

ANALISIS PENGARUH JENIS PEKERJAAN DAN SALDO TERHADAP RESPON PADA KAMPANYE BANK MENGUNAKAN ALGORITMA *DECISION TREE*

Aldo Bonifasius Simbolon^{1*}, Dean Siregar², Sybil Auzi³,
Febe Gracia Sembiring⁴, Arnita⁵, Fanny Ramadhani⁶

Universitas Negeri Medan, Medan, Sumatera Utara, Indonesia¹

Email*: aldosimbolon017@gmail.com

Universitas Negeri Medan, Medan, Sumatera Utara, Indonesia²

Email: deansiregar1609@gmail.com

Universitas Negeri Medan, Medan, Sumatera Utara, Indonesia³

Email: saibil1492@gmail.com

Universitas Negeri Medan, Medan, Sumatera Utara, Indonesia⁴

Email: febegracia45@gmail.com

Universitas Negeri Medan, Medan, Sumatera Utara, Indonesia⁵

Email: arnita@unimed.ac.id

Universitas Negeri Medan, Medan, Sumatera Utara, Indonesia⁶

Email: fannyr@unimed.ac.id

(*) *Corresponding Author*

ABSTRAK

Dalam era digitalisasi, data menjadi aset penting dalam pengambilan keputusan strategis di sektor perbankan. Institusi keuangan menghadapi tantangan dalam memahami preferensi serta perilaku nasabah secara personal dalam menjalankan kampanye promosi produk. Penelitian ini bertujuan untuk menganalisis dan mengungkap pola pengaruh jenis pekerjaan dan saldo rekening terhadap respon nasabah pada kampanye bank menggunakan algoritma *Decision Tree*. Pendekatan analitis berbasis data dibutuhkan untuk menggali pola tersembunyi dari informasi nasabah yang telah dikumpulkan. Algoritma *Decision Tree* dipilih karena kemampuannya memberikan prediksi dan menyajikan logika pengambilan keputusan dalam bentuk pohon yang mudah dipahami, serta fleksibilitasnya dalam menangani data kategorikal dan numerik. Dataset disederhanakan berfokus pada fitur 'job' dan 'balance', yang kemudian diolah menggunakan teknik *label encoding* untuk fitur kategorikal. Model dilatih menggunakan metode *fit* dan dievaluasi menggunakan metrik akurasi, *precision*, *recall*, dan *F1-score*. Hasil evaluasi menunjukkan akurasi model mencapai 92%. Untuk kelas respon 'ya', *recall* mencapai 97% dan *precision* 88%, mengindikasikan kemampuan model yang baik dalam mengidentifikasi nasabah berpotensi positif. Interpretasi hasil dari struktur *Decision Tree* secara eksplisit menunjukkan bahwa jenis pekerjaan dan saldo nasabah memiliki pola kecenderungan yang berbeda terhadap kampanye pemasaran, di mana nasabah dengan saldo tinggi cenderung merespon positif. Penelitian ini diharapkan dapat memberikan kontribusi dalam pengembangan sistem pendukung keputusan berbasis data yang lebih akurat, efektif, dan efisien untuk aktivitas pemasaran di sektor perbankan, berdasarkan pola pengaruh yang terungkap.

Kata kunci: Algoritma, Bank, *Big Data*, *Decision Tree*, Respon Kampanye

ABSTRACT

In the era of digitalization, data has become a crucial asset in strategic decision-making within the banking sector. Financial institutions face challenges in understanding customer preferences and behavior more personally when conducting product promotion campaigns. This research aims to analyze and uncover the influence patterns of job type and account balance on customer responses to bank campaigns using the Decision Tree algorithm. A data-driven analytical approach is needed to uncover hidden patterns from collected customer information. The Decision Tree algorithm was chosen for its ability to provide predictions and present decision-making logic in an easily understandable tree structure, as well as its flexibility in handling both categorical and numerical data. The simplified dataset focuses on the 'job' and 'balance' features, which are then processed using label encoding for categorical features. The model was trained using the fit method and evaluated using Accuracy, precision, recall, and F1-score metrics. The evaluation results show that the model's overall Accuracy reached 92%. For the 'yes' response class, recall reached 97% and precision 88%, indicating the model's good ability to identify potentially positive customers. Interpretation of the results from the Decision Tree structure explicitly indicates that job type and customer balance have different patterns of inclination towards marketing campaigns, where customers with high balances tend to respond positively. This research is expected to contribute to the development of a more accurate, effective, and efficient data-driven decision support system for marketing activities in the banking sector, based on the revealed influence patterns.

Keywords: Algorithm, Bank, Big Data, Decision Tree, Campaign Response

1. PENDAHULUAN

Dalam era digitalisasi yang semakin berkembang, data menjadi aset penting dalam pengambilan keputusan strategis, termasuk di sektor perbankan. Institusi keuangan saat ini menghadapi tantangan dalam memahami preferensi dan perilaku nasabah secara lebih personal, terutama dalam menjalankan kampanye promosi produk. Oleh karena itu, dibutuhkan pendekatan analitis berbasis data yang mampu menggali pola tersembunyi dari informasi nasabah yang telah dikumpulkan selama ini [1]. Data mining merupakan suatu proses yang digunakan untuk mengekstraksi informasi bermakna dari sekumpulan data yang besar dan kompleks dengan memanfaatkan berbagai teknik, seperti statistik, pembelajaran mesin, dan teknologi basis data [2]. Dalam dunia perbankan, penerapan data mining telah terbukti sangat luas dan signifikan untuk mendeteksi penipuan, memprediksi kemungkinan kegagalan pembayaran pinjaman, melakukan segmentasi nasabah, hingga memberikan rekomendasi produk keuangan yang sesuai dengan kebutuhan individu [3].

Salah satu metode yang umum digunakan dalam pemodelan analisis tersebut adalah teknik data mining, khususnya algoritma klasifikasi seperti Decision Tree [4]. Decision Tree merupakan salah satu metode klasifikasi yang tidak hanya mampu memberikan prediksi, namun juga menyajikan logika pengambilan keputusan dalam bentuk pohon yang mudah dipahami. Metode ini pertama kali diperkenalkan oleh Quinlan melalui algoritma ID3, yang kemudian dikembangkan lebih lanjut menjadi algoritma C4.5 dengan peningkatan efisiensi dan akurasi [5]. Keunggulan utama dari algoritma Decision Tree terletak pada fleksibilitasnya dalam menangani baik data bertipe kategorikal maupun numerik, serta kemampuannya menghasilkan model yang mudah dipahami dan diinterpretasikan. Penelitian yang dilakukan oleh Ratniasih (2019) juga menunjukkan bahwa data historis dalam kapasitas besar lebih mudah dipahami jika dikonversi menjadi

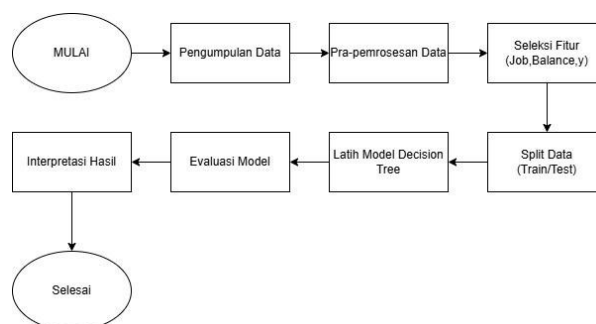
bentuk pohon keputusan, karena menyajikan informasi secara lebih terstruktur dan cepat diserap dibandingkan bentuk tabulasi yang rumit [6]. Dalam bidang keuangan, Decision Tree telah diaplikasikan secara luas untuk mendeteksi aktivitas penipuan, mengklasifikasikan tingkat risiko pinjaman, serta melakukan segmentasi pelanggan [7]. Studi yang dilakukan oleh Prameswari et al. (2024) menunjukkan bahwa penggunaan metode stacking ensemble yang merupakan pendekatan gabungan dari beberapa algoritma machine learning, termasuk Decision Tree, mampu meningkatkan akurasi dalam memprediksi status pinjaman bank hingga mencapai 90,7% [8].

Jenis pekerjaan dan saldo rekening merupakan dua variabel yang memiliki peranan krusial dalam proses pengambilan keputusan keuangan oleh nasabah. Jenis pekerjaan tidak hanya mencerminkan sumber pendapatan utama seseorang, tetapi juga memberikan gambaran mengenai tingkat kestabilan ekonomi dan kemampuan finansial jangka panjang [9]. Penelitian yang dilakukan oleh Hendarsyah (2020) menunjukkan bahwa jenis pekerjaan termasuk dalam kategori faktor pribadi yang dominan dalam mempengaruhi perilaku konsumen, khususnya dalam penggunaan kartu kredit [10]. Sementara itu, saldo rekening nasabah menjadi indikator penting yang mencerminkan kapasitas daya beli serta potensi finansial seseorang. Perilaku konsumen adalah proses pengambilan keputusan dalam membeli dan menggunakan produk, yang dipengaruhi oleh faktor pribadi, sosial, budaya, dan psikologis [11]. Rohmah et al. (2024) menyebutkan bahwa masyarakat Indonesia cenderung konsumtif dan mudah terpengaruh oleh promosi, khususnya yang berkaitan dengan tren dan gaya hidup [12]. Beberapa faktor pribadi seperti pendapatan, pekerjaan, dan usia juga telah terbukti memengaruhi keputusan konsumen terhadap layanan perbankan [13].

Dengan mengolah variabel seperti jenis pekerjaan dan saldo rekening, Decision Tree dapat mengidentifikasi segmen nasabah yang lebih berpotensi merespons promosi secara positif [14]. Penelitian ini bertujuan untuk menganalisis pengaruh jenis pekerjaan dan saldo rekening terhadap respon nasabah pada kampanye bank menggunakan algoritma Decision Tree. Dengan pendekatan ini, hasil analisis diharapkan dapat memberikan kontribusi dalam pengembangan sistem pendukung keputusan berbasis data yang lebih akurat, efektif, dan efisien untuk aktivitas pemasaran di sektor perbankan [8].

2. METODE PENELITIAN

Proses pelatihan dan pengujian dilakukan menggunakan algoritma *Decision Tree* yang diimplementasikan dengan bantuan pustaka *scikit-learn* pada Python. Dataset disederhanakan untuk fokus pada dua fitur utama, yaitu *job* dan *balance*, yang kemudian diolah menggunakan teknik *label encoding* untuk fitur kategorikal. Model dilatih menggunakan metode *fit*, kemudian diuji dengan data uji untuk mengukur performa. Evaluasi dilakukan dengan menggunakan metrik akurasi, *precision*, *recall*, dan *F1-score*.



Gambar 1. Alur metode penelitian

Penelitian ini dimulai dengan merumuskan masalah mengenai pengaruh jenis pekerjaan dan saldo nasabah terhadap keputusan mereka dalam menanggapi kampanye pemasaran bank. Untuk mengatasi masalah ini, digunakan model klasifikasi *Decision Tree* karena kemampuannya menangani fitur kategorikal dan numerik serta menyajikan hasil yang mudah dipahami. Data kampanye pemasaran bank dikumpulkan dari platform Kaggle, yang merupakan repositori data publik yang banyak digunakan untuk proyek-proyek *data science* dan *machine learning*. Dataset yang digunakan difokuskan pada atribut '*job*' (jenis pekerjaan) dan '*balance*' (saldo rekening) sebagai variabel prediktor, serta '*y*' (respon kampanye) sebagai variabel target.

Tabel 1. Contoh dataset yang digunakan

job	marital	education	balance	housing	loan	campaign	y
unemployed	married	primary	1787	no	no	1	no
services	married	secondary	4789	yes	yes	1	no
management	single	tertiary	1350	yes	no	1	no
management	married	tertiary	1476	yes	yes	4	no
blue-collar	married	secondary	0	yes	no	1	no
management	single	tertiary	747	no	no	2	no
self-employed	married	tertiary	307	yes	no	1	yes
technician	married	secondary	147	yes	no	2	yes
entrepreneur	married	tertiary	221	yes	no	2	yes
services	married	primary	-88	yes	yes	1	yes

Tahap pra-pemrosesan data merupakan langkah krusial untuk mempersiapkan dataset agar siap digunakan dalam pemodelan. Proses diawali dengan pembersihan data seperti penghapusan duplikat dan pengisian nilai kosong untuk memastikan kualitas dan integritas data.

Tabel 2. Dataset setelah disederhanakan

job	balance	y
unemployed	1787	no
services	4789	no
management	1350	no
management	1476	no
blue-collar	0	no
management	747	no
self-employed	307	yes
technician	147	yes
entrepreneur	221	yes
services	-88	yes

Selanjutnya, untuk mengolah fitur kategorikal, teknik *Label Encoding* diterapkan. Atribut '*job*', yang merupakan data kategorikal, diubah menjadi representasi numerik. Proses ini melibatkan pemetaan setiap kategori unik dalam fitur '*job*' ke nilai integer yang berbeda. Demikian pula, variabel target '*y*', yang awalnya memiliki nilai '*yes*' atau '*no*', dikonversi menjadi format biner, di mana '*yes*' dipetakan menjadi 1 dan '*no*' menjadi 0. Proses *encoding* ini memungkinkan algoritma *Decision Tree* yang bersifat matematis

untuk memproses data kategorikal secara efektif. Setelah proses *encoding*, hanya fitur '*job*' (setelah label *encoding*) dan '*balance*' yang dipilih sebagai input, dengan '*y*' sebagai target.

Tabel 3. Dataset setelah *one hot label encoding*

job	balance	y
10	1787	0
7	4789	0
4	1350	0
4	1476	0
1	0	0
4	747	0
6	307	1
9	147	1
2	221	1
7	-88	1

Dataset kemudian dibagi menjadi 80% data pelatihan dan 20% data pengujian. Pemilihan rasio pembagian ini didasarkan pada hasil uji coba yang menunjukkan bahwa konfigurasi 80% data pelatihan menghasilkan akurasi model tertinggi, yaitu 92%. Sebagai perbandingan, pembagian data 70% dan 75% untuk pelatihan hanya mencapai akurasi sebesar 90%. Oleh karena itu, rasio 80/20 dipilih untuk memaksimalkan performa model dalam memprediksi respon nasabah, sekaligus memastikan kemampuan generalisasi model yang optimal.

Model *Decision Tree* dilatih, membangun pohon keputusan berdasarkan fitur terbaik yang memisahkan data menggunakan nilai *Entropy* dan *Information Gain*. Evaluasi model dilakukan menggunakan metrik seperti *Akurasi*, *Precision*, *recall*, dan *F1-Score*, dengan akurasi model mencapai 92%. Hasil interpretasi menunjukkan bahwa jenis pekerjaan dan saldo nasabah memiliki pola kecenderungan yang berbeda terhadap kampanye pemasaran, di mana nasabah dengan saldo tinggi cenderung merespon positif. Proyek diakhiri dengan kesimpulan bahwa jenis pekerjaan dan saldo nasabah berpengaruh signifikan terhadap respon kampanye bank, dan model *Decision Tree* memberikan hasil yang baik serta mudah diinterpretasikan, yang dapat dimanfaatkan bank untuk strategi pemasaran yang lebih efektif.

Proses Kerja *Decision Tree*

Proses pembangunan *Decision Tree* untuk menganalisis pengaruh jenis pekerjaan (*job*) dan saldo nasabah (*balance*) terhadap respon kampanye bank (*y*) diawali dengan pemilihan fitur yang paling relevan. Algoritma memilih fitur yang memberikan informasi paling tinggi dalam membedakan antara nasabah yang merespon ("yes") dan tidak merespon ("no"). Pemilihan ini dihitung menggunakan metode seperti *Information Gain* dan *Entropy*.

Dalam penelitian ini, algoritma *Decision Tree* dibangun berdasarkan dua fitur utama, yaitu *job* (jenis pekerjaan) dan *balance* (saldo rekening), dengan target klasifikasi *y* yang merepresentasikan respon nasabah terhadap kampanye. Proses diawali dengan pemilihan fitur yang memiliki nilai *Information Gain* tertinggi, yaitu fitur yang paling

informatif dalam memisahkan data antara respon “yes” dan “no”. Sebagai contoh, jika hasil pelatihan model menunjukkan bahwa fitur balance memiliki pengaruh paling besar, fitur ini akan digunakan sebagai root node.

Sebagai contoh, model dapat memutuskan:

1. Jika $balance > 1000$ dan job adalah admin, maka probabilitas respon “yes” tinggi (nasabah cenderung merespon kampanye).
2. Jika $balance < 0$ dan job adalah blue-collar, maka kemungkinan besar respon adalah “no”.

Setiap node dan cabang dalam pohon mewakili keputusan berdasarkan ambang tertentu atau kategori spesifik. Fitur job yang merupakan data kategorikal sebelumnya telah diubah ke format numerik menggunakan label encoding. Setelah itu, Decision Tree terus melakukan pemisahan data secara rekursif sampai seluruh cabang memiliki data homogen (semua “yes” atau “no”) atau kedalaman pohon tercapai. Proses ini memungkinkan pembentukan pola-pola keputusan yang terstruktur dan mudah ditelusuri.

Sebagai langkah akhir, dilakukan pruning untuk memangkas cabang yang terlalu spesifik terhadap data latih dan tidak berdampak signifikan pada akurasi. Hal ini bertujuan untuk menghindari overfitting dan memastikan model tetap optimal saat diterapkan pada data baru. Dengan pendekatan ini, sistem klasifikasi tidak hanya menghasilkan prediksi, tetapi juga memvisualisasikan logika keputusan yang dapat digunakan sebagai dasar rekomendasi strategi pemasaran bank secara lebih terarah.

$$Entropy(S) = \sum_{j=1}^k -p_j \log_2 p_j$$

$$Gain(S, A) = Entropy(S) - \sum_{v \in \text{Value}(A)} \frac{|S_v|}{|S|} Entropy(S_v) \quad (2)$$

Pembuatan Node dan Cabang Keputusan

Setelah fitur terbaik ditentukan, pohon keputusan membentuk *node* berdasarkan nilai-nilai unik dari fitur tersebut. Dalam konteks judul, jika fitur *job* menjadi *root node*, maka akan dibuat cabang berdasarkan jenis pekerjaan seperti *student*, *management*, *services*, dan sebagainya. Untuk fitur numerik seperti *balance*, pemisahan dilakukan berdasarkan rentang nilai tertentu, misalnya $balance > 1000$ atau $balance < 0$. Setiap cabang tersebut kemudian mengarah ke *node* berikutnya atau ke hasil akhir prediksi jika data dalam cabang tersebut sudah homogen.

Proses Rekursif pada Subset Data

Proses pemilihan fitur dan pembentukan *node* dilakukan secara rekursif pada setiap subset data yang terbentuk dari percabangan sebelumnya. Dalam konteks kampanye bank, hal ini berarti sistem akan terus mencari pola-pola baru dalam jenis pekerjaan dan saldo nasabah untuk membedakan respon mereka. Misalnya, setelah membagi berdasarkan pekerjaan, model dapat melanjutkan dengan memeriksa nilai saldo (*balance*) untuk masing-masing jenis pekerjaan guna memperhalus prediksi respon kampanye.

Kondisi Berhenti

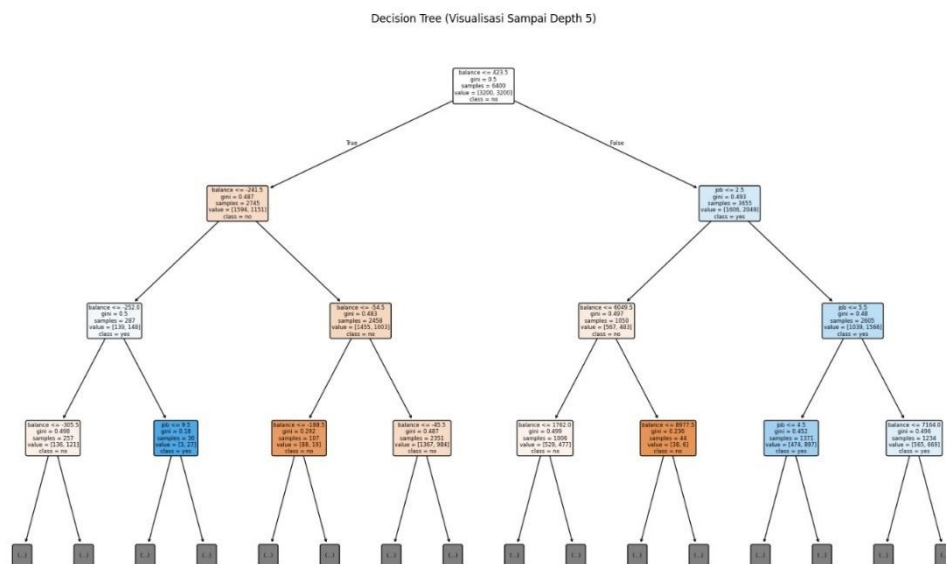
Pohon keputusan akan berhenti membagi data lebih lanjut jika semua data dalam suatu cabang memiliki kelas target yang sama (misalnya semuanya menjawab "no"), jika tidak ada fitur yang tersisa untuk dibagi, atau jika kedalaman pohon telah mencapai batas yang ditentukan. Dalam studi pengaruh pekerjaan dan saldo terhadap respon kampanye, kondisi ini mencerminkan bahwa semua nasabah dengan kriteria tertentu menunjukkan respon yang seragam, sehingga tidak perlu dilakukan pemisahan lebih lanjut.

Pruning untuk Mencegah *Overfitting*

Setelah pohon selesai dibentuk, dilakukan proses pruning atau pemangkasan cabang-cabang yang tidak memberikan kontribusi signifikan terhadap akurasi model. Dalam konteks kampanye bank, pruning membantu mencegah model menjadi terlalu kompleks dan hanya cocok untuk data pelatihan, tapi buruk dalam memprediksi data baru. Dengan memangkas *node* yang terlalu spesifik atau tidak umum, model menjadi lebih sederhana dan general, sehingga dapat lebih akurat dalam memprediksi apakah nasabah dari jenis pekerjaan tertentu dan dengan saldo tertentu akan merespon terhadap kampanye atau tidak.

3. HASIL DAN PEMBAHASAN

Untuk memahami proses pengambilan keputusan oleh model klasifikasi, dilakukan visualisasi pohon keputusan (*Decision Tree*) hingga kedalaman tertentu. Visualisasi ini memberikan gambaran mengenai bagaimana fitur-fitur seperti *balance* dan *job* memengaruhi keputusan akhir model dalam mengklasifikasikan data ke dalam kelas 'yes' atau 'no'. Gambar berikut menunjukkan struktur pohon keputusan hingga kedalaman 3 (depth 3), yang mempermudah interpretasi aturan-aturan klasifikasi yang terbentuk dari data latih.



Gambar 2. Visualisasi *Decision Tree* hingga kedalaman 5

1. Warna Biru (Biru Muda sampai Biru Tua)
Menunjukkan bahwa mayoritas class di *node* tersebut adalah "yes". Semakin gelap warna birunya, maka semakin dominan jumlah data berkelas "yes" dalam *node* tersebut (probabilitas "yes" tinggi).
2. Warna Oranye (Oranye Muda sampai Oranye Tua)

Menandakan bahwa mayoritas class di *node* tersebut adalah "no". Semakin gelap warna oranye, maka semakin dominan jumlah data berkelas "no".

3. Warna Putih atau Netral

Kadang digunakan di *node* awal (root) atau *node* yang memiliki pembagian cukup seimbang antara kelas "yes" dan "no" (nilai gini tinggi, dekat 0.5). Warna lebih netral atau terang menandakan kelas mayoritas tidak terlalu dominan.

Analisis mendalam terhadap struktur *decision tree* pada gambar 2 mengungkapkan bagaimana fitur 'balance' dan 'job' secara berjenjang memengaruhi respon nasabah terhadap kampanye. Pada *root node*, pemisahan utama terjadi berdasarkan $\text{balance} \leq 425.5$. Hal ini secara jelas mengindikasikan bahwa saldo rekening merupakan faktor penentu awal yang paling signifikan dalam memprediksi respon nasabah. Cabang 'True' (yaitu, saldo kurang dari atau sama dengan 425.5) mayoritas mengarah ke kelas 'no' (ditunjukkan dengan warna oranye gelap pada node berikutnya seperti $\text{balance} \leq 143.5$ atau $\text{balance} \leq -43.5$), menunjukkan bahwa nasabah dengan saldo rendah memiliki kecenderungan kuat untuk tidak merespon kampanye.

Sebaliknya, cabang 'False' dari *root node* (yaitu, saldo lebih dari 425.5) sebagian besar mengarah pada respon 'yes' (ditunjukkan dengan warna biru pada node-node selanjutnya), menegaskan bahwa nasabah dengan saldo tinggi cenderung merespon positif terhadap kampanye pemasaran. Pada segmen nasabah dengan saldo tinggi ($\text{balance} > 425.5$), pengaruh jenis pekerjaan mulai terlihat. Contohnya, pada salah satu cabang kanan, terdapat pemisahan berdasarkan $\text{job} \leq 4.5$. Pemisahan ini menunjukkan bahwa di antara nasabah dengan saldo di atas 425.5, jenis pekerjaan (setelah dilakukan *label encoding*) menjadi faktor diferensiasi selanjutnya dalam menentukan respon. Pola ini mengkonfirmasi bahwa Decision Tree tidak hanya mengklasifikasikan, tetapi juga secara eksplisit menguraikan aturan-aturan keputusan yang mencerminkan bagaimana kombinasi dari saldo dan jenis pekerjaan secara langsung memengaruhi perilaku respon nasabah.

Evaluasi Kinerja Model *Decision Tree*

Tabel dibawah menyajikan ringkasan komprehensif dari kinerja model klasifikasi *Decision Tree*. Tabel ini memberikan performa model tidak hanya secara keseluruhan tetapi juga untuk setiap kelas target secara spesifik, menggunakan beberapa metrik standar dalam evaluasi klasifikasi. Pemahaman terhadap setiap komponen dalam tabel ini penting untuk menginterpretasikan kekuatan dan kelemahan model secara objektif sebelum melangkah ke pembahasan implikasi kinerjanya.

Tabel 4. Laporan klasifikasi model *Decision Tree*

	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>	<i>Support</i>
Kelas 0 (Tidak)	0.97	0.87	0.92	800
Kelas 1 (Ya)	0.88	0.97	0.92	800
<i>Accuracy</i>			0.92	1600
<i>Macro avg</i>	0.92	0.92	0.92	1600
<i>Weighted avg</i>	0.92	0.92	0.92	1600

Kelas 0 (Tidak) dan Kelas 1 (Ya), menampilkan metrik evaluasi individual untuk masing-masing kelas prediksi:

Precision (Presisi), Mengukur proporsi prediksi positif yang benar dari total prediksi positif yang dibuat oleh model untuk kelas tersebut. Sebagai contoh, untuk 'Kelas 1 (Ya)', nilai 0.88 berarti 88% dari nasabah yang diprediksi oleh model akan merespon 'ya' memang benar-benar merespon 'ya'.

Recall (Sensitivitas atau True Positive Rate), Mengukur proporsi kasus positif aktual yang berhasil diidentifikasi dengan benar oleh model. Misalnya, untuk 'Kelas 1 (Ya)', nilai 0.97 menunjukkan bahwa model berhasil mengidentifikasi 97% dari semua nasabah yang seharusnya merespon 'ya' dalam data uji.

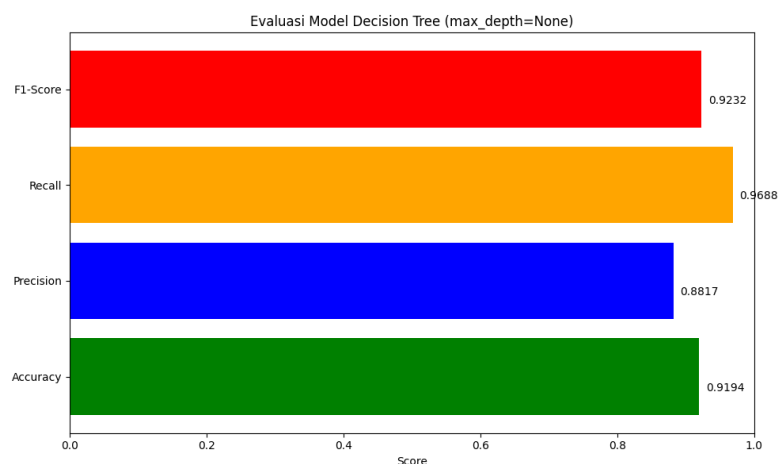
f1-Score, Metrik ini memberikan skor tunggal yang menyeimbangkan kedua aspek tersebut. Untuk 'Kelas 1 (Ya)', nilai 0.92 menunjukkan keseimbangan yang baik antara presisi dan *recall* untuk kelas tersebut.

Support, Menunjukkan jumlah aktual sampel atau observasi untuk masing-masing kelas dalam data uji. Dalam kasus ini, terdapat 800 sampel untuk 'Kelas 0 (Tidak)' dan 800 sampel untuk 'Kelas 1 (Ya)'.

Accuracy, menunjukkan proporsi total prediksi yang benar (baik True Positives maupun True Negatives) dari keseluruhan sampel data uji. Dalam model ini, akurasi keseluruhan adalah 0.92 atau 92%.

Macro avg, nilai *Macro avg* untuk presisi, *recall*, dan *F1-score* masing-masing adalah 0.92. Angka ini memberikan gambaran kinerja model secara keseluruhan dengan asumsi bahwa setiap kelas (Kelas 0 dan Kelas 1) memiliki tingkat kepentingan yang setara.

Weighted avg, nilai *Weighted avg* untuk presisi, *recall*, dan *F1-score* juga masing-masing adalah 0.92. Karena data uji Anda dalam kasus ini seimbang (dengan 800 sampel untuk Kelas 0 dan 800 sampel untuk Kelas 1), maka nilai *Macro avg* dan *Weighted avg* akan identik atau sangat mirip. Kedua rata-rata ini, yang menunjukkan angka 0.92, mengkonfirmasi kinerja model yang tinggi dan konsisten di kedua kelas prediksi.



Gambar 3. Grafik plot metrik evaluasi model *Decision Tree*

Analisis Kinerja Model

Berdasarkan data pada Tabel Laporan klasifikasi model dan Grafik plot metrik evaluasi, model *Decision Tree* utama (*max_depth=None*) menunjukkan kinerja yang sangat baik. Akurasi keseluruhan model mencapai 92.00%, mengindikasikan bahwa model ini mampu memprediksi dengan tepat respon mayoritas nasabah hanya dengan menggunakan fitur '*job*' dan '*balance*'. Lebih lanjut, untuk Kelas 1 (respon '*ya*'), yang sering menjadi fokus dalam kampanye pemasaran, model menghasilkan *recall* sebesar 97%. Nilai *recall* yang sangat tinggi ini krusial, karena menunjukkan kemampuan model untuk mengidentifikasi hampir semua nasabah yang sebenarnya berpotensi merespon positif, sehingga meminimalkan peluang kehilangan prospek. Presisi untuk kelas '*ya*' adalah 88%, yang berarti mayoritas dari mereka yang diprediksi akan merespon '*ya*' memang benar demikian. Keseimbangan antara kedua metrik ini tercermin dalam F1-Score sebesar 92% untuk kelas '*ya*'. Untuk Kelas 0 (respon '*tidak*'), presisi model juga sangat tinggi (97%), dengan *recall* 87%. Metrik-metrik ini secara kolektif menegaskan bahwa model *Decision Tree* dengan kedalaman penuh mampu mempelajari pola signifikan dari fitur '*job*' dan '*balance*' untuk melakukan klasifikasi respon nasabah dengan efektif dan seimbang. Kinerja tinggi inilah yang menjadi dasar untuk pembahasan lebih lanjut mengenai bagaimana model ini memanfaatkan fitur yang ada.

Kinerja model yang sangat baik ini, dengan akurasi keseluruhan mencapai 92.00% dan *recall* 97% untuk kelas '*ya*', bukan hanya menunjukkan kemampuan prediksi yang kuat, tetapi juga mengindikasikan bahwa algoritma *Decision Tree* telah berhasil mengidentifikasi pola-pola pengaruh yang signifikan dari fitur '*job*' dan '*balance*' terhadap respon kampanye nasabah. Kemampuan model untuk secara efektif membedakan antara nasabah yang berpotensi merespon positif dan negatif didasarkan pada pemahaman akan bagaimana kombinasi dan ambang batas nilai dari kedua fitur tersebut *memengaruhi* keputusan respon mereka. Oleh karena itu, metrik kinerja yang tinggi ini menjadi validasi atas kemampuan *Decision Tree* dalam memberikan analisis yang valid mengenai hubungan antara jenis pekerjaan, saldo, dan respon kampanye.

4. KESIMPULAN DAN SARAN

Kesimpulan

Penelitian ini menyimpulkan bahwa model klasifikasi *Decision Tree* sangat efektif dalam menganalisis pengaruh jenis pekerjaan dan saldo rekening nasabah terhadap respon mereka pada kampanye bank. Visualisasi pohon keputusan dengan jelas menunjukkan bagaimana fitur '*balance*' dan '*job*' secara spesifik memengaruhi klasifikasi respon nasabah, dengan warna biru mengindikasikan dominasi respon "*yes*" dan warna oranye untuk respon "*no*".

Kinerja model ini sangat baik, mencapai akurasi keseluruhan 92%. Khususnya untuk Kelas 1 (respon '*ya*'), model menunjukkan *recall* yang tinggi sebesar 97% dan *precision* 88%, menandakan kemampuannya yang kuat dalam mengidentifikasi prospek positif dan meminimalkan kehilangan potensi nasabah. Sementara itu, untuk Kelas 0 (respon '*tidak*'), model juga memberikan *precision* yang sangat tinggi (97%) dan *recall* 87%. Secara keseluruhan, metrik evaluasi seperti Accuracy, precision, recall, dan F1-Score, dengan nilai rata-rata 0.92 untuk Macro avg dan Weighted avg, mengkonfirmasi kinerja model yang tinggi dan konsisten untuk kedua kelas prediksi.

Hasil interpretasi lebih lanjut dari struktur *Decision Tree* secara eksplisit menunjukkan bahwa jenis pekerjaan dan saldo nasabah memiliki pola kecenderungan yang berbeda terhadap kampanye pemasaran. Secara spesifik, nasabah dengan saldo

tinggi cenderung memberikan respon positif, sebuah pola pengaruh yang teridentifikasi melalui ambang batas dan cabang keputusan dalam model. Oleh karena itu, model Decision Tree ini terbukti menjadi alat yang berharga bagi bank untuk mengidentifikasi segmen nasabah yang paling berpotensi dan merancang strategi pemasaran yang lebih akurat, efektif, dan efisien berdasarkan pola pengaruh yang terungkap.

Saran

Berdasarkan hasil penelitian, saran untuk pengembangan penelitian selanjutnya adalah sebagai berikut:

1. Peningkatan Kedalaman Analisis Fitur: Meskipun penelitian ini berfokus pada jenis pekerjaan dan saldo rekening, penelitian selanjutnya dapat mengeksplorasi pengaruh variabel lain yang mungkin relevan, seperti usia, status pernikahan, riwayat kredit, atau produk perbankan yang sudah dimiliki, untuk mendapatkan pemahaman yang lebih komprehensif tentang perilaku nasabah.
2. Eksplorasi Algoritma Klasifikasi Lain: Selain *Decision Tree*, dapat dilakukan perbandingan dengan algoritma *machine learning* lain seperti Random Forest, K-Nearest Neighbor (KNN), *Support Vector Machine* (SVM), atau *Logistic Regression* untuk melihat potensi peningkatan akurasi dan performa model.
3. Implementasi Teknik *Ensemble Learning*: Mengingat metode *stacking ensemble* telah terbukti meningkatkan akurasi dalam prediksi status pinjaman bank hingga 90.7%, penelitian selanjutnya dapat menerapkan teknik *ensemble learning* (misalnya *bagging* atau *boosting*) pada algoritma *Decision Tree* untuk lebih meningkatkan stabilitas dan akurasi model dalam memprediksi respon kampanye.
4. Validasi Model dengan Data yang Lebih Besar dan Beragam: Untuk meningkatkan generalisasi model, disarankan untuk menguji dan memvalidasi model dengan dataset yang lebih besar dan lebih beragam dari berbagai bank atau wilayah, sehingga hasil penelitian dapat lebih relevan dan dapat diterapkan secara luas.
5. Analisis Faktor Non-Demografis: Selain faktor demografis dan finansial, faktor psikologis atau gaya hidup nasabah juga dapat dipertimbangkan dalam analisis, mengingat perilaku konsumen dipengaruhi oleh berbagai faktor termasuk sosial, budaya, dan psikologis.
6. Pengembangan Sistem Pendukung Keputusan: Hasil analisis dari model *Decision Tree* yang mudah diinterpretasikan dapat menjadi dasar untuk mengembangkan sistem pendukung keputusan *data-driven* yang lebih akurat, efektif, dan efisien untuk aktivitas pemasaran di sektor perbankan.

5. DAFTAR PUSTAKA

- [1] S. D. Kurniawan *et al.*, *Big Data: Mengenal Big Data & Implementasinya di Berbagai Bidang*. PT. Sonpedia Publishing Indonesia, 2024. [Online]. Available: <https://books.google.co.id/books?id=uCP2EAAAQBAJ>
- [2] R. Swastika, S. Mukodimah, F. Susanto, M. Muslihudin, and S. Ipinuwati, *IMPLEMENTASI DATA MINING (Clustering, Association, Prediction, Estimation, Classification)*. Penerbit Adab. [Online]. Available: <https://books.google.co.id/books?id=LsOqEAAAQBAJ>
- [3] S. K. Pradhan, N. V Krishna Rao, N. M. Deepika, P. Harish, M. A. P. Kumar, and P. S. Kumar, "Credit Card Fraud Detection Using Artificial Neural Networks and Random Forest Algorithms," in *2021 5th International Conference on Electronics*,

- Communication and Aerospace Technology (ICECA)*, 2021, pp. 1471–1476. doi: 10.1109/ICECA52323.2021.9676142.
- [4] S. A. Arnomo, F. A. Annurrullah, S. Yuli, and S. S. F. Nur, “Evaluasi Model Decision Tree Pada Keputusan Kelayakan Kredit,” *J. Desain Dan Anal. Teknol.*, vol. 2, no. 2, pp. 200–206, 2023, doi: 10.58520/jddat.v2i2.39.
 - [5] S. Sukmawati, S. Sulastri, H. Februariyanti, and A. Jananto, “Perbandingan Algoritma C4.5 Dan Algoritma Naive Bayes Untuk Klasifikasi Pekerja Migran Indonesia,” *Informatika*, vol. 14, no. 1, 2022, doi: 10.36723/juri.v14i1.280.
 - [6] N. L. Ratniasih, “Optimasi Data Mining Menggunakan Algoritma Naïve Bayes Dan C4.5 Untuk Klasifikasi Kelulusan Mahasiswa,” *JUTIK*, vol. 5, 2020, doi: 10.36002/jutik.v5i1.634.
 - [7] J. Jufri *et al.*, *Dasar Dasar Data Mining: Konsep, Teknik Dan Aplikasi*. Yayasan Tri Edukasi Ilmiah, 2025. [Online]. Available: <https://books.google.co.id/books?id=4HBREQAAQBAJ>
 - [8] M. Prameswari, P. Kania, I. G. Ayu, and S. N. Harnoko, “Penerapan Metode Stacking Ensemble Untuk Klasifikasi Status Pinjaman Nasabah Bank,” *Pros. Semin. Nas. SAINS DATA*, vol. 4, no. 1, pp. 802–811, Oct. 2024, doi: 10.33005/senada.v4i1.339.
 - [9] A. Adriyanto, D. Prasetyo, and R. Khodijah, “Angkatan Kerja dan Faktor yang Mempengaruhi Pengangguran,” *J. Ilmu Ekon. Sos. Unmus*, vol. 11, no. 2, 2020, doi: 10.35724/jies.v11i2.2965.
 - [10] D. Hendarsyah, “Analisis Perilaku Konsumen Dan Keamanan Kartu Kredit Perbankan,” *JPS (Jurnal Perbank. Syariah)*, vol. 1, no. 1, pp. 85–96, 2020, doi: 10.46367/jps.v1i1.204.
 - [11] M. S. Mauludin, “Analisis Perilaku Konsumen Dalam Transaksi Di e-Commerce,” *Proc. Islam. Econ. Business, Philanthr.*, vol. 1, no. 1, pp. 108–123, 2022, [Online]. Available: <https://jurnalfebi.iainkediri.ac.id/index.php/proceedings/article/view/225>
 - [12] F. T. Rohmah, F. Silviahana, A. Titasyfa, Z. Ibrahim, and W. Hidayat, “Pengaruh Gaya Hidup dan Perilaku Konsumtif Belanja Online Pada Remaja,” *J. Intelek Insa. Cendikia*, vol. 1, no. 4 SE-Articles, pp. 1199–1210, [Online]. Available: <https://jicnusantara.com/index.php/jiic/article/view/619>
 - [13] Z. S. Ritonga, "Pengaruh Pendidikan dan Pendapatan Terhadap Preferensi Perbankan Syariah di Kalangan Masyarakat Urban (Studi Kasus Kecamatan Percut Sei Tuan, Kabupaten Deli Serdang, Sumatera Utara)," Doctoral dissertation, Universitas Islam Negeri Sumatera Utara Medan, 2024 [Online]. Available: http://repository.uinsu.ac.id/22982/1/Cover_Zernita_Sari.pdf
 - [14] N. P. Amandari and N. S. ER, “Penerapan Algoritma Decision Tree dalam Segmentasi Customer,” *JELIKU (Jurnal Elektron. Ilmu Komput. Udayana)*, vol. 12, no. 2, pp. 315–322, 2023, doi: 10.24843/JLK.2023.v12.i02.p09.