasi seperti:

- Confusion Matrix*: Tabel yang menunjukkan hasil prediksi vs aktual.
- Accuracy*: Persentase prediksi yang benar dari keseluruhan data.
- Precision*: Proporsi data yang benar-benar positif dari seluruh prediksi positif.
- Recall*: Kemampuan model dalam menemukan data positif sebenarnya.

Alur Penelitian

Tahapan dalam penelitian ini disusun secara sistematis agar hasil yang diperoleh valid dan sesuai dengan tujuan penelitian. Berikut alur penelitiannya :



Gambar 1. Flowchart Alur Penelitian.

Langkah awal dilakukan melalui proses *pre-processing* data, yaitu pembersihan data dari nilai kosong atau duplikat, konversi format data numerik, serta seleksi atribut yang paling relevan. Setelah itu, data dibagi menjadi dua bagian: data latih (60%) dan data uji (40%) dengan metode split data.

Model klasifikasi dibangun menggunakan algoritma *Decision Tree* pada aplikasi RapidMiner. Evaluasi dilakukan dengan mengukur akurasi, precision, recall, dan confusion matrix dari hasil prediksi pada data uji. Selain itu, untuk memperkuat validitas temuan, dilakukan analisis statistik tambahan menggunakan perangkat lunak SPSS. Uji korelasi

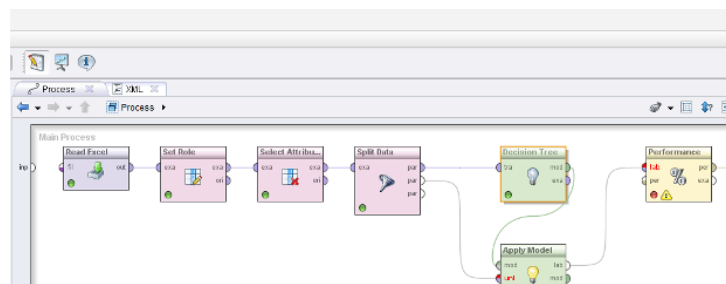
Pearson digunakan untuk melihat kekuatan hubungan antar variabel, sedangkan regresi linear berganda digunakan untuk mengetahui kontribusi masing-masing variabel terhadap total honor yang diterima.

Dengan pendekatan gabungan ini, penelitian tidak hanya menghasilkan model prediktif yang akurat, tetapi juga memberikan justifikasi statistik atas pengaruh variabel yang digunakan, sehingga mendukung keputusan berbasis data dalam pengelolaan sistem penggajian guru.

4. HASIL DAN PEMBAHASAN

Penelitian ini bersumber dari data primer dari administrasi keuangan SMA YPKPP Bandung, khususnya data honorium guru di bulan Mei. Data ini mempresentasikan sistem penggajian guru selama satu bulan berjalan dan digunakan sebagai dasar untuk melakukan analisis klasifikasi terhadap besaran gaji (honor) yang diterima oleh masing-masing guru.

Proses di Rapidminer



Gambar 2. Proses di Rapidminer.

Alur ini terdiri dari beberapa operator utama yang saling terhubung untuk membentuk *pipeline* klasifikasi berbasis algoritma Decision Tree. Berikut penjelasan tiap tahapan:

Read Excel

Operator ini digunakan untuk membaca dataset dari file Excel yang berisi data honor guru. Dataset ini memuat atribut seperti Jumlah Jam, Honor per Jam, Mapel, dan Honor Kelayakan.

Set Role

Pada tahap ini, atribut Honor Kelayakan ditetapkan sebagai label (target) klasifikasi. Label ini berupa kategori “Layak” atau “Tidak Layak”.

Select Attributes

Digunakan untuk memilih atribut-atribut yang relevan sebagai faktor input. Atribut Total tidak disertakan karena merupakan hasil perhitungan, bukan penyebab. Atribut yang digunakan antara lain: Jumlah Jam, Honor per Jam, dan Mapel.

Split Data

Dataset dibagi menjadi dua bagian: 60% untuk training dan 40% untuk testing, dengan metode *shuffled sampling*. Output pertama (data 4. Split Data

Dataset dibagi menjadi dua bagian: 60% untuk training dan 40% untuk testing, dengan metode *shuffled sampling*. Output pertama (data latih) dikirim ke Decision Tree, dan output kedua (data uji) ke Apply Model.

Decision Tree

Model klasifikasi dibangun menggunakan algoritma *Decision Tree* berdasarkan data training. Model ini bertugas mengenali pola dari kombinasi faktor-faktor input terhadap label Honor Kelayakan.

Apply Model

Model hasil training kemudian diterapkan pada data testing untuk memprediksi apakah guru masuk kategori “Layak” atau “Tidak Layak”.

Performance

Operator ini menghitung evaluasi performa model berdasarkan hasil prediksi, menggunakan metrik seperti akurasi, precision, dan *recall*.

Untuk mengetahui seberapa baik algoritma *Decision Tree* dapat memprediksi kelayakan honor guru berdasarkan atribut yang dimiliki, serta mengidentifikasi faktor-faktor yang paling berpengaruh.

	true Layak	true Tidak Layak	class precision
pred Layak	10	0	100.00%
pred Tidak Layak	2	4	66.67%
class recall	83.33%	100.00%	

Gambar 3. Hasil Klasifikasi.

Hasil evaluasi model klasifikasi menunjukkan tingkat akurasi sebesar 93,75%, yang mengindikasikan bahwa model mampu memprediksi kelas dengan cukup baik.

Dari confusion matrix yang disajikan, dapat dilihat bahwa model berhasil mengklasifikasikan 11 data sebagai kelas "Layak" dengan benar, sedangkan untuk kelas "Tidak Layak", terdapat 4 data yang diklasifikasikan secara salah dan 1 data yang diklasifikasikan dengan benar.

Nilai classification error yang dihasilkan sebesar 6,25%, menunjukkan persentase kesalahan klasifikasi yang masih rendah. Selain itu, nilai kappa sebesar 0,846 mengindikasikan

kesepakatan yang tinggi antara prediksi model dengan data sebenarnya, sehingga model tergolong reliabel.

Akurasi yang dihasilkan mencapai sekitar 93,75%, yang mengindikasikan bahwa model cukup efektif dalam memprediksi kategori Layak dan Tidak Layak. Namun, meskipun hasilnya sudah baik, nilai akurasi ini masih perlu diperhatikan dengan seksama. Adanya nilai akurasi yang cukup tinggi ini bisa disebabkan oleh jumlah data yang relatif terbatas, khususnya data uji yang hanya beberapa baris. Hal ini berpotensi menyebabkan model mengalami *overfitting*, di mana model lebih menghafal data latih daripada belajar pola yang general. Sehingga, performa model pada data baru bisa menurun.

Metode ini membantu menguji model secara lebih menyeluruh dengan menggunakan berbagai subset data sehingga hasil evaluasi menjadi lebih objektif dan dapat dipercaya.

Secara keseluruhan, model Decision Tree ini menunjukkan potensi yang baik untuk mengklasifikasikan kelayakan honor guru dengan tingkat akurasi yang memuaskan, namun hasil ini tetap perlu didukung dengan evaluasi data yang lebih luas .

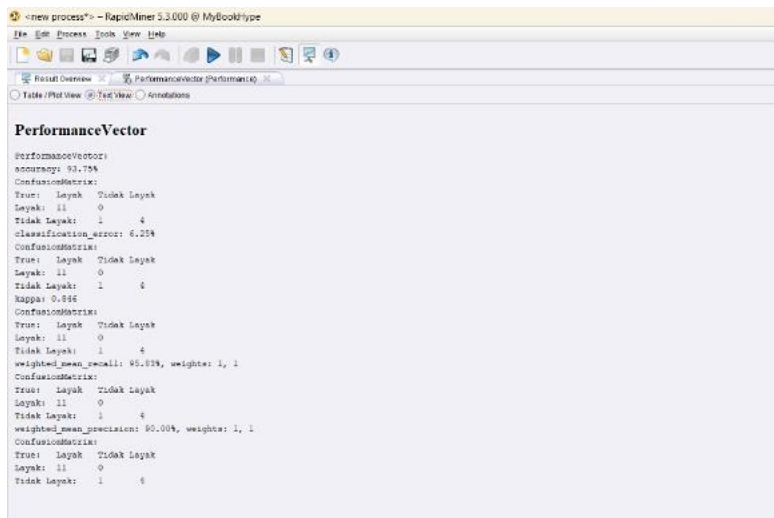
Perbandingan RapidMiner dengan SPSS dalam Kelayakan Honor Guru

Dalam penelitian ini, penggunaan RapidMiner dipilih sebagai alat bantu utama karena kemampuannya dalam membangun model klasifikasi berbasis algoritma Decision Tree secara visual dan interaktif. Namun, untuk memperluas wawasan dan validasi metode, perbandingan dengan perangkat lunak statistik lain seperti SPSS (*Statistical Package for the Social Sciences*) perlu dikemukakan.

RapidMiner merupakan platform data mining yang mengintegrasikan berbagai metode klasifikasi, regresi, dan clustering.

Kelebihan utamanya adalah pada kemudahan penggunaan antarmuka visual (*drag-and-drop*), kemampuan integrasi berbagai algoritma *machine learning* seperti Decision Tree, SVM, Naive Bayes, serta mendukung evaluasi kinerja model melalui metrik seperti akurasi, precision, recall, dan confusion matrix.

Sementara itu, SPSS merupakan perangkat lunak statistik yang lebih fokus pada analisis deskriptif, inferensial, dan regresi statistik. SPSS dapat digunakan untuk uji-uji statistik seperti uji t, ANOVA, korelasi Pearson, regresi linear, dan crosstab analysis. Dalam konteks kelayakan honor guru, SPSS dapat digunakan untuk mencari hubungan antar variabel seperti tingkat pendidikan dan masa kerja terhadap jumlah honor melalui analisis regresi atau uji korelasi, tetapi tidak dirancang untuk klasifikasi otomatis seperti dalam model Decision Tree.



Gambar 4. Hasil Klasifikasi SPSS.

Hasil dari *classification table* menunjukkan bahwa dari 15 guru yang dikategorikan Layak, sebanyak 12 guru berhasil diprediksi dengan benar, dan 3 guru diprediksi salah, sehingga akurasi untuk kelas "Layak" mencapai 80%. Sementara itu, dari 17 guru yang tergolong Tidak Layak, terdapat 11 guru yang diprediksi dengan benar dan 6 guru yang diprediksi salah, menghasilkan akurasi untuk kelas "Tidak Layak" sebesar 64,7%. Secara keseluruhan, akurasi model SPSS mencapai 71,9%, yang berarti bahwa model ini mampu mengklasifikasikan data dengan tingkat ketepatan sedang.

Perbandingan Hasil Rapidminer dengan SPSS

Penelitian ini membandingkan hasil klasifikasi kelayakan honor guru menggunakan dua pendekatan yang berbeda, yaitu metode data mining melalui RapidMiner dan metode statistik konvensional melalui SPSS. Tujuannya adalah untuk menilai efektivitas masing-masing alat dalam mengklasifikasikan guru ke dalam kategori "Layak" atau "Tidak Layak" berdasarkan data yang tersedia. Hasil analisis menunjukkan bahwa:

RapidMiner, dengan algoritma Decision Tree, menghasilkan akurasi sebesar 93,75%, dan mampu mengenali pola-pola klasifikasi secara visual dan otomatis berdasarkan atribut seperti jumlah jam, honor per jam, dan sertifikasi.

Model ini menunjukkan kinerja yang stabil, presisi tinggi, dan recall yang seimbang, sehingga cocok digunakan sebagai sistem pendukung keputusan (*decision support system*).

Sementara itu, model klasifikasi di SPSS menggunakan pendekatan regresi logistik menghasilkan akurasi sebesar 71,9%. Meskipun cukup layak, model ini tidak seakurat RapidMiner dan memerlukan interpretasi manual terhadap hasil, seperti signifikansi variabel dan nilai korelasi antar atribut.

5. KESIMPULAN DAN SARAN

Kesimpulan dari penelitian ini menunjukkan bahwa penggunaan algoritma Decision Tree dalam mengklasifikasikan kelayakan honor guru memberikan hasil yang sangat memuaskan, dengan tingkat akurasi mencapai 93,75%.

Model klasifikasi yang dibangun melalui RapidMiner mampu mengenali pola-pola penting dari data guru, terutama pada atribut jumlah jam mengajar dan honor per jam. Kedua variabel tersebut terbukti menjadi penentu utama dalam membedakan guru ke dalam kategori “Layak” dan “Tidak Layak” menerima honor.

Sebagai pelengkap dan pembanding, analisis menggunakan SPSS juga dilakukan untuk mengetahui hubungan dan pengaruh antar variabel secara statistik. Hasil uji korelasi Pearson menunjukkan bahwa terdapat hubungan yang sangat kuat dan signifikan antara jumlah jam mengajar serta honor per jam terhadap total honor. Sementara itu, hasil regresi linear berganda menunjukkan bahwa kedua variabel tersebut berkontribusi signifikan dalam menjelaskan total honor guru.

UCAPAN TERIMA KASIH

Pertama-tama penulis mengucapkan puji syukur ke hadirat Allah SWT atas segala nikmat berupa kesehatan, kekuatan, dan inspirasi yang melimpah dalam proses penyelesaian penelitian ini. Sholawat serta salam senantiasa tucurahkan kepada Nabi Muhammad SAW. Skripsi ini dipersembahkan sebagai bukti semangat usaha serta wujud cinta kasih penulis kepada orang-orang yang berharga dalam kehidupan.

Karya sederhana ini penulis persembahkan kepada Ayahanda Hilmi sebagai cinta pertama penulis, yang meskipun tidak menyaksikan langsung perjuangan penulis dalam menempuh pendidikan, namun doa yang senantiasa dipanjatkan menjadi kekuatan hingga karya ini dapat diselesaikan. Penulis juga menyampaikan terima kasih kepada Ibunda Karlina yang memiliki peran sangat besar dalam menyelesaikan program studi ini. Meskipun beliau tidak sempat menempuh pendidikan hingga bangku perkuliahan, semangat dan doa beliau mampu mengantarkan penulis hingga meraih gelar sarjana. Ucapan terima kasih turut disampaikan kepada adik-adik tercinta, Kamilla Sari Bella dan Muhammad Ridho Alfiansyah, atas dukungan, semangat, dan hiburan yang diberikan selama proses penulisan skripsi.

Selanjutnya, penulis mengucapkan terima kasih kepada Prada Murdaka Wijaya yang kehadirannya sangat berarti dalam perjalanan hidup penulis dan telah banyak berkontribusi dalam penyusunan karya ini, baik tenaga maupun waktu, serta senantiasa memberikan dukungan, hiburan, dan motivasi untuk tidak menyerah. Penulis juga berterima kasih kepada

Siti Fauziah Dawami sebagai sahabat sekaligus keluarga yang selalu memberikan dukungan dan penguatan. Ucapan terima kasih disampaikan kepada seluruh pihak yang tidak dapat disebutkan satu per satu atas doa dan dukungan yang diberikan. Terakhir, penulis mengapresiasi diri sendiri, Femmy Novica Ramadanis, yang telah berjuang, bekerja keras, dan mampu bertahan menghadapi berbagai tekanan hingga skripsi ini dapat diselesaikan dengan sebaik-baiknya, serta kepada motor Beat hitam yang setia menemani aktivitas penulis dalam bekerja, penelitian, perkuliahan, dan bimbingan hingga karya ini terselesaikan.

DAFTAR REFERENSI

- Ayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (2020). From data mining to knowledge discovery in databases. *AI Magazine*, 17(3), 37–54.
- Han, J., Kamber, M., & Pei, J. (2020). *Data mining: Concepts and techniques*. Elsevier.
- Larose, D. T. (2005). *Discovering knowledge in data: An introduction to data mining*. John Wiley & Sons.
- Prasetyo, E., & Setiawan, B. (2023). Prediksi prestasi belajar siswa SD menggunakan algoritma C4.5. *Jurnal Teknologi Informasi dan Pendidikan*, 16(2), 115–122.
- Putra, A. (2023). Model prediksi lama studi mahasiswa menggunakan RapidMiner. *Jurnal Informatika dan Sistem Cerdas*, 9(1), 44–52.
- Putri, S. M., & Mahendra, R. (2023). Rekomendasi jurusan pendidikan menggunakan algoritma K-medoids. *Jurnal Informatika dan Sistem Informasi*, 12(3), 140–150.
- Rahmawati, T., & Susanti, R. (2020). Pengaruh sertifikasi guru terhadap honor di sekolah swasta. *Jurnal Ilmu Pendidikan*, 6(2), 112–119.
- Ramadhan, A., & Widodo, A. (2024). Penerapan decision tree dalam prediksi kemampuan hafalan santri. *Jurnal Pendidikan Islam*, 10(1), 45–53.
- Rosmanida, R., et al. (2022). Ketimpangan penggajian guru honorer di Indonesia. *Jurnal Kebijakan Pendidikan Nasional*, 7(2), 88–96.
- Sanjaya, R., Ramdhani, Y., & Prasetio, R. (2016). Implementasi algoritma neural network untuk klasifikasi kualitas kain tenun. *RESPONSIF*, 3, 3–10.
- Sari, et al. (2022). Clustering-based salary analysis for educational institutions. *Journal of Educational Management and Technology*, 6(4), 211–220.
- Witten, I. H., Frank, E., & Hall, M. A. (2016). *Data mining: Practical machine learning tools and techniques* (4th ed.). Morgan Kaufmann.
- Wulandari, N., & Nugroho, A. (2023). Prediksi kelulusan mahasiswa menggunakan algoritma C4.5. *Jurnal Sistem Informasi Akuntansi*, 8(1), 55–61.
- Wulandari, R. (2018). Pengaruh tingkat pendidikan dan masa kerja terhadap gaji guru. *Jurnal Ekonomi dan Pendidikan*, 5(3), 140–146.