

IMPLEMENTASI DATA MINING MENGGUNAKAN ALGORITMA C4.5 UNTUK PREDIKSI EVALUASI ANGGOTA SATUAN PENGAMANAN STUDI KASUS PT. YIMM PULOGADUNG

Fery Pirmansyah ^{1*}, Tri Wahyudi ²

^{1,2} STIKOM Cipta Karya Informatika, Kota Jakarta Timur, Daerah Khusus Ibukota Jakarta, Indonesia.

Email: feripirmansyah03@gmail.com^{1*}, triwahyudi199003@gmail.com²

Histori Artikel:

Dikirim 15 Juli 2023; Diterima dalam bentuk revisi 6 Agustus 2023; Diterima 15 Agustus 2023; Diterbitkan 10 September 2023. Semua hak dilindungi oleh Lembaga Penelitian dan Pengabdian Masyarakat (LPPM) STMIK Indonesia Banda Aceh.

Abstrak

Ketidakmampuan dalam melaksanakan tugas jaga merupakan masalah yang memiliki dampak serius baik bagi perusahaan maupun anggota satuan pengamanan. Di tengah persaingan yang semakin ketat dalam era saat ini, di mana perusahaan harus mencari individu yang memiliki kompetensi yang tinggi, diperlukan pendekatan teknis yang lebih canggih untuk mengatasi masalah ini. Salah satu teknik yang digunakan dalam analisis data adalah data mining. Peneliti mengimplementasikan metode klasifikasi menggunakan Decision tree. Decision tree dipilih karena merupakan salah satu teknik yang paling umum digunakan dalam analisis data dan menghasilkan aturan yang dapat diterapkan dengan mudah. Penelitian ini menggunakan Algoritma C4.5 untuk menghasilkan aturan klasifikasi yang memutuskan apakah seorang anggota satuan pengamanan harus dipertahankan di perusahaan atau tidak. Hasil dari penelitian ini menunjukkan tingkat akurasi sebesar 99,84%, menggambarkan kemampuan yang sangat baik dalam memprediksi ketidakmampuan anggota satuan pengamanan. Dengan pendekatan ini, data yang terkait dengan ketidakmampuan anggota satuan pengamanan dapat disajikan secara lebih profesional dan efisien, membantu perusahaan dalam mengambil keputusan yang lebih tepat terkait dengan personel mereka.

Kata Kunci: Kompetensi Anggota Satuan Pengamanan; Data Mining; Decision Tree; Algoritma C4.5.

Abstract

The inability to perform security duties is a problem that has serious consequences for both the company and the members of the security force. In today's increasingly competitive environment, where companies must search for highly competent people, a more sophisticated technical approach is needed to overcome this problem. One of the techniques used in data analysis is data mining. The researchers implemented a classification method using decision trees. Decision trees were chosen because they are one of the most used techniques in data analysis and provide rules that are easy to apply. This study uses the C4.5 algorithm to generate classification rules that decide whether a member of a security unit should be retained in a company. The results of this study showed an accuracy rate of 99.84%, demonstrating excellent ability to predict the incompetence of security unit members. Through this approach, data regarding the incompetence of security force members can be presented in a more professional and effective manner, helping companies make more informed decisions about personnel. their affairs.

Keyword: Capacity of Security Unit Members; Data Mining; Decision Tree; Algorithms C4.5.

1. Pendahuluan

Perkembangan dunia usaha saat ini semakin bertambah pesat, sehingga perusahaan di dalam mengelola usaha diharapkan mampu menggunakan sumber daya manusia dengan baik dan benar. Sumber daya manusia merupakan salah satu faktor kunci dalam reformasi ekonomi, yakni bagaimana menciptakan sumber daya manusia yang berkualitas dan memiliki keterampilan serta berdaya saing tinggi dalam persaingan global yang selama ini kita abaikan. Karyawan merupakan salah satu sumber daya yang digunakan sebagai alat penggerak dalam memajukan perusahaan. Peranan sumber daya manusia dalam perusahaan sangatlah penting karena sebagai penggerak utama seluruh kegiatan atau aktivitas perusahaan dalam mencapai tujuannya, baik untuk memperoleh keuntungan maupun untuk mempertahankan kelangsungan hidup perusahaan. Berhasil tidaknya suatu perusahaan dalam mempertahankan eksistensi perusahaan dimulai dari manusia itu sendiri dalam mempertahankan perusahaan dalam meningkatkan efektivitas dan efisiensi secara maksimal. Jumlah anggota satuan pengamanan di PT. YIMM Pulogadung pada tahun 2023 berjumlah 106 orang yang terdiri dari 101 anggota, 4 Komandan Regu dan 1 *Chief Security*, merupakan jumlah yang besar, demi menciptakan tenaga kerja yang profesional dan terampil di suatu perusahaan, maka dari itu bagaimana perusahaan mencari anggota satuan pengamanan yang sesuai dengan kategori perusahaan yang diinginkan. PT. YIMM Pulogadung setiap tahun menerima calon anggota satuan pengaman belum lagi mencari pengganti anggota yang sudah masuk masa pensiun, di PHK/dirumahkan, namun mencari anggota satuan pengamanan yang dibutuhkan sangat sulit mengambil keputusan dan terutama butuh waktu yang cukup lama. Satuan pengamanan atau disingkat satpam, yang juga secara tidak resmi disebut sekuriti, adalah bentuk pengamanan swakarsa yang bertugas membantu POLRI dalam bidang penyelenggaraan keamanan dan ketertiban masyarakat, terbatas pada lingkungan kerjanya [1]. Penentuan anggota satuan pengamanan disuatu perusahaan mempunyai penilaian yang berbeda-beda seperti penyeleksian berkas, tes lisan atau tertulis maupun wawancara. Cara tersebut biasa dan umum yang dilakukan dalam penentuan karyawan disuatu perusahaan.

Namun penilaian tersebut oleh tersebut intinya mencari karyawan yang berkompeten dalam menjalani atau jabatan yang dibebankan kepada karyawan tersebut. Secara etimologi kompetensi dari kata "*competency*" yang artinya kecakapan atau kemampuan [2]. Namun kompetensi bukan hanya keahlian maupun kecakapan dalam berkerja tetapi sebagai karakteristik yang mendasari seseorang dan berkaitan dengan efektifitas kinerja individu dalam pekerjaannya [3]. Perusahaan yang sehat memiliki manajemen sumber daya yang baik di dalamnya. Terlebih lagi pada sumber daya manusia yang membantu berjalan lancarnya dari tingkat manajerial hingga tingkat operasional. Sumber daya manusia yang baik pula perlu dibentuk pula pembentukan daya saing untuk meningkatkan kinerja dari sumber daya tersebut. Pada masa sekarang ini, banyak perusahaan yang melakukan pemutusan hubungan kerja (PHK) maupun merumahkan karyawannya untuk menstabilkan kondisi perusahaan. Menurut Soetjipto (2008), dalam suatu organisasi seringkali terjadi PHK, tidak hanya anggota satuan pengamanan saja, karyawan biasa pun banyak yang menjadi korban [4]. Dengan jumlah individu yang semakin sedikit, seorang atasan kini rela untuk tidak memiliki bawahan sebanyak dulu. Pengurangan karyawan seperti itu terkadang tidak diketahui atas dasar apa menentukan karyawan yang akan dilakukan PHK/dirumahkan. Selain melakukan pengurangan anggota satuan pengaman, perusahaan terkadang melakukan penerimaan anggota sebagai pengganti anggota yang di PHK/dirumahkan. Bagian SDM melakukan penilaian terhadap hal tersebut.

Penilaian dilakukan untuk menjaring karyawan yang berkompeten. Bagian SDM mengalami beberapa kendala dalam proses penilaian penetapan status dalam segi teknis ataupun efektifitas. Karena penilaian dinilai secara manual oleh kepala/pimpinan Satuan Pengaman atau biasa disebut *Chief Security*. Pada proses ini memakan waktu lama dan keputusannya sering dinilai secara subyektif. Untuk menghindari adanya subjektifitas dalam penilaian kinerja anggota satuan pengamanan dan untuk mengatasi kendala-kendala tersebut maka diperlukan data mining yang merupakan teknik baru dimana terdapat penggabungan antara metode analisis tradisional dengan algoritma untuk memproses data berjumlah besar. Salah satu manfaat data mining dalam perusahaan yaitu dapat digunakan sebagai bahan untuk mendukung pengambilan keputusan dan menunjang program

kegiatan perusahaan. Salah satunya untuk menentukan dipertahankan atau tidaknya anggota satuan pengaman di perusahaan. Data mining merupakan proses ataupun kegiatan untuk mengumpulkan data yang berukuran besar kemudian mengekstraksi data tersebut menjadi informasi – informasi yang nantinya dapat digunakan. Data mining merupakan sebuah inti dari proses *Knowledge Discovery* in Database, meliputi dugaan Algoritma yang mengeksplor data, membangun model dan menemukan pola yang belum diketahui. *Knowledge Discovery* in Database (KDD) bersifat otomatis, dapat didefinisikan sebagai pengorganisasian proses untuk pengidentifikasian yang benar, berguna dan penemuan pola dari kumpulan data yang besar dan komplek. Data mining juga bisa disebut sebagai proses pengekstrakan informasi dari kumpulan data yang besar. Pengekstrakan informasi dilakukan berdasarkan metode data mining yang akan digunakan. Ada beberapa metode serta algoritma data mining yang digunakan untuk mengekstrak informasi antara lain: metode klasifikasi, metode asosiasi, metode clustering, metode prediksi, dan metode estimasi. Dengan Kelebihan yang dimiliki metode Algoritma C4.5 yaitu sebuah Algoritma yang digunakan untuk membangun *decision tree* (pengambilan keputusan). Algoritma C4.5 adalah salah satu Algoritma induksi pohon keputusan, yaitu ID3 (*Iterative Dichotomiser 3*) [5]. ID3 dikembangkan oleh J. Ross Quinlan. Dalam prosedur Algoritma ID3, input berupa sampel training, label training, dan atribut. Algoritma C4.5 merupakan pengembangan dari ID3. Pohon dibangun dengan cara membagi data secara rekursif hingga tiap bagian terdiri dari data yang berasal dari kelas yang sama. Bentuk pemecahan (*split*) yang digunakan untuk membagi data tergantung dari jenis atribut yang digunakan dalam *split*. Algoritma C4.5 dapat menangani data numerik (kontinyu) dan diskret. Split untuk atribut numerik yaitu mengurutkan contoh berdasarkan atribut kontinyu A, kemudian membentuk minimum permulaan (*threshold*) M dari contoh - contoh yang ada dari kelas mayoritas pada setiap partisi yang bersebelahan, lalu menggabungkan partisi-partisi yang bersebelahan tersebut dengan kelas mayoritas yang sama. *Split* untuk atribut diskret a mempunyai bentuk value (A) X dimana $X \subset \text{domain}(A)$. Jika suatu set data mempunyai variabel tidak ada, Jika jumlah pengamatan terbatas maka atribut dengan missing value dapat diganti dengan nilai rata - rata dari variabel yang bersangkutan. Dalam melakukan pemisahan obyek (*split*) dilakukan tes terhadap atribut dengan mengukur tingkat ketidakmurnian pada sebuah simpul (*node*) [6]. Algoritma C4.5 adalah salah satu algoritma dari metode klasifikasi yang digunakan untuk membuat pohon keputusan. Pohon keputusan merupakan metode klasifikasi dan prediksi yang sangat kuat dan terkenal. Metode pohon keputusan mengubah fakta yang sangat besar menjadi pohon keputusan yang merepresentasikan aturan. Aturan dapat dengan mudah dipahami dengan bahasa alami. Dan metode ini dirasa dapat digunakan untuk memecahkan masalah pada klasifikasi kepuasan pegawai menggunakan data kuisioner.

Penelitian terkait dengan penerapan algoritma C4.5 dalam berbagai bidang telah menjadi fokus penelitian yang signifikan. Sebagai contoh, dalam bidang pendidikan, penelitian oleh Abdurraghib Segaf Suweleh (2020) menggunakan algoritma C4.5 untuk menentukan penerima beasiswa [7]. Di sisi lain, Alkhairi dan Situmorang (2022) menerapkan data mining dengan algoritma C4.5 untuk menganalisis kepuasan karyawan terhadap layanan sumber daya manusia [8]. Dalam penelitian pemilihan jurusan siswa, Anestiviya, Ferico, dan Pasaribu (2021) melakukan analisis pola dengan metode C4.5 untuk memahami minat siswa berdasarkan kurikulum [9]. Di dunia bisnis, Ardiansyah dan Walim (2018) menggunakan algoritma C4.5 untuk mengklasifikasikan calon peserta lomba cerdas cermat siswa SMP [10]. Penerapan algoritma C4.5 juga ditemukan dalam analisis keuangan, seperti yang dilakukan oleh Ginting, Kusrini, dan Taufiq (2020) untuk memprediksi keterlambatan pembayaran sumbangan pendidikan [13]. Begitu juga dalam konteks pelayanan perbankan, Hakim Tanjung (2021) mengoptimalkan algoritma C4.5 untuk memprediksi kerusakan mesin ATM [14]. Dalam sektor kesehatan, Irawan (2021) menggunakan algoritma Decision Tree C4.5 untuk memprediksi kelayakan calon pendonor darah [15]. Selain itu, Irnanda, Hartama, dan Windarto (2021) menganalisis faktor-faktor penyebab penurunan prestasi belajar mahasiswa selama pandemi [16]. Terakhir, penelitian Nofitri dan Eska (2018) mencoba mengimplementasikan algoritma C4.5 dalam menentukan kelayakan pengambilan kredit [19]. Semua penelitian ini menunjukkan beragam penerapan algoritma C4.5 dalam berbagai domain untuk analisis dan pengambilan keputusan. Dalam lingkup PT. YIMM Pulogadung, penggunaan data mining dengan algoritma C4.5 menjanjikan sebagai

alat efektif untuk memprediksi evaluasi anggota Satuan Pengamanan. Penelitian ini memiliki potensi besar untuk meningkatkan pemahaman tentang faktor-faktor yang memengaruhi evaluasi tersebut dan dapat memberikan wawasan berharga yang dapat digunakan dalam pengambilan keputusan yang lebih baik dalam manajemen anggota Satuan Pengamanan PT. YIMM Pulogadung.

Selain penelitian yang telah disebutkan sebelumnya, terdapat penelitian-penelitian relevan lainnya yang berkaitan dengan aplikasi data mining dalam beragam konteks. Sebagai contoh, dalam kajian yang dilakukan oleh Lajawala (2020), algoritma klasifikasi digunakan untuk menganalisis kesehatan mental dengan melakukan prediksi berbasis data mining [26]. Penelitian ini secara khusus menyoroti peran penting data mining dalam mendukung pemahaman dan evaluasi kondisi psikologis individu serta menyediakan wawasan berharga dalam bidang kesehatan mental. Selanjutnya, Manullang, Buulolo, dan Lubis (2020) mengimplementasikan data mining dalam rangka memprediksi jumlah pinjaman dengan menggunakan algoritma C4.5, terutama dalam konteks Kopdit CU Damai Sejahtera [27]. Penelitian ini mengilustrasikan bagaimana data mining dapat diaplikasikan dalam domain keuangan untuk meningkatkan pemahaman tentang tren pinjaman dan memberikan kontribusi positif dalam pengambilan keputusan yang lebih efektif. Selain itu, Rambola, Varshney, dan Vishwakarma (2018) melakukan percobaan untuk menerapkan teknik data mining guna mendeteksi potensi tindakan kecurangan di sektor perbankan [28]. Penelitian ini menunjukkan potensi data mining dalam mendukung sektor perbankan dalam mengidentifikasi perilaku yang mencurigakan dan mengurangi risiko keuangan. Semua penelitian ini mencerminkan bagaimana data mining, termasuk penggunaan algoritma C4.5, telah menjadi alat yang sangat berharga dalam berbagai konteks, mulai dari kesehatan mental hingga keuangan dan perbankan. Data mining membantu menganalisis data, melakukan prediksi, dan memberikan dukungan yang penting dalam pengambilan keputusan yang lebih efisien dan efektif.

2. Metode Penelitian

Data penelitian adalah kumpulan informasi atau fakta yang dikumpulkan atau dihasilkan dalam sebuah studi atau penelitian. Data ini biasanya terdiri dari fakta, angka, atau informasi yang terukur dan dapat diuji secara empiris. Data penelitian dapat berasal dari berbagai sumber, seperti *survei*, wawancara, observasi, atau pengukuran. Berikut adalah penjelasan untuk data penelitian yang diperoleh, sebagai berikut:

- 1) Data penelitian yang digunakan dalam penelitian ini diperoleh dari PT. YIMM Pulogadung. Data tersebut mencakup informasi tentang anggota satuan pengaman, termasuk atribut-atribut yang relevan dengan evaluasi kinerja dan kualitas mereka.
- 2) Data tersebut mencakup variabel-variabel seperti riwayat pelatihan, pengalaman kerja, hasil tes atau evaluasi sebelumnya, dan informasi lainnya yang relevan dengan kinerja anggota satuan pengaman.
- 3) Sumber data ini dapat berupa data internal PT. YIMM Pulogadung yang bersifat privat, yang mungkin termasuk data personal, kinerja, dan profil anggota satuan pengaman.
- 4) Untuk menjaga kerahasiaan dan privasi, penggunaan data penelitian ini akan mengikuti persyaratan dan regulasi yang berlaku serta memperoleh izin dari PT. YIMM Pulogadung untuk mengakses dan menganalisis data tersebut.

2.1 Teknik Pengumpulan Data

Pada tahap ini pengumpulan data ini, akan ada dua langkah yang dilakukan yaitu :

- 1) Studi Observasi

Observasi adalah salah satu proses yang dilakukan secara mengamati langsung ke lokasi dan pencatatan secara langsung dan sistematis mengenai masalah-masalah yang akan di teliti. *Observasi* adalah salah satu sebagai satu tempat 23 pengumpulan data yang sesuai dengan tujuan si peneliti yang direncanakan dan di catat dengan

sistematis, dan di kontrol dengan keadaan (*reliabilitas*) dan kebenarannya (*validation*) dan kali ini penulis selaku observasi telah melakukan observasi pada STMIK Budi Darma Medan.

2) Studi Literatur

Studi literatur yaitu studi yang dilakukan penelitian dalam mencari teoriteori yang bersangkutan dengan penelitian yang dilakukan. Seperti teori-teori data mining, teori metode-metode, ataupun yang bersangkutan lainnya yang mana bias berupa jurnal, majalah atau buku.

2.2 Penerapan Metodologi

1) Penerapan Metodologi

Langkah pertama dalam penerapan metodologi adalah pengumpulan data yang relevan dari PT. YIMM Pulogadung. Hal ini dapat melibatkan akses ke sistem atau basis data perusahaan yang menyimpan informasi tentang anggota satuan pengamanan.

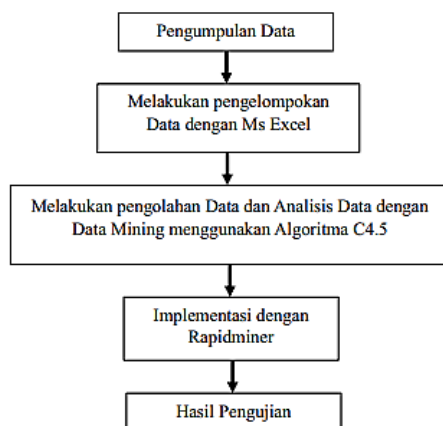
2) Pra-Pemrosesan

Selanjutnya, data yang telah dikumpulkan akan melalui tahap pra-pemrosesan. Langkah ini meliputi pembersihan data, pemilihan atribut yang relevan, dan penanganan nilai yang hilang atau outlier. Identifikasi atribut-atribut yang paling relevan atau memiliki pengaruh yang signifikan terhadap klasifikasi anggota Satuan Pengamanan. Gunakan teknik pemilihan fitur seperti analisis korelasi, analisis informasi, atau pemilihan berbasis model untuk memilih subset atribut yang paling informatif untuk proses klasifikasi.

3) Penerapan Algoritma C4.5

Setelah pra-pemrosesan, algoritma C4.5 akan diterapkan pada data yang telah dipersiapkan. Langkah ini melibatkan pembuatan model berdasarkan data training dan melakukan proses pengklasifikasian untuk mengklasifikasi evaluasi anggota satuan pengamanan di PT. YIMM Pulogadung. dapat di gambarkan dalam blok diagram sebagai berikut :

- a) Pengumpulan data yang dilakukan dengan memprediksi hasil penilaian karyawan anggota pengamanan, Prediksi dilakukan dari hasil kinerja serta perilaku atau akhlak dan absensi kehadiran. Yang bertepatan dengan evaluasi satu tahun pertama.
- b) Melakukan pengelompokan data dengan Microsoft Excel untuk memudahkan proses analisa maka data terlebih dahulu dikelompokkan berdasarkan bulan yang akan menjadi atribut pada Microsoft excel.
- c) Melakukan proses pengolahan data dan proses analisis data dengan menggunakan algoritma C4.5 setelah data diolah pada excel selanjutnya data karyawan anggota satuan pengamanan diolah dan diklasifikasikan menggunakan algoritma C4.5 untuk menghasilkan pohon keputusan (*decision tree*).
- d) Implementasi dengan *rapidminer*, data yang sudah diolah akan diuji dengan aplikasi *rapidminer* untuk menentukan output penelitian.
- e) Hasil penelitian yang di peroleh yaitu penentuan masa evaluasi kinerja karyawan satuan pengamanan dalam satu tahun, dari klasifikasi data sebelumnya dengan aplikasi *rapidminer* dalam bentuk pohon keputusan dan aturan (*rule*).



Gambar 1. Blok diagram algoritma C4.5

Evaluasi kinerja dan akurasi model akan dilakukan menggunakan metrik yang relevan seperti akurasi, presisi, recall, dan F1-score. Hal ini akan membantu dalam mengevaluasi sejauh mana algoritma C4.5 dapat mengklasifikasi dan mengevaluasi dengan akurat. Setelah model C4.5 dibangun menggunakan data yang telah melalui pra-pemrosesan, langkah selanjutnya adalah melakukan klasifikasi dan evaluasi kinerja serta akurasi model. Berikut adalah penjelasan mengenai klasifikasi dan evaluasi model:

2.2.1 Klasifikasi

- 1) Gunakan model C4.5 yang telah dilatih untuk melakukan klasifikasi pada data anggota Satuan Pengamanan yang baru. Setiap anggota akan diberikan atribut-atribut yang relevan sebagai input ke dalam model.
- 2) Model C4.5 akan menerapkan aturan-aturan yang telah dipelajari pada saat pelatihan untuk mengklasifikasikan anggota menjadi kelas atau kategori yang sesuai. Misalnya, dapat memprediksi apakah seorang anggota memiliki performa baik atau buruk, berdasarkan atribut-atribut yang dimiliki.

2.2.2 Evaluasi Kinerja

- 1) Gunakan subset pengujian (testing set) yang telah disiapkan sebelumnya untuk menguji model. Setiap anggota pada subset pengujian akan diberikan atribut-atribut sebagai input ke dalam model C4.5 yang telah dilatih.
- 2) Bandingkan hasil prediksi yang diberikan oleh model dengan kelas sebenarnya dari anggota pada subset pengujian.
- 3) Metrik evaluasi yang umum digunakan untuk mengukur kinerja model klasifikasi adalah sebagai berikut:
 - a) Akurasi (Accuracy): Mengukur persentase jumlah prediksi yang benar dibandingkan dengan jumlah total prediksi. Rumus: $(\text{Jumlah prediksi benar} / \text{Jumlah total prediksi}) * 100\%$.
 - b) Presisi (Precision): Mengukur tingkat ketepatan dari prediksi positif yang dilakukan oleh model. Rumus: $(\text{Jumlah prediksi positif benar} / \text{Jumlah prediksi positif}) * 100\%$.
 - c) Recall: Mengukur tingkat keseluruhan prediksi positif yang berhasil diidentifikasi oleh model. Rumus: $(\text{Jumlah prediksi positif benar} / \text{Jumlah anggota dengan kelas positif}) * 100\%$.
 - d) F1-Score: Menggabungkan presisi dan recall menjadi satu metrik yang mencerminkan keseimbangan antara keduanya. Rumus: $2 * [(\text{Presisi} * \text{Recall}) / (\text{Presisi} + \text{Recall})]$.

2.2.3 Akurasi Model

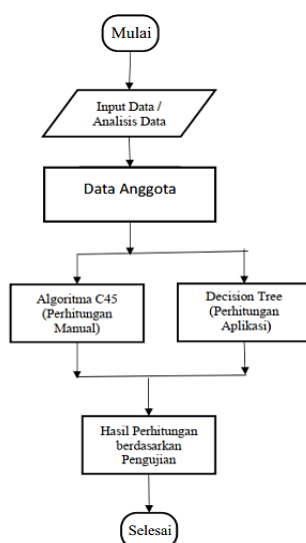
- 1) Akurasi model mengacu pada persentase prediksi yang benar yang dihasilkan oleh model. Ini dapat dihitung menggunakan rumus akurasi yang telah disebutkan di atas.
- 2) Akurasi yang tinggi menunjukkan bahwa model C4.5 memiliki kemampuan yang baik dalam mengklasifikasikan anggota Satuan Pengamanan.

Data yang digunakan adalah data penilaian anggota selama 1 tahun yang mana akan dilakukan awal tahun 2023. Setelah data-data diperoleh, maka data tersebut digunakan untuk mengetahui anggota mana saja yang dipertahankan atau tidak dipertahankan.

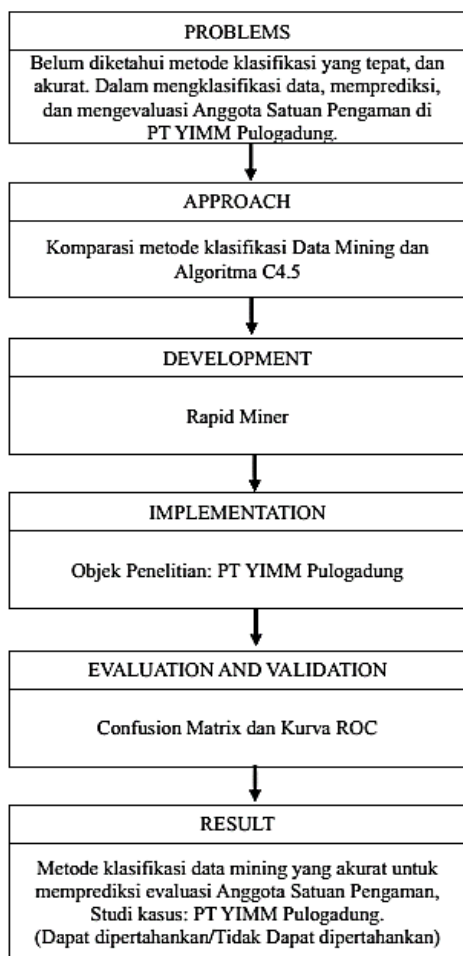
2.3 Rancangan Pengujian

Pada rancangan penelitian ini akan dibuat perincian atau tahap keseluruhan proses penelitian, mulai dari pengumpulan data hingga memperoleh kesimpulan dan hasil akhir yang didapat oleh peneliti, sehingga menghasilkan sebuah keputusan dipertahankan atau tidak dipertahakannya dalam penilaian anggota satuan pengamanan. Adapun alur rancangan dari penelitian ini dapat dilihat pada gambar 3.3 dibawah ini.

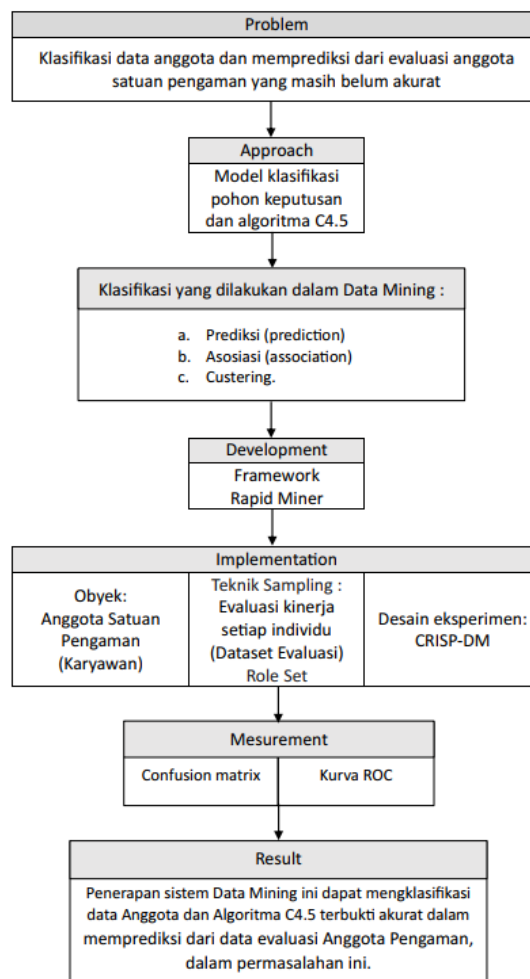
- 1) Pengujian akan dilakukan dengan menggunakan metode validasi silang (cross-validation). Data yang telah dikumpulkan akan dibagi menjadi subset/subsampil yang saling terpisah.
- 2) Bagian dari data akan digunakan sebagai data pelatihan (training set) untuk mengembangkan model klasifikasi dengan menggunakan algoritma C4.5. Model ini akan dilatih dengan memanfaatkan pola-pola yang ada pada data pelatihan.
- 3) Subset data yang tersisa akan digunakan sebagai data uji (test set) untuk menguji kinerja model yang telah dikembangkan. Model akan diterapkan pada data uji untuk mengklasifikasi evaluasi anggota satuan pengaman.
- 4) Selanjutnya, metrik evaluasi seperti akurasi, presisi, recall, dan F1-score akan dihitung berdasarkan hasil klasifikasi model. Metrik ini akan digunakan untuk mengevaluasi sejauh mana model dapat mengklasifikasi evaluasi anggota satuan pengaman dengan akurat.
- 5) Pengujian juga akan melibatkan perbandingan hasil klasifikasi dengan metode evaluasi tradisional yang digunakan di PT. YIMM Pulogadung. Hal ini akan membantu dalam mengevaluasi keunggulan dan kelemahan masing-masing pendekatan.
- 6) Pengujian akan dilakukan dalam beberapa iterasi untuk memastikan konsistensi dan reliabilitas hasil. Analisis statistik juga dapat diterapkan untuk menguji signifikansi perbedaan antara hasil klasifikasi menggunakan metode data mining dengan algoritma C4.5 dan metode evaluasi tradisional.



Gambar 2. Desain Flowchart Penelitian



Gambar 3. Struktur Rancangan Penelitian



Gambar 4. Kerangka Pemikiran

Tugas-tugas yang di lakukan dalam penelitian, pada penerapan data mining untuk diklasifikasikan menjadi tiga kategori utama: prediksi (*prediction*), asosiasi (*association*), dan clustering. Berikut adalah penjelasan tentang masing-masing kategori.

2.3.1 Prediksi (Prediction)

Tugas prediksi dalam data mining bertujuan untuk membuat model atau algoritma yang dapat digunakan untuk memprediksi dan mengklasifikasikan suatu kejadian atau objek berdasarkan pola dan atribut yang terdapat dalam data. Beberapa teknik yang digunakan dalam prediksi ini adalah pohon keputusan (*decision tree*), dan regresi (*regression*).

2.3.2 Asosiasi (Association)

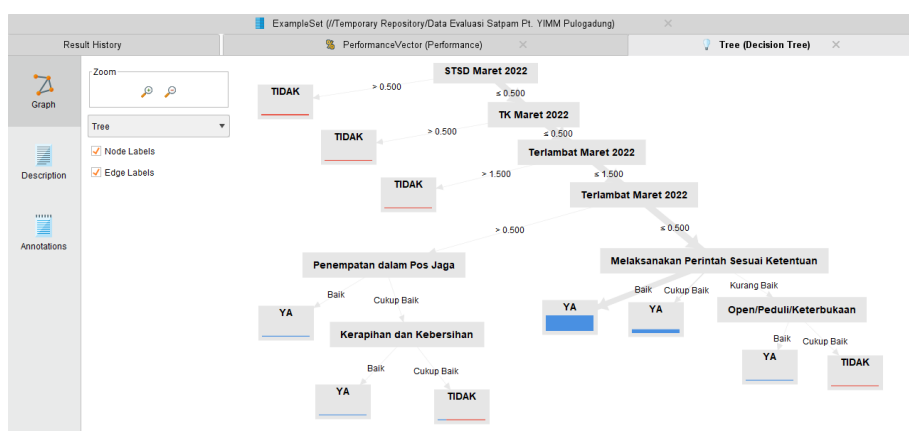
Tugas asosiasi dalam data mining bertujuan untuk menemukan hubungan atau keterkaitan antara item atau atribut dalam data. Teknik asosiasi sering digunakan untuk menemukan pola atau perilaku konsumen yang sering terjadi bersamaan. Asosiasi sering disebut sebagai analisis keranjang belanja (*market basket analysis*) merupakan metode dalam Data Mining untuk menemukan aturan untuk mengukur hubungan antara dua atau lebih atribut Algoritma yang umum digunakan dalam asosiasi adalah algoritma Apriori dan algoritma FP-Growth.

2.3.3 Clustering

Tugas clustering dalam data mining adalah untuk mengelompokkan objek-objek yang serupa berdasarkan kesamaan atribut atau karakteristik tertentu. Tujuannya adalah untuk menemukan kelompok-kelompok yang bermakna tanpa adanya label kelas sebelumnya. Algoritma clustering yang umum digunakan adalah algoritma k-means dan algoritma hierarkis.

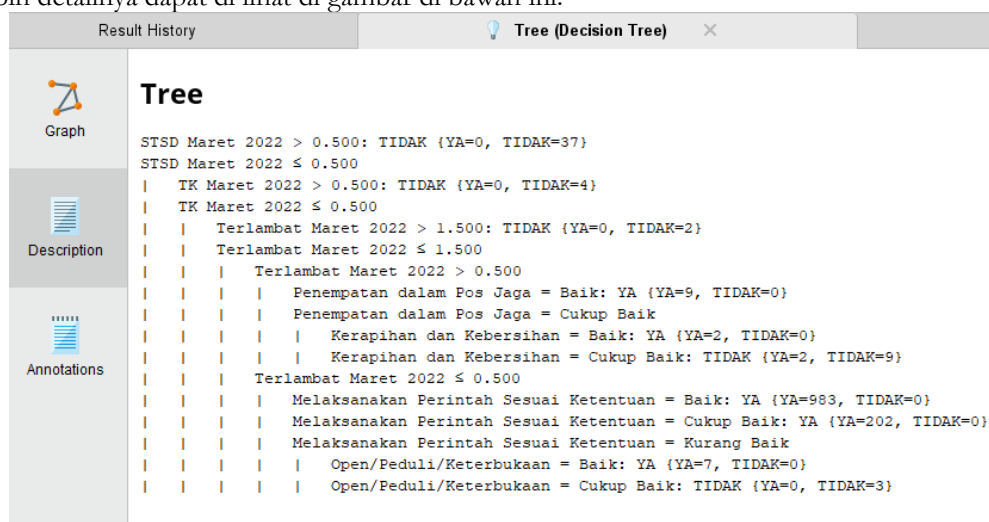
3. Hasil dan Pembahasan

Hasil akhir atau output dari aplikasi Rapid Miner adalah berupa pohon keputusan atau *decision tree*. Dari pohon keputusan inilah akan dihasilkan rules atau aturan-aturan yang dapat membantu dalam mengambil keputusan. Pohon keputusan akhir dapat dilihat pada Gambar 3.1 Tampilan Hasil Tree Akhir berikut.



Gambar 5. Tampilan Hasil Tree Akhir

Dari Gambar diatas kita sudah sedikit melihat gambaran hasil dari data yang sebelumnya di olah, untuk lebih detailnya dapat di lihat di gambar di bawah ini.



Gambar 6. Tampilan Description *Decision tree*

Dari Gambar di atas maka bisa di simpulkan bahwa *decision tree* yang terbentuk, didapatkan informasi bahwa faktor utama yang mempengaruhi anggota dipertahankan atau tidak dipertahankan

adalah salah satunya dari hasil akurasi berdasarkan kategori Kehadiran dengan penjelasan sebagai berikut:

- 1) STSD (Sakit tanpa surat dokter) dengan point > 0.500 diantaranya ada 37 kategori anggota yang tidak dapat dipertahankan.
- 2) STSD dengan point < 0.500 Jika ada TK (Tanpa keterangan) maka ada 4 kategori anggota yang tidak dapat dipertahankan.
- 3) Jika point TK maret < 0.500 dan point Terlambat maret > 1.500 maka ada 2 kategori anggota yang tidak dapat dipertahankan.
- 4) Jika point Terlambat > 0.500 dan Penempatan dalam pos jaga nilainya Baik maka ada 9 kategori anggota yang dapat dipertahankan.
- 5) Jika penempatan dalam pos jaga nilainya Cukup Baik, kerapihan dan kebersihannya nilainya Baik maka ada 2 kategori anggota yang dapat dipertahankan.
- 6) Jika penempatan dalam pos jaga nilainya Cukup Baik, kerapihan dan kebersihannya nilainya Cukup Baik maka ada 2 kategori anggota yang dapat dipertahankan dan 9 kategori anggota tidak dapat dipertahankan.
- 7) Jika point Terlambat < 0.500 dan Melaksanakan Perintah Sesuai Ketentuan nilainya Baik maka ada 983 kategori anggota yang dapat dipertahankan.
- 8) Jika point Terlambat < 0.500 dan Melaksanakan Perintah Sesuai Ketentuan nilainya Cukup Baik maka ada 202 kategori anggota yang dapat dipertahankan.
- 9) Jika point Terlambat < 0.500 dan Melaksanakan Perintah Sesuai Ketentuan nilainya Kurang Baik, Serta Open/Peduli/Keterbukaan nilainya Baik maka ada 7 kategori anggota yang dapat dipertahankan.
- 10) Jika point Terlambat < 0.500 dan Melaksanakan Perintah Sesuai Ketentuan nilainya Kurang Baik, Serta Open/Peduli/Keterbukaan nilainya Cukup Baik maka ada 3 kategori anggota yang dapat dipertahankan.

ExampleSet (/Temporary Repository/Data Evaluasi Satpam PT. YIMM Pufogadung)

Result History | Tree (Decision Tree) | PerformanceVector (Performance)

Open in Turbo Prep | Auto Model

Filter (1,260 / 1,260 examples): all

Row No.	Kriteria Penil...	Nama	Lokasi Kerja	Jabatan	Usia	Jenis Kelamin	Pelaksanaa...	Jumlah Abs...	SSD Maret 2...	STSD Maret ...	Izin mu
1	YA	ABDUL LATIF	PT. YIMM Pul...	Anggota Satp...	17 - 25 Tahun	Laki-Laki	Cukup Baik	20	0	0	0
2	YA	ABDUL LATIF	PT. YIMM Pul...	Anggota Satp...	17 - 25 Tahun	Laki-Laki	Cukup Baik	20	0	0	0
3	YA	ABDUL LATIF	PT. YIMM Pul...	Anggota Satp...	17 - 25 Tahun	Laki-Laki	Cukup Baik	20	0	0	0
4	YA	ABDUL LATIF	PT. YIMM Pul...	Anggota Satp...	17 - 25 Tahun	Laki-Laki	Cukup Baik	20	0	0	0
5	YA	ABDUL LATIF	PT. YIMM Pul...	Anggota Satp...	17 - 25 Tahun	Laki-Laki	Cukup Baik	20	0	0	0
6	YA	ABDUL LATIF	PT. YIMM Pul...	Anggota Satp...	17 - 25 Tahun	Laki-Laki	Cukup Baik	20	0	0	0
7	YA	ABDUL LATIF	PT. YIMM Pul...	Anggota Satp...	17 - 25 Tahun	Laki-Laki	Cukup Baik	20	0	0	0
8	YA	ABDUL LATIF	PT. YIMM Pul...	Anggota Satp...	17 - 25 Tahun	Laki-Laki	Cukup Baik	20	0	0	0
9	YA	ABDUL LATIF	PT. YIMM Pul...	Anggota Satp...	17 - 25 Tahun	Laki-Laki	Cukup Baik	20	0	0	0
10	YA	ABDUL LATIF	PT. YIMM Pul...	Anggota Satp...	17 - 25 Tahun	Laki-Laki	Cukup Baik	20	0	0	0
11	YA	ABDUL LATIF	PT. YIMM Pul...	Anggota Satp...	17 - 25 Tahun	Laki-Laki	Cukup Baik	20	0	0	0
12	YA	ABDUL LATIF	PT. YIMM Pul...	Anggota Satp...	17 - 25 Tahun	Laki-Laki	Cukup Baik	18	0	0	0
13	YA	ABDUL MUNIF	PT. YIMM Pul...	Anggota Satp...	26 - 35 Tahun	Laki-Laki	Baik	21	0	0	0

ExampleSet (1,260 examples, 1 special attribute, 31 regular attributes)

Gambar 7. Tampilan ExampleSet

Pada gambar di atas dapat kita lihat tampilan dataset anggota satuan pengamanan yang mana sebelumnya telah diimport atau di masukan ke dalam rapidminer.



Result History

ExampleSet (Retrieve Data Evaluasi Satpam PL YIMM Pulogadung) PerformanceVector (Performance) Tree (Decision Tree)

Plot

Plot 1

Plot type

Histogram

Value columns

SSD Maret 2022, Jum...

Color

-

Number of Bins

Number of bins

Plot style

Add new plot

General

X-Axis

Y-Axis

Title

ExampleSet

Frequency

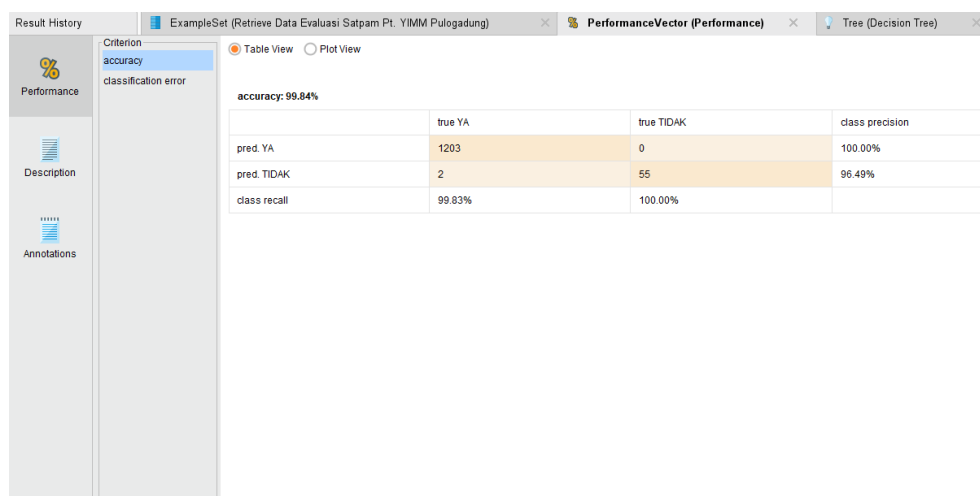
SSD Maret 2022, Jumlah Absensi Hari Kerja Bulan Maret 2022, STSD Maret 2022(...)

Legend:

- SSD Maret 2022
- Izin maret 2022
- Terlambat Maret 2022
- Jumlah Absensi Hari Kerja Bula...
- Cuti Maret 2022
- Hadir Maret 2022
- STSD Maret 2022
- TK Maret 2022

Gambar 9. Tampilan Charts ExampleSet

Jurnal Indonesia : Manajemen Informatika dan Komunikasi, Volume 4 No 3, September (2023), pp. 1566-1580
DOI : <https://doi.org/10.35870/jimik.v4i3.370>



Result History

ExampleSet (Retrieve Data Evaluasi Satpam Pt. YIMM Pulogadung) PerformanceVector (Performance) Tree (Decision Tree)

Criterion: accuracy, classification error

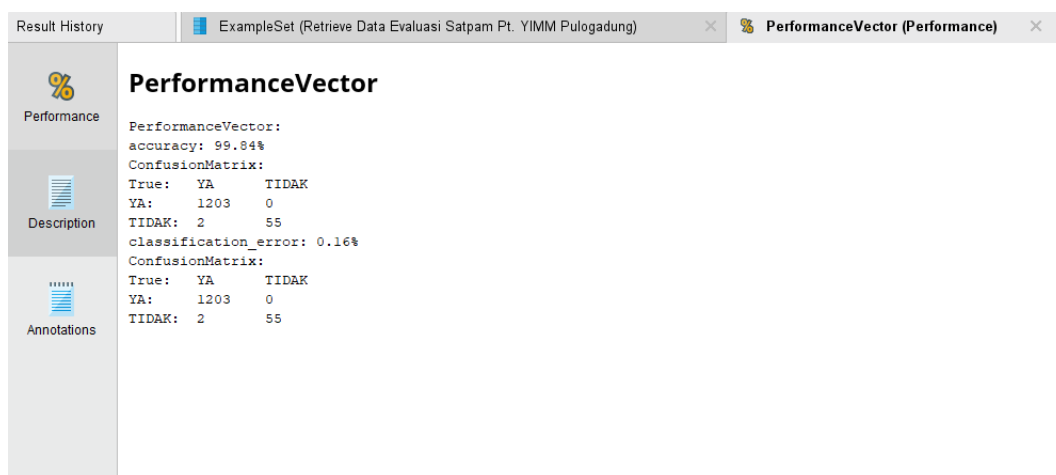
Table View Plot View

accuracy: 99.84%

	true YA	true TIDAK	class precision
pred. YA	1203	0	100.00%
pred. TIDAK	2	55	96.49%
class recall	99.83%	100.00%	

Gambar 10. Tampilan PerformanceVector

Dari gambar di atas menunjukkan bahwa akurasi data adalah 99.84%, yang di peroleh dari data yang sebelumnya diolah. untuk lebih jelasnya bisa di lihat pada gambar di bawah ini.



Result History

ExampleSet (Retrieve Data Evaluasi Satpam Pt. YIMM Pulogadung) PerformanceVector (Performance)

Performance

PerformanceVector

PerformanceVector:
accuracy: 99.84%

ConfusionMatrix:
True: YA TIDAK
YA: 1203 0
TIDAK: 2 55
classification_error: 0.16%

ConfusionMatrix:
True: YA TIDAK
YA: 1203 0
TIDAK: 2 55

Gambar 11. Tampilan Description PerformanceVector

4. Kesimpulan

Berdasarkan rumusan masalah yang telah diajukan dan pembahasan yang telah dilakukan, beberapa kesimpulan dapat diambil. Pertama, melalui analisis *decision tree*, ditemukan bahwa terdapat faktor utama yang memiliki dampak signifikan terhadap keputusan apakah seorang anggota akan dipertahankan atau tidak, yaitu tingkat akurasi mencapai 99, 84%. Hal ini menunjukkan betapa pentingnya tingkat akurasi dalam proses pengambilan keputusan terkait nasib anggota. Kedua, hasil yang diperoleh dari pendekatan manual dalam pembentukan pohon keputusan ternyata konsisten dengan hasil yang dihasilkan oleh Aplikasi Rapid Miner Classification *Decision tree*. Kesesuaian ini menggarisbawahi bahwa metode manual tetap valid dan mampu menghasilkan hasil yang serupa dengan metode otomatis yang lebih canggih.

5. Ucapan Terima Kasih

Puji dan syukur saya panjatkan kepada Tuhan Yang Maha Esa, karena atas berkat dan rahmat-Nya, saya dapat menyelesaikan karya tulis ilmiah ini. Penulisan karya tulis ilmiah ini dilakukan dalam rangka memenuhi salah satu syarat untuk mendapatkan gelar Sarjana (S1) Jurusan Teknik Informatika Sekolah Tinggi Ilmu Komputer Cipta Karya Informatika (STIKOM CKI). Saya menyadari bahwa tanpa bantuan dan bimbingan dari berbagai pihak, cukup sulit bagi saya untuk menyelesaikan karya tulis ilmiah ini. Pada kesempatan ini penulis menyampaikan terimakasih setulus-tulusnya kepada:

- 1) Ibu Mesra Betty Yel, MM, M.Kom selaku Ketua Sekolah Tinggi Ilmu Komputer Cipta Karya Informatika (STIKOM CKI).
- 2) Bapak Yuma Akbar, M.Kom selaku Wakil Ketua Bidang Akademik dan Kemahasiswaan.
- 3) Bapak. Kiki Setiawan, M.Kom selaku Wakil Ketua III Bidang Kemahasiswaan.
- 4) Bapak Dadang Iskandar Mulyana, M.Kom selaku Ketua Program Studi Teknik Informatika
- 5) Bapak Veri Arinal, M.Kom selaku Ketua Program Studi Sistem Informasi.
- 6) Management dan Seluruh Dosen Sekolah Tinggi Ilmu Komputer Cipta Karya Informatika.
- 7) Kedua orang tua saya dan kedua mertua saya yang selalu memberikan dukungan moril maupun materil dan tentunya doa yang tiada henti kepada saya.
- 8) Kepada Istri tercinta yang selalu mendukung saya untuk segera menyelesaikan pendidikan ini.
- 9) Seluruh Jajaran Anggota dan Pimpinan PT. Secom Bhayangkara Divisi PT. YIMM Pulogadung.
- 10) Untuk teman-teman Mahasiswa/i seperjuangan Sekolah Tinggi Ilmu Komputer Cipta Karya Informatika. Terimakasih atas pertemanan selama ini.
- 11) Seluruh responden yang telah memberikan waktu dan informasi untuk membantu penyelesaian skripsi ini.
- 12) Semua pihak yang tidak dapat disebutkan satu per satu, yang telah dengan tulus ikhlas memberikan doa dan motivasi sehingga dapat terselesaikannya skripsi ini.

Penulis menyadari dalam penulisan karya tulis ilmiah ini masih terdapat kekurangan, untuk itu diharapkan kritik dan saran yang membangun untuk dapat menyempurnakan karya tulis ilmiah ini. Akhir kata, penulis mengucapkan terima kasih dan semoga karya tulis ilmiah ini dapat bermanfaat bagi semua pihak yang membutuhkan.

6. Daftar Pustaka

- [1] Wibowo, A. P., & Hartati, S. (2016). Sistem Klasifikasi Kinerja Satpam Menggunakan Metode Naïve Bayes Classifier. *INOVTEK Polbeng-Seri Informatika*, 1(2), 192-201.
- [2] Echols, M. (2005). John dan Shadily. *Hasan Kamus Inggris Indonesia*.
- [3] Spencer, L., & Spencer, S. (1993). *Competence at Work*, Canada, Jhon Wiley & Sons.
- [4] Soetjipto, B. W. (2008). *Kisah sukses para kampiun SDM*. Penerbit Salemba.
- [5] Luvia, Y.S. *et al.* (2017) 'Penerapan Algoritma C4.5 Untuk Klasifikasi Predikat Keberhasilan Mahasiswa Di Amik Tunas Bangsa', *Jurasik (Jurnal Riset Sistem Informasi dan Teknik Informatika)*, 1(1), p. 75. Available at: <https://doi.org/10.30645/jurasik.v1i1.12>.
- [6] Rani, L.N. (2016) 'Klasifikasi Nasabah Menggunakan Algoritma C4.5 Sebagai Dasar Pemberian Kredit', *INOVTEK Polbeng - Seri Informatika*, 1(2), p. 126. Available at: <https://doi.org/10.35314/isi.v1i2.131>.

- [7] Abdurraghib Segaf Suweleh, D.S. (2020) 'Aplikasi Penentuan Penerima Beasiswa Menggunakan Algoritma C4.5', *Jurnal BITE*, 2(1), pp. 12–21. Available at: <https://doi.org/10.30812/bite.v2i1.798>.
- [8] Alkhairi, P. and Situmorang, Z. (2022) 'Application of Data Mining to Analyze Employee Satisfaction with HR Services with the C4.5 Algorithm', *Jurasik (Jurnal Riset Sistem Informasi dan Teknik Informatika)*, 7(1), p. 40. Available at: <https://tunasbangsa.ac.id/ejurnal/index.php/jurasik/article/view/414>.
- [9] Anestiviya, V., Ferico, A. and Pasaribu, O. (2021) 'Analisis Pola Menggunakan Metode C4.5 Untuk Peminatan Jurusan Siswa Berdasarkan Kurikulum (Studi Kasus : Sman 1 Natar)', *Jurnal Teknologi dan Sistem Informasi (JTSI)*, 2(1), pp. 80–85. Available at: <http://jim.teknokrat.ac.id/index.php/JTSI>.
- [10] Ardiansyah, D. and Walim, W. (2018) 'Algoritma c4.5 untuk klasifikasi calon peserta lomba cerdas cermat siswa smp dengan menggunakan aplikasi rapid miner', *Jurnal Inkofer*, 1(2), pp. 5–12.
- [11] Azwanti, N. (2018) 'Analisa Algoritma C4.5 Untuk Memprediksi Penjualan Motor Pada Pt. Capella Dinamik Nusantara Cabang Muka Kuning', *Informatika Mulawarman : Jurnal Ilmiah Ilmu Komputer*, 13(1), p. 33. Available at: <https://doi.org/10.30872/jim.v13i1.629>.
- [12] Azwanti, N. and Elisa, E. (2020) 'Analisa Kepuasan Konsumen Menggunakan Algoritma C4.5', *Prosiding Seminar Nasional Ilmu Sosial dan Teknologi*, (3), pp. 126–131.
- [13] Ginting, V.S., Kusri, K. and Taufiq, E. (2020) 'Implementasi Algoritma C4.5 untuk Memprediksi Keterlambatan Pembayaran Sumbangan Pembangunan Pendidikan Sekolah Menggunakan Python', *Inspiration: Jurnal Teknologi Informasi dan Komunikasi*, 10(1), pp. 36–44. Available at: <https://doi.org/10.35585/inspir.v10i1.2535>.
- [14] Hakim Tanjung, D.Y. (2021) 'Optimalisasi Algoritma C4.5 untuk Prediksi Kerusakan Mesin ATM', *Infosys (Information System) Journal*, 6(1), p. 12. Available at: <https://doi.org/10.22303/infosys.6.1.2021.12-21>.
- [15] Irawan, Y. (2021) 'Penerapan Algoritma Decision Tree C4.5 Untuk Memprediksi Kelayakan Calon Pendorong Melakukan Donor Darah Dengan Klasifikasi Data Mining', *JTIM: Jurnal Teknologi Informasi dan Multimedia*, 2(4), pp. 181–189. Available at: <https://doi.org/10.35746/jtim.v2i4.75>.
- [16] Irnanda, K.F., Hartama, D. and Windarto, A.P. (2021) 'Analisa Klasifikasi C4.5 Terhadap Faktor Penyebab Menurunnya Prestasi Belajar Mahasiswa Pada Masa Pandemi', *Jurnal Media Informatika Budidarma*, 5(1), p. 327. Available at: <https://doi.org/10.30865/mib.v5i1.2763>.
- [17] Marlina, D. and Bakri, M. (2021) 'Penerapan Data Mining Untuk Memprediksi Transaksi Nasabah Dengan Algoritma C4.5', *Jurnal Teknologi dan Sistem Informasi (JTSI)*, 2(1), pp. 23–28.
- [18] Nasrullah, A.H. (2018) 'Penerapan Metode C4.5 untuk Klasifikasi Mahasiswa Berpotensi Drop Out', *ILKOM Jurnal Ilmiah*, 10(2), pp. 244–250. Available at: <https://doi.org/10.33096/ilkom.v10i2.300.244-250>.

- [19] Nofitri, R. and Eska, J. (2018) 'Implementasi Data Mining Klasifikasi C4.5 Dalam Menentukan Kelayakan Pengambilan Kredit', *Seminar Nasional Royal (SENAR)*, p.
- [20] Puspa, M.A. (2019) 'Implementasi Data Mining Klasifikasi Algoritma C4.5 Dalam Perekrutan Perangkat Desa', *Building of Informatics, Technology and Science (BITS)*, 1(2), pp. 92–97.
- [21] Takalapeta, S. (2018) 'Penerapan Data Mining Untuk Menganalisis Kepuasan Konsumen Menggunakan Metode Algoritma C4.5', *J I M P - Jurnal Informatika Merdeka Pasuruan*, 3(3), pp. 34–38. Available at: <https://doi.org/10.37438/jimp.v3i3.186>.
- [22] Tumanggor, A. and Hasugian, P.S. (2021) 'Penerapan Data Mining Untuk Memprediksi Tingkat Kemampuan Anak Dalam Mengikuti Mata Pelajaran Dengan Metode C4.5 Pada SDN 105351 Bakaran Batu', *Jurnal Nasional Komputasi dan Teknologi Informasi (JNKTI)*, 4(1), pp. 57–63. Available at: <https://doi.org/10.32672/jnkti.v4i1.2718>.
- [23] Widiastari, A., Solikhun, S. and Irawan, I. (2021) 'Analisa Datamining dengan Metode Klasifikasi C4.5 Sebagai Faktor Penyebab Tanah Longsor', *Journal of Computer System and Informatics (JoSYC)*, 2(3), pp. 247–255. Available at: <http://ejurnal.seminar-id.com/index.php/josyc/article/view/741>.
- [24] Romli, I. and Zy, A.T. (2020) 'Penentuan Jadwal Overtime Dengan Klasifikasi Data Karyawan Menggunakan Algoritma C4.5', *Jurnal Sains Komputer & Informatika (J-SAKTI)*, 4(2), pp. 694–702.
- [25] Siahaan, N.V.B. *et al.* (2022) 'Penerapan Data Mining Algoritma C4.5 Terhadap Prediksi Faktor Menurunnya Hasil Panen Padi', *Jurasik (Jurnal Riset Sistem Informasi dan Teknik Informatika)*, 7(1), p. 27. Available at: <https://doi.org/10.30645/jurasik.v7i1.412>.
- [26] Laijawala, V. (2020) 'Algoritma Klasifikasi Kesehatan Mental Berbasis Algoritma Prediksi menggunakan Data Mining', (Icces), pp. 1174–1178.
- [27] Manullang, S.D., Buulolo, E. and Lubis, I. (2020) 'Implementasi Data Mining Dalam Memprediksi Jumlah Pinjaman Dengan Algoritma C4.5 Pada Kopdit CU Damai Sejahtera', *Jurnal Sistem Komputer dan Informatika (JSON)*, 1(3), p. 265. Available at: <https://doi.org/10.30865/json.v1i3.2153>.
- [28] Rambola, R., Varshney, P. and Vishwakarma, P. (2018) 'Teknik Data Mining untuk Deteksi Kecurangan di Sektor perbankan', pp. 1–5.