KLASIFIKASI PELANGGAN DI KARAKATAU BILIAR MENGGUNAKAN *DECISION TREE*

TUGAS AKHIR

Disusun Untuk Memenuhi Syarat Kelulusan Program Strata I
pada Sekolah Tinggi Manajemen Informatika Dan Komputer
(STMIK) Palangkaraya



OLEH :

SOFINA GRACIA FLODARIANI
C2155201088
PROGRAM STUDI TEKNIK INFORMATIKA

SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER
(STMIK) PALANGKARAYA
2026

KLASIFIKASI PELANGGAN DI KARAKATAU BILIAR MENGGUNAKAN *DECISION TREE*

TUGAS AKHIR

Disusun Untuk Memenuhi Syarat Kelulusan Program Strata I
pada Sekolah Tinggi Manajemen Informatika Dan Komputer
(STMIK) Palangkaraya

OLEH :

SOFINA GRACIA FLODARIANI
C2155201088
PROGRAM STUDI TEKNIK INFORMATIKA

SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER
(STMIK) PALANGKARAYA
2026

PERSETUJUAN

KLASIFIKASI PELANGGAN DI KARAKATAU BILIAR MENGGUNAKAN *DECISION TREE*

Tugas Akhir Ini Telah Disetujui dan Disahkan

Pembimbing I,

Hotmian Sithoang, M.Kom.
NIK. 198503282008002

Pembimbing II,

Amaya Andri Damaini, S.Kom,. MT.
NIK. 199309042018001

Mengetahui

Ketua STMIK Palangkaraya,

Ir. Hj. Siti Maryamah, M.M.
NIK. 196901041995105

Mengesahkan

Ketua Prodi Teknik Informatika

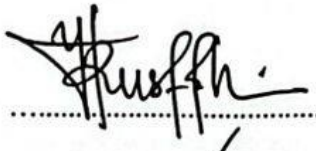
Ir. Abdul Hadi, S.T, M.Kom.
NIK. 198505102021101

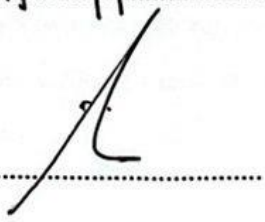
PENGESAHAN

KLASIFIKASI PELANGGAN DI KARAKATAU BILIAR MENGGUNAKAN *DECISION TREE*

Tugas Akhir ini Telah Diujikan, Dinilai, dan Disahkan
Oleh Tim Penguji pada Tanggal 3 Januari 2026

Tim Penguji Tugas Akhir :

1. Sulistyowati, S.Kom., M.Cs. ..... 
   Ketua

2. Rudini, M.Pd. ..... 
   Sekretaris

3. Lili Rusdiana, M.Kom. ..... 
   Anggota

4. Hotmian Sitohang, M. Kom ..... 
   Anggota

5. Amaya Andri Damaini, S.Kom,. MT. ..... 
   Anggota

PERNYATAAN KEASLIAN

Yang bertanda tangan di bawah ini:

Nama Mahasiswa : **SOFINA GRACIA FLODARIANI**
N I M : **C2155201088**

menyatakan bahwa Tugas Akhir dengan judul :

KLASIFIKASI PELANGGAN DI KARAKATAU BILIAR MENGGUNAKAN *DECISION TREE*

adalah hasil karya saya dan bukan merupakan duplikasi sebagian atau seluruhnya dari karya orang lain, kecuali bagian yang sumber informasi dicantumkan.

Pernyataan ini dibuat dengan sebenar-benarnya secara sadar dan bertanggungjawab dan saya bersedia menerima sanksi pembatalan Tugas Akhir apabila terbukti melakukan duplikasi terhadap Tugas Akhir atau karya ilmiah lain yang sudah ada.

Palangka Raya, 21 Januari 2026

Membuat Pernyataan,



SOFINA GRACIA FLODARIANI

MOTTO DAN PERSEMBAHAN

MOTTO :

"Kesuksesan bukan tentang seberapa cepat kita sampai, tetapi tentang kesabaran, ketekunan, dan keikhlasan dalam setiap proses yang dijalani".
- Penulis

PERSEMBAHAN :

Tugas Akhir ini saya persembahkan dengan penuh rasa syukur kepada:

- Tuhan Yesus Kristus, yang telah memberikan kasih, kekuatan, kesehatan, dan penyertaan-Nya dalam setiap proses sehingga Tugas Akhir ini dapat diselesaikan dengan baik.
- Kepada kedua orang tua tercinta, terima kasih atas doa yang tidak pernah putus, kasih sayang yang tulus, pengorbanan, serta kesabaran yang selalu menjadi sumber kekuatan dalam setiap langkah hidup dan pendidikan saya.
- Kepada Teman-teman seperjuangan, terima kasih atas kebersamaan, dukungan, dan semangat yang telah diberikan selama masa perkuliahan hingga penyusunan Tugas Akhir ini.

INTISARI

SOFINA GRACIA FLODARIANI, C2155201088, 2025. *"Klasifikasi Pelanggan Di Karakatau Biliar Menggunakan Decision Tree"*. Pembimbing I, Hotmian Sitohang, M. Kom. dan Pembimbing II Amaya Andri Damaini, S.Kom,. MT.

Krakatau Biliar sebagai salah satu penyedia layanan hiburan, Menghadapi tantangan dalam mengenali karakteristik dan pola perilaku pelanggan secara akurat untuk menentukan strategi layanan yang tepat. Selama ini, belum adanya sistem klasifikasi otomatis menyebabkan sulitnya membedakan antara pelanggan yang hanya datang sesekali dengan pelanggan yang memiliki loyalitas tinggi. Penelitian ini bertujuan untuk menegtahui pelanggan yang Baru, Pelanggan Tetap, dan Pelanggan Aktif, berdasarkan intensitas aktivitas bermain mereka.

Metode yang digunakan dalam penelitian ini adalah algoritma *Decision Tree (C4.5)* yang diimplementasikan menggunakan perangkat lunak *RapidMiner*. Data penelitian bersumber dari rekam jejak transaksi pelanggan selama periode satu tahun, terhitung mulai dari 1 Maret 2024 hingga 28 Februari 2025. Variabel yang dianalisis difokuskan pada atribut perilaku, khususnya frekuensi kunjungan dan durasi bermain, untuk mendapatkan pola klasifikasi yang objektif. Proses pengolahan data dilakukan melalui tahapan preprocessing, pembentukan model, serta pengujian validitas model menggunakan teknik *10-Fold Cross Validation*.

Hasil penelitian menunjukkan bahwa algoritma *Decision Tree* berhasil mengklasifikasikan pelanggan dengan tingkat akurasi yang sangat baik sebesar 95.00%. Secara spesifik, model mampu mengenali kategori Pelanggan yang Baru dengan tingkat keberhasilan 100%, Pelanggan Aktif sebesar 95.24%, dan Pelanggan Tetap sebesar 94.74%. Berdasarkan struktur pohon keputusan yang terbentuk, atribut Frekuensi Kunjungan teridentifikasi sebagai faktor penentu utama *(Root Node)*, dengan aturan keputusan: jika kunjungan ≤ 3 kali dikategorikan sebagai Pelanggan Baru, kunjungan antara 4 sampai 13 kali sebagai Pelanggan Tetap, dan kunjungan di atas 13 kali sebagai Pelanggan Aktif. Hasil ini diharapkan dapat menjadi landasan bagi manajemen Krakatau Biliar dalam merancang strategi retensi pelanggan yang lebih efektif.

Kata Kunci : ***Data Mining, Decision Tree, Klasifikasi Pelanggan, RapidMiner.***

ABSTRACT

SOFINA GRACIA FLODARIANI, C2155201088, 2025. *"Customer Classification at Krakatau Billiards Using Decision Tree"*. Advisor I, Hotmian Sitohang, M. Kom., and Advisor II, Amaya Andri Damaini, S.Kom,. MT.

Krakatau Biliar, as an entertainment service provider, encounters challenges in accurately identifying customer characteristics and behavioral patterns to determine appropriate service strategies. The absence of an automated classification system has made it difficult to distinguish between occasional customers and those with high loyalty. This study aims to apply data mining techniques to classify Krakatau Biliar customers into three main categories—New Customers, Regular Customers, and Active Customers—based on the intensity of their playing activity.

The method employed in this research is the Decision Tree (C4.5) algorithm, implemented using RapidMiner software. The research data is sourced from customer transaction records over a one-year period, ranging from March 1, 2024, to February 28, 2025. The analyzed variables focus on behavioral attributes, specifically visit frequency and playing duration, to obtain an objective classification pattern. The data processing involves several stages, including preprocessing, model formation, and model validity testing using the 10-Fold Cross Validation technique.

The results indicate that the Decision Tree algorithm successfully classified customers with an excellent accuracy rate of 96.00%. Specifically, the model achieved a recognition success rate of 100% for New Customers, 95.24% for Active Customers, and 94.74% for Regular Customers. Based on the generated decision tree structure, the Visit Frequency attribute was identified as the primary determining factor (Root Node), with the following decision rules: visits ≤ 3 times are categorized as New Customers, visits between 4 to 13 times as Regular Customers, and visits exceeding 13 times as Active Customers. These findings are expected to serve as a foundation for Krakatau Biliar's management in designing more effective customer retention strategies.

Keywords: Data Mining, Decision Tree, Customer Classification, RapidMiner.

KATA PENGANTAR

Puji syukur penulis panjatkan kehadirat Tuhan Yang Maha Esa yang telah melimpahkan rahmat dan hidayah-Nya. Sehingga penulis dapat menyelesaikan Tugas Akhir ini yang berjudul "**Klasifikasi Pelanggan Di Karakatau Biliar Menggunakan *Decision Tree***" yang merupakan salah satu syarat kelulusan program Sarjana Komputer pada Sekolah Tinggi Manajemen Informatika dan Komputer (STMIK) Palangkaraya.

Pada kesempatan ini penulis mengucapkan banyak terimakasih atas bantuan serta dukungan kepada pihak-pihak yang sudah membantu menyelesaikan Tugas Akhir, penulis mengucapkan terimakasih kepada :

1. Bapak Feri, selaku pemilik dan pengelola Krakatau Biliar.
2. Ir. Hj. Siti Maryamah, M.M. selaku ketua Sekolah Tinggi Manajemen Informatika dan Komputer (STMIK) Palangkaraya.
3. Hotmian Sitohang, M. Kom, sebagai dosen pembimbing I Tugas Akhir.
4. Amaya Andri Damaini, S.Kom,. MT, sebagai dosen pembimbing II Tugas Akhir.

Penulis mengharapkan saran serta kritikan yang dapat membangun dalam penyempurnaan Tugas Akhir ini, penulis juga menyampaikan banyak terimakasih kepada semua pihak yang telah membantu dan mendukung dalam penulisan Tugas Akhir, Akhir kata penulis mengucapkan terimakasih.

Palangka Raya, 21 Januari 2026

Penulis

DAFTAR ISI

PERSETUJUAN ..... **iii**
PENGESAHAN ..... **iv**
PERNYATAAN KEASLIAN ..... **v**
MOTTO DAN PERSEMBAHAN ..... **vi**
INTISARI ..... **vii**
ABSTRACT ..... **viii**
KATA PENGANTAR ..... **ix**
DAFTAR ISI ..... **x**
DAFTAR TABEL ..... **xii**
DAFTAR GAMBAR ..... **xiii**
DAFTAR LAMPIRAN ..... **xiv**
BAB I PENDAHULUAN **1**
1.1 Latar Belakang 1
1.2 Perumusan Masalah 2
1.3 Batasan Masalah 2
1.4 Tujuan dan Manfaat 3
1.5 Sistematika Penulisan 5
BAB II TINJAUAN PUSTAKA **6**
2.1 Kajian Teori 6
2.2 Penelitian Yang Relevan 26
BAB III METODOLOGI PENELITIAN **30**
3.1 Tinjauan Umum 30
3.2 Jenis Penelitian 30
3.3 Desain Penelitian 31
3.4 Instrumen Penelitian 33
3.5 Analisis Kebutuhan 34
3.6 Desain 37
3.7 Jadwal Penelitian 42
BAB IV HASIL DAN PEMBAHASAN **43**

4.1 Hasil 43
4.2 Pembahasan 67
BAB V KESIMPULAN DAN SARAN 74
5.1. Kesimpulan 74
5.2. Saran 75
DAFTAR PUSTAKA
LAMPIRAN

DAFTAR TABEL

Tabel 2. 1 Tabel Data untuk Klasifikasi Resiko Kredit 15
Tabel 2. 2 Nilai *Entropy* dan *Gain* untuk Menentukan Simpul Akar 16
Tabel 2. 3 Simbol *Flowchart* 24
Tabel 2. 4 Penelitian Yang Relevan 26
Tabel 3. 1 Variabel Klasifikasi 35
Tabel 3. 2 Klasifikasi Pelanggan 35
Tabel 3. 3 Format Dataset 36
Tabel 3. 4 Perangkat Keras 36
Tabel 3. 5 Perangkat Lunak. 36
Tabel 3. 6 Jadwal Penelitian 42
Tabel 4. 1 Hasil *Confusion Matrix* 70
Tabel 4. 2 Ringkasan Performa Model *Decision Tree* 72

DAFTAR GAMBAR

Gambar 2. 1 Permainan Billiard 6
Gambar 2. 2 Konsep *Decision Tree* 10
Gambar 2. 3 Pohon Keputusan *C4.5* dengan Simpul Akar Aset 17
Gambar 2. 4 *Confusion Matrix* 18
Gambar 2. 5 Tahapan Model Proses *CRISP-DM* 22
Gambar 2. 6 Tampilan *Rapid Miner* 26
Gambar 3. 1 Krakatau Billiard 30
Gambar 3. 2 Desain Proses 38
Gambar 4. 1 *RapidMiner* Studio 45
Gambar 4. 2 *Impot Data RapidMiner Studio* 46
Gambar 4. 3 Lokasi File 47
Gambar 4. 4 File Dataset 48
Gambar 4. 5 Tambahkan Label 49
Gambar 4. 6 Simpan Dataset 50
Gambar 4. 7 Hasil Dataset 51
Gambar 4. 8 Operator Set Role 52
Gambar 4. 9 Operator *Split Data* 53
Gambar 4. 10 Operator *Cross Validation* 54
Gambar 4. 11 Operator *Decision Tree* 55
Gambar 4. 12 Operator *Apply Model* 56
Gambar 4. 13 *Operator Performance (Classification)* 57
Gambar 4. 14 *Set Role* Dataset 58
Gambar 4. 15 Split Dataset (80:20) 59
Gambar 4. 16 Pengujian Validitas Model 60
Gambar 4. 17 *Design Cross Validation* 61
Gambar 4. 18 Evaluasi Model 62
Gambar 4. 19 Deskripsi Evaluasi 62
Gambar 4. 20 Deskripsi Statistik 63
Gambar 4. 21 Pohon Keputusan 64
Gambar 4. 22 Visualisasi Kategori Pelanggan 65
Gambar 4. 23 Visualisasi Frekuensi Kunjungan 66

DAFTAR LAMPIRAN

Lampiran 1. Surat Izin Penelitian
Lampiran 2. Surat Balasan Izin Penelitian
Lampiran 3. Surat tugas pembimbing Tugas Akhir
Lampiran 4. Lembar konsultasi bimbingan Tugas Akhir
Lampiran 5. Surat Tugas Penguji Seminar Proposal Tugas Akhir
Lampiran 6. Berita Acara Penilaian Seminar Proposal Tugas Akhir
Lampiran 7. Daftar Hadir Peserta Seminar Proposal Tugas Akhir
Lampiran 8. Lembar Wawancara
Lampiran 9. Dokumentasi
Lampiran 10. Surat Tugas Penguji Tugas Akhir
Lampiran 11. Berita Acara Penilaian Sidang Tugas Akhir

BAB I

PENDAHULUAN

1.1 Latar Belakang

Karakatau Biliar adalah salah satu tempat hiburan yang menyediakan permainan biliar untuk berbagai kalangan, mulai dari pemula hingga pemain profesional. Namun, dalam menjalankan usahanya, Karakatau Biliar menghadapi beberapa masalah, terutama dalam memahami kebiasaan pelanggan. Tanpa sistem yang jelas untuk mengelompokkan pelanggan, pengelola kesulitan menentukan strategi pemasaran yang tepat, mengoptimalkan layanan, dan mempertahankan pelanggan, yang pada akhirnya bisa memengaruhi pendapatan bisnis.

Saat ini, strategi pemasaran Karakatau Biliar masih bersifat umum dan belum mempertimbangkan perbedaan antara jenis pelanggan. Misalnya, promosi yang diberikan sama untuk semua orang, baik yang hanya datang sesekali maupun yang sering bermain. Akibatnya, promosi kurang efektif, banyak diskon dan paket layanan tidak sesuai dengan kebutuhan pelanggan, dan biaya pemasaran menjadi lebih besar tanpa hasil yang maksimal.

Dari segi operasional, pengelola juga mengalami kesulitan dalam mengatur penggunaan meja biliar dan jadwal layanan. Misalnya, pada jam sibuk, sering terjadi antrean panjang yang membuat pelanggan tidak nyaman, sementara di jam sepi banyak meja yang kosong dan tidak dimanfaatkan dengan baik. Hal ini menunjukkan bahwa

diperlukan analisis lebih lanjut untuk memahami pola kunjungan pelanggan agar pengelolaan tempat bisa lebih efisien.

Untuk mengatasi masalah ini, algoritma *Decision Tree* dapat digunakan sebagai solusi. *Decision Tree* adalah metode klasifikasi yang digunakan untuk mengelompokkan data berdasarkan pola tertentu dari data historis. Metode ini membentuk struktur pohon, di mana setiap cabang merepresentasikan keputusan berdasarkan atribut, dan setiap daun menunjukkan hasil klasifikasi. Dalam konteks analisis pelanggan, *Decision Tree* digunakan untuk mengenali pola perilaku dan membantu pengambilan keputusan seperti segmentasi atau strategi pemasaran.

Dengan menerapkan *Decision Tree*, pelanggan dapat diklasifikasikan menjadi beberapa kategori, seperti pelanggan baru (yang datang hanya sekali-kali), pelanggan

1.2 Perumusan Masalah

Bagaimana penerapan algoritma *Decision Tree* dalam mengklasifikasikan pelanggan di Karakatau Biliar?.

1.3 Batasan Masalah

Batasan masalah untuk penelitian Klasifikasi Pelanggan di Karakatau Biliar Menggunakan *Decision Tree* dapat dijabarkan sebagai berikut:

a. Data pelanggan yang digunakan dalam penelitian ini terbatas pada pelanggan Karakatau Biliar dalam 1 Tahun yaitu dari 1 Maret 2024 sampai 28 Februari 2025, Periode ini mencakup 12 bulan penuh, yang memungkinkan peneliti untuk

menangkap pola musiman, tren kunjungan, dan fluktuasi pelanggan secara menyeluruh dari awal hingga akhir tahun.

b. Data yang dianalisis meliputi variabel seperti frekuensi kunjungan, durasi bermain, dan jumlah transaksi.
c. Penelitian ini dibatasi pada penggunaan *C4,5* dengan algoritma *Decision Tree* untuk mengelompokkan pelanggan.
d. Pelanggan akan diklasifikasikan ke dalam beberapa kategori utama, seperti: Pelanggan baru, pelanggan tetap, dan pelanggan aktif.
e. Seluruh proses pengolahan data dan implementasi *Decision Tree* dilakukan menggunakan *RapidMiner*.

1.4 Tujuan dan Manfaat

1.4.1 Tujuan

Penelitian ini bertujuan untuk menggunakan algoritma *Decision Tree* untuk mengelompokkan pelanggan Karakatau Biliar berdasarkan kebiasaan kunjungan, durasi bermain dan jumlah transaksi.

1.4.2 Manfaat

Adapun manfaat dari penelitian ini ialah :

a. Bagi Pembaca

   Hasil penelitian ini bisa menjadi sumber informasi dan referensi bagi siapa saja yang ingin memahami cara mengelompokkan pelanggan menggunakan *Decision Tree*.

b. Bagi Penulis

Penelitian ini membantu penulis dalam menambah ilmu, pengalaman, dan keterampilan dalam menggunakan algoritma *Decision Tree* serta meningkatkan pencapaian pribadi dalam bidang penelitian.

c. Bagi STMIK Palangkaraya

Penelitian ini bisa dijadikan referensi bagi kampus untuk memperkaya wawasan akademik tentang klasifikasi pelanggan dan penerapan algoritma *Decision Tree* dalam bisnis.

d. Bagi Pemilik Usaha

Pemilik usaha Karakatau Biliar dapat memahami pola pelanggan, mengelompokan berdasarkan kebiasaan, serta meningkatkan layanan dan strategi bisnis berdasarkan hasil klasifikasi.

e. Bagi Dunia Bisnis

Penelitian ini bisa menjadi contoh bagaimana teknologi dapat membantu bisnis memahami pelanggan dengan lebih baik dan mengambil keputusan berdasarkan data, sehingga lebih efektif dan efisien.

f. Bagi Pengembang Teknologi

Penelitian ini dapat menjadi inspirasi bagi pengembang teknologi untuk menciptakan sistem otomatis yang bisa membantu bisnis dalam menganalisis data pelanggan dengan lebih mudah.

1.5 Sistematika Penulisan

Adapun sistematika penulisan dalam penulisan tugas akhir ini terdiri dari beberapa bab dan masing-masing bab membahas dan menguraikan pokok permasalahan yang berbeda, sebagai gambaran disini penulis menyertakan garis-garis besarnya yaitu :

BAB I PENDAHULUAN

Bab ini berisikan latar belakang masalah, rumusan masalah, Batasan masalah, tujuan dan manfaat, serta sistematika penulisan.

BAB II TINJAUAN PUSTAKA

Bab ini berisikan tentang tinjauan pustaka yang diambil dari penelitian yang relevan beserta susunan kajian teori yang disesuaikan dengan tema Tugas Akhir.

BAB III METODOLOGI PENELITIAN

Bab ini berisikan tentang tahapan yang dilakukan peneliti dalam mengumpulkan informasi atau data yang dibutuhkan.

BAB IV HASIL DAN PEMBAHASAN

Bab ini berisikan tentang hasil dari penelitian dan pembahasan dari hasil yang didapat.

BAB V KESIMPULAN DAN SARAN

Bab ini berisikan tentang kesimpulan dan saran yang diberikan untuk pengembangan selanjutnya.

BAB II

TINJAUAN PUSTAKA

2.1 Kajian Teori

2.1.1 Permainan *Billiar*

Billiar merupakan permainan ketangkasan memasukan bola ke dalam lobang di sudut meja billiar dimana meja billiar memiliki model yang sudah dimodifikasi dengan bentuk ukuran yang telah ditentukan yaitu 3,7m x 1,8m, dengan cara menyodok bola induk/putih dengan stik kepada bola yang ingin dimasukkan itu. Bola induk berwarna putih, bola billiard bernomor 1 sampai dengan 15 (MARCHAMAH dalam Lutfi & Situmorang, 2023).



Gambar 2. 1 Permainan Billiard
Sumber: https://surabaya.uph.edu/id/student_club/billiard/

Di Indonesia sendiri permainan billiar pertama kali muncul dari kalangan masyarakat bawah. Hal ini sangat jauh berbeda dengan asal perjalanan billiard yang di temukan abad ke 15 di Eropa yang mengalami kemajuan pesat, sehingga menjadi kegiatan olahraga yang dilakukan oleh semua kalangan baik Raja, Presiden,

Pengusaha, dan anggota masyarakat lainnya. Olahraga Billiar menempati sebuah status atau alat yang dapat melepas jenuh atau bosan, billiar juga dipilih karena tidak mengeluarkan biaya yang sangat mahal dan tidak ribet untuk memainkannya, dengan menyusuaikan dengan konteks-konteks sosial yang ada.

Billiar adalah salah satu jenis permainan bola sodok yang dapat membangun sebuah nama dan dapat mengangkat harkat martabat bangsa di manca negara (Lutfi & Situmorang, 2023).

2.1.2 Klasifikasi

Dalam *Big Data Analytics*, pengumpulan data pada algoritma *decision tree* dapat dilakukan berdasarkan klasifikasi. Klasifikasi adalah pengelompokan dari suatu rekaman (set pelatihan), di mana setiap rekaman berisi kumpulan atribut. Tujuan klasifikasi adalah untuk menemukan model dari atribut kelas, sehingga kemudian dapat memberikan nilai terhadap atribut lainnya (Ramadhon, et al., 2024).

Saat mengklasifikasikan data, perhatian harus diberikan terhadap apakah data tersebut termasuk dalam data latihan atau data pengujian. Data latihan merupakan data yang memiliki proporsi lebih besar daripada data pengujian. Data-data tersebut kemudian diolah menggunakan bantuan komputer sehingga menghasilkan suatu model.

2.1.3 Data Mining

Data mining dapat diartikan sebagai analisi otomatis dari data yang berjumlah besar atau kompleks dengan tujuan untuk menemukan pola dan relasi-relasi yang tersembunyi dalam sejumlah data yang besar dan bertujuan untuk melakukan

klasifikasi, estimasi, prediksi, assosiasi *rule*, *clustering*, deskripsi dan visualisasi. Secara sederhana data mining merupakan sebagai proses menyaring atau menambang pengetahuan dari sejumlah data yang besar (Wahyudi, et al., 2021).

Tujuan Dari Adanya *Data Mining* adalah:

a. *Explanatory*, yaitu untuk menjelaskan beberapa kegiatan observasi atau suatu kondisi.
b. *Confirmatory*, yaitu untuk mengkonfirmasi suatu hipotesis yang telah ada.
c. *Exploratory*, yaitu untuk menganalisa data baru suatu relasi yang janggal.

Data mining adalah proses menemukan pola yang bermakna dari kumpulan data berukuran besar, yang dapat disimpan dalam database, data warehouse, atau media penyimpanan lainnya. Bidang ini berkaitan dengan berbagai disiplin ilmu, seperti sistem *database*, data *warehousing*, statistik, *machine learning, information retrieval*, dan komputasi tingkat tinggi. Selain itu, data mining juga didukung oleh bidang lain seperti jaringan saraf tiruan, pengenalan pola, analisis data spasial, database gambar, serta pemrosesan sinyal. Secara umum, data mining merupakan proses otomatis atau semiotomatis untuk mengidentifikasi pola dalam data. Pola yang ditemukan harus memiliki nilai signifikan dan memberikan manfaat, terutama dalam aspek ekonomi. Proses ini memerlukan data dalam jumlah besar untuk menghasilkan wawasan yang berharga.

Karakteristik data mining adalah sebagai berikut:

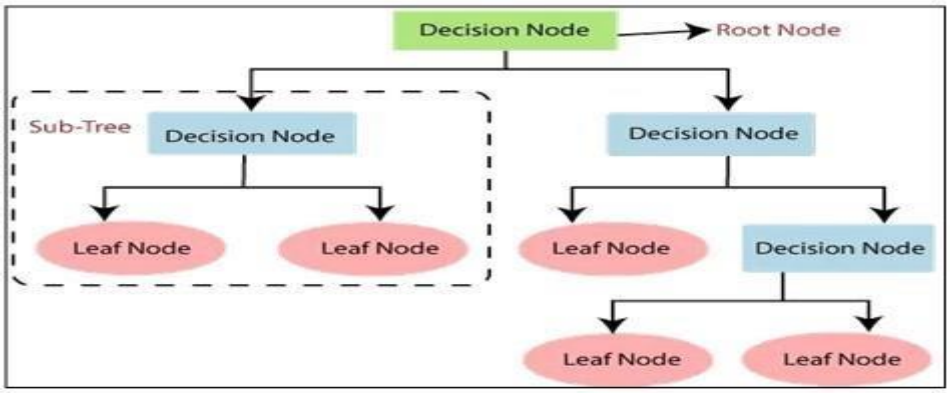
a. *Data mining* berhubungan dengan penemuan sesuatu yang tersembunyi dan pola data tertentu yang tidak diketahui sebelumnya.

b. *Data mining* biasa menggunakan data yang sangat besar. Biasanya data yang besar digunakan untuk membuat hasil lebih dipercaya.
c. *Data mining* berguna untuk membuat keputusan yang kritis, terutama dalam strategi (Abdillah dalam Wahyudi, et al., 2021).

2.1.4 Decision Tree

Decision Tree merupakan salah satu metode *supervised learning* yang digunakan dalam klasifikasi dan prediksi, berbentuk struktur pohon yang terdiri dari simpul-simpul *(nodes)* dan cabang *(branches)* yang merepresentasikan proses pengambilan keputusan dan konsekuensinya (Ramadhon, et al., 2024). Model ini memerlukan training dataset sebagai dasar dalam membentuk pola keputusan, menggantikan pengalaman manusia di masa lalu (Kurniawan dalam Ramadhon et al., 2024).

Struktur utama dalam *Decision Tree* terdiri dari beberapa komponen penting, yaitu *root node, decision node, dan leaf node. Root node* merupakan simpul awal yang berfungsi sebagai dasar dalam pengambilan keputusan pertama, biasanya dipilih dari atribut dengan nilai *information gain* tertinggi. *Decision node* adalah simpul internal yang menguji nilai dari suatu atribut tertentu untuk melanjutkan proses klasifikasi. Sementara itu, *leaf node* atau simpul daun merupakan simpul akhir yang menandakan hasil klasifikasi atau keputusan akhir dari pohon tersebut. Agar lebih jelas dapat kita lihat visualisasi *decision tree* pada gambar 2.2.



Gambar 2. 2 Konsep Decision Tree
Sumber: Ramadhon, et al., 2024

Dalam membangun *decision tree*, proses pemilihan atribut sangat penting. Atribut terbaik dipilih berdasarkan seberapa baik atribut tersebut memisahkan *data* ke dalam kelas yang berbeda. Untuk itu, biasanya digunakan konsep *Entropy* untuk mengukur tingkat ketidakpastian data dan *Information Gain* untuk menentukan atribut yang memberikan pembagian terbaik.

Rumus matematika *Entropy* adalah sebagai berikut.

Entropy(S) = - Entropy = - $\sum n\ i = pi * log2\ (pi)$ (1)

Sumber: Ramadhon, et al., 2024

Keterangan:

S: Himpunan kasus n: Jumlah partisi s

pi: Jumlah kasus pada partisi ke-i

Setelah pemilihan atribut, algoritma *Decision Tree* akan memulai proses pembentukan pohon dengan memilih atribut yang memiliki *information gain* tertinggi sebagai akar *(root node).* Proses ini dilanjutkan dengan pemisahan data *(splitting)*

berdasarkan atribut-atribut yang relevan. *Information gain* sendiri merupakan nilai yang menunjukkan seberapa besar suatu atribut dapat mengurangi ketidakpastian *(entropy)* dalam data. Atribut dengan nilai *gain* tertinggi akan dipilih, sementara atribut lain yang kurang relevan dapat diabaikan dalam proses pembentukan pohon keputusan.

Rumus Matematika Gain adalah sebagai berikut.

$$\text{Gain}(S,A) = \text{Entropy}(S) - \sum_{i=1}^{n} |Si| \; x \; \overline{Entropy}(Si)|S| \qquad (2)$$

Sumber: Ramadhon, et al., 2024

Keterangan:

S: Himpunan kasus

n: Jumlah partisi atribut

A|Si|: Jumlah kasus pada partisi ke -i

|S|: Jumlah kasus dalam S

Dengan memahami *formula entropy* dan *information gain*, data dapat diproses menggunakan algoritma *Decision Tree* untuk membentuk pohon keputusan. Algoritma ini memiliki keunggulan dalam hal interpretasi hasil, mampu menangani data kategorik maupun numerik, serta tidak memerlukan proses normalisasi. Namun, *Decision Tree* juga memiliki kelemahan, seperti kecenderungan mengalami *overfitting* jika terlalu sesuai dengan data pelatihan. Model yang dihasilkan dapat dipahami sebagai kumpulan aturan keputusan, di mana setiap jalur dari akar ke daun merepresentasikan logika klasifikasi tertentu.

Karena setiap area keputusan bersifat terpisah, pohon ini juga dapat dianggap sebagai representasi alternatif yang saling melengkapi (disjungsi).

a. Keunggulan Metode Pohon Keputusan

Kelebihan dari pendekatan pohon keputusan melibatkan:

1) Simpulan daerah pengambilan keputusan yang sebelumnya kompleks dan sangat global, dapat disederhanakan menjadi lebih ringkas dan spesifik.
2) Penghapusan perhitungan yang tidak perlu, karena saat menggunakan metode pohon keputusan, sampel diuji hanya berdasarkan kriteria atau kelas tertentu.
3) Fleksibilitas untuk memilih fitur dari node internal yang berbeda, dengan fitur yang dipilih membedakan suatu kriteria dari kriteria lain dalam node yang sama. Kefleksibelan ini meningkatkan kualitas keputusan dibandingkan dengan metode penghitungan satu tahap yang lebih konvensional.
4) Dalam analisis multivariat, terutama dengan jumlah kriteria dan kelas yang tinggi, pohon keputusan dapat menghindari masalah dengan menggunakan kriteria yang lebih sedikit pada setiap node internal tanpa mengurangi kualitas keputusan.

b. Kekurangan Metode Pohon Keputusan

Adapun kekurangan dari penerapan algoritma *Decision Tree* adalah sebagai berikut:

a. Terjadi tumpang tindih terutama ketika terdapat banyak kelas dan kriteria yang digunakan, yang dapat menyebabkan peningkatan waktu pengambilan keputusan dan penggunaan memori.
b. Akumulasi kesalahan dari setiap tingkat dalam pohon keputusan yang besar.
c. Kesulitan dalam merancang pohon keputusan yang optimal.
d. Kualitas keputusan yang dihasilkan sangat tergantung pada desain pohon tersebut.

2.1.5 C4,5

Salah satu algoritma yang dapat digunakan untuk membuat pohon keputusan *(decision tree)* adalah algoritma *C4.5*. Algoritma *C4.5* merupakan algoritma yang sangat populer yang digunakan oleh banyak peneliti di dunia, hal ini dijelaskan oleh Xindong Wu dan Vipin Kumar dalam bukunya yang berjudul *The Top Ten Algorithms in Data Mining*. Algoritma *C4.5* merupakan pengembangan dari algoritma *ID3* yang di ciptakan oleh J. Rose Quinlan (Mardi, 2021).

Menurut Mardi (2021) secara umum algoritma *C4.5* untuk membangun pohon keputusan adalah sebagai berikut:

a. Pilih atribut sebagai akar
b. Buat cabang untuk tiap-tiap nilai
c. Bagi kasus dalam cabang
d. Ulangi proses untuk setiap cabang sampai semua kasus pada cabang memiliki kelas yang sama.

Untuk memilih atribut sebagai akar, didasarkan pada nilai gain tertinggi dari atribut-atribut yang ada. Untuk menghitunggaindigunakan persamaan 3.

Gain(S,A) = Entropy(S) - $\sum_{i=1}^{n}|Si|\ x\ \overline{Entropy}(Si)|S|$ (3)

Sumber: Mardi, 2021

Keterangan:

S: Himpunan kasus A: Atribut

n: Jumlah partisi dalam atribut A

|Si|: Jumlah kasus pada partisi ke -i

|S|: Jumlah kasus dalam S

Sementara itu, perhitungan nilai entropy dapat dilihat pada persamaan 4.

Entropy (S) = $\sum_{i=1}^{n}$ - pi * log2 pi (4)

Sumber: Mardi, 2021

Keterangan:

S = himpunan kasus A = fitur

n = jumlah partisi S

pi = proporsi dari Si terhadap S

Untuk lebih jelasnya langkah-langkah dalam pembuatan pohon keputusan, berikut diberikan contoh kasus yang dapat diklasifikasi menggunakan algoritma C4.5. Dari Tabel 2.1, hitung apakah seorang nasabah bermasalah dalam kredit atau tidak.

Data training pada Tabel 2.1 adalah untuk menentukan apakah seorang nasabah bermasalah atau tidak yang ditentukan oleh kolom *predictor*, simpanan, aset, dan pendapatan. Kolom resiko kredit adalah kelas dari masing-masing record.

Tabel 2. 1 Tabel Data untuk Klasifikasi Resiko Kredit

Pelanggan	**Simpanan**	**Aset**	**Pendapatan**	**Resiko**
1	*Medium*	*High*	75	*Good*
2	*Low*	*Low*	50	*Bad*
3	*High*	*Medium*	25	*Bad*
4	*Medium*	*Medium*	50	*Good*
5	*Low*	*Medium*	100	*Good*
6	*High*	*High*	25	*Good*
7	*Low*	*Low*	25	*Bad*
8	*Medium*	*Medium*	75	*Good*

Sumber: Mardi, 2021

Adapun langkah untuk membuat pohon keputusan, yaitu :

1) Untuk atribut pendapatan yang bernilai angka, dibuat dalam bentuk kategori, yaitu pendapatan <=25, pendapatan >25, pendapatan <=50, pendapatan >50, pendapatan ,=75, dan pendapatan >75
2) Hitung nilai entropy. Dari data training diketahui jumlah kasus ada 8, yang beresiko kredit *good* 5 record dan *bad* 3 *record* sehingga didapat entropy :

   Entropy (S) = $\sum_{i=1}^{n}$- pi * log2 pi (5)

   = (-5/8 * log2(5/8)) + (-3/8 * log2(3/8))

   = 0,9544
3) Hitung nilai *gain* untuk tiap atribut, lalu tentukan nilai *gain* tertinggi. Yang mempunyai nilai gain tertinggi itulah yang akan dijadikan akar dari pohon. Misalkan untuk atribut simpanan dengan nilai low didapat /nilai gain:

Gain(S,A) = Entropy(S) - $\sum_{i=1}^{n}$ (6)

= 0,9544 - (2/8(1) + 3/8(0) + 3/8(0,9183))

= 0,3601

Hasil perhitungan gain untuk tiap atribut dapat terlihat pada Tabel 2.2, nilai *gain* tertinggi akan menjadi akar dari pohon.

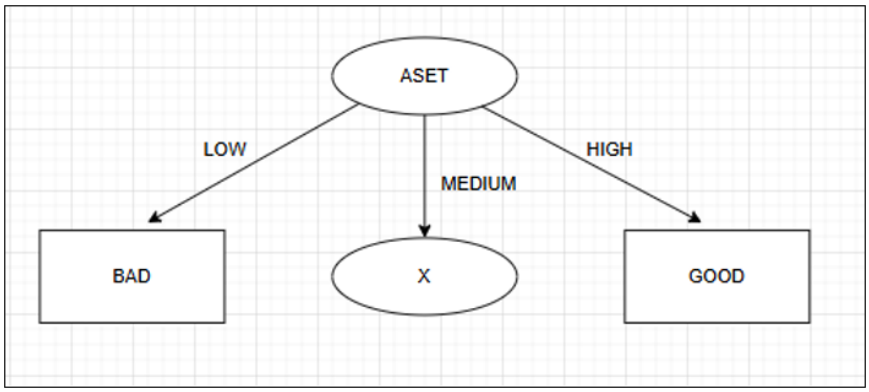
Tabel 2. 2 Nilai *Entropy* dan *Gain* untuk Menentukan Simpul Akar

Simpul			**Data**	**Good**	**Bad**	**Entropy**	**Gain**
Akar	**Total**		8	5	3	0,9544	
	Simpanan						0,3601
		Low	3	1	2	0,9183	
		Medium	3	3	0	0	
		High	2	1	1	1	
	Aset						0,5488
		Low	2	0	2	0	
		Medium	4	3	1	0,8113	
		High	2	2	0	0	
	Pendapatan					0,1589	
		<=25	3	1	2	0,9183	
		>25	5	4	1	0,7219	
							0,3478
		<=50	5	2	3	0,971	
						0	
		>50	3	3	0	0	
							0,0924
		<=75	7	4	3	0,9852	
		>75	1	1	0	0	

Sumber: Mardi, 2021

Terlihat dari Tabel 2.2 bahwa atribut aset mempunyai nilai *low, medium,* dan *high*. Nilai *low* dan *high* masing-masing sudah menjadi satu klasifikasi karena pada data training, semua aset menghasilkan keputusan yang sama yaitu *bad* untuk nilai *low*

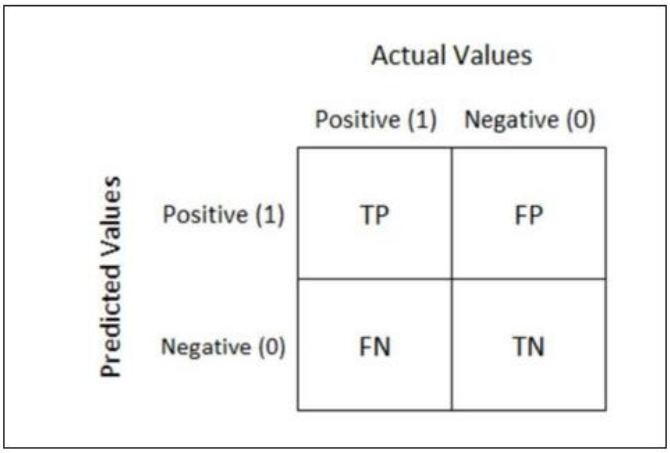
dan *good* untuk nilai *high*. Sedangkan untuk simpul dengan nilai *medium* perlu dipartisi lagi. Gambar 2.9 adalah hasil pembentukan pohon keputusan berdasarkan perhitungan yang terdapat pada Tabel 2.2. Dari hasil perhitungan didapat nilai *gain* tertinggi untuk atribut aset, maka aset menjadi akar pohon keputusan. Untuk menentukan akar dari atribut *medium*, dilakukan lagi perhitungan *gain*.



Gambar 2. 3 Pohon Keputusan C4.5 dengan Simpul Akar Aset
Sumber: Mardi, 2021

2.1.6 Confusion Matrix

Salah satu metode yang efektif untuk mengevaluasi performa model klasifikasi adalah *confusion matrix*. *Confusion matrix* merupakan tabel yang merepresentasikan jumlah prediksi yang benar dan salah dari data uji, sehingga memudahkan dalam menilai tingkat akurasi sistem klasifikasi. Melalui *confusion matrix*, kinerja model dapat dianalisis secara lebih rinci, termasuk dalam mengidentifikasi kesalahan klasifikasi yang terjadi. Metode ini tergolong sederhana namun sangat informatif dalam mengukur efektivitas sistem klasifikasi (Nurhidayat & Dewi, 2023).



Gambar 2. 4 *Confusion Matrix*
Sumber: Nurhidayat & Dewi, 2023

Confusion Matrix digunakan untuk mengevaluasi tingkat akurasi dari suatu proses klasifikasi yang telah dilakukan. Tingkat akurasi ini mengindikasikan proporsi jumlah prediksi yang benar. Untuk menghitung akurasi, presisi, dan *recall*, digunakan rumus sebagai berikut:

Total = TPosPos + PFNon + PFNeg + NonFP + TNonNon + NonFNNeg+ NegFP + NegFNon + TNegNeg

$$Accuracy = \frac{\text{TPosPos+TNonNon+TNegNeg}}{\text{Total}} \quad (6)$$

$$Precision = \frac{\text{TPosPos}}{\text{PosPos+NonFP+NegFP}}$$

$$Recall = \frac{\text{TPosPos}}{\text{TPosPos+PFNon+PFNeg}}$$

2.1.7 Penelitian *Kuantitatif*

Menurut Sugiyono penelitian *kuantitatif* adalah metode yang berlandaskan positivisme, digunakan untuk meneliti populasi atau sampel tertentu dengan teknik pengambilan sampel secara acak. Data dikumpulkan menggunakan instrumen penelitian dan dianalisis secara statistik untuk menguji hipotesis (Lestari, et al., 2022).

Pendekatan ini menganggap bahwa fenomena dapat dikategorikan, diukur, dan memiliki hubungan sebab akibat yang jelas. Penelitian kuantitatif bersifat deduktif, yaitu menggunakan teori sebagai dasar untuk merumuskan hipotesis dan menjawab pertanyaan penelitian.

Menurut Moleong (dalam Lestari, et al., 2022) penelitian kuantitatif memiliki beberapa karakteristik utama, yaitu:

a. Latar alami sebagai sumber data utama, fenomena yang diteliti harus terjadi secara nyata.
b. Fokus pada proses, teknik dan tahapan dalam pengumpulan data lebih diutamakan daripada hasil akhir.
c. Manusia sebagai instrumen penelitian, data dikumpulkan melalui observasi atau wawancara langsung.
d. Berbasis teori dari dasar (*grounded theory*), konsep dan teori ditemukan berdasarkan data empiris.
e. Deskriptif, hasil penelitian disajikan dalam bentuk narasi untuk menggambarkan peristiwa atau objek yang diteliti.

f. Analisis data induktif, kesimpulan penelitian diambil dari fakta-fakta yang ditemukan di lapangan.
g. Desain fleksibel, penelitian dapat diperbarui secara berkala sesuai dengan perkembangan data di lapangan.
h. Validasi hasil dengan informan, interpretasi data harus dirundingkan dengan informan untuk memastikan kesepahaman.
i. Analisis data sejak awal penelitian, proses analisis dilakukan sejak data pertama dikumpulkan hingga tahap interpretasi akhir.

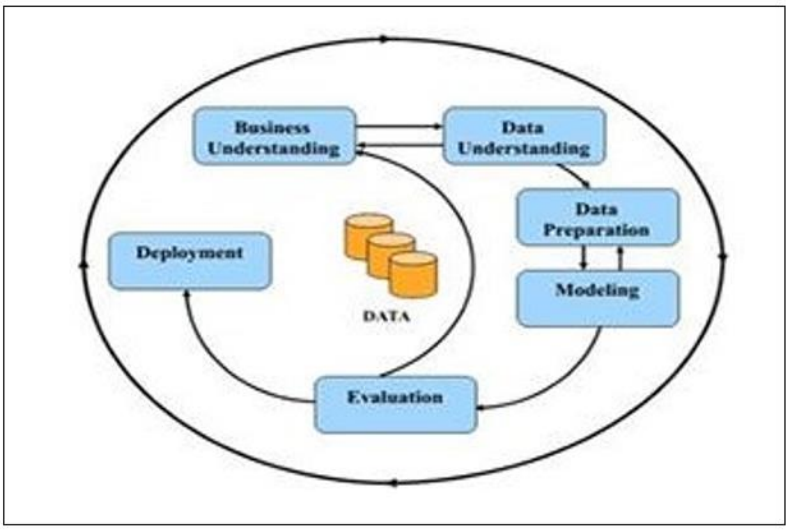
Penelitian ini menggunakan pendekatan kuantitatif karena bertujuan untuk mengklasifikasikan pelanggan berdasarkan data yang dapat diukur secara objektif. Metode ini memungkinkan analisis yang sistematis dan statistik, sehingga hasil yang diperoleh lebih akurat dan dapat diuji keabsahannya. Selain itu, pendekatan kuantitatif sesuai dengan tujuan penelitian yang ingin menemukan pola atau hubungan dalam data pelanggan tanpa keterlibatan subjektivitas peneliti.

2.1.8 CRISP-DM

CRISP-DM merupakan model proses iteratif yang kemudian dikembangkan pada tahun 1996 (Herlambang, et al., 2024). *CRISP-DM* merupakan singkatan dari *Cross-Industry Standard Process for Data Mining* dan mempunyai enam tahapan yaitu:

a. *Business Understanding* : Tahapan ini berfokus pada pemahaman tujuan dan kebutuhan bisnis secara menyeluruh. Proses ini bertujuan untuk merumuskan permasalahan bisnis yang ingin diselesaikan melalui pendekatan data mining serta menentukan tujuan analisis yang sesuai.

b. *Data Understanding*: Setelah tujuan bisnis ditetapkan, langkah selanjutnya adalah memahami data yang tersedia. Kegiatan dalam tahap ini meliputi pengumpulan data awal, eksplorasi struktur data, serta identifikasi kualitas data. Tujuannya adalah memperoleh wawasan awal yang dapat mendukung proses analisis.
c. *Data Preparation*: Tahap ini mencakup semua kegiatan yang berkaitan dengan pembersihan dan transformasi data agar dapat digunakan dalam proses pemodelan. Proses ini mencakup seleksi atribut, penanganan data kosong, penggabungan data, serta format ulang data.
d. *Modeling*: Pada tahap ini, teknik dan algoritma data mining diterapkan untuk membangun model analisis sesuai dengan tujuan yang telah ditentukan. Pemilihan metode modeling biasanya disesuaikan dengan jenis data dan permasalahan yang dihadapi.
e. *Evaluation*: Setelah model dibangun, dilakukan evaluasi untuk menilai seberapa baik model tersebut menjawab permasalahan bisnis. Selain mengukur performa model secara teknis, penting juga untuk meninjau kesesuaian model terhadap kebutuhan dan konteks bisnis.
f. *Deployment*: Tahap akhir adalah penerapan hasil analisis ke dalam sistem nyata. Implementasi ini dapat berupa pembuatan laporan, dashboard, atau integrasi ke dalam sistem operasional, sehingga hasilnya dapat dimanfaatkan oleh pihak terkait.



Gambar 2. 5 Tahapan Model Proses *CRISP-DM*
Sumber: Herlambang, et al., 2024

2.1.9 Metode Pengumpulan Data

Pada saat melakukan pengumpulan data penulis menggunakan 2 metode pengumpulan data sebagai berikut.

a. Observasi

   Observasi adalah teknik pengumpulan data dengan cara mengamati secara langsung objek atau fenomena yang diteliti dalam lingkungan alaminya. Observasi dapat dilakukan dengan atau tanpa alat bantu dan dapat bersifat partisipatif atau non-partisipatif (Trivaika & Senubekti, 2022).

b. Wawancara

c. Wawancara merupakan teknik pengumpulan data yang dilakukan melalui tatap muka dan Tanya jawab langsung antara pengumpul data terhadap narasumber atau sumber data (Trivaika & Senubekti, 2022). Dari penjelasan tersebut, dapat disimpulkan bahwa wawancara merupakan bentuk komunikasi dua arah yang

melibatkan dua atau lebih individu, di mana satu pihak mengajukan pertanyaan atau membahas suatu topik dengan pihak lain untuk memperoleh informasi atau pandangan.

d. Studi Pustaka

Dilaksanakan dengan cara mempelajari beberapa jurnal, penelitian maupun dokumen yang terkait atau memiliki hubungan dengan penelitian yang sedang dilaksanakan (Trivaika & Senubekti, 2022).

Berdasarkan penjelasan di atas, dapat disimpulkan bahwa studi pustaka merupakan metode penelitian yang melibatkan penelaahan dan analisis terhadap berbagai sumber tertulis, seperti literatur, jurnal ilmiah, dan buku yang relevan dengan topik penelitian. Metode ini membantu peneliti memahami perkembangan pengetahuan yang sudah ada serta menjadikannya sebagai dasar dalam penelitian mereka.

2.1.10 Flowchart

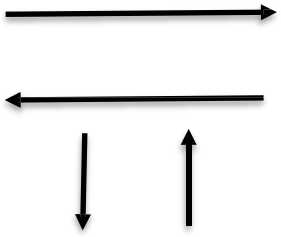
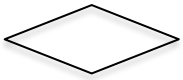
Flowchart adalah bentuk gambar/diagram yang mempunyai aliran satu atau dua arah secara sekuensial. *Flowchart* digunakan untuk merepresentasikan maupun mendesain program. Oleh karena itu *flowchart* harus bisa merepresentasikan komponen-komponen dalam Bahasa pemrograman. Baik *flowchart* maupun algoritma bisa dibuat sebelum maupun setelah pembuatan program. *Flowchart* dan *Algoritma* yang dibuat sebelum membuat program digunakan untuk mempermudah pembuat program untuk menentukan alur logika program, sedangkan yang dibuat setelah

pembuatan program digunakan untuk menjelaskan alur program kepada orang lain (Pratiwi, 2020).

Dari penjelasan tersebut, dapat disimpulkan bahwa *flowchart* merupakan metode untuk memvisualisasikan serta mendeskripsikan tahapan atau alur suatu proses dengan menggunakan simbol-simbol dan ilustrasi khusus.

Adapun Simbol-simbol *flowchart* yang digunakan seperti pada Tabel 2.3.

Tabel 2. 3 Simbol *Flowchart*

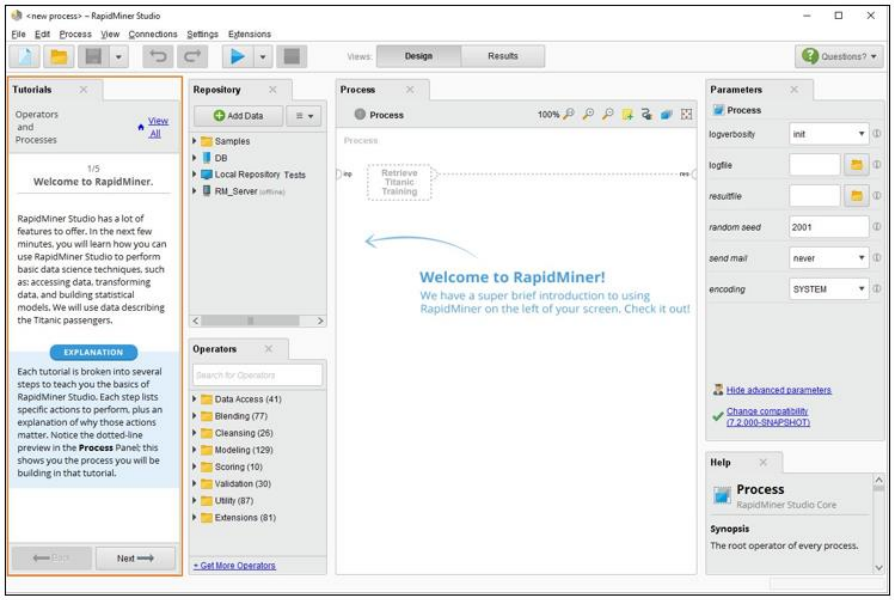
No	Simbol	Nama	Keterangan
1		Terminal	Mulai / selesai
2		*Input / Output*	Memasukan data *input / output*
3		*Process*	Proses pengolahan dan perhitungan data
4		*Flow*	Arus suatu proses
5		*Decision*	Perbandingan pernyataan, penyeleksian data yang memberikan pilihan untuk Langkah selanjutnya

Sumber: Pratiwi, 2020

2.1.11 Rapid Miner

Rapid Miner adalah *platform* perangkat lunak yang kuat untuk ilmu data dan pembelajaran mesin. Perangkat lunak ini menyediakan berbagai alat untuk persiapan data, pemodelan, evaluasi, dan implementasi. *Rapid Miner* dirancang agar mudah digunakan, memungkinkan pengguna membangun dan menguji berbagai model bahkan tanpa pengalaman pemrograman (Nahjan, et al., 2023).

Rapid Miner menawarkan antarmuka *drag-and-drop* yang mempermudah pengguna dalam membangun alur kerja untuk memproses dan menganalisis data. Platform ini mendukung berbagai sumber data, termasuk file datar, basis data, serta platform big data seperti *Hadoop* dan *Spark*. Selain itu, *Rapid Miner* dilengkapi dengan berbagai operator bawaan yang berfungsi sebagai blok bangunan dalam alur kerja, mencakup semua tahap proses *data mining*, seperti pembersihan data**,** pemilihan fitur, dan pemodelan.



Gambar 2. 6 Tampilan *Rapid Miner*
Sumber: https://docs.rapidminer.com

2.2 Penelitian Yang Relevan

Dalam penyusunan penelitian ini, penulis sedikit banyak terinspirasi dan mereferensi dari penelitian-penelitian sebelumnya. Penelitian terdahulu yang berhubungan dengan penelitian ini akan dijabarkan pada tabel 2.4.

Tabel 2. 4 Penelitian Yang Relevan

No	Penulis/Tahun	Topik Penelitian	Metode	Pembahasan	Hasil
1	Ramadhon, et al (2024).	Implementasi Algoritma Decision Tree untuk Klasifikasi Pelanggan Aktif atau Tidak Aktif pada Data Bank.	Decision Tree (CART)	Menggunakan dataset pelanggan bank untuk mengklasifikasikan apakah pelanggan aktif atau tidak berdasarkan saldo, frekuensi transaksi, dan jenis produk	Model Decision Tree menunjukkan akurasi sebesar 89%, dengan hasil yang kurang lebih sama antara entropy dan gini index

No	Penulis/Tahun	Topik Penelitian	Metode	Pembahasan	Hasil
				yang digunakan.	
2	Hartama, et al (2020).	Analisa Kepuasan Pelanggan Menggunakan Klasifikasi Data Mining	Algoritma C4.5	Mengukur kepuasan pelanggan di Indiis Cafe berdasarkan 5 variabel: Product, Price, Place, Promotion, dan Service.	- Ditemukan 15 aturan (rules) untuk klasifikasi kepuasan pelanggan. - Akurasi model mencapai 98%.
3	Fihir, et al (2023).	Klasifikasi Tingkat Kepuasan Pelanggan Kopi Kenangan Menggunakan Metode Decision Tree pada Aplikasi Kopi Kenangan	Decision Tree	Menganalisis data kepuasan pelanggan berdasarkan transaksi melalui aplikasi Kopi Kenangan menggunakan algoritma Decision Tree	Decision Tree menunjukkan akurasi 86,96%, dengan presisi prediksi kepuasan pelanggan sebesar 90% dan ketidakpuasan sebesar 66,67%
4	Faisal (2020)	Klasifikasi Data Minning Menggunakan Algoritma C4.5 Terhadap Kepuasan	Algoritma C4.5	Mengukur kepuasan pelanggan Sewa Kamera Cikarang berdasarkan 4 variabel: harga, fasilitas,	Akurasi model mencapai 95%. - Precision = 92.31%, Recall = 100%.

No	Penulis/Tahun	Topik Penelitian	Metode	Pembahasan	Hasil
		Pelanggan Sewa Kamera Cikarang		pelayanan, dan loyalitas.	
5	Wijayanti (2020).	Prediksi Loyalitas Pelanggan Menggunaka n Decision Tree	Decision Tree (C4.5)	Menggunakan dataset pelanggan untuk mengklasifikas ikan loyalitas berdasarkan jenis kelamin, jumlah transaksi, dan frekuensi kunjungan	Model Decision Tree menunjukkan akurasi sebesar 86%, dengan hasil klasifikasi pelanggan ke dalam kategori loyal dan churn

Penelitian ini memiliki beberapa perbedaan dibandingkan dengan penelitian sebelumnya yang juga menggunakan *decision tree*. Perbedaan utama terletak pada objek penelitian, variabel yang digunakan, tujuan penelitian, metode analisis, dan hasil yang diharapkan.

Berbeda dengan penelitian sebelumnya yang banyak berfokus pada industri perbankan, ritel, atau restoran, penelitian ini berfokus pada klasifikasi pelanggan di Karakatau Biliar, sebuah pusat hiburan yang menyediakan layanan biliar. Variabel yang digunakan dalam penelitian ini meliputi frekuensi kunjungan, jumlah transaksi, jenis paket yang dipilih, dan durasi bermain, yang lebih spesifik dibandingkan penelitian lain yang umumnya menggunakan faktor seperti saldo rekening, riwayat transaksi belanja, atau demografi pelanggan.

Selain itu, penelitian ini bertujuan untuk mengoptimalkan strategi pemasaran dan efisiensi operasional Karakatau Biliar melalui klasifikasi pelanggan, sedangkan penelitian sebelumnya umumnya hanya bertujuan memprediksi churn pelanggan atau mengelompokkan tingkat kepuasan pelanggan tanpa mengarah pada strategi bisnis yang lebih lanjut.

Dari segi metode, penelitian ini menggunakan decision tree dalam *RapidMiner*, sementara sebagian besar penelitian sebelumnya menggunakan decision tree dengan algoritma seperti *C4.5* atau *CART*. Perbedaan lainnya terletak pada hasil yang diharapkan.

Dengan demikian, penelitian ini memiliki keunikan karena diterapkan dalam industri hiburan yang belum banyak diteliti, menggunakan variabel spesifik yang relevan dengan pola pelanggan biliar.

BAB III

METODOLOGI PENELITIAN

3.1 Tinjauan Umum

Peneliti membuat penelitian klasfikasi pelanggan dengan menggunakan Decision Tree ini berlokasi di Krakatau Billiar. Karakatau Biliar adalah salah satu pusat hiburan yang menyediakan fasilitas permainan biliar bagi berbagai kalangan, mulai dari pemula hingga pemain *profesional*. Berlokasi di Jl. Krakatau, Palangka, Kec. Jekan Raya, Kota Palangka Raya, Kalimantan Tengah, bisnis ini beroperasi dengan menyediakan layanan sewa meja biliar, penjualan makanan dan minuman. Dengan meningkatnya popularitas biliar sebagai olahraga rekreatif dan kompetitif, Karakatau Biliar terus berupaya meningkatkan layanan untuk memenuhi kebutuhan pelanggan yang beragam.



Gambar 3. 1 Krakatau Billiard

3.2 Jenis Penelitian

Jenis penelitian yang digunakan dalam penelitian ini adalah penelitian *kuantitatif*, karena berfokus pada pengolahan dan analisis *data numerik* secara sistematis.

Penelitian ini bertujuan untuk mengklasifikasikan pelanggan di Krakatau Billiard menggunakan metode *Decision Tree*.

Melalui klasifikasi ini, penelitian berupaya mengidentifikasi pola kunjungan, preferensi layanan, dan tingkat loyalitas pelanggan. Informasi tersebut diharapkan dapat membantu pengelola Krakatau Billiard dalam merancang strategi pemasaran yang lebih efektif, menyusun promosi yang sesuai dengan kebutuhan pelanggan, serta mengoptimalkan alokasi sumber daya.

Dengan pendekatan berbasis data ini, diharapkan Krakatau Billiard dapat meningkatkan efisiensi operasional dan mempertahankan pelanggan melalui segmentasi yang lebih terarah.

3.3 Desain Penelitian

Dalam melakukan proses klasifikasi pelanggan menggunakan *Decision Tree* , peneliti menggunakan 5 tahapan *CRISP DM* dari 6 tahapan yang ada, 5 tahapan tersebut adalah:

a. *Business Understanding*

   Tahap ini bertujuan untuk memahami tujuan bisnis dari penelitian, yaitu mengklasifikasikan pelanggan Krakatau Billiard berdasarkan pola kunjungan dan karakteristik lainnya. Hasil klasifikasi ini diharapkan dapat digunakan oleh manajemen untuk menyusun strategi pemasaran yang lebih tepat sasaran, meningkatkan layanan, dan mendorong loyalitas pelanggan.

b. *Preprocessing Data*

   Pada tahap ini, peneliti melakukan pengumpulan dan eksplorasi awal terhadap data pelanggan yang tersedia. Data yang digunakan meliputi frekuensi

kunjungan, durasi bermain, dan total transaksi. Proses ini mencakup identifikasi kualitas data, pengecekan nilai kosong atau tidak konsisten, serta pemahaman struktur data.

c. *Data Preparation*

Data yang telah dikumpulkan kemudian dipersiapkan untuk proses pemodelan. Tahap ini meliputi pembersihan data, konversi data kategorikal menjadi format numerik (jika diperlukan), penghapusan data duplikat, serta pembagian data menjadi data latih dan data uji (misalnya dengan rasio 80:20). Tahapan ini sangat penting untuk memastikan kualitas input pada proses klasifikasi.

d. *Modeling*

Pada tahap ini, algoritma *Decision Tree* digunakan untuk membangun model klasifikasi pelanggan. Proses modeling dilakukan menggunakan perangkat lunak *RapidMiner*, yang menyediakan antarmuka visual *(drag and drop*) untuk membangun *workflow* klasifikasi tanpa perlu menulis kode pemrograman. Model dilatih menggunakan data latih dan diuji menggunakan data uji untuk mengevaluasi performa klasifikasinya dalam mengelompokkan pelanggan ke dalam kategori tertentu

e. *Evaluation*

Model yang telah dibangun dievaluasi menggunakan metrik seperti akurasi, *precision*, *recall*, dan *F1-score*. Tujuan dari tahap ini adalah untuk memastikan bahwa model memiliki performa yang baik dalam

mengelompokkan pelanggan secara tepat dan dapat digunakan sebagai dasar pengambilan keputusan bisnis.

3.4 Instrumen Penelitian

Peneliti memiliki peran sebagai instrumen dalam pengumpulan data. Dalam hal ini, peneliti melakukan wawancara dengan Bapak Feri, selaku pemilik dan pengelola Krakatau Biliar. Wawancara ini bertujuan untuk memahami pola kunjungan pelanggan, serta tantangan dalam operasional bisnis.

Dalam proses pengumpulan data, peneliti juga menggunakan perangkat bantu. Perangkat bantu yang digunakan meliputi lembar wawancara tidak terstruktur yang tetap menggunakan bahasa formal untuk memastikan komunikasi yang jelas antara kedua belah pihak. Selain itu, handphone digunakan sebagai alat dokumentasi selama wawancara berlangsung.

a. Metode Observasi

   Peneliti melakukan observasi dengan mendatangi secara langsung lokasi penelitian di Krakatau Billiard di Jl. Krakatau, Palangka, Kec. Jekan Raya, Kota Palangka Raya, Kalimantan Tengah. peneliti mengamati secara langsung aktivitas operasional tempat usaha, mencatat perilaku dan pola kunjungan pelanggan, serta berinteraksi dengan staf untuk memperoleh informasi mengenai layanan, sistem pencatatan transaksi, dan kebijakan promosi. Peneliti juga mendokumentasikan kondisi fisik tempat, fasilitas yang tersedia, serta jam operasional untuk mendukung pemahaman terhadap data yang dianalisis.

b. Metode Wawancara

Peneliti melakukan wawancara dengan Bapak Feri, selaku penanggung jawab di Krakatau Biliar, untuk memperoleh informasi terkait pola kunjungan pelanggan, serta tantangan dalam pengelolaan bisnis biliar.

c. Studi Pustaka

Dalam penelitian ini, penulis melakukan studi pustaka dengan mengumpulkan literatur, jurnal, buku, dan dokumen yang berhubungan dengan klasifikasi pelanggan, data mining, dan penggunaan algoritma *Decision Tree* dalam analisis bisnis.

3.5 Analisis Kebutuhan

Dalam penelitian ini, penulis membutuhkan data dan sejumlah perangkat yang diperlukan untuk mendukung proses klasifikasi pelanggan menggunakan *Decision Tree* di *RapidMiner*. Kebutuhan tersebut sangat penting untuk memastikan hasil klasifikasi akurat dan dapat digunakan untuk analisis lebih lanjut. Kebutuhan tersebut adalah sebagai berikut:

3.5.1. Analisis Data

Dalam proses klasfikasi menggunakan *Decision Tree* terdapat beberapa data yang dibutuhkan oleh penulis. Kebutuhan tersebut adalah sebagai berikut:

a. Variabel Klasifikasi

Dalam proses klasifikasi pelanggan menggunakan *Decision Tree* di RapidMiner, terdapat beberapa variabel yang digunakan untuk mengelompokkan pelanggan berdasarkan pola kunjungan dan transaksi

mereka. Variabel yang digunakan dalam penelitian ini akan dijabarkan pada tabel 3.2.

Tabel 3. 1 Variabel Klasifikasi

Variabel	Keterangan
Frekuensi Kunjungan	Jumlah kunjungan pelanggan ke Krakatau Biliar dalam periode 1 tahun.
Durasi Bermain	Lama waktu rata-rata pelanggan menghabiskan waktu bermain dalam satu sesi.
Jumlah Transaksi	Total pengeluaran pelanggan dalam satu tahun berdasarkan layanan yang digunakan.

b. Klasifikasi Pelanggan

Pelanggan Krakatau Biliar akan diklasifikasikan menjadi beberapa kategori utama berdasarkan frekuensi kunjungan, durasi bermain, dan jumlah transaksi. Adapun kategori pelanggan akan dijabarkan pada tabel 3.2.

Tabel 3. 2 Klasifikasi Pelanggan

Kategori	Frekuensi	Durasi (jam)	Transaksi Total (Rp)
Pelanggan Baru	1–3 kali	< 5 jam	< 250.000
Pelanggan Tetap	4–12 kali	5–25 jam	250.000 – 2.000.000
Pelanggan Aktif	> 12 kali	> 25 jam	> 2.000.000

c. Dataset

Dataset yang digunakan dalam penelitian ini adalah data kunjungan pelanggan di Krakatau Biliar dalam periode 1 tahun, yaitu dari 1 Maret 2024 hingga 28 Februari 2025 yang didapat dari pihak Krakatau Biliar. *Dataset* ini berisi informasi mengenai perilaku pelanggan berdasarkan transaksi dan pola kunjungan mereka.

Format *Dataset:*

1) *Dataset* disimpan dalam format *CSV (Comma-Separated Values)* agar mudah diolah menggunakan *RapidMiner*. Setiap baris dalam dataset merepresentasikan satu transaksi pelanggan, dengan atribut-atribut berikut:

Tabel 3. 3 Format Dataset

Nama Kolom	**Keterangan**
ID Pelanggan	Nomor unik yang mengidentifikasi setiap pelanggan.
Tanggal Kunjungan	Tanggal pelanggan berkunjung ke Krakatau Biliar.
Frekuensi Kunjungan	Jumlah kunjungan pelanggan dalam periode tertentu (per bulan).
Durasi Bermain	Lama waktu pelanggan bermain dalam satu sesi (dalam jam).
Jumlah Transaksi	Total pengeluaran pelanggan dalam satu kunjungan (dalam Rupiah).
Kategori Pelanggan	Hasil klasifikasi pelanggan berdasarkan pola kunjungan **(**Baru, Tetap, Aktif**).**

3.5.2. Analisis Kebutuhan Perangkat Keras

Pada penelitian ini penulis menggunakan perangkat keras *(Hardware)* komputer dengan spesifikasi yaitu:

Tabel 3. 4 Perangkat Keras

Spesifikasi	**Keterangan**
Type	*Laptop asus notebook S410U*
Processor	*Intel Core i5 8250U (1.60 Ghz)*
Memory	*16 GB (Ram)*
Hardisk	*1 TB SATA SSD*

3.5.3. Analisis Kebutuhan Perangkat Lunak

Beberapa perangkat lunak yang akan digunakan penulis pada penelitian ini:

Tabel 3. 5 Perangkat Lunak.

Perangkat Lunak	**Keterangan**
Windows 10 64-bit	Sebagai Sistem Operasi

Perangkat Lunak	Keterangan
Microsoft Excel	Digunakan untuk mengelola *Dataset*
Rapid Miner	Digunakan untuk melakukan klasifikasi

3.5.4. Analisis Kelemahan

Meskipun penelitian ini memberikan wawasan yang berguna dalam mengklasifikasikan pelanggan Krakatau Biliar menggunakan *Decision Tree*, terdapat beberapa keterbatasan. Salah satu kelemahan utama adalah jangka waktu pengambilan data yang terbatas, yaitu hanya dalam periode 1 tahun. Hal ini dapat menyebabkan pola klasifikasi yang dihasilkan belum sepenuhnya mencerminkan kebiasaan pelanggan dalam jangka panjang, terutama jika ada faktor musiman atau promosi khusus yang memengaruhi jumlah kunjungan.

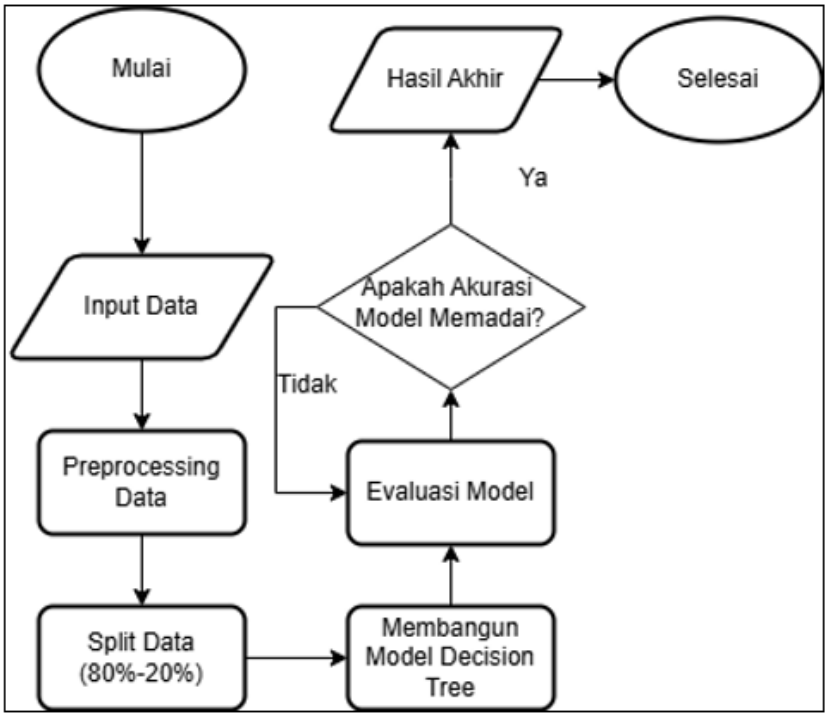
Selain itu, penelitian ini hanya menggunakan *Decision Tree* sebagai metode klasifikasi tanpa membandingkannya dengan algoritma lain seperti *Random Forest* atau *Naïve Bayes*, yang mungkin dapat memberikan akurasi lebih tinggi atau hasil yang lebih stabil. Meskipun demikian, penelitian ini tetap valid untuk dilanjutkan karena dapat memberikan gambaran awal yang cukup jelas mengenai pola pelanggan, yang nantinya bisa dikembangkan lebih lanjut dalam penelitian berikutnya.

3.6 Desain

3.6.1. Desain Proses

Dalam perancangan proses, penulis memanfaatkan *Flowchart* untuk menjelaskan tahapan klasifikasi pelanggan di Krakatau Biliar menggunakan metode *Decision Tree* di *RapidMiner*. *Flowchart* ini menggambarkan alur dari pengumpulan data pelanggan, preprocessing data, pembuatan *model Decision Tree*,

hingga evaluasi hasil klasifikasi. Penjelasan dan ilustrasi lebih lanjut mengenai perancangan proses ini akan dijabarkan pada Gambar 3.3.



Gambar 3. 2 Desain Proses

Berikut merupakan penjelasan terkait desain proses pada gambar 3.2:

a. Mulai (*Start*)

   Proses klasifikasi pelanggan dimulai. Ini merupakan titik awal sebelum data pelanggan diolah dan dianalisis.

b. Input Data Pelanggan

   Pada tahap ini, data pelanggan dikumpulkan dari sistem transaksi Krakatau Biliar. Data ini mencakup informasi seperti frekuensi kunjungan, durasi bermain, jenis layanan yang digunakan, dan jumlah transaksi pelanggan. Data ini kemudian diimpor ke *RapidMiner* dalam format *CSV* agar dapat diproses lebih lanjut.

c. *Preprocessing* Data

Sebelum model *Decision Tree* dibuat, data harus dibersihkan dan disiapkan agar hasil klasifikasi lebih akurat. Langkah-langkah dalam *preprocessing* data meliputi:

1) Mengatasi Data Kosong → Jika ada data yang hilang, operator *Replace*
2) Missing Values digunakan untuk mengisi atau menghapus data kosong.
3) Memilih Variabel Penting → Hanya variabel yang berpengaruh dalam klasifikasi yang dipilih menggunakan *Select Attributes* di *RapidMiner*.
4) Menstandarisasi Format Data → Pastikan bahwa semua data telah sesuai dengan format yang dapat diproses oleh model *Decision Tree.*

d. *Split Data* (80%-20%)

Setelah data bersih, langkah berikutnya adalah membagi dataset menjadi dua bagian:

1) 80% digunakan sebagai data training → Digunakan untuk melatih model *Decision Tree.*
2) 20% digunakan sebagai data *testing* → Digunakan untuk menguji akurasi model. Pembagian ini dilakukan menggunakan operator Split Data di *RapidMiner* untuk memastikan model dapat menggeneralisasi data dengan baik.

e. Membangun Model *Decision Tree*

Pada tahap ini, model *Decision Tree* dibangun menggunakan algoritma *C4.5* melalui operator *Decision Tree* di RapidMiner. *C4.5* dipilih karena memiliki keunggulan dibanding *ID3*, terutama dalam menangani atribut numerik, *missing values*, serta penerapan *pruning* otomatis untuk mencegah *overfitting.* Model ini dilatih menggunakan data training, dengan memperhatikan parameter berikut:

1) *Criterion (Gini Index atau Gain Ratio)* → *C4.5* menggunakan Gain
2) *Ratio* sebagai dasar pemilihan atribut terbaik pada setiap node.
3) *Max Depth* → Menentukan batas maksimal kedalaman pohon, untuk mengendalikan kompleksitas model.
4) *Pruning* → *C4.5* secara otomatis memangkas cabang yang tidak signifikan untuk meningkatkan generalisasi model.

f. Evaluasi Model

Setelah model *Decision Tree* terbentuk, langkah berikutnya adalah mengevaluasi performanya menggunakan data testing. Evaluasi dilakukan dengan menggunakan operator *Performance* (*Classification***)** untuk mendapatkan metrik berikut:

1. *Confusion Matrix* → Menunjukkan jumlah prediksi benar dan salah dalam klasifikasi pelanggan.
2. *Akurasi* → Mengukur seberapa sering model membuat prediksi yang benar.
3. *Precision & Recall* → Mengukur kualitas klasifikasi untuk setiap kategori pelanggan.

4. *F1-Score* → Menyeimbangkan *Precision* dan *Recall* untuk memastikan model bekerja optimal.

g. Apakah Akurasi Model Memadai?

Pada tahap ini, model diuji untuk melihat apakah akurasi yang diperoleh sudah mencukupi:

1. Jika Ya → Lanjut ke tahap interpretasi hasil dan klasifikasi pelanggan.
2. Jika Tidak → Model perlu dioptimalkan dengan *pruning Decision Tree*, menyesuaikan parameter, atau menambah data training.

3. Hasil Klasifikasi Pelanggan

Setelah model dinyatakan memiliki akurasi yang baik, pelanggan dapat dikelompokkan ke dalam kategori pelanggan baru, pelanggan tetap, dan pelanggan aktif berdasarkan pola transaksi pelanggan. Hasil klasifikasi ini dapat digunakan untuk analisis lebih lanjut atau sebagai dasar untuk meningkatkan strategi layanan.

4. Selesai (*End*)

Proses klasifikasi selesai. Hasil analisis dapat digunakan untuk membantu manajemen Krakatau Biliar memahami pola pelanggan dengan lebih baik.

3.7 Jadwal Penelitian

Jadwal penelitian dibuat untuk mengatur waktu penelitian sehingga tahapan-tahapan dapat dilakukan dan selesai tepat waktu. Oleh karena itu, penulis membuat jadwal penelitian agar tahapan-tahapan penelitian dapat dilakukan sesuai waktu yang ditentukan. Tabel 3.6 menunjukkan jadwal penelitian Tugas Akhir.

Tabel 3. 6 Jadwal Penelitian

No	Tahapan	Maret 2025				April 2025				Mei 2025				Juni 2025				Juli 2025				Agustus 2025			
		1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
1	*Business Understanding*																								
2	*Preprocessing data*																								
3	*Data Preperation*																								
4	*Modeling*																								
6	*Evaluation*																								

BAB IV

HASIL DAN PEMBAHASAN

4.1 Hasil

Hasil dari penelitian ini adalah implementasi data yang telah dikumpulkan menggunakan *RapidMiner Studio*, serta serta klasifikasi pelanggan pada Billiar menggunakan metode *Decision Tree*.

4.1.1 Implementasi

Dalam bab 4 mengenai implementasi, penulis mendeskripsikan proses penerapan algoritma *Decision Tree (C4.5)* dalam *platform RapidMiner* untuk mengklasifikasikan pelanggan di Krakatau Biliar. Implementasi ini melibatkan beberapa langkah penting dalam melakukan analisis data. Tahap pertama adalah memuat data transaksi pelanggan periode 2024-2025 ke dalam *RapidMiner*, diikuti oleh tahap pra-pemrosesan data. Pada tahap ini, kami menetapkan peran atribut *(Set Role)* dengan menentukan "Kategori Pelanggan" sebagai label target, serta menyeleksi fitur yang paling berpengaruh seperti Frekuensi Kunjungan dan Durasi Bermain untuk mempersiapkannya bagi analisis lebih lanjut.

Selanjutnya, algoritma *Decision Tree* diterapkan, dengan menggunakan kriteria *Gain Ratio* untuk membangun model klasifikasi yang optimal. Proses pembentukan model dilakukan dengan memecah data *(splitting)* berdasarkan nilai atribut terbaik, sehingga terbentuk struktur pohon keputusan yang memuat aturan-aturan *(rules)* logis. Akhirnya, hasil model dievaluasi menggunakan *Cross Validation* dan visualisasi pohon

keputusan diinterpretasikan untuk mendapatkan pemahaman yang lebih dalam mengenai karakteristik pelanggan Baru, Tetap, dan Aktif.

a. Pengumpulan Data

Tahap pengumpulan data dalam penelitian ini dilakukan dengan memanfaatkan data sekunder yang telah tersedia dan disediakan secara resmi oleh pihak manajemen Krakatau Biliar. Dataset yang digunakan merupakan rekam jejak historis aktivitas pelanggan selama periode satu tahun penuh, terhitung mulai dari 1 Maret 2024 hingga 28 Februari 2025. Data mentah tersebut diambil dari arsip operasional harian yang mencakup informasi detail mengenai identitas pelanggan, frekuensi kedatangan, durasi bermain, serta riwayat transaksi, yang kemudian diekstraksi dan dikompilasi ke dalam format digital untuk dipersiapkan sebagai bahan analisis utama.

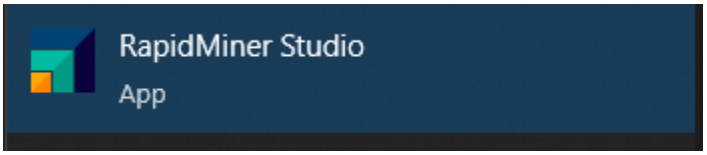
b. Pemilihan Data

Tahap ini bertujuan menentukan variabel mana yang paling relevan untuk menjawab tujuan penelitian, yaitu mengklasifikasikan pelanggan berdasarkan intensitas bermain. Penulis memilih variabel Frekuensi Kunjungan dan Durasi Bermain sebagai atribut utama, serta mengabaikan variabel nominal transaksi agar model berfokus pada pola perilaku kunjungan, bukan sekadar besaran pembayaran.

c. Import Data (Pra-pemrosesan data)

1) Buka *RapidMiner Studio*

Pada gambar 4.1 memulai konfigurasi dan analisis data dalam penelitian ini menggunakan *RapidMiner*, langkah awal yang perlu dilakukan adalah membuka *RapidMiner Studio*. Dengan membuka *RapidMiner Studio*, akan mendapatkan akses ke berbagai alat dan fitur yang diperlukan untuk *impor*, *preprocessing*, analisis, dan visualisasi data secara efisien. Setelah aplikasi terbuka, dapat melanjutkan dengan langkah-langkah berikutnya, seperti mengimpor dataset, melakukan *preprocessing data*, dan menerapkan metode *Decision Tree* untuk menentukan klasifikasi pelanggan. Langkah pertama ini merupakan fondasi yang penting sebelum melakukan langkah-langkah analisis yang lebih kompleks. Dapat dilihat pada gambar 4.1.

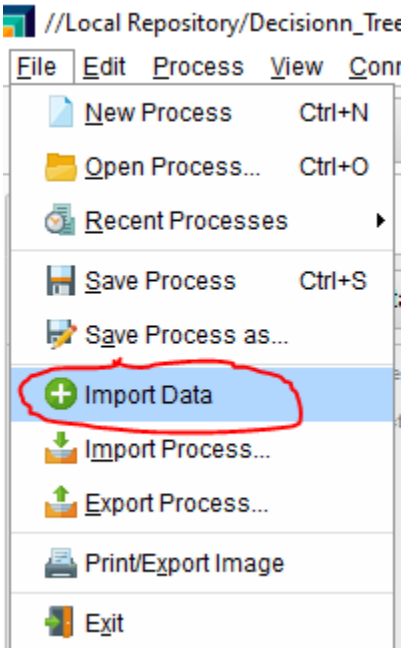


Gambar 4. 1 *RapidMiner Studio*

2) Menu *File*

Pada Menu *File* melakukan konfigurasi dan analisis data menggunakan *RapidMiner* dalam penelitian ini, langkah pertama yang perlu dilakukan adalah membuka *RapidMiner Studio*. Hal ini dilakukan untuk memperoleh akses ke berbagai fungsi dan alat analisis yang disediakan oleh *RapidMiner*. Setelah *RapidMiner Studio* terbuka, perhatikan bahwa menu utama *RapidMiner* terletak di pojok kiri atas layar.

dapat membuka menu ini dengan mengarahkan kursor ke sana. Di dalam menu utama tersebut, pilihan "*File*" menyediakan opsi untuk mengelola *file* dan data. Dari *menu* "*File*", dapat memilih opsi "*Import Data*" untuk memulai proses impor *dataset* yang akan digunakan dalam klasifikasi pelanggan. Langkah ini sangat penting karena dataset yang tepat dan relevan diperlukan untuk melakukan analisis dengan akurat dan efektif. Dapat dilihat pada gambar 4.2.

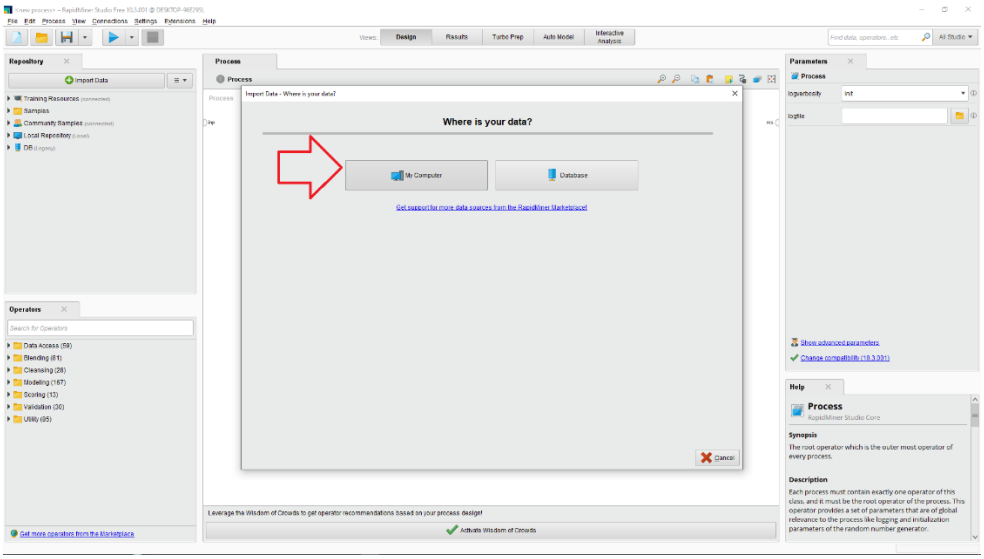


Gambar 4. 2 *Import Data RapidMiner Studio*

3) Import Data

Pada *Import Data* memulai analisis data di *RapidMiner*, langkah awalnya adalah dengan mengimpor dataset yang akan digunakan. Setelah *RapidMiner Studio* terbuka, kita perlu menuju ke menu utama di pojok kiri atas layar. Di sana, kita akan melihat opsi "*File*" yang dapat kita pilih. Setelah memilih "*File*", langkah selanjutnya adalah

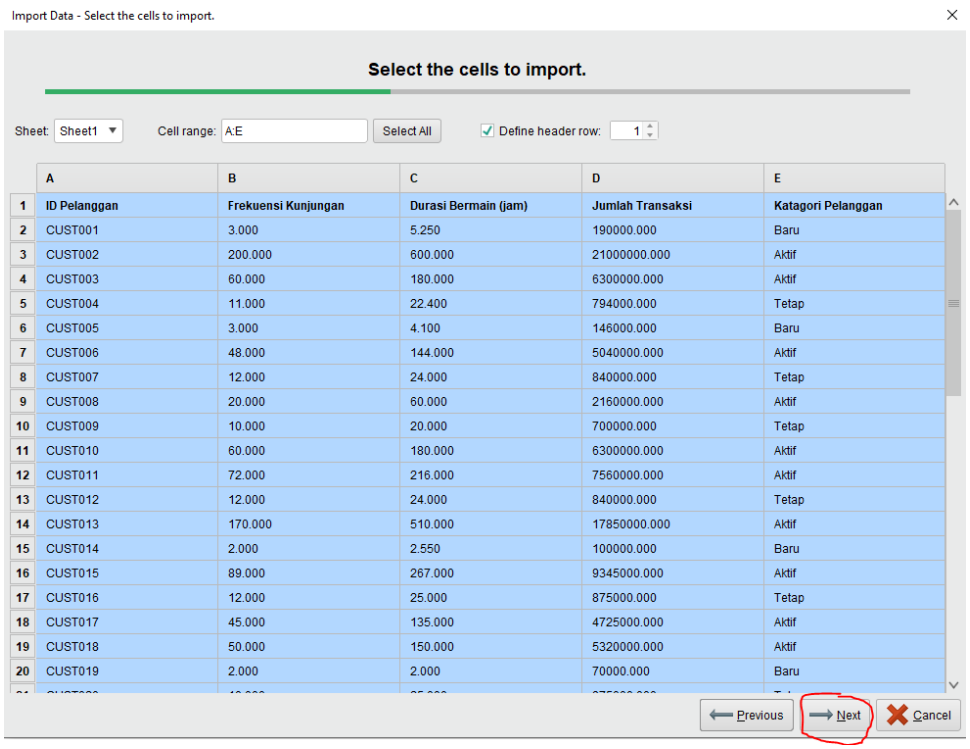
memilih opsi *"Import Data"*. Proses ini akan membuka jendela di mana kita bisa memilih lokasi penyimpanan *dataset* kita. Pilih opsi *"My Computer"* atau "Penyimpanan Lokal" untuk mengarahkan *RapidMiner* ke lokasi di mana *file dataset* kita disimpan. Setelah memilih file dataset yang diinginkan, *RapidMiner* akan memuat dataset tersebut ke dalam *workspace* untuk digunakan dalam analisis data selanjutnya. Langkah ini krusial karena memastikan *RapidMiner* memiliki akses ke data yang diperlukan untuk analisis yang tepat dan efektif. Dapat dilihat pada gambar 4.3.



Gambar 4. 3 Lokasi *File*

Pada gambar 4.4 menunjukkan tahap *Import Data*, di mana Penulis sedang menyeleksi seluruh dataset pelanggan Krakatau Biliar yang terdiri dari atribut *ID*, Frekuensi Kunjungan, Durasi Bermain, dan Jumlah Transaksi sebagai fitur penentu, serta kolom Katagori Pelanggan (Baru, Aktif, Tetap) sebagai target label yang akan dipelajari oleh komputer. Tombol *Next* yang dilingkari merah berfungsi untuk mengonfirmasi pilihan sel tersebut agar data dimuat masuk ke dalam sistem

(*workspace*); setelah diklik, langkah berikutnya biasanya adalah mengecek tipe data (memastikan kolom 'Katagori' terbaca sebagai data *categorical*) lalu melanjutkannya ke proses *Training Model* menggunakan algoritma *Decision Tree* agar sistem bisa menciptakan pola aturan ("pohon") untuk mengklasifikasikan pelanggan secara otomatis. Dapat dilihat pada gambar 4.4.

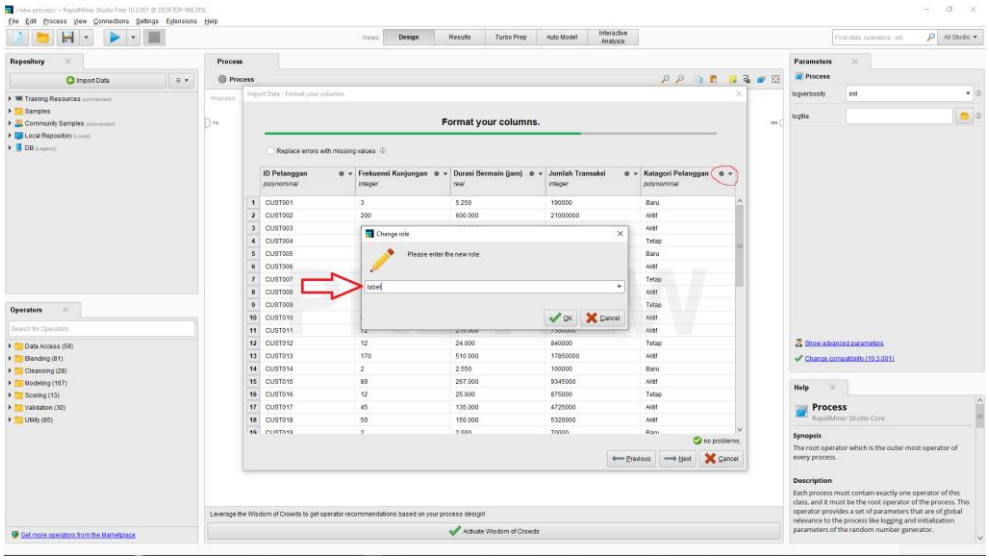


	A	B	C	D	E
1	ID Pelanggan	Frekuensi Kunjungan	Durasi Bermain (jam)	Jumlah Transaksi	Katagori Pelanggan
2	CUST001	3.000	5.250	190000.000	Baru
3	CUST002	200.000	600.000	21000000.000	Aktif
4	CUST003	60.000	180.000	6300000.000	Aktif
5	CUST004	11.000	22.400	794000.000	Tetap
6	CUST005	3.000	4.100	146000.000	Baru
7	CUST006	48.000	144.000	5040000.000	Aktif
8	CUST007	12.000	24.000	840000.000	Tetap
9	CUST008	20.000	60.000	2160000.000	Aktif
10	CUST009	10.000	20.000	700000.000	Tetap
11	CUST010	60.000	180.000	6300000.000	Aktif
12	CUST011	72.000	216.000	7560000.000	Aktif
13	CUST012	12.000	24.000	840000.000	Tetap
14	CUST013	170.000	510.000	17850000.000	Aktif
15	CUST014	2.000	2.550	100000.000	Baru
16	CUST015	89.000	267.000	9345000.000	Aktif
17	CUST016	12.000	25.000	875000.000	Tetap
18	CUST017	45.000	135.000	4725000.000	Aktif
19	CUST018	50.000	150.000	5320000.000	Aktif
20	CUST019	2.000	2.000	70000.000	Baru

Gambar 4. 4 *File Dataset*

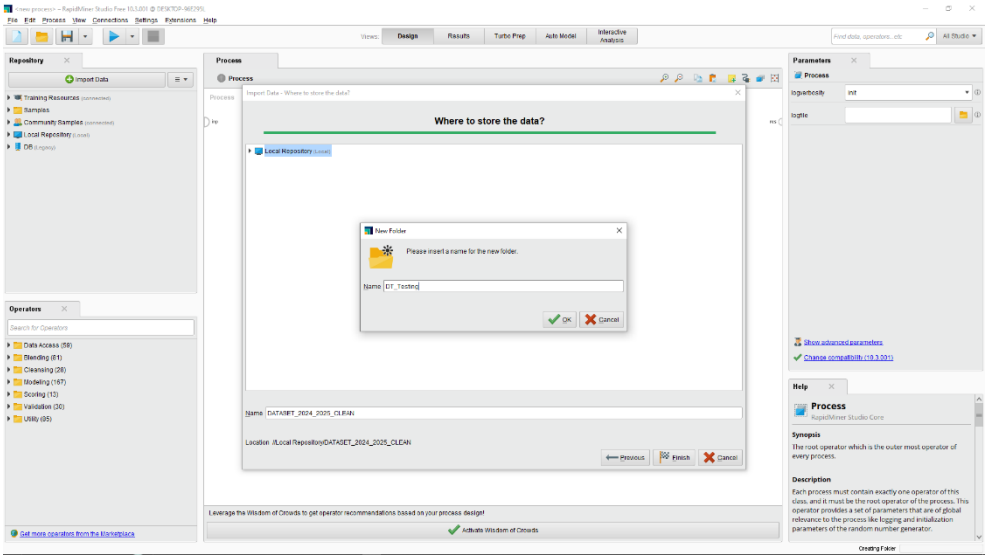
Pada gambar 4.5 menunjukkan tahap krusial di *RapidMiner* yang disebut Set *Role*, di mana Penulis sedang mengubah status kolom Katagori Pelanggan menjadi Label (target prediksi) untuk memberi tahu algoritma bahwa kolom inilah "kunci jawaban" yang harus diprediksi, sementara kolom lainnya (seperti Frekuensi dan Durasi) hanya dianggap sebagai atribut pendukung biasa. Setelah Penulis mengetik

"label" dan menekan *OK/Next*, data ini akan siap digunakan dalam *workspace* utama untuk dihubungkan dengan *operator Decision Tree*, sehingga komputer bisa mulai memproses data tersebut (tahap *training*) dan menghasilkan pohon keputusan yang memetakan pola pelanggan secara otomatis. Dapat dilihat pada gambar 4.5.



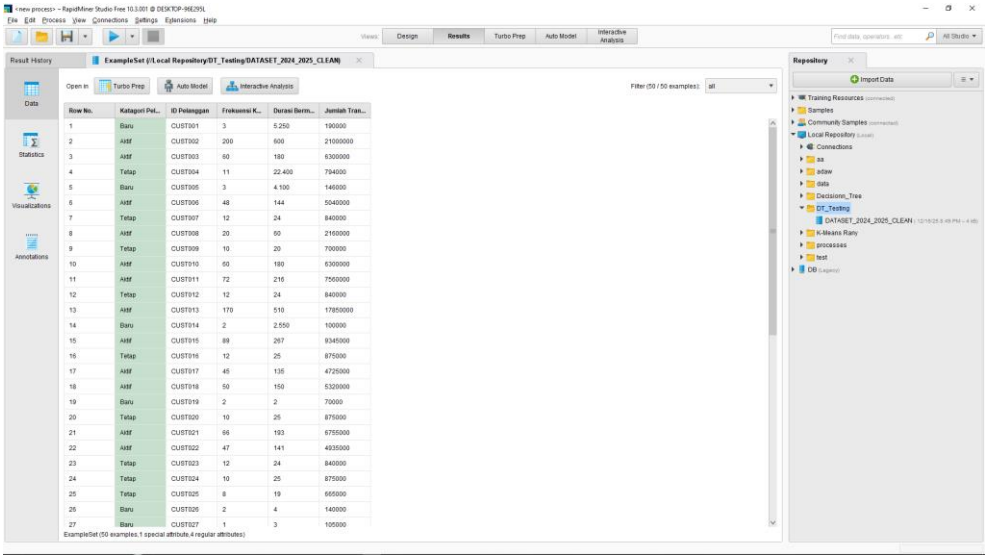
Gambar 4. 5 Tambahkan Label

Pada gambar 4.6 menunjukkan tahap Finalisasi Penyimpanan *Data (Storing Data)* di *RapidMiner*, di mana Penulis menentukan lokasi "rumah" bagi dataset yang baru saja diimpor agar bisa digunakan berulang kali tanpa harus upload ulang. Di situ terlihat Penulis sedang membuat *folder* pengelompokan baru bernama *DT_Testing* di dalam *Local Repository* untuk menyimpan file *DATASET_2024_2025_CLEAN*, yang bertujuan agar manajemen file penelitianmu tetap rapi dan terstruktur. Setelah Penulis klik tombol *Finish*, dataset tersebut akan resmi masuk ke *panel Repository* di sebelah kiri layar dan siap untuk ditarik (*drag* & *drop*) ke *canvas* utama guna memulai proses pembuatan *model Decision Tree*. Dapat dilihat pada gambar 4.6.



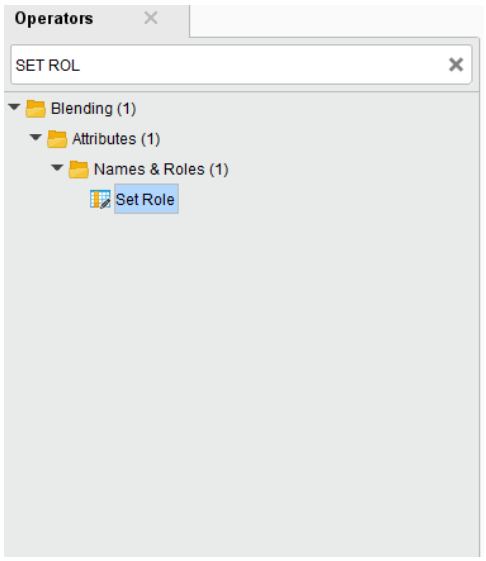
Gambar 4. 6 Simpan Dataset

Pada gambar 4.7 menampilkan *ExampleSet* di *tab Results*, yang mengonfirmasi bahwa dataset pelanggan Krakatau Biliar telah berhasil dimuat sepenuhnya ke dalam *RapidMiner* dan siap diolah. Indikator paling penting di sini adalah kolom Katagori Pelanggan yang kini berwarna hijau, menandakan bahwa atribut tersebut sudah sukses dikenali sistem sebagai Label (*target*) untuk prediksi, berbeda dengan kolom lain yang berwarna putih (atribut biasa/reguler). Keterangan di pojok kiri bawah "1 *special attribute*" semakin mempertegas hal ini, sehingga langkah selanjutnya Penulis tinggal berpindah kembali ke tab *Design* untuk melakukan *Split Data* (membagi data latih dan uji) atau langsung menarik operator algoritma *Decision Tree* untuk memulai pembuatan model. Dapat dilihat pada gambar 4.7.



Gambar 4. 7 Hasil *Dataset*

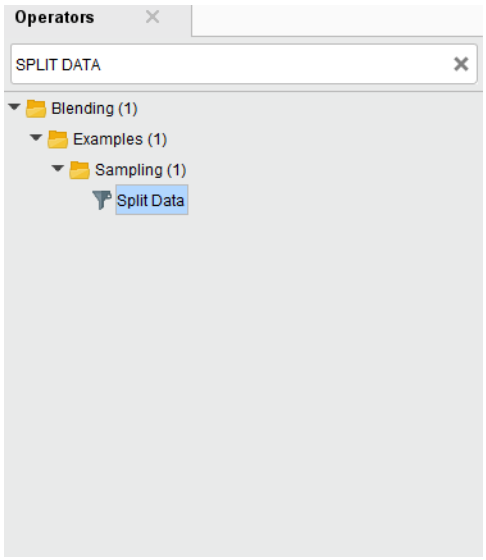
Pada gambar 4.8 menampilkan panel *Operators* di *RapidMiner*, di mana Penulis sedang melakukan pencarian (*searching*) untuk menemukan alat bernama *Set Role* dengan cepat. Di situ terlihat Penulis mengetik "*SET ROL*" pada kolom pencarian, dan sistem langsung memfilter ribuan operator untuk memunculkan operator Set Role yang berada di bawah kategori *Blending* > *Attributes* > *Names & Roles*. Langkah ini penting karena operator tersebutlah yang nantinya harus Penulis tarik (*drag and drop*) ke lembar kerja utama (*Process Canvas*) untuk dihubungkan dengan datasetmu, guna menegaskan secara teknis bahwa kolom "Katagori Pelanggan" berfungsi sebagai target prediksi (*Label*) sebelum *data* masuk ke *algoritma Decision Tree*. Dapat dilihat pada gambar 4.8.



Gambar 4. 8 *Operator Set Role*

4) Split Data

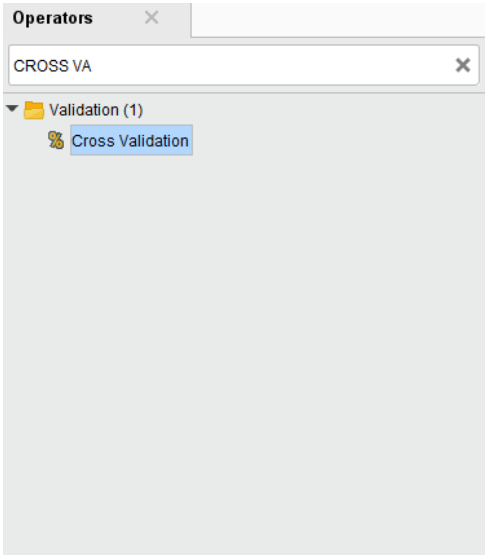
Pada gambar 4.9 menampilkan panel *Operators* di mana Penulis sedang mencari *operator Split Data*. Dalam penelitian *machine learning* seperti klasifikasi pelanggan ini, langkah tersebut sangat vital karena berfungsi untuk memecah total data Penulis menjadi dua bagian terpisah: Data Latih (*Training*) untuk mengajarkan pola pada model, dan Data Uji (*Testing*) untuk menguji seberapa akurat tebakan model tersebut nantinya. Di gambar terlihat operator ini berada di bawah kategori *Blending* > *Examples* > *Sampling*, dan setelah ditemukan, Penulis perlu menariknya ke *canvas* utama untuk mengatur rasio pembagiannya (biasanya 70:30 atau 80:20) agar hasil evaluasi model valid dan objektif. Dapat dilihat pada gambar 4.9.



Gambar 4. 9 *Operator Split Data*

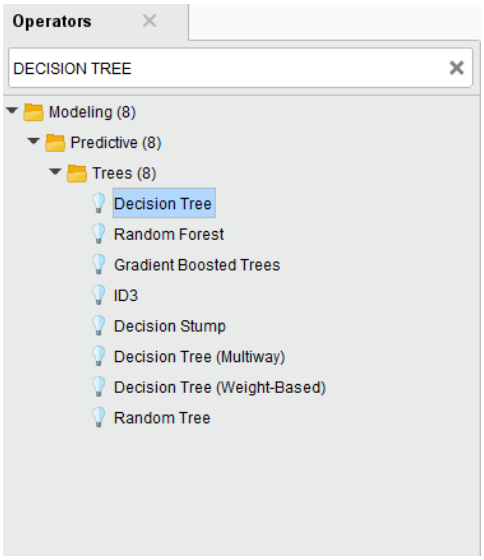
d. Penerapan Algoritma *Decision Tree* (*Modeling*)

Pada gambar 4.10 memperlihatkan panel *Operators* di mana penulis melakukan pencarian operator *Cross Validation*. Metode evaluasi ini dipilih karena lebih terpercaya dibandingkan metode *Split Data* biasa. *Operator* tersebut berfungsi untuk membagi data menjadi beberapa bagian (misalnya 10-*fold*) dan melakukan pengujian berulang secara otomatis, sehingga hasil akurasi klasifikasi pelanggan Krakatau Biliar yang dihasilkan menjadi lebih valid dan objektif. Setelah operator ditemukan pada kategori *Validation*, langkah selanjutnya adalah menariknya ke lembar kerja utama (*canvas*), kemudian penulis dapat masuk ke dalam sub-proses *operator* tersebut untuk menyusun alur training dan testing agar model *Decision Tree* dapat diuji secara menyeluruh. Dapat dilihat pada gambar 4.10.



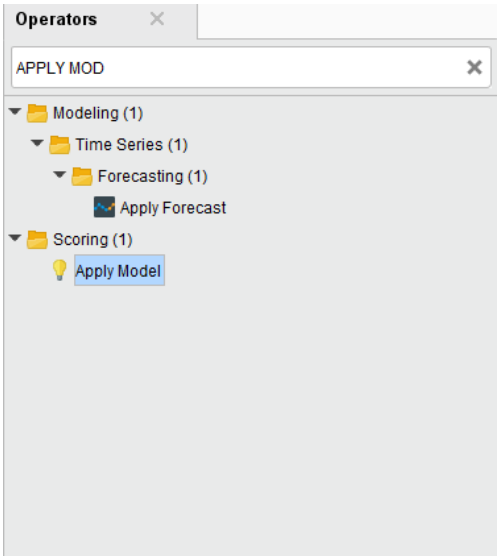
Gambar 4. 10 *Operator Cross Validation*

Pada gambar 4.11 menampilkan langkah pemilihan algoritma utama dalam penelitian ini, di mana penulis mencari operator *Decision Tree* melalui kolom pencarian pada panel *Operators*. Terlihat bahwa operator tersebut berada di bawah kategori *Modeling* > *Predictive* > *Trees* bersama dengan varian algoritma pohon lainnya. Pemilihan operator ini merupakan tahap inti dari pemodelan, karena operator inilah yang akan berfungsi memproses data latih untuk membentuk struktur pohon keputusan (*tree structure*) yang berisi aturan-aturan (*rules*) klasifikasi pelanggan Krakatau Biliar secara otomatis. Selanjutnya, penulis akan menempatkan operator ini ke dalam rangkaian proses (biasanya diletakkan di dalam sub-proses training pada operator *Cross Validation*) untuk mulai melatih model mengenali pola data. Dapat dilihat pada gambar 4.11.



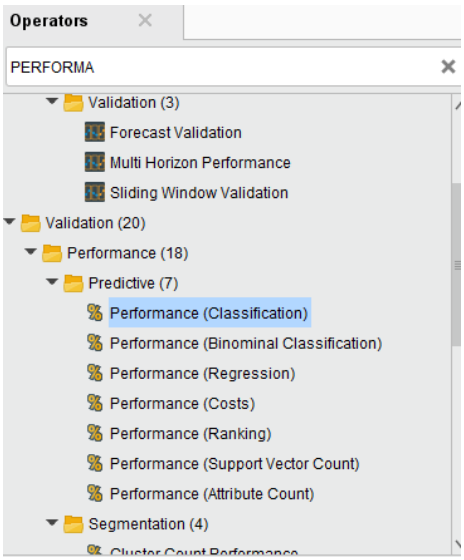
Gambar 4. 11 *Operator Decision Tree*

Pada gambar 4.12 memperlihatkan langkah penulis dalam mencari dan memilih *operator Apply Model* melalui *panel Operators*. *Operator* yang berada di bawah kategori Scoring ini berfungsi vital sebagai jembatan untuk mengaplikasikan model *Decision Tree* yang telah dilatih (*trained model*) ke dalam data uji (*testing set*). Melalui operator ini, sistem akan memproyeksikan aturan-aturan klasifikasi yang telah dipelajari sebelumnya untuk memprediksi kategori pelanggan pada data pengujian, yang mana hasil prediksinya nanti akan dibandingkan dengan label asli guna mengukur performa dan akurasi model yang dikembangkan. Dapat dilihat pada gambar 4.12.



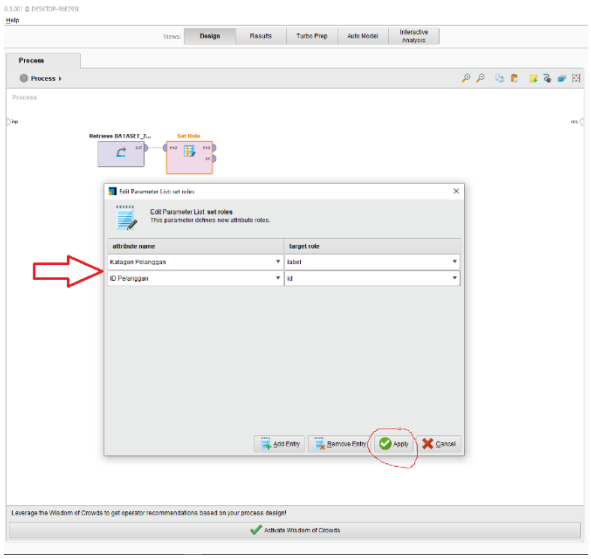
Gambar 4. 12 *Operator Apply Model*

Pada gambar 4.13 menampilkan tahap seleksi *operator* evaluasi akhir, di mana penulis memilih operator *Performance* (*Classification*) melalui panel *Operators* untuk mengukur keberhasilan model. Operator yang terletak di dalam kategori *Validation* > *Performance* > *Predictive* ini dipilih secara spesifik karena studi kasus di Krakatau Biliar melibatkan klasifikasi dengan lebih dari dua kelas label (*multinominal*/*polinomial*), bukan sekadar *binomial*. Fungsi utama dari operator ini adalah untuk menghitung dan menampilkan statistik performa model, seperti nilai akurasi (*accuracy*) dan matriks kebingungan (*confusion matrix*), dengan cara membandingkan hasil prediksi dari operator *Apply Model* terhadap label data aktual, sehingga penulis dapat menyimpulkan seberapa presisi algoritma *Decision Tree* dalam mengelompokkan pelanggan. Dapat dilihat pada gambar 4.13.



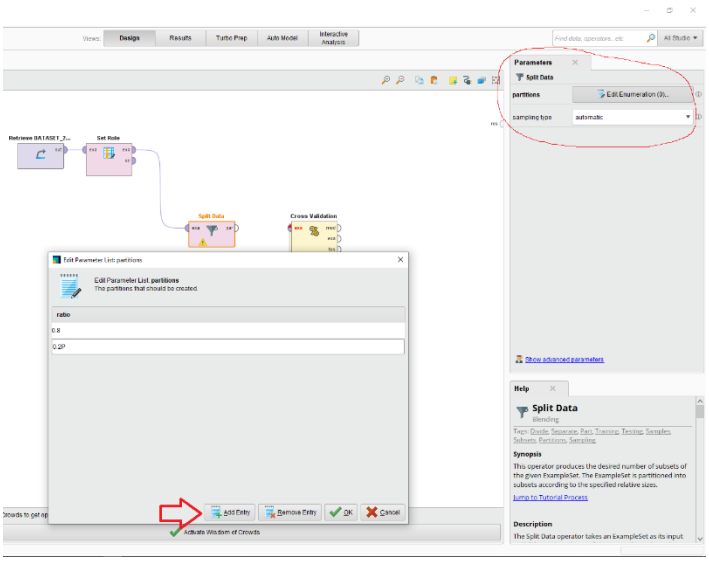
Gambar 4. 13 *Operator Performance* (*Classification*)

Pada gambar 4.14 memperlihatkan konfigurasi parameter di dalam *operator Set Role*, di mana penulis menetapkan peran khusus untuk atribut-atribut tertentu guna mempersiapkan data sebelum proses pemodelan. Pada tahap ini, atribut Katagori Pelanggan diatur sebagai label (target prediksi), sedangkan *ID* Pelanggan diubah perannya menjadi id agar tidak ikut diproses sebagai fitur analisis karena hanya berfungsi sebagai penanda unik. Langkah ini diakhiri dengan menekan tombol *Apply* untuk menyimpan pengaturan tersebut, memastikan algoritma *Decision Tree* nantinya hanya mempelajari pola dari atribut yang relevan dan mengabaikan kolom identitas. Dapat dilihat pada gambar 4.14.



Gambar 4. 14 *Set Role Dataset*

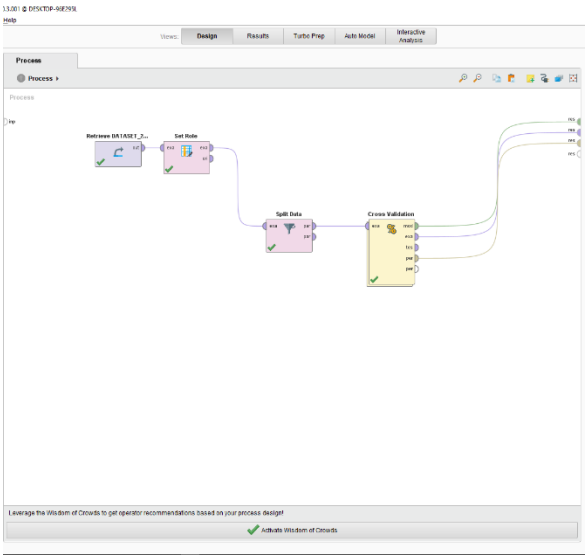
Pada gambar 4.15 mendemonstrasikan langkah konfigurasi teknis pada operator *Split Data*, di mana penulis menetapkan proporsi pembagian dataset untuk keperluan *eksperimen*. Pada jendela parameter *partitions*, terlihat penulis memasukkan nilai rasio 0.8 dan 0.2, yang mengindikasikan bahwa *dataset* akan dipecah menjadi dua bagian: 80% dialokasikan sebagai data latih (*training set*) untuk membangun model klasifikasi, sedangkan 20% sisanya digunakan sebagai data uji (*testing set*) untuk mengevaluasi performa model secara objektif. Pengaturan ini dilakukan dengan menambahkan entri rasio melalui tombol *Add Entry* dan diakhiri dengan menekan tombol *OK*, memastikan bahwa validasi model dilakukan pada data yang terpisah dari proses pembelajaran. Dapat dilihat pada gambar 4.15.



Gambar 4. 15 *Split Dataset* (80:20)

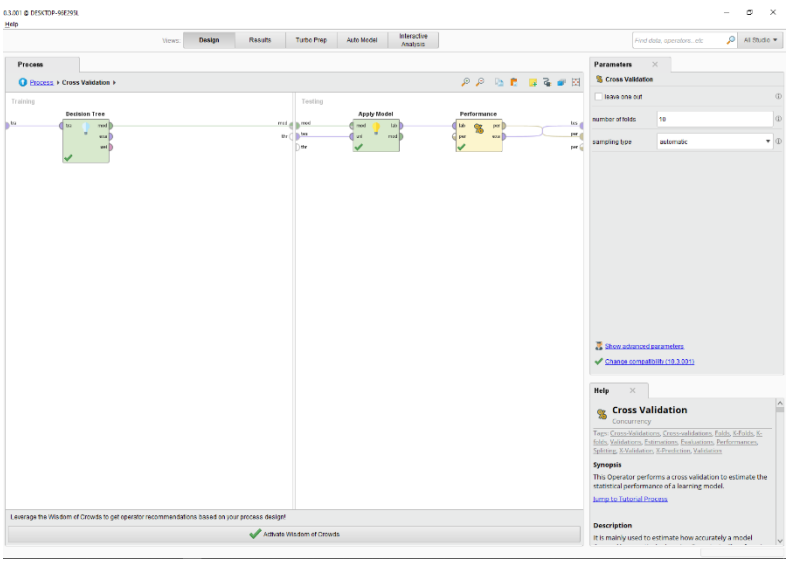
e. Pengujian *Validitas Model*

Pada gambar 4.16 menampilkan rancangan akhir alur proses (*workflow*) pengujian *validitas* model secara keseluruhan di dalam *Process Canvas*. Pada tahap ini, penulis telah menghubungkan seluruh operator secara berurutan, dimulai dari pengambilan dataset, pengaturan peran atribut (*Set Role*), hingga pembagian data (*Split Data*) yang bermuara pada operator *Cross Validation*. Terlihat bahwa *port* keluaran dari *operator Cross Validation*—khususnya *Model* (mod) untuk menampilkan struktur pohon keputusan dan *Performance* (*per*) untuk hasil akurasi telah disambungkan ke *port* hasil (*res*) di sisi kanan layar. Koneksi final ini menandakan bahwa sistem siap dieksekusi (*run*) untuk melakukan validasi silang pada data klasifikasi pelanggan Krakatau Biliar guna menghasilkan laporan performa dan aturan keputusan yang valid. Dapat dilihat pada gambar 4.16.



Gambar 4. 16 Pengujian *Validitas Model*

Pada gambar 4.17 memperlihatkan struktur internal dari *sub-proses operator Cross Validation*, yang secara otomatis membagi alur kerja menjadi dua panel utama: *panel* kiri (*Training*) dan *panel* kanan (*Testing*). Pada sisi *Training*, penulis menempatkan operator *Decision Tree* yang berfungsi untuk melatih dan membangun model menggunakan data latihan dari setiap lipatan (*fold*) validasi. Sementara itu, pada sisi Testing, penulis merangkai operator *Apply Model* dan *Performance* secara berurutan; di mana model yang dihasilkan dari tahap training akan langsung diaplikasikan ke data uji untuk diprediksi, kemudian hasil prediksinya dievaluasi oleh operator *Performance* guna mendapatkan nilai akurasi rata-rata yang objektif. Dapat dilihat pada gambar 4.17.



Gambar 4. 17 *Design Cross Validation*

f. Evaluasi Kinerja

Pada gambar 4.18 menyajikan hasil evaluasi akhir model dalam bentuk *Confusion Matrix*, di mana penulis dapat melihat ringkasan performa klasifikasi dengan tingkat akurasi rata-rata yang sangat tinggi mencapai 95.00%. Tabel ini memetakan perbandingan antara hasil prediksi model (baris) terhadap data aktual (kolom); sebagai contoh, model berhasil memprediksi kategori 'Aktif' dengan presisi sempurna (100%), meskipun masih terdapat kesalahan minor di mana satu data 'Tetap' salah diprediksi sebagai 'Baru' dan satu data 'Aktif' terprediksi sebagai 'Tetap'. Berdasarkan nilai class recall dan precision yang dominan di atas 88%, penulis dapat menyimpulkan bahwa algoritma *Decision Tree* yang dibangun sangat andal untuk diterapkan dalam mengelompokkan pelanggan Krakatau Biliar. Dapat dilihat pada gambar 4.18.

Table View Plot View

accuracy: 95.00% +/- 10.54% (micro average: 95.00%)

	true Baru	true Aktif	true Tetap	class precision
pred. Baru	8	0	1	88.89%
pred. Aktif	0	16	0	100.00%
pred. Tetap	0	1	14	93.33%
class recall	100.00%	94.12%	93.33%	

Gambar 4. 18 Evaluasi Model

Pada gambar 4.19 menampilkan deskripsi tekstual dari hasil evaluasi model yang disebut *Performance Vector*, yang berfungsi sebagai rincian statistik formal dari matriks kebingungan yang telah dibahas sebelumnya. Pada tampilan ini, penulis dapat melihat konfirmasi nilai akurasi total sebesar 95.00% beserta standar *deviasinya* (+/- 10.54%), yang menunjukkan tingkat kepercayaan dan konsistensi model dalam melakukan prediksi. Data teks ini merinci distribusi prediksi di mana baris merepresentasikan kelas prediksi (Baru, Aktif, Tetap) terhadap kolom nilai sebenarnya (True), memperlihatkan secara jelas bahwa mayoritas data berhasil diklasifikasikan dengan tepat pada diagonal utamanya (seperti 16 data 'Aktif' dan 14 data 'Tetap'), sehingga menjadi bukti validitas *kuantitatif* yang kuat bagi penelitian penulis. Dapat dilihat pada gambar 4.19.

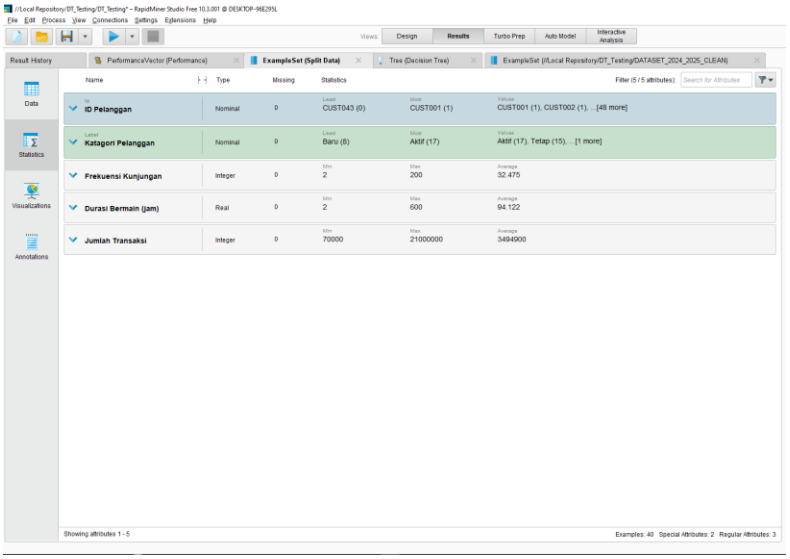
PerformanceVector

```
PerformanceVector:
accuracy: 95.00% +/- 10.54% (micro average: 95.00%)
ConfusionMatrix:
True:   Baru    Aktif   Tetap
Baru:   8       0       1
Aktif:  0       16      0
Tetap:  0       1       14
```

Gambar 4. 19 Deskripsi Evaluasi

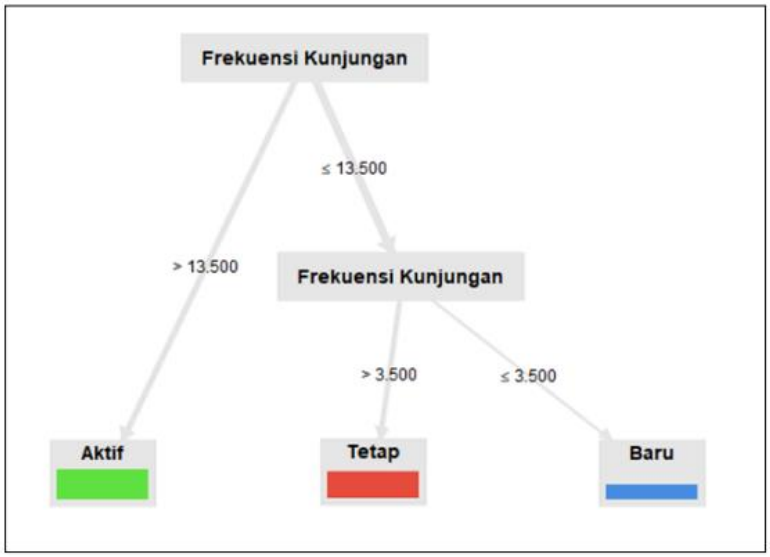
g. Interpretasi Hasil dan Pola Keputusan

Pada gambar 4.20 menyajikan tampilan *Statistics* pada *tab Results*, yang digunakan oleh penulis untuk meninjau ringkasan statistik deskriptif dari seluruh atribut dalam dataset. Dalam tampilan ini, atribut Katagori Pelanggan ditandai dengan warna hijau yang menegaskan perannya sebagai Label (target klasifikasi), sedangkan atribut numerik lainnya seperti Frekuensi Kunjungan, Durasi Bermain, dan Jumlah Transaksi ditampilkan lengkap dengan nilai minimum, maksimum, serta rata-ratanya. Tahap ini sangat penting bagi penulis untuk memvalidasi kualitas data, memastikan distribusi nilai yang wajar, dan mengonfirmasi bahwa tidak ada data yang hilang (*missing values*) sebelum algoritma *Decision Tree* bekerja memproses pola-pola tersebut. Dapat dilihat pada gambar 4.20.



Gambar 4. 20 Deskripsi Statistik

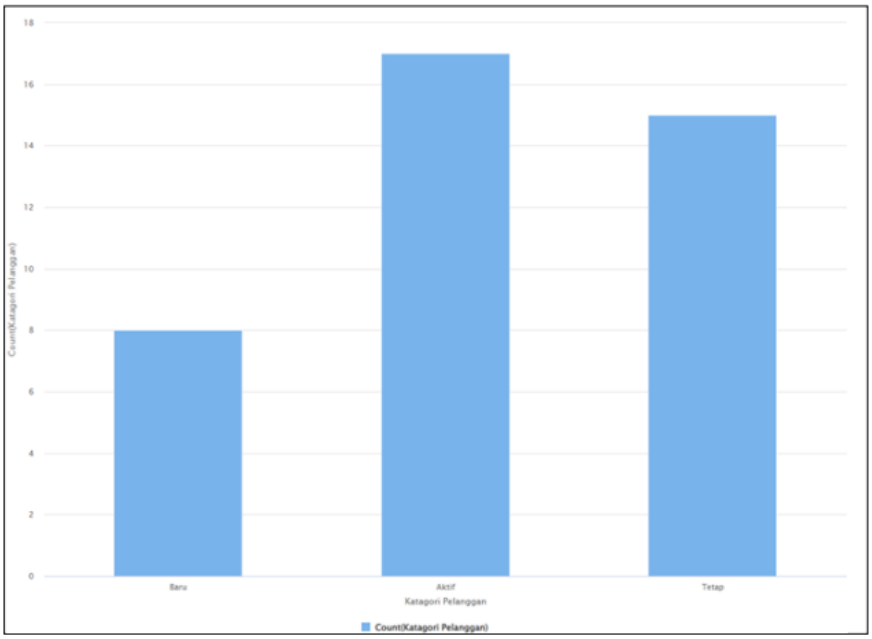
Pada gambar 4.21 memvisualisasikan hasil akhir pemodelan berupa grafik Pohon Keputusan (*Decision Tree*), di mana atribut Frekuensi Kunjungan teridentifikasi sebagai akar utama (*root node*) yang paling berpengaruh dalam membedakan jenis pelanggan di Krakatau Biliar. Berdasarkan visualisasi pohon keputusan di atas, klasifikasi pelanggan terbentuk melalui tiga aturan logika utama: *IF* Frekuensi Kunjungan ≤ 3.5 kali, *THEN* sistem mengidentifikasinya sebagai Pelanggan Baru; selanjutnya *IF* kunjungan berada di rentang > 3.5 namun ≤ 13.5 kali, *THEN* masuk kategori Pelanggan Tetap; dan terakhir *IF* Frekuensi Kunjungan > 13.5 kali, *THEN* pelanggan tersebut secara otomatis diklasifikasikan sebagai Pelanggan Aktif, yang menunjukkan bahwa intensitas kedatangan menjadi satu-satunya indikator penentu dalam model ini. Dapat dilihat pada gambar 4.21.



Gambar 4. 21 Pohon Keputusan

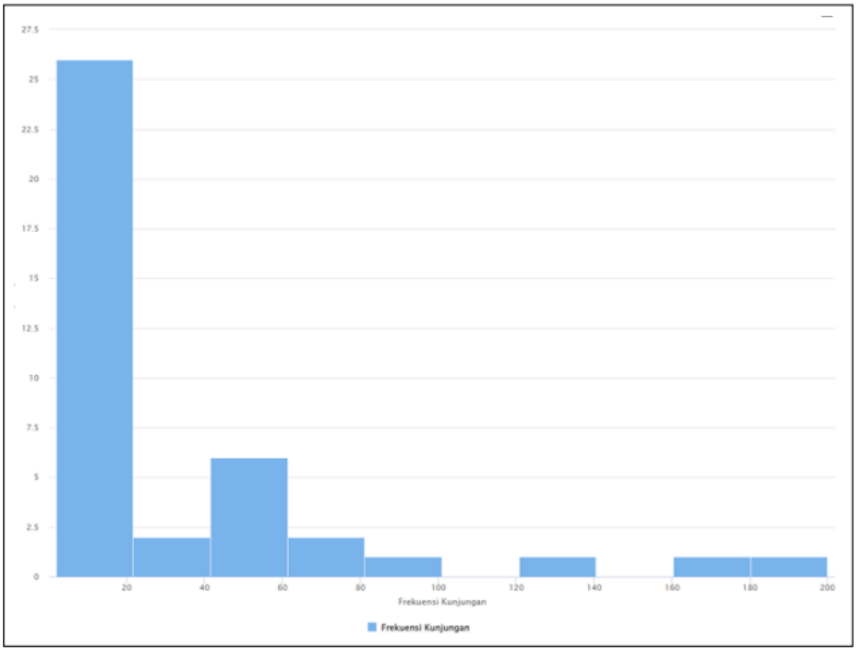
h. Visualisasi

Pada gambar 4.22 menyajikan visualisasi distribusi data dalam bentuk grafik batang (bar chart), di mana penulis dapat mengamati proporsi jumlah pelanggan berdasarkan kelas labelnya secara intuitif. Pada grafik tersebut, terlihat jelas bahwa kategori pelanggan 'Aktif' merupakan kelompok yang paling dominan dengan jumlah tertinggi (mencapai 17 data), diikuti oleh kategori 'Tetap' di urutan kedua, dan kategori 'Baru' dengan frekuensi paling sedikit. Visualisasi ini memberikan pemahaman kepada penulis mengenai peta sebaran pelanggan di Krakatau Biliar, sekaligus menjadi penutup analisis hasil yang mengonfirmasi bahwa model telah berhasil memetakan seluruh data ke dalam tiga segmen target yang ditentukan. Dapat dilihat pada gambar 4.22.



Gambar 4. 22 Visualisasi Kategori Pelanggan

Pada gambar 4.23 menampilkan histogram yang memvisualisasikan distribusi data pada atribut Frekuensi Kunjungan, di mana penulis dapat mengamati pola kedatangan pelanggan secara mendetail. Grafik tersebut menunjukkan pola distribusi yang sangat miring ke kiri (*skewed left*), di mana mayoritas pelanggan memiliki frekuensi kunjungan yang rendah (kisaran 0 hingga 20 kali), ditandai dengan batang grafik yang menjulang tinggi di sisi paling kiri. Sebaliknya, jumlah pelanggan dengan frekuensi kunjungan yang sangat tinggi (di atas 100 kali) terlihat sangat sedikit dan menyebar tipis ke arah kanan, yang mengindikasikan bahwa pelanggan setia dengan intensitas kedatangan ekstrem merupakan segmen minoritas namun spesifik dibandingkan dengan pelanggan reguler lainnya. Dapat dilihat pada gambar 4.23



Gambar 4. 23 Visualisasi Frekuensi Kunjungan

4.2 Pembahasan

Berdasarkan dari semua tahapan yang sudah dilakukan, baik dari tahap perancangan sampai tahap implementasi, menghasilkan sebuah kesimpulan bahwa semua proses-proses analisis data yang sudah dilakukan berjalan dengan baik, dan menghasilkan sebuah klasifikasi pelanggan di Karakatau Billiar. Pembahasan lebih lanjut mengenai proses klasifikasi *Decision Tree RapidMiner* adalah sebagai berikut:

4.2.1. Pembahasan Hasil Implementasi

a. *Import Data*

Tahap awal dimulai dengan proses *Import Data*, di mana penulis memuat dataset pelanggan Krakatau Biliar yang tersimpan dalam format *Excel* ke dalam lingkungan kerja *RapidMiner*. Penulis menyeleksi rentang sel yang relevan—mencakup atribut seperti *ID*, Frekuensi Kunjungan, Durasi Bermain, dan Jumlah Transaksi, serta memastikan baris pertama terbaca sebagai *header* atribut yang benar. Data tersebut kemudian disimpan secara permanen ke dalam *Local Repository* pada *folder* khusus yang telah dibuat penulis dengan nama *DT_Testing*, agar data dapat diakses kembali dengan mudah untuk proses pengujian berulang tanpa perlu melakukan konfigurasi ulang.

b. *Preprocessing Data*

Setelah data berhasil dimuat, penulis melakukan tahapan *Preprocessing Data* untuk mempersiapkan dataset agar sesuai dengan kebutuhan algoritma klasifikasi. Langkah krusial pada tahap ini adalah penggunaan *operator Set Role*

untuk menetapkan atribut 'Katagori Pelanggan' sebagai label (*target* prediksi) dan mengubah peran *'ID* Pelanggan' menjadi id agar tidak memengaruhi perhitungan *model*. Selanjutnya, penulis menggunakan operator *Split Data* untuk membagi *dataset* dengan *rasio* 0.8 (80%) sebagai data latih (*training*) untuk pembentukan model, dan 0.2 (20%) sisanya dialokasikan sebagai data uji (*testing*) guna memvalidasi performa model secara objektif.

c. Model klasifikasi *Decision Tree*

Pada tahap pemodelan utama, penulis menerapkan algoritma *Decision Tree* yang diintegrasikan ke dalam operator *Cross Validation* (validasi silang) untuk menjamin keandalan hasil pengujian. Di dalam sub-proses validasi ini, operator Decision Tree ditempatkan pada panel *training* untuk mempelajari pola data dan membentuk struktur pohon keputusan, yang kemudian diuji kemampuannya menggunakan operator *Apply Model* pada panel testing. Penulis juga menambahkan operator *Performance* (*Classification*) di akhir rangkaian untuk mengukur tingkat keberhasilan model berdasarkan metrik statistik seperti *accuracy*, *precision*, dan *recall* setelah proses pelatihan selesai.

d. Visualisasi Data

Hasil dari pemodelan tersebut kemudian diterjemahkan ke dalam bentuk Visualisasi Data yang intuitif, utamanya berupa grafik Pohon Keputusan (*Tree Graph*) yang menampilkan aturan klasifikasi berbasis if-then yang mudah dipahami. Grafik ini memperlihatkan bahwa atribut 'Frekuensi Kunjungan' menjadi akar penentu utama, di mana pelanggan dengan frekuensi kunjungan di

atas 13,5 kali secara otomatis diklasifikasikan sebagai pelanggan 'Aktif'. Selain itu, visualisasi statistik juga disajikan dalam bentuk grafik batang yang menunjukkan dominasi kategori pelanggan 'Aktif' dan 'Tetap', serta histogram yang menggambarkan sebaran data frekuensi kunjungan yang memiliki pola kemiringan ke kiri (*left-skewed*).

e. Eksekusi Proses

Tahap akhir adalah Eksekusi Proses, di mana penulis menjalankan (*run*) seluruh rangkaian operator yang telah disambungkan pada Process *Canvas* untuk mendapatkan hasil kalkulasi akhir. Eksekusi ini menghasilkan output evaluasi berupa matriks kebingungan (*Confusion Matrix*) yang menunjukkan bahwa model mampu memprediksi kategori pelanggan dengan tingkat akurasi rata-rata yang sangat tinggi, yakni sebesar 95.00%. Tingginya nilai akurasi ini membuktikan bahwa model *Decision Tree* yang dibangun oleh penulis sangat efektif dalam memetakan karakteristik pelanggan Krakatau Biliar berdasarkan data historis yang tersedia.

4.2.2. Pembahasan Hasil Evaluasi

Setelah tahap pelatihan model selesai, evaluasi kinerja algoritma *Decision Tree* dilakukan menggunakan metode *Cross Validation* dan diukur melalui *Confusion Matrix*. Evaluasi ini bertujuan untuk mengukur seberapa akurat model dalam memprediksi kategori pelanggan (Baru, Aktif, Tetap) dibandingkan dengan data aktualnya.

a. *Confusion Matrix*

Hasil pengujian model ditampilkan dalam bentuk *matriks* kebingungan (*Confusion Matrix*) seperti terlihat pada Tabel 4.1 di bawah ini. *Matriks* ini merepresentasikan perbandingan antara kelas prediksi *model* (baris) dengan kelas sebenarnya aktual (kolom).

Tabel 4. 1 Hasil *Confusion Matrix*

	True Baru	True Aktif	True Tetap	Total Prediksi
Pred. Baru	8	0	1	9
Pred. Aktif	0	16	0	16
Pred. Tetap	0	1	14	15
Total Aktual	8	17	15	Total Data: 40

Sumber: Hasil Pengolahan *RapidMiner* (2025)

Berdasarkan tabel di atas, total data uji yang digunakan berjumlah 40 data. Diagonal utama (angka yang ditebalkan: 8, 16, 14) menunjukkan jumlah prediksi yang benar, sedangkan angka di luar diagonal menunjukkan prediksi yang salah (*miss-classification*).

b. Perhitungan *Manual* Akurasi, Presisi, dan *Recall*

Untuk memvalidasi hasil otomatis dari *RapidMiner*, penulis melakukan perhitungan manual berdasarkan nilai-nilai dalam *Confusion Matrix* di atas.

1. Perhitungan Akurasi (*Accuracy*) Akurasi mengukur rasio prediksi benar terhadap keseluruhan data.

$$Akurasi = \frac{Jumlah\ Prediksi\ Benar}{Total\ Data}$$

$$Akurasi = \frac{TP_{Baru} + TP_{Aktiif} + TP_{Tetap}}{Total\ Data}$$

$$Akurasi = \frac{8 + 16 + 14}{40} = \frac{38}{40} = 0{,}95$$

Hasil: Tingkat akurasi model adalah 95,00%.

2. Perhitungan Presisi (*Precision*) Presisi mengukur seberapa akurat model saat menyatakan suatu data masuk ke kelas tertentu (Fokus pada Baris Prediksi). Rumus: *TP / (TP + FP).*

   - Kelas Baru:

$$Presisi = \frac{8}{(8 + 0 + 1} = \frac{8}{9} = 0{,}8880 \rightarrow 88{,}89\%$$

   (Artinya: Dari 9 data yang diprediksi sebagai 'Baru', 8 di antaranya benar).

   - Kelas Aktif:

$$Presisi = \frac{16}{(0 + 16 + 0} = \frac{16}{16} = 1{,}00 \rightarrow 100{,}00\%$$

   (Artinya: Model tidak pernah salah saat memprediksi pelanggan 'Aktif').

   - Kelas Tetap:

$$Presisi = \frac{14}{(0 + 1 + 14} = \frac{14}{15} = 0{,}9333 \rightarrow 93{,}33\%$$

3. Perhitungan *Recall* (*Class Recall*) Recall mengukur kemampuan model dalam menemukan kembali seluruh data positif dari kelas sebenarnya (Fokus pada Kolom Aktual). Rumus: *TP / (TP + FN).*

   - Kelas Baru:

$$Recall = \frac{8}{(8 + 0 + 0} = \frac{8}{8} = 1{,}00 \rightarrow 100{,}00\%$$

(Artinya: Seluruh data pelanggan 'Baru' yang ada berhasil ditemukan oleh model).

- Kelas Aktif:

$$Recall = \frac{16}{(0 + 16 + 1} = \frac{16}{17} = 0{,}9411 \rightarrow 94{,}12\%$$

(Keterangan: Ada 1 data 'Aktif' yang meleset diprediksi sebagai 'Tetap').

- Kelas Tetap:

$$Recall = \frac{14}{(1 + 0 + 14} = \frac{14}{15} = 0{,}9333 \rightarrow 93{,}33\%$$

(Keterangan: Ada 1 data 'Tetap' yang meleset diprediksi sebagai 'Baru').

c. Tabel Hasil Akhir Evaluasi

Berdasarkan perhitungan di atas, ringkasan performa model klasifikasi pelanggan Krakatau Biliar disajikan dalam Tabel 4.2 berikut:

Tabel 4. 2 Ringkasan Performa *Model Decision Tree*

Kategori Pelanggan	***Class Precision***	***Class Recall***
Baru	88.89%	100.00%
Aktif	100.00%	94.12%
Tetap	93.33%	93.33%
Rata-rata Akurasi	**95.00% +/- 10.54%**	

d. Analisis Hasil

Dari hasil evaluasi tersebut, dapat disimpulkan bahwa algoritma *Decision Tree* memiliki kinerja yang sangat baik dengan akurasi 95,00%.

1). Kekuatan Model

Model memiliki kemampuan sempurna (*Precision* 100%) dalam mengidentifikasi pelanggan 'Aktif', yang berarti jika sistem melabeli pelanggan sebagai 'Aktif', maka kepastiannya sangat tinggi. Selain itu, model juga sempurna dalam Recall (100%) untuk pelanggan 'Baru', yang berarti tidak ada pelanggan baru yang terlewat dari deteksi sistem.

2). Kelemahan

Terjadi sedikit kesalahan klasifikasi (misklasifikasi) sebanyak 2 data, yaitu satu pelanggan aktual 'Tetap' yang diprediksi sebagai 'Baru', dan satu pelanggan aktual 'Aktif' yang diprediksi sebagai 'Tetap'. Namun, tingkat kesalahan ini sangat minim (5%) dan masih dalam batas toleransi yang wajar untuk implementasi bisnis.

BAB V

KESIMPULAN DAN SARAN

5.1. Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan mengenai klasifikasi pelanggan Krakatau Biliar menggunakan algoritma *Decision Tree* dengan bantuan *RapidMiner Studio*, maka dapat disimpulkan beberapa hal sebagai berikut:

a. Penelitian ini berhasil mengolah data pelanggan Krakatau Biliar periode 2024–2025 dengan menerapkan algoritma *Decision Tree (C4.5)* untuk mengelompokkan pelanggan ke dalam tiga kategori, yaitu pelanggan baru, pelanggan aktif, dan pelanggan tetap.
b. Seluruh tahapan penelitian, mulai dari proses pengumpulan data, import data ke *RapidMiner*, *preprocessing*, pembagian data, hingga pemodelan dan evaluasi, dapat dijalankan dengan baik dan sesuai dengan metode data mining yang digunakan.
c. Hasil pengujian model menggunakan metode *Cross Validation* menunjukkan bahwa *algoritma Decision Tree* mampu menghasilkan tingkat akurasi yang tinggi, yaitu sebesar 95,00%, sehingga model dinilai cukup andal dalam melakukan klasifikasi pelanggan.
d. Dari hasil interpretasi pohon keputusan, diketahui bahwa frekuensi kunjungan menjadi faktor utama dalam menentukan kategori pelanggan,

sedangkan durasi bermain berperan sebagai faktor pendukung dalam proses pengambilan keputusan.

e. Dengan hasil tersebut, dapat disimpulkan bahwa penerapan algoritma *Decision Tree* efektif digunakan sebagai alat bantu dalam pengelompokan pelanggan dan dapat dimanfaatkan oleh pihak Krakatau Biliar untuk mendukung pengambilan keputusan bisnis.

5.2. Saran

Berdasarkan hasil penelitian dan kesimpulan yang telah diperoleh, maka beberapa saran yang dapat diberikan adalah sebagai berikut:

a. Penelitian selanjutnya disarankan untuk menggunakan data dengan jumlah yang lebih banyak serta periode pengamatan yang lebih panjang agar hasil klasifikasi yang diperoleh semakin akurat dan stabil.
b. Untuk pengembangan penelitian, algoritma lain seperti *Naïve Bayes*, *Random Forest*, atau *KNN* dapat digunakan sebagai pembanding untuk mengetahui metode klasifikasi yang paling optimal.
c. Penambahan atribut lain, seperti jenis permainan, waktu kunjungan, atau riwayat promo, diharapkan dapat meningkatkan kualitas model klasifikasi yang dihasilkan.
d. Model yang telah dibuat pada penelitian ini dapat dikembangkan lebih lanjut menjadi sistem pendukung keputusan yang terintegrasi dengan sistem operasional Krakatau Biliar agar dapat digunakan secara berkelanjutan.

e. Pihak manajemen Krakatau Biliar diharapkan dapat memanfaatkan hasil penelitian ini sebagai bahan pertimbangan dalam menentukan strategi pelayanan dan pemasaran pelanggan.

DAFTAR PUSTAKA

Faisal, S., 2020. Klasifikasi Data Minning Menggunakan Algoritma C4.5 Terhadap Kepuasan Pelanggan Sewa Kamera Cikarang. *Jurnal Ilmu Komputer & Teknologi Informasi*, 4(1), pp. 38-45.

F., Hartama, D. & Damanik, I. S., 2020. Analisa Kepuasan Pelanggan Menggunakan Klasifikasi Data Mining. BRAHMANA: *Jurnal Penerapan Kecerdasan Buatan , 2(1), pp. 41-48.*

Fihir, M., M. & Hayati, U., 2023. Klasifikasi Tingkat Kepuasan Pelanggan Kopi Kenangan Menggunakan Metode Decision Tree Pada Aplikasi Kopi Kenangan. *JATI (Jurnal Mahasiswa Teknik Informatika)* , 7(6), pp. 3830- 3833.

Fihir, M., M. & Hayati, U., 2023. Klasifikasi Tingkat Kepuasan Pelanggan Kopi Kenangan Menggunakan Metode Decision Tree Pada Aplikasi Kopi Kenangan. JATI (Jurnal Mahasiswa Teknik Informatika), 7(6), pp. 3830- 3833.

Herlambang, E., Basuki, C., Purnomo, L. & S., 2024. Peningkatan Donasi dengan Strategi Pemasaran Digital PeduliSehat.id Menggunakan Framework CRISP-DM. *Jurnal Bisnis DarmaJaya*, 10(1), pp. 86-100.

Hasanah , M. A., Soim, S. & Handayani, A. S., 2021. Implementasi CRISP-DM Model Menggunakan Metode Decision Tree dengan Algoritma CART untuk Prediksi Curah Hujan Berpotensi Banjir. Journal of Applied Informatics and Computing (JAIC) , 5(2), pp. 103-108.

Ilhami, M. W. et al., 2024. Penerapan Metode Studi Kasus Dalam Penelitian Kualitatif. *Jurnal Ilmiah Wahana Pendidikan,* 10(9), pp. 462-469.

Lestari, A., Fitrisia, A. & O., 2022. Metodologi Ilmu Pengetahuan : Kuantitatif Dan Kualitatif Dalam Bentuk Implementasi. *Jurnal Pendidikan dan Konseling*, 4(6), pp. 8558-8563.

Lutfi, M. I. & Situmorang, L., 2023. Stigma Masyarakat Tentang Permainan Billiard di Kelurahan Timbau Kecamatan Tenggarong Kabupaten Kutai Kartanegara. *eJournal Pembangunan Sosial,* 1(1), pp. 60-75.

Mardi, Y., 2021. Data Mining : Klasifikasi Menggunakan Algoritma C4.5. *Jurnal Edik Informatika*, 2(2), pp. 213-219.

Nahjan, M. R., Heryana, N. & Voutama, A., 2023. Implementasi RapidMiner dengan Metode Clustering K-Means untuk Analisa Penjualan pada Toko OJ Cell. JATI *(Jurnal Mahasiswa Teknik Informatika),* 7(1), pp. 101-104.

Nurhidayat, R. & Dewi, K. E., 2023. Penerapan Algoritma K-Nearest Neighbor dan Fitur Ekstraksi N-Gram dalam Analisis Sentimen Berbasis Aspek. *KOMPUTA : Jurnal Ilmiah Komputer dan Informatika*, 12(1), pp. 91-100.

Pratiwi, E. L., 2020. Konsep Dasar Algoritma dan Pemrograman Dengan Bahasa Java.. Banjarmasin: POLIBAN PRESS Anggota APPTI (Anggota Penerbit Perguruan Tinggi Indonesia).

Ramadhon, R. N. et al., 2024. Implementasi Algoritma Decision Tree untuk Klasifikasi Pelanggan Aktif atau Tidak Aktif pada Data Bank. Karimah Tauhid, 3(2), pp. 1860-1874.

Shafarindu, A. I. et al., 2021. Klasifikasi Data Penjualan pada Supermarket dengan Metode Decision Tree. Seminar Nasional Mahasiswa Ilmu Komputer dan Aplikasinya (SENAMIKA), pp. 660-667.

Trivaika, E. & Senubekti, M. A., 2022. Perancangan Aplikasi Pengelola Keuangan Pribadi Berbasis Android. *Jurnal Nuansa Informatika*, 16(1), pp. 33-40.

Wahyudi, R., Orisa, M. & Vendyansyah , N., 2021. Penerapan Algoritma K-Nearest Neighbors Pada Klasifikasi Penentuan Gizi Balita (Studi Kasus di Posyandu Desa Bluto). *JATI (Jurnal Mahasiswa Teknik Informatika)*, 5(2), pp. 750- 757.

Wijayanti, S., 2020. Klasifikasi Pelanggan Produk Angklung Customer Churn Prediction Menggunakan Decision Tree di Ikm Kampung Angklung Panyingkiran – Ciamis. *JURNAL MAHASISWA INDUSTRI GALUH*, 1(1), pp. 76-86.

LAMPIRAN

SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER
(STMIK) PALANGKARAYA
Jl. G. Obos No.114 Telp.0536-3224593, 3225515 Fax.0536-3225515 Palangkaraya
email : humas@stmikpk.ac.id - website : www.stmikpk.ac.id

Nomor : 837 /STMIK-C.1/05 /III/2025
Lampiran : -
Perihal : Permohonan Izin Penelitian dan Pengumpulan Data untuk Tugas Akhir

Kepada
Yth. **Pimpinan Karakatau Biliar**
Jl. Krakatau,Palangka,Kec.Jekan Raya,Kalimantan Tengah,Kota Palangka Raya

Dengan hormat,

Sehubungan dengan penyusunan Tugas Akhir mahasiswa sebagai persyaratan kelulusan Program Studi Teknik Informatika (S1) pada STMIK Palangkaraya, maka dengan ini kami sampaikan permohonan izin penelitian dan pengumpulan data bagi mahasiswa kami berikut:

Nama	:	SOFINA GRACIA FLODARIANI
NIM	:	C2155201088
Prodi (Jenjang)	:	Teknik Informatika (S1)
Thn. Akad. (Semester)	:	2024/2025 (8)
Lama Penelitian	:	19 Maret 2025 s.d 19 April 2025
Tempat Penelitian	:	Krakatau Biliar

Dengan judul Tugas Akhir:

Klasifikasi Pelanggan di Karakatau Biliar Menggunkan Decision Tree

Adapun ketentuan dan aturan pemberian informasi dan data yang diperlukan dalam penelitian tersebut menyesuaikan dengan ketentuan/peraturan pada instansi Bapak/Ibu.

Demikian permohonan ini disampaikan, atas perhatian dan kerja samanya diucapkan terima kasih.

Palangka Raya, 19 Maret 2025
Ketua,

Ir. Hj. Siti Maryamah, MM.
NIK. 196902021995004

KARAKATAU BILIAR

Jl. Krakatau, Palangka Raya, Kalimantan Tengah

Perihal : **Balasan Permohonan Izin Penelitian**

Kepada

Yth. Ketua Sekolah Tinggi Manajemen Informatika dan Komputer (STMIK)Palangkaraya

di Tempat

Dengan hormat,

Sehubungan dengan surat permohonan izin penelitian dan pengumpulan data untuk Tugas Akhir mahasiswa STMIK Palangkaraya, maka dengan ini kami menyampaikan bahwa pihak Karakatau Biliar memberikan izin kepada mahasiswa berikut:

Nama	:	Sofina Gracia Flodariani
NIM	:	C2155201088
Program Studi	:	Teknik Informatika (S1)

Untuk melaksanakan penelitian dan pengumpulan data di Karakatau Biliar dengan judul Tugas Akhir:

Klasifikasi Pelanggan di Karakatau Biliar Menggunakan Decision Tree

Penelitian tersebut dapat dilaksanakan sesuai dengan waktu dan ketentuan yang telah ditetapkan, dengan tetap menjaga etika, tata tertib, serta kerahasiaan data perusahaan.

Demikian surat balasan ini kami sampaikan. Atas perhatian dan kerja samanya, kami ucapkan terima kasih.

Palangka Raya, 19 Maret 2025

Pimpinan Karakatau Biliar,

Feri

SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER
(STMIK) PALANGKARAYA
Jl G Obos No 114 Telp 0536-3224593, 3225515 Fax 0536-3225515 Palangka Raya
email : humas@stmikpk.ac.id - website : www.stmikpk.ac.id

SURAT TUGAS

No.54/STMIK-3.C.2/KP/II/2025

Ketua Program Studi Teknik Informatika Sekolah Tinggi Manajemen Informatika dan Komputer (STMIK) Palangkaraya menugaskan nama- nama tersebut di bawah ini :

1. Nama : Hotmian Sitohang, M.Kom.
   NIK : 198503282008002
   Sebagai Pembimbing I Dalam Pembuatan Program
2. Nama : Amaya Andri Damaini, S.Kom., M.T
   NIK : 199309042018001
   Sebagai Pembimbing II Dalam Penulisan Tugas Akhir

Untuk membimbing Tugas Akhir mahasiswa :

Nama	: Sofina Gracia Flodariani
NIM	: C2155201088
Program Studi	: TEKNIK INFORMATIKA (55201)
Tanggal Daftar	: 17 Februari 2025
Judul Tugas Akhir	: Klasifikasi Pelanggan di Krakatau Billiar Menggunakan Decision Tree

Demikian surat ini dibuat agar dapat dipergunakan sebagaimana mestinya dan dilaksanakan dengan penuh tanggung jawab.

Palangka Raya, 24 Februari 2025
Ketua Program Studi Teknik Informatika,

Lili Resdiana, M.Kom
NIK. 198707282011007

Tembusan :
1. Pembimbing I dan II
2. Mahasiswa yang bersangkutan
3. Arsip

SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER
(STMIK) PALANGKARAYA
Jl. G. Obos No.114 Telp.0536-3224593, 3225515 Fax.0536-3225515 Palangkaraya
email : humas@stmikplk.ac.id - website : www.stmikplk.ac.id

KARTU KEGIATAN KONSULTASI TUGAS AKHIR

Nama Mahasiswa : Sofira Gracia Fiodarani
NIM : C21552010BB
Tanggal Persetujuan Judul :
Judul Tugas Akhir : KLASIFIKASI PELANGGAN DI KARAKATAU BILIAR MENGGUNAKAN DECION TREE

No.	Tanggal Konsultasi		Uraian	Tanda Tangan
	Terima	Kembali		
	9/3/25		- Atur Indentasi paragraf sesuai penomoran sub-bab	
			- Atur urutan Landasan teori	
			- Lanjutkan ke bab III	
	10/3/25		- Untuk bagian latar belakang batasannya	
	13/3/25		- di ubah agar tercinta menggunakan bahasa kamu	
			- Bagian Batasan masalah masukan yang inginya saja.	
	11/3/25		- untuk bagian manfaat ditambahkan lagi agar lebih rinci	
			- untuk contoh Pohon keputusan di ganti menjadi YA atau TIDAK	
			- untuk Penelitian yang Relevan cari topik yang Penelitianya tentang Pelanggan	
			- Lanjutkan	
	19/3/25		- Lanjutkan ke bab III	
	26/3/25		- masukan Contoh Perhitungan	
	9/4/25		- Rapikan penulisan sesuai pedoman	
			- Lengkapi Lampiran	
	10/4/25		Perbaiki daftar isi x daftar pustaka	
	11/04/25		ACC seminar	

No.	Tanggal Konsultasi		Uraian	Tanda Tangan
	Terima	Kembali		
	19/2/25		ACC seminar	
	17/11/25		- Sesuaikan sistematika penulisan dengan pedoman terbaru	
	30/10/25		- Perbaiki: hitungan metode * Data dimasukkan & diubah jd CSV * Lanjut Bab IV	
	05/11/25		*. Pembahasan dibahas semua yg dibuat. cara menjalankan. Upload data	
	10/12/25		*. Hasil recall, Akurasi dan kesimpulan ditulis semua dan hasil program.	
	18/12/25		*. lengkapi lampiran dan daftar pustaka * Acc sidang	
	17/12/25		lengkapi lampiran.	
	18/12/25		Acc sidang	

Menyetujui :

Dosen Pembimbing I,

Hotmian Situmorang, M.Kom

Dosen Pembimbing II,

AMAYA ANDRI P.

SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER
(STMIK) PALANGKARAYA
Jl. G. Obos No 114 Telp 0536-3224593, 3225515 Fax 0536-3225515 Palangka Raya
email : humas@stmikpk.ac.id - website : www.stmikpk.ac.id

SURAT TUGAS
PENGUJI SEMINAR PROPOSAL TUGAS AKHIR

No 90/STMIK-3 C.2/KP/IV/2025

Ketua Program Studi Teknik Informatika Sekolah Tinggi Manajemen Informatika dan Komputer (STMIK) Palangkaraya menugaskan kepada nama-nama berikut :

1. Nama : Lili Rusdiana, M.Kom
   NIK : 198707282011007
   Sebagai Ketua
2. Nama : Hotmian Sitohang, M.Kom
   NIK : 198503282008002
   Sebagai Sekretaris
3. Nama : Amaya Andri Damaini, S.Kom., M.T
   NIK : 199309042018001
   Sebagai Anggota

Tim Penguji Seminar Proposal Tugas Akhir :

Nama	:	Sofina Gracia Flodariani
NIM	:	C2155201088
Hari/Tanggal	:	Rabu, 30 April 2025
Waktu	:	08.00 WIB
Judul Proposal	:	Klasifikasi Pelanggan di Krakatau Biliar Menggunakan Decision Tree

Demikian surat ini dibuat agar dapat dipergunakan sebagaimana mestinya dan dilaksanakan dengan penuh tanggung jawab.

Palangka Raya, 15 April 2025
Ketua Program Studi Teknik Informatika

Abdul Hadi, S.T., M.Kom
NIK. 198505102021101

Tembusan :
1. Dosen Penguji
2. Mahasiswa yang Bersangkutan
3. Arsip Prodi

SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER
(STMIK) PALANGKARAYA
Jl. G. Obos No.114 Telp.0536-3224593, 3225515 Fax.0536-3225515 Palangka Raya
email : humas@stmikplk.ac.id - website : www.stmikplk.ac.id

BERITA ACARA
SEMINAR PROPOSAL TUGAS AKHIR

Periode (Bulan) : .....Tahun

1. Hari/Tanggal Seminar : Rabu / 30 April 2025
2. Waktu (Jam) : 08.00 WIB sampai dengan ..... WIB
3. Nama Mahasiswa : Sofina Gracia Floridiati
4. Nomor Induk Mahasiswa : C2155201088
5. Program Studi : Teknik Informatika
6. Tahun Angkatan : 2021
7. Judul Tugas Akhir : Klasifikasi Pelanggan di Karakatau Biliar menggunakan Decision Tree
8. Dosen Penguji :

	Nama	Nilai	Tanda Tangan
1.	Lili Ruliana	=	( )
2.	Hotmian	=	( )
3.	Amaya A.D	=	( )

9. Hasil Ujian : **LULUS** / ~~TIDAK LULUS~~ *) NILAI = 75
   Dengan Perbaikan/ ~~Tanpa Perbaikan~~ *)
10. Catatan Penting : 1. Lama Perbaikan : 12 hari (Maks. 15 hari)
    2. Jika lebih dari 15 hari s/d 1 (satu) bulan dikenakan sanksi berupa denda sebesar Rp. 300.000,- (Tiga ratus ribu rupiah), dan jika lebih dari 1 (satu) bulan dikenakan denda Rp. 600.000,- (Enam Ratus ribu rupiah) per bulan dari tanggal ujian
    3. Jika lebih dari 3 (tiga) bulan dari tanggal ujian maka hasil ujian dibatalkan dan wajib mengajukan judul dan pembimbing baru. Wajib membayar Denda dan membayar biaya seminar ulang.

Palangka Raya, 30 April 2025

Mengetahui :
Ketua Program Studi Teknik Informatika,

Abdul Hadi, S.T., M.Kom.
NIK. 198505102021101

Ketua Penguji,

Lili Ruliana
NIK. 198702282021007

Tembusan :
1. Arsip Prodi Teknik Informatika
2. Mahasiswa yang bersangkutan

Dibawa saat konsultasi perbaikan dengan dosen penguji

*) *Coret yang tidak perlu*

DAFTAR HADIR PESERTA
SEMINAR PROPOSAL TUGAS AKHIR

1. Nama Penyaji : Safina Gracia Fiodariani
2. Hari/ Tanggal : Rabu, 30-April-2025
3. Waktu : 08.00-Selesai
4. Judul Proposal : Klasifikasi Pelanggan di Karakatau Billiar Menggunakan Decision Tree

No.	Nama Mahasiswa	NIM	Tanda Tangan
1	Dicky	C2155201092	
2	Agriva Abel Patewuli	C2255201089	
3	Cleyrin Paskatabitha	C2255201086	
4	STEVENSENSIUS	C2255201108	
5	Okdilia	C2255201112	
6	Daniel Christoper Pemas	C2255201026	
7	Angella Karismanti	C2255201021	
8	Alif Rahmatullah L.	C2255201029	
9	M. khairan Jamil	C2255201038	
10	Mutya Talitha Tanaah	C2255201080	
11	Ryan Fazri Rahman	C2255201005	
12	WILDAN HAMDANI MAHMUDA	C2155201098	
13	Dahlia Febriani	C2255201081	
14	Pro [illegible] Sipahutar	C2255201075	
15			
16			
17			
18			
19			
20			
21			
22			
23			
24			
25			
26			
27			
28			
29			
30			

Palangka Raya, 30-April-2025

Mengetahui :
Ketua ~~Tim Penguji~~

Lili Rusdiana, M.Kom.

Mahasiswa Penyaji,

Safina Gracia Fiodariani

LEMBAR WAWANCARA

Nama Peneliti : Sofina Gracia Floridini

Tanggal : 19-Maret-2025

Tempat : Krakatau Billiard

Narasumber : Bapak Feri – Pengawas Krakatau Billiard

1. Sejak kapan Krakatau Billiard berdiri, Pak?

   Jawaban: Sejak Tahun 2015

2. Apakah Krakatau Billiard memiliki data pelanggan dan bagaimana pencatatannya?

   Jawaban: Iya kami memiliki data pelanggan,untuk pencatatannya kami masih secara

   Manual dengan menulis di kertas untuk laporan harian.

3. Apakah Bapak membedakan jenis pelanggan, misalnya pelanggan tetap atau baru?

   Jawaban: Iya kami membedakanya dengan beberapa tipe pelanggan,misalkan pelanggan tetap untuk pelanggan tetap sendiri kami meliat seberapa sering dia bermain misalkan 4-12 kali dengan durasi bermain 5-25 jam dan mengeluarkan transaksi sebesar Rp.250.000-2.000.000 ini kami masukan ke data pelanggan tetap dan untuk pelanggan baru yang hanya bermain 1-3 kali dengan durasi 5 jam dan mengelurkan transaksi sebesar Rp.250.000

4. Apakah Bapak tertarik untuk mengetahui pola pelanggan menggunakan sistem klasifikasi?

   Jawaban: Iya tertarik,siapa tau dengan ada nya sistem klasifikasi di billiard kami bisa sangat Membantu dalam mengkategorikan pelanggan.

5. Apa harapan Bapak dari hasil penelitian ini untuk Krakatau Billiard?

   Jawaban: Harapan saya dengan ada nya penelitian ini membatu kami dalam mengkategorikan dan mengambil keputusan berdasarkan data,sehingga lebih efektif dan efisien.

Narasumber	Pewawancara
Feri	SofinaGracia


9
KRAKATAU
BILLIARD


9

SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER
(STMIK) PALANGKARAYA
Jl. G. Obos No.114 Telp.0536-3224593, 3225515 Fax.0536-3225515 Palangka Raya
email : humas@stmikplk.ac.id - website : www.stmikplk.ac.id

SURAT TUGAS PENGUJI TUGAS AKHIR

370/STMIK-3.C.2/AU/XII/2025

Ketua Program Studi Teknik Informatika Sekolah Tinggi Manajemen Informatika dan Komputer (STMIK) Palangkaraya menugaskan kepada nama- nama berikut :

1. Nama : Sulistyowati, S.Kom., M.Cs.
   NIK : 198212162007002
   Sebagai Ketua
2. Nama : Rudini, M.Pd
   NIK : 198709172015105
   Sebagai Sekretaris
3. Nama : Lili Rusdiana, M.Kom
   NIK : 198707282011007
   Sebagai Anggota
4. Nama : Hotmian Sitohang, M.Kom
   NIK : 198503282008002
   Sebagai Anggota
5. Nama : Amaya Andri Damaini, S.Kom., M.T
   NIK : 199309042018001
   Sebagai Anggota

Untuk menguji Tugas Akhir mahasiswa :

Nama Mahasiswa	:	Sofina gracia flodariani
N I M	:	C2155201088
Hari / Tanggal Ujian	:	Sabtu, 3 Januari 2026
Waktu	:	07.30 WIB
Judul TA	:	KLASIFIKASI PELANGGAN DI KARAKATAU BILIAR MENGGUNAKAN DECISION TREE
Ruang	:	Sidang 1

Demikian surat ini dibuat agar dapat dipergunakan sebagaimana mestinya dan dilaksanakan dengan penuh tanggung jawab.

Palangka Raya, 29 Desember 2025

Ketua Program Studi Teknik Informatika,

29 Desember 2025

Ir. ABDUL HADI, S.T., M.Kom
NIK. 198505102021101

Tembusan :
1. Dosen Penguji
2. Mahasiswa yang Bersangkutan
3. Arsip

SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER
(STMIK) PALANGKARAYA
Jl. G. Obos No.114 Telp.0536-3224593, 3225515 Fax.0536-3225515 Palangka Raya
email : humas@stmikplk.ac.id - website : www.stmikplk.ac.id

BERITA ACARA
UJIAN TUGAS AKHIR

Periode : Januari 2026

Pada hari Ini Sabtu, 3 Januari 2026 Pukul 07.30 WIB sampai dengan 08.30 WIB telah dilaksanakan Ujian Tugas Akhir mahasiswa atas nama :

Nama Mahasiswa	:	Sofina gracia flodariani
N I M	:	C2155201088
Program Studi	:	Teknik Informatika
Tahun Angkatan	:	2021
Judul TA	:	KLASIFIKASI PELANGGAN DI KARAKATAU BILIAR MENGGUNAKAN DECISION TREE

Dosen Penguji :

	Nama		Nilai	Tanda Tangan
1.	Sulistyowati, S.Kom., M.Cs	=	.....	(.....)
2.	Rudini, M.Pd	=	.....	(.....)
3.	Lili Rusdiana, M.Kom	=	.....	(.....)
4.	Hotmian Sitohang, M.Kom	=	.....	(.....)
5.	Amaya Andri Damaini, S.Kom., M.T	=	.....	(.....)

Hasil Ujian : **LULUS / ~~TIDAK LULUS~~ *)** NILAI = 73

Revisi (~~Tidak Ada / Minor /~~ Mayor *)

Catatan Penting :
1. Lama Perbaikan : 14 hari
2. Jika lebih dari 1 (satu) bulan dikenakan sanksi berupa denda sebesar Rp. 600.000,- (Enam ratus ribu rupiah) per bulan dari tanggal ujian.
3. Jika lebih dari 3 (tiga) bulan dari tanggal ujian maka hasil ujian dibatalkan dan wajib mengajukan judul dan pembimbing baru.

Demikian Berita Acara ini dibuat, agar dapat digunakan sebagaimana semestinya.

Palangka Raya, 02 Januari 2026

Ketua Penguji,

Sulistyowati, S.Kom., M.Cs
NIK. 198212162007002

Mahasiswa,

Sofina gracia flodariani
NIM. C2155201088

Mengetahui :
Ketua Program Studi Teknik Informatika,

Ir. Abdul Hadi, S.T., M.Kom
NIK. 198505102021101

Tembusan:
Mahasiswa yang bersangkutan.
Berita acara ini diberikan setelah menyelesaikan semua revisi dari semua penguji
**) Coret yang tidak perlu*

Dipindai dengan CamScanner