

IMPLEMENTASI ALGORITMA *DECISION TREE* DALAM SISTEM PENDUKUNG KEPUTUSAN SELEKSI CALON KARYAWAN PADA PT G-TEXTILE INDONESIA MENGGUNAKAN *FRAMEWORK DJANGO PYTHON*

Aditya Irawan, Dewi Handayani Untari Ningsih
Teknik Informatika, Universitas Stikubank Semarang
Jl. Tri Lomba Juang 1, Semarang, Indonesia
sousukeshirogane@gmail.com

ABSTRAK

Proses rekrutmen karyawan merupakan tahapan penting dalam manajemen sumber daya manusia karena berpengaruh langsung terhadap kualitas tenaga kerja yang dihasilkan. Namun, pada banyak perusahaan, termasuk PT G-Textile Indonesia, proses seleksi masih dilakukan secara manual sehingga berpotensi menimbulkan kesalahan, membutuhkan waktu yang lama, dan bersifat subjektif. Permasalahan tersebut mendorong perlunya sistem rekrutmen yang lebih efisien, objektif, dan terintegrasi. Penelitian ini bertujuan mengembangkan sistem informasi rekrutmen berbasis web yang terintegrasi dengan algoritma *Decision Tree* sebagai sistem pendukung keputusan seleksi calon karyawan. Sistem dibangun menggunakan *framework Django Python*, basis data *SQLite*, dan template *AdminLTE*, dengan fitur pengelolaan departemen, permintaan tenaga kerja, data pelamar, monitoring seleksi, dan manajemen pengguna. Algoritma *Decision Tree* digunakan untuk mengklasifikasikan pelamar ke dalam kategori “Diterima” atau “Ditolak” berdasarkan variabel usia, pendidikan, pengalaman kerja, hasil psikotes, dan wawancara. Hasil evaluasi model menunjukkan performa yang sangat baik dengan nilai *accuracy* 98,8%, *precision* 98,75%, *recall* 100%, dan *F1-score* 98,1%. Temuan ini membuktikan bahwa sistem mampu mempercepat proses evaluasi, mengurangi subjektivitas penilaian, serta meningkatkan transparansi dan kualitas keputusan rekrutmen.

Kata kunci : *Decision Tree, Sistem Rekrutmen, Django, Klasifikasi, Sistem Pendukung Keputusan*

1. PENDAHULUAN

Sumber daya manusia merupakan elemen strategis dalam mendukung keberhasilan operasional dan daya saing perusahaan, khususnya pada industri tekstil yang memiliki kebutuhan tenaga kerja bersifat fluktuatif dan proses produksi yang berlangsung cepat. Kondisi tersebut menuntut perusahaan memiliki sistem rekrutmen yang mampu menyediakan kandidat yang tepat secara efisien dan terukur. Seiring meningkatnya jumlah pelamar serta tingginya tingkat turnover, proses seleksi karyawan menjadi semakin kompleks dan membutuhkan dukungan sistem yang andal berbasis teknologi informasi [1]. Namun, pada praktiknya, proses rekrutmen di banyak perusahaan, termasuk PT G-Textile Indonesia, masih dilakukan secara manual melalui seleksi berkas, pencatatan menggunakan spreadsheet, serta penilaian subjektif berdasarkan wawancara. Proses manual tersebut berpotensi menimbulkan ketidakkonsistenan keputusan, meningkatkan risiko kesalahan, serta mengurangi transparansi dalam penentuan hasil seleksi [2]. Selain itu, evaluasi pelamar yang sangat bergantung pada penilaian individu pihak *Human Resource Development (HRD)* menyebabkan proses pengambilan keputusan sulit diukur secara objektif dan konsisten [3].

Meningkatnya volume pelamar dan tingginya tingkat *turnover* pada sektor tekstil menambah kompleksitas proses rekrutmen. Perusahaan membutuhkan mekanisme seleksi yang mampu menghasilkan keputusan secara lebih objektif, cepat, dan terukur. Sejalan dengan perkembangan teknologi

informasi, pendekatan berbasis data melalui metode *machine learning* mulai banyak dimanfaatkan untuk meningkatkan kualitas pengambilan keputusan. Metode klasifikasi, khususnya *Decision Tree*, dinilai efektif dalam mengidentifikasi pola penilaian kandidat karena mampu menyajikan model yang mudah dipahami oleh HRD serta dapat menunjukkan atribut yang paling berpengaruh terhadap kelulusan pelamar [4].

Sejumlah penelitian sebelumnya telah menerapkan algoritma *Decision Tree* dalam proses seleksi karyawan. Salah satu penelitian menggunakan algoritma *C4.5* untuk menilai kelayakan pelamar melalui tahapan screening, wawancara, dan psikotes, dan menghasilkan akurasi 99,43% [5]. Penelitian lain pada Lotte Grosir Batam menunjukkan bahwa *Decision Tree* mampu mengklasifikasikan calon karyawan berdasarkan nilai tes masuk dan pendidikan dengan akurasi 94,8% [6]. Studi terkait lainnya menegaskan bahwa metode ini efektif dalam mengidentifikasi variabel penentu keputusan HRD dengan akurasi 98,44% [7]. Hasil tersebut menunjukkan potensi kuat *Decision Tree* dalam mendukung proses seleksi yang lebih objektif dan sistematis.

Berbagai penelitian terdahulu menunjukkan bahwa penerapan algoritma *Decision Tree* mampu meningkatkan objektivitas dan akurasi dalam proses seleksi karyawan. Meskipun demikian, sebagian besar penelitian tersebut masih berfokus pada pembangunan dan evaluasi model klasifikasi semata, tanpa mengintegrasikannya ke dalam sistem rekrutmen

berbasis web yang dapat digunakan secara langsung dalam proses bisnis perusahaan. Selain itu, implementasi sistem pendukung keputusan berbasis Decision Tree pada konteks industri tekstil di Indonesia masih relatif terbatas. Kondisi ini menunjukkan adanya research gap, yaitu perlunya sistem rekrutmen terintegrasi yang tidak hanya mengelola data pelamar secara digital, tetapi juga mampu memberikan rekomendasi kelulusan secara real-time berdasarkan pola data historis yang akurat.

Berdasarkan permasalahan dan celah penelitian tersebut, penelitian ini bertujuan mengembangkan sistem informasi rekrutmen berbasis web yang terintegrasi dengan algoritma *Decision Tree* sebagai sistem pendukung keputusan seleksi calon karyawan di PT G-Tekstil Indonesia. Penelitian ini diharapkan dapat memberikan kontribusi baik secara akademik maupun praktis, khususnya dalam penerapan metode machine learning yang dapat diinterpretasikan dengan mudah oleh HRD serta diimplementasikan langsung dalam proses rekrutmen perusahaan.

Sebagai rencana penyelesaian masalah, penelitian ini merancang dan membangun sistem rekrutmen berbasis web menggunakan framework *Django Python*, basis data *SQLite*, dan template AdminLTE. Algoritma *Decision Tree* diterapkan untuk mengklasifikasikan calon karyawan ke dalam kategori “Diterima” dan “Ditolak” berdasarkan variabel usia, tingkat pendidikan, pengalaman kerja, hasil psikotes, dan wawancara. Model yang dihasilkan dievaluasi menggunakan metrik *accuracy*, *precision*, *recall*, dan *F1-score*, kemudian diintegrasikan ke dalam modul monitoring seleksi untuk mendukung pengambilan keputusan HRD secara lebih objektif, konsisten, dan transparan.

2. TINJAUAN PUSTAKA

2.1. Penelitian Terdahulu

Penelitian sebelumnya membahas tentang Klasifikasi Seleksi Penerimaan Karyawan Baru Menggunakan Algoritma C4.5. Penelitian ini menerapkan algoritma C4.5 untuk proses seleksi calon karyawan baru melalui tahapan screening CV, wawancara, psikotes, dan *interview user*. Tujuannya membuat proses seleksi lebih efisien dan objektif. Metode Data Mining menggunakan algoritma C4.5 *Decision Tree* dengan tahapan data *preprocessing*, model *training*, dan *evaluation* (*accuracy*, *precision*, *recall*, *AUC*). Model C4.5 menghasilkan akurasi 99,43%, *precision* 99,29%, *recall* 97,69%, dan *AUC* 0,750. Terbukti efektif meningkatkan objektivitas dan efisiensi seleksi karyawan [5].

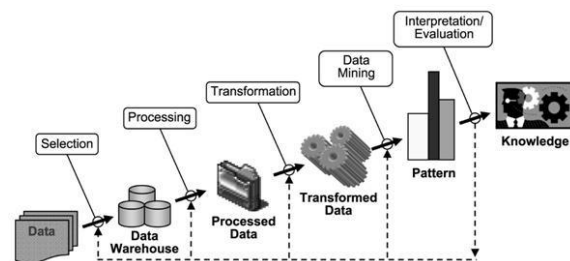
Penelitian terdahulunya membahas tentang bagaimana menentukan penerimaan Karyawan Super Market di Lotte Grosir Batam Penggunaan algoritma *Decision Tree* (C4.5) untuk menentukan calon karyawan yang memenuhi standar perusahaan berdasarkan nilai tes masuk, prestasi kerja, dan pendidikan. Algoritma C4.5 *Decision Tree* diterapkan untuk membentuk model klasifikasi kelayakan

pelamar berdasarkan nilai tes dan riwayat pendidikan. Model menghasilkan akurasi 94,8%. Sistem klasifikasi membantu mempercepat dan mempermudah proses seleksi karyawan sesuai kriteria Lotte Grosir Batam [6].

Penelitian lainnya terkait membahas tentang membahas penentuan kelayakan calon pegawai melalui data mining. Menunjukkan atribut mana yang paling memengaruhi keputusan HRD sehingga proses seleksi menjadi terstruktur dan objektif. Hasil penelitian menunjukkan bahwa nilai akurasi yang diperoleh 98,44% [7]

2.2. Data Mining dan Knowledge Discovery In Database (KDD)

Knowledge Discovery in Databases (KDD) merupakan suatu metode untuk menemukan pengetahuan atau informasi baru yang sebelumnya tidak tampak dari sebuah basis data [8]. Istilah ini sering dianggap sama dengan data mining, meskipun keduanya sebenarnya memiliki konsep yang berbeda. *KDD* mencakup proses yang lebih luas, sementara data mining adalah salah satu tahap penting dalam keseluruhan proses tersebut [9].



Gambar 1. Proses KDD

Data mining sendiri adalah aktivitas menganalisis kumpulan data berukuran besar guna mengidentifikasi pola, kecenderungan, serta hubungan tersembunyi di dalamnya [10]. Proses ini tidak sebatas mengambil data, tetapi juga melibatkan pembangunan model prediktif yang dapat digunakan untuk memperkirakan perilaku atau kejadian di masa mendatang [11]. Dengan menerapkan teknik statistik serta algoritma pembelajaran mesin, data mining banyak dimanfaatkan dalam berbagai bidang seperti pemasaran, kesehatan, dan deteksi penipuan untuk membantu menghasilkan keputusan yang lebih akurat [12].

2.3. Sistem Informasi Rekrutmen

Sistem informasi rekrutmen merupakan aplikasi terintegrasi yang digunakan untuk mengelola proses seleksi calon karyawan secara digital [13]. Sistem ini mencakup pengelolaan data pelamar, penjadwalan seleksi, penilaian, serta dokumentasi hasil keputusan. Penggunaan sistem rekrutmen berbasis web mampu meningkatkan efisiensi operasional karena data dapat diakses secara *real-time* dan terpusat. Studi-studi terdahulu menunjukkan bahwa digitalisasi proses rekrutmen dapat mengurangi beban administratif,

meningkatkan transparansi, serta meminimalkan kesalahan manusia dalam penilaian

2.4. Decision Tree

Decision Tree merupakan salah satu algoritma *supervised learning* yang digunakan untuk menyelesaikan masalah klasifikasi dan regresi [14]. Algoritma ini bekerja dengan membentuk struktur pohon yang tersusun dari simpul akar (*root node*), simpul keputusan (*decision node*), dan simpul daun (*leaf node*) [15], [16]. Setiap simpul keputusan merepresentasikan proses pengujian terhadap suatu atribut, sedangkan simpul daun menggambarkan hasil keputusan atau kelas yang diprediksi [17].

Rumus perhitungan *Gain*:

$$Gain(S, A) = Entropy(S) - \sum_{i=1}^n *$$

$$Entropy(S_i) \quad (1)$$

Keterangan :

S: himpunan

A: atribut

n: Jumlah Partisi atribut A

|Si| : Jumlah kasus pada partikel ke-i

|S| : Jumlah kasus dalam S

Rumus perhitungan Entropy:

$$Entropy(S) = \sum_{i=1}^n -p_i * \log_2 p_i \quad (2)$$

Keterangan :

S: himpunan

A: Fitur

n: Jumlah Partisi S

pi : Proporsi dari S

Proses pembentukan *Decision Tree* dilakukan melalui mekanisme pembagian data secara berulang (*recursive partitioning*). Pada setiap tahap, algoritma memilih atribut terbaik yang dapat memisahkan data berdasarkan tingkat homogenitas kelas. Pemilihan atribut ini biasanya didasarkan pada ukuran seperti *entropy*, *information gain*, *gain ratio*, atau *Gini index*, tergantung jenis algoritma *Decision Tree* yang digunakan. Semakin tinggi nilai *gain* atau semakin rendah *impurity* suatu atribut, semakin besar peluang atribut tersebut dijadikan simpul keputusan.

3. METODE PENELITIAN

3.1. Jenis dan Objek Penelitian

Penelitian ini menggunakan pendekatan *eksperimental* untuk menguji efektivitas algoritma *Decision Tree* dalam proses klasifikasi calon karyawan pada PT G-Textile Indonesia. Pendekatan *eksperimental* dipilih karena penelitian ini melibatkan pemberian *treatment* berupa penerapan model klasifikasi terhadap dataset pelamar, kemudian mengevaluasi hasilnya menggunakan metrik tertentu. Pada pendekatan ini, algoritma *Decision Tree* diimplementasikan sebagai variabel perlakuan untuk melihat sejauh mana model mampu mengklasifikasikan pelamar ke dalam kategori “Diterima” atau “Ditolak” berdasarkan atribut-atribut yang dinilai relevan, seperti usia, tingkat pendidikan, pengalaman kerja, hasil psikotes, dan nilai wawancara.

Dalam penelitian ini, model *Decision Tree* dievaluasi melalui serangkaian pengujian yang terkontrol, yaitu dengan membagi data menjadi data latih dan data uji, kemudian mengukur performanya menggunakan *accuracy*, *precision*, *recall*, dan *F1-score*. Proses eksperimen dilakukan berulang untuk memastikan konsistensi hasil dan reliabilitas model. Selanjutnya, model yang telah dievaluasi diintegrasikan ke dalam sistem pendukung keputusan berbasis web menggunakan *framework Django Python* agar dapat digunakan secara langsung oleh bagian HRD dalam proses seleksi. Dengan pendekatan *eksperimental* ini, penelitian mampu menunjukkan efektivitas penerapan algoritma *Decision Tree* secara empiris dalam mendukung pengambilan keputusan rekrutmen karyawan di PT G-Textile Indonesia.

3.2. Studi Literatur

Tinjauan literatur berfungsi sebagai landasan teoretis untuk mengidentifikasi referensi yang relevan dengan topik penelitian. Sumber-sumber literatur diperoleh dari buku, jurnal ilmiah, dan publikasi daring yang mendukung proses penyusunan studi. Fokus utama tinjauan literatur pada penelitian ini mencakup konsep sistem informasi rekrutmen, teori algoritma *Decision Tree*, serta pendekatan-pendekatan klasifikasi yang digunakan dalam sistem pendukung keputusan berbasis data. Selain itu, tinjauan literatur juga mencakup penelitian terdahulu yang mengimplementasikan *Decision Tree* atau metode serupa dalam proses seleksi karyawan sehingga dapat memberikan gambaran mengenai keunggulan, kelemahan, dan celah penelitian yang dapat dikembangkan lebih lanjut. Dengan adanya tinjauan literatur yang komprehensif, penelitian ini mendapatkan landasan konseptual yang kuat untuk mengembangkan sistem pendukung keputusan seleksi calon karyawan pada PT G-Textile Indonesia.

3.3. Pengumpulan dan Praproses Data

Data dalam penelitian ini diperoleh dari bagian *Human Resource Development (HRD)* PT G-Textile Indonesia, yang menyediakan data historis pelamar yang pernah mengikuti proses seleksi karyawan. Data tersebut mencakup berbagai informasi penting yang digunakan dalam penilaian kandidat, seperti usia, tingkat pendidikan, pengalaman kerja, hasil psikotes, dan nilai wawancara. Proses pengumpulan data dilakukan melalui beberapa metode, yaitu dokumentasi terhadap arsip seleksi yang tersimpan pada sistem internal perusahaan, wawancara dengan staf HRD untuk memahami indikator penilaian yang digunakan dalam proses rekrutmen, serta observasi langsung terhadap alur seleksi mulai dari tahap *screening* berkas hingga penentuan hasil akhir. Pendekatan ini memastikan bahwa data yang digunakan dalam pengembangan sistem benar-benar mencerminkan kondisi seleksi karyawan yang terjadi di PT G-Textile Indonesia.

Sebelum data digunakan untuk pembangunan model, dilakukan tahap praproses untuk memastikan data bersih, terstruktur, dan siap dianalisis. Tahap praproses mencakup pembersihan data untuk menghapus duplikasi dan mengatasi nilai kosong menggunakan metode imputasi yang sesuai dengan karakteristik atribut. Selanjutnya dilakukan transformasi data dengan mengonversi atribut kategorikal, seperti tingkat pendidikan, hasil psikotes, dan hasil wawancara, ke dalam bentuk numerik menggunakan teknik *label encoding*. Penyesuaian format juga dilakukan untuk memastikan konsistensi nilai antaratribut. Setelah itu, data dibagi menjadi data latih dan data uji dengan perbandingan 80:20 menggunakan fungsi *train_test_split* dari *scikit-learn*, sehingga proses pengujian model dapat dilakukan secara objektif. Validasi data juga dilakukan pada tahap ini guna memastikan bahwa seluruh atribut telah sesuai dengan kebutuhan model *Decision Tree*.

Dataset yang digunakan dalam penelitian ini terdiri atas 400 pelamar, bergantung pada jumlah data yang berhasil dihimpun dari HRD. Dataset memiliki enam atribut utama yang berperan sebagai variabel prediktor, yaitu usia, pendidikan terakhir, pengalaman kerja, hasil psikotes, hasil wawancara. Atribut target dari dataset ini adalah status seleksi yang dibagi menjadi dua kategori, yaitu “Diterima” dan “Ditolak.” Struktur dataset ini memungkinkan penerapan algoritma *Decision Tree* secara optimal dalam mengidentifikasi pola keputusan berdasarkan data historis perusahaan.

Melalui proses pengumpulan, praproses, dan deskripsi dataset yang terstruktur, penelitian ini memastikan bahwa model klasifikasi yang dibangun memiliki dasar data yang valid, representatif, dan sesuai dengan kebutuhan implementasi sistem pendukung keputusan seleksi calon karyawan pada PT G-Textile Indonesia.

3.4. Pembagian Data Latih dan Data Uji

Pembagian data latih dan data uji dalam penelitian ini dilakukan menggunakan metode *train-test split* dengan komposisi 80% sebagai data latih dan 20% sebagai data uji. Proses pemisahan dilakukan menggunakan fungsi dari *scikit-learn* secara acak untuk memastikan distribusi kelas tetap seimbang antara pelamar yang “Diterima” dan “Ditolak.” Pembagian ini dilakukan untuk melatih model pada sebagian besar data, kemudian mengevaluasi kemampuan model menggunakan data yang belum pernah digunakan pada proses pelatihan.

Model klasifikasi dibangun menggunakan algoritma *Decision Tree* dengan memanfaatkan atribut usia, pendidikan terakhir, pengalaman kerja, hasil psikotes, dan hasil wawancara sebagai variabel prediktor. Sementara itu, status seleksi pelamar digunakan sebagai variabel target. Pada tahap pembangunan model, beberapa parameter seperti *criterion*, *max_depth*, dan *min_samples_split* disesuaikan untuk memperoleh struktur pohon yang

optimal. Teknik *pruning* juga diterapkan untuk mencegah *overfitting* sehingga model tetap mampu melakukan prediksi yang baik pada data baru.

Setelah model selesai dilatih, evaluasi dilakukan menggunakan *Confusion Matrix*. *Confusion Matrix* digunakan untuk mengukur kinerja model berdasarkan empat komponen utama, yaitu *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN). Dari nilai-nilai ini dihitung metrik akurasi, presisi, recall, dan F1-score yang menunjukkan tingkat ketepatan dan keandalan model dalam mengklasifikasikan pelamar.

Model *Decision Tree* yang dihasilkan kemudian diintegrasikan ke dalam sistem berbasis *Django* sehingga proses klasifikasi pelamar dapat dilakukan secara otomatis pada modul Monitoring Seleksi. Integrasi ini membantu HRD dalam mempercepat proses penilaian dan menghasilkan keputusan yang lebih objektif dan konsisten.

3.5. Perbandingan dan Evaluasi Model

Evaluasi model *Decision Tree* dilakukan menggunakan *Confusion Matrix* untuk menilai kemampuan model dalam mengklasifikasikan pelamar ke dalam kategori “Diterima” dan “Ditolak.” *Confusion Matrix* memberikan informasi mengenai jumlah prediksi benar dan salah melalui nilai *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN). Berdasarkan nilai tersebut, dihitung metrik *accuracy*, *precision*, *recall*, dan *F1-score* menggunakan fungsi evaluasi pada *scikit-learn*. Metrik-metrik ini digunakan untuk menilai ketepatan dan keandalan model sebelum diintegrasikan ke dalam sistem pendukung keputusan seleksi karyawan di PT G-Textile Indonesia.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (3)$$

$$Precision = \frac{TP}{TP + FP} \quad (4)$$

$$Recall = \frac{TP}{TP + FN} \quad (5)$$

$$F1 - Score = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \quad (6)$$

4. HASIL DAN PEMBAHASAN

4.1. Pengambilan Data

Pengambilan data dalam penelitian ini dilakukan melalui dua tahap utama, yaitu wawancara dengan pihak *Human Resource Development* (HRD) dan pengumpulan dataset pelamar kerja di PT G-Textile Indonesia. Wawancara dilakukan bersama staf HRD untuk memahami alur proses rekrutmen, kriteria penilaian, serta atribut yang digunakan dalam menentukan kelulusan pelamar. Informasi yang diperoleh dari wawancara ini menjadi dasar dalam menentukan variabel yang relevan untuk membangun model klasifikasi.

Setelah memperoleh pemahaman tersebut, penelitian kemudian mengumpulkan dataset pelamar yang berisi data historis seleksi karyawan. Dataset mencakup beberapa atribut penting yang digunakan perusahaan, seperti usia, tingkat pendidikan, pengalaman kerja, hasil psikotes, dan nilai wawancara.

Atribut-atribut ini merupakan variabel yang menjadi dasar dalam proses evaluasi kelayakan pelamar. Data yang berhasil dikumpulkan kemudian diolah dan dianalisis lebih lanjut pada tahap berikutnya dalam pembangunan model *Decision Tree*.

4.2. Pra-Proses Data

Pada tahap praproses data, dilakukan beberapa langkah penting untuk memastikan dataset siap digunakan dalam proses pelatihan model *Decision Tree*. Langkah pertama adalah penanganan nilai kosong (*missing values*) dengan menggunakan metode imputasi. Atribut numerik diisi menggunakan nilai rata-rata untuk menjaga konsistensi distribusi data, sedangkan atribut kategorikal diisi menggunakan nilai yang paling sering muncul (*modus*). Pendekatan ini digunakan agar tidak terjadi kehilangan data yang berpotensi memengaruhi kinerja model.

Selain penanganan nilai kosong, atribut yang tidak relevan terhadap proses klasifikasi, seperti kolom penomoran atau informasi administratif yang tidak berpengaruh pada keputusan seleksi, dihapus dari dataset. Penghapusan ini dilakukan untuk menyederhanakan struktur data dan fokus pada variabel yang memiliki kontribusi terhadap hasil klasifikasi, seperti usia, pendidikan, pengalaman kerja, hasil psikotes, dan hasil wawancara.

Seluruh proses praproses data dilakukan dalam lingkungan *Django Python*. Data pelamar diimpor ke dalam database *SQLite*, kemudian diproses menggunakan fungsi bawaan *Python* dan *library* seperti *pandas* serta modul utilitas *Django*. Transformasi data, seperti pengubahan atribut kategorikal menjadi nilai numerik menggunakan *Label Encoding*, dilakukan sebelum data digunakan dalam pelatihan model. Melalui tahapan praproses ini, dataset yang bersih, konsisten, dan terstruktur dapat digunakan secara optimal oleh model *Decision Tree* untuk mendukung proses klasifikasi pelamar pada PT G-Textile Indonesia.

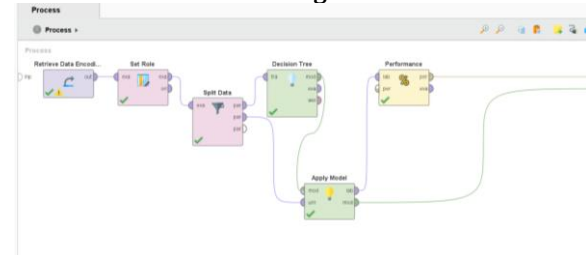
4.3. Pembagian Data Latih dan Data Uji

Pembagian data latih dan data uji dalam penelitian ini dilakukan menggunakan metode *train-test split* untuk menilai performa algoritma *Decision Tree* secara objektif. Data dibagi menjadi dua bagian, yaitu 80% sebagai data latih dan 20% sebagai data uji. Pembagian ini dilakukan secara acak untuk memastikan setiap kelas, baik “Diterima” maupun “Ditolak,” tetap terdistribusi secara seimbang pada kedua subset data.

Metode ini dipilih karena sederhana, efisien, dan sesuai untuk penelitian yang berfokus pada satu algoritma, yaitu *Decision Tree*. Dengan pembagian ini, model dapat dilatih menggunakan sebagian besar data, sementara data uji digunakan untuk mengevaluasi kemampuan model dalam melakukan prediksi pada data yang belum pernah dilihat sebelumnya. Pendekatan ini memastikan proses evaluasi lebih

realistik dan tidak hanya bergantung pada satu skenario pelatihan.

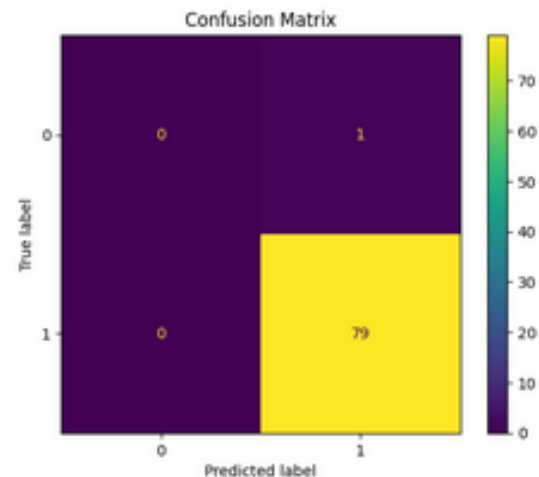
4.4. Pembuatan Model Algoritma *Decision Tree*



Gambar 2. Pembentukan Model Algoritma

Pembangunan model klasifikasi dilakukan menggunakan algoritma *Decision Tree* pada Rapid Miner. Tahapan pemodelan dimulai dengan penetapan atribut target melalui operator *Set Role*, kemudian data dibagi menjadi data latih dan data uji menggunakan operator *Split Data*. Selanjutnya, operator *Decision Tree* digunakan untuk membentuk model berdasarkan atribut-atribut seleksi seperti usia, pendidikan, pengalaman kerja, psikotes, dan wawancara. Model yang terbentuk diuji menggunakan operator *Apply Model*, sedangkan evaluasi performa dilakukan melalui operator *Performance* untuk memperoleh nilai *accuracy*, *precision*, *recall*, dan *F1-score*. Alur lengkap proses pemodelan ini ditunjukkan pada Gambar 8 dan menjadi dasar dalam menilai efektivitas algoritma *Decision Tree* sebagai metode klasifikasi dalam sistem pendukung keputusan seleksi calon karyawan.

4.5. Hasil Pembentukan Model *Decision Tree*

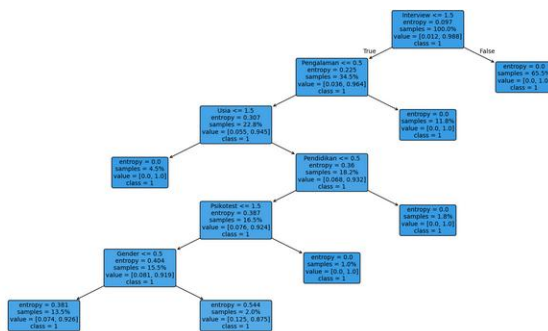


Gambar 3. Hasil Model *Decision Tree*

Evaluasi kinerja model *Decision Tree* dilakukan menggunakan Confusion Matrix untuk menilai kemampuan model dalam membedakan kelas “Diterima” dan “Ditolak.” Berdasarkan hasil pada Gambar, model menghasilkan 0 *True Negative*, 1 *False Positive*, 0 *False Negative*, dan 79 *True Positive*. Hasil ini menunjukkan bahwa model sangat tepat

dalam memprediksi pelamar yang termasuk pada kelas Diterima.

Dari *Confusion Matrix* tersebut diperoleh nilai *akurasi* sebesar 98,8%, *precision* 98,75%, *recall* 100%, dan *F1-score* 98,1%. Nilai evaluasi ini mengindikasikan bahwa model *Decision Tree* memiliki kemampuan klasifikasi yang sangat baik dan mampu memberikan prediksi yang akurat dalam proses penentuan kelulusan calon karyawan di PT G-Textile Indonesia.



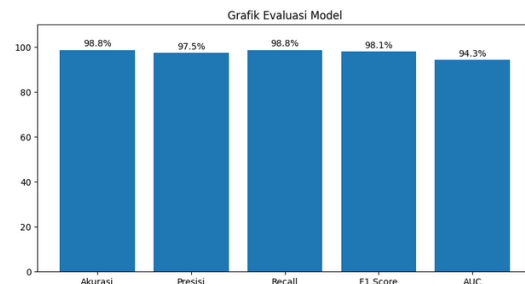
Gambar 4. Pohon Keputusan

Hasil pembentukan model *Decision Tree* menunjukkan bahwa atribut *Interview* menjadi node utama (*root node*) dalam proses klasifikasi kelulusan calon karyawan. Hal ini mengindikasikan bahwa variabel tersebut memiliki *information gain* terbesar dan merupakan faktor paling dominan dalam menentukan keputusan akhir. Pada cabang *Interview* ≤ 1.5 , mayoritas data berstatus Diterima dengan nilai *entropy* sangat rendah, sehingga menunjukkan tingkat homogenitas yang tinggi. Sementara itu, cabang *Interview* > 1.5 juga langsung mengarah ke kelas Diterima, menegaskan kuatnya pengaruh hasil wawancara terhadap prediksi model.

Setelah pemisahan berdasarkan *Interview*, atribut *Pengalaman Kerja*, *Usia*, *Pendidikan*, *Psikotest*, dan *Gender* menjadi pemecahan lanjutan. Atribut-atribut tersebut tetap menghasilkan cabang yang didominasi oleh kelas Diterima, meskipun beberapa node memiliki nilai *entropy* lebih tinggi akibat variasi data. Atribut seperti *Psikotest* dan *Gender* terlihat memiliki kontribusi yang relatif kecil karena nilai *entropy* yang rendah menunjukkan tidak banyak perbedaan dalam kelas output.

Pada struktur pohon memperlihatkan bahwa sebagian besar node berakhir pada kelas Diterima, yang mengindikasikan bahwa dataset pelatihan cenderung didominasi oleh pelamar yang lulus. Meskipun demikian, model berhasil menggambarkan pola keputusan yang terjadi pada proses seleksi di PT G-Textile Indonesia, di mana *Interview* muncul sebagai faktor paling berpengaruh, diikuti oleh pengalaman kerja, usia, dan pendidikan sebagai pendukung.

4.6. Evaluasi Model



Gambar 5. Grafik Evaluasi Model

Berdasarkan hasil evaluasi yang ditunjukkan pada Gambar, algoritma *Decision Tree* memberikan performa yang sangat baik dalam mengklasifikasikan kelulusan calon karyawan. Model ini berhasil mencapai *akurasi* sebesar 98,8%, yang menunjukkan bahwa hampir seluruh prediksi sesuai dengan kondisi sebenarnya. Nilai *precision* sebesar 97,5% mengindikasikan bahwa model mampu memberikan prediksi positif secara tepat pada mayoritas kasus. Sementara itu, *recall* sebesar 98,8% menunjukkan kemampuan model dalam mengenali hampir seluruh data pada kelas positif dengan sangat baik.

Keseimbangan antara *precision* dan *recall* tercermin melalui *F1-Score* sebesar 98,1%, menandakan bahwa model memiliki konsistensi performa yang stabil dalam menilai kandidat. Selain itu, nilai *AUC* sebesar 94,3% memperlihatkan bahwa model memiliki kemampuan pemisahan kelas yang kuat dan dapat diandalkan dalam membedakan antara pelamar yang diterima dan ditolak. Secara keseluruhan, hasil tersebut menunjukkan bahwa algoritma *Decision Tree* memiliki performa yang sangat baik dan layak digunakan sebagai pendukung keputusan dalam proses seleksi calon karyawan di PT G-Textile Indonesia.

5. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian dan implementasi yang telah dilakukan, sistem informasi rekrutmen berbasis web yang dikembangkan menggunakan *Django*, *SQLite*, dan template *AdminLTE* terbukti mampu mendukung proses seleksi karyawan di PT G-Textile Indonesia secara lebih terstruktur, efisien, dan objektif. Integrasi algoritma *Decision Tree* pada modul Monitoring Seleksi menghasilkan performa model yang sangat baik dengan *accuracy* 98,8%, *precision* 97,5%, *recall* 98,8%, *F1-Score* 98,1%, dan *AUC* 94,3%, sehingga mampu memberikan rekomendasi kelulusan kandidat secara konsisten dan akurat. Untuk pengembangan selanjutnya, penelitian ini menyarankan penambahan variabel penilaian yang lebih beragam, penggunaan dataset yang lebih besar, penerapan algoritma pembandingan, serta pemanfaatan *cross validation* guna meningkatkan stabilitas model. Selain itu, sistem dapat dikembangkan dengan fitur otomatisasi, dashboard analitik, dan peningkatan keamanan data, serta dilakukan uji coba langsung

dalam proses rekrutmen nyata agar sistem semakin optimal dalam mendukung pengambilan keputusan HRD.

DAFTAR PUSTAKA

- [1] D. Mahameru, H. Mubayinna, dan W. Lestari, "Implementasi Sistem E-Recruitment Karyawan Dengan Memanfaatkan Teknologi Informasi Menggunakan Metode Waterfall," *JEKIN - Jurnal Teknik Informatika*, vol. 4, no. 3, hal. 440–446, 2024, doi: 10.58794/jekin.v4i3.736.
- [2] P. H. Safarina Syah Alhaura, Hasbi Firmansyah, Wahyu Asriyani, "Klasifikasi Tingkat Pendidikan Karyawan Menggunakan Algoritma C4.5," *Journal of Multidisciplinary Inquiry in Science, Technology and Educational Research*, vol. 3, no. 1, hal. 943–954, 2026, doi: 10.32672/mister.v3i1.xxx.
- [3] F. umil B. Firda dan A. Gunawan, "Strategi Rekrutmen dan Seleksi yang Efektif untuk Meningkatkan Kualitas Tenaga Kerja," *GLOBAL: Jurnal Lentera BITEP*, vol. 1, no. 02, hal. 60–65, 2023, doi: 10.59422/global.v1i02.145.
- [4] D. A. Sahputra dan M. R. R. Saelan, "Penentuan Faktor Kelayakan Penerimaan Karyawan Menggunakan Algoritma Decision Tree pada Perusahaan PT . Personel Alih Daya," *Jurnal Sains dan Informatika*, vol. 6, no. November, hal. 148–156, 2020, doi: 10.34128/jsi.v6i2.225.
- [5] D. Mayangsari dan T. Wahyudi, "Klasifikasi Seleksi Penerimaan Karyawan Baru Menggunakan Algoritma C4.5," *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 5, no. 1, hal. 26–34, 2024, doi: 10.57152/malcom.v5i1.1525.
- [6] E. E. Mhd Isrok Aryanto, "DECISION TREE TECHNIQUE DALAM MENENTUKAN PENERIMAAN KARYAWAN SUPER MARKET DI LOTTE GROSIR BATAM," *Jurnal Comasie*, vol. 01, no. 1, hal. 96–103, 2022.
- [7] H. Kadar dan A. Budiyantra, "Penentuan Kelayakan Karyawan Baru Menggunakan Data Mining Dengan Algoritma Decision Tree (C4 . 5)," *Jurnal Riset Sistem Informasi dan Teknik Informasi*, vol. 2, no. 6, hal. 29–41, 2024, doi: 10.61132/mercurius.v2i6.389.
- [8] D. S. Untung Suropati, "IMPLEMENTASI ALGORITMA DECISION TREE C4.5 UNTUK KLASIFIKASI PENGANGKATAN KARYAWAN TETAP (STUDI KASUS) DI PT INTINUSA TEKNIK SEJAHTERA)," *Journal of Business Management*, vol. 7, hal. 427–433, 2025, doi: 10.31539/68x3rd85.
- [9] A. M. Husein dan M. Brutu, "Prediksi Penerimaan Calon Karyawan Dengan Menggunakan Algoritma C4 . 5 Pada Biro Kesejahteraan Rakyat Provinsi Sumatera Utara," *Digital Transformation Technology (Digitech)*, vol. 2, no. 1, hal. 16–20, 2022, doi: 10.47709/digitech.v2i1.1769.
- [10] A. Ardhy, T. Maulana, dan D. Arman, "Penerapan Metode Decision Tree C4 . 5 untuk Klasifikasi Data Kandidat Tenaga Kerja pada Perusahaan Outsourcing," *Jurnal Aplikasi Sains, Informasi, Elektronika dan Komputer*, vol. 6, no. 1, hal. 41–48, 2024, doi: 10.26905/jasiek.v6i1.12670 Jasiek.unmer.ac.id.
- [11] E. Devia, "Penerapan Decision Tree Dengan Algoritma C4 . 5 Untuk Menentukan Rekomendasi Kenaikan Jabatan Karyawan," *Journal of Information Systems and Informatics*, vol. 3, no. 1, hal. 28–37, 2023, doi: 10.61488/jis.v3i1.257.
- [12] C. Science, U. Nugraha, D. D. Rochman, A. Anwar, U. Widyatama, dan U. Widyatama, "Integrating Data Mining and Decision Support Systems for Optimizing New Employee Selection," *International Journal of Engineering and Computer Science*, vol. 13, no. 07, hal. 27579–27585, 2025, doi: 10.18535/ijecs/v14i07.5183.
- [13] A. Ajiz dan E. Wahyudin, "Penerapan Algoritma Decision Tree dalam Penentuan Karyawan Kontrak," *Jurnal MEANS - Media Informasi Analisa dan Sistem*, vol. 7, no. 1, hal. 124–129, 2022, doi: 10.54367/means.v7i1.1844.
- [14] M. A. K. Nur Roiz Azzam Naufal, "SISTEM PENDUKUNG KEPUTUSAN PENILAIAN KINERJA KARYAWAN MENGGUNAKAN ALGORITMA DECISION TREE," *SPIRIT - Journal Of Computing and Cybernetic System*, vol. 16, no. 2, hal. 328–333, 2024, doi: 10.53567/spirit.v16i2.346.
- [15] I. Jayanto, "Analisis Perbandingan Algoritma Decision Tree untuk Prediksi Karyawan dengan Potensi Atrisi di PT . XYZ," *FAHMA - Jurnal Informatika Komputer, Bisnis dan Manajemen*, vol. 22, no. 1, hal. 49–59, 2024, doi: 10.61805/fahma.v22i1.112.
- [16] J. M. Polgan et al., "Analisis Penerimaan Karyawan Konveksi Posisi Tukang Jahit Menggunakan Algoritma Decision Tree Pada PT . Nilosa Rama Buana," *Jurnal Minfo Polgan*, vol. 14, no. 1, hal. 1406–1411, 2025, doi: 10.33395/jmp.v14i1.15060.
- [17] I. Machine, L. Dengan, D. Tree, dan D. Jakarta, "Jurnal Sosial dan Teknologi (SOSTECH) Implementasi Machine Learning dengan Decision Tree e-ISSN 2774-5155 Algoritma C4 . 5 dalam Penerimaan Karyawan Baru pada p-ISSN 2774-5147 PT . Gitareksa Dinamika Jakarta," *Jurnal Sosial dan Teknologi*, vol. 1, no. 7, hal. 615–623, 2021, doi: 10.59188/jurnalsostech.v1i7.126.