



Penerapan Metode Algoritma C4.5 Untuk Menentukan Kelayakan Calon Nasabah Pemegang Kartu Kredit Bank Mega Card center Kuningan

Wahyu Hidayat^{1*}, Ayu Marlina Utami²

¹ Sistem Informasi/Universitas Binaniaga Indonesia

Email: wahyu.hidayat@unbin.ac.id

²Sistem Informasi/Universitas Binaniaga Indonesia

Email: ayumarliana@gmail.com

ABSTRACT

PT Bank Mega Card Center Kuningan, which is the center for making Bank Mega credit cards for the Jakarta region, is a subsidiary of PT CT Corpora. Bank Mega Card Center Kuningan's role in credit card manufacturing makes it easy for customers to find out more information about Bank Mega credit cards. The procedure for prospective customers who wish to apply for a credit card will greatly affect the growth and development of Bank Mega's credit card marketing. The purpose of this study is to determine whether or not prospective credit cardholders are eligible. The method used in this research is the C4.5 Algorithm method. The pattern of data generated is to determine the determination of prospective credit card owners. From the C4.5 Algorithm, the results are in the form of a decision tree, and to analyze the data using the Confusion Matrix. The purpose of this study was to find out the results scientifically to be owned by credit card owners, from a sample data of 200 taken and the results of the decision tree and the accuracy of the Confusion Matrix were 100%.

Keywords: Prospective Credit Card Holders, C4.5 Algorithm, Confusion Matrix.

ABSTRAK

PT Bank Mega Card center Kuningan, merupakan pusat dari pembuatan kartu kredit Bank mega untuk wilayah regional Jakarta, merupakan salah satu anak perusahaan dari PT CT Corpora. Peranan Bank Mega Card Center Kuningan di dalam bidang pembuatan kartu kredit memudahkan nasabah untuk mengetahui informasi lebih lanjut tentang kartu kredit Bank Mega. Prosedur calon nasabah yang ingin mengajukan kartu kredit akan sangat berpengaruh pada tumbuh kembangnya pemasaran kartu kredit Bank Mega. Tujuan dari penelitian ini untuk mengetahui layak atau tidak nya calon nasabah pemegang kartu kredit. Metode yang digunakan dalam penelitian ini adalah metode Algoritma C4.5. Pola data yang di hasilkan untuk menentukan klasifikasi kelayakan calon nasabah pemegang kartu kredit. Dari Algoritma C4.5 didapatkan hasil berupa pohon keputusan, dan untuk analisa data menggunakan *Confusion Matrix*. Tujuan penelitian ini adalah untuk mengetahui hasil secara ilmiah untuk kelayakan nasabah pemegang kartu kredit, dari data sample yang diambil sebanyak 200 dan didapatkan hasil pohon keputusan serta hasil *accuracy* dari *Confusion Matrix* sebesar 100 %.

Kata kunci: Calon Nasabah Pemegang Kartu Kredit, Algoritma C4.5, *Confusion Matrix*.

A. PENDAHULUAN

1. Latar Belakang

Pada zaman yang serba modern sekarang ini banyak dari masyarakat Indonesia menggunakan pembayaran dengan cara menggunakan kartu kredit untuk lebih memudahkan pembayaran atau ingin mencicil suatu barang tertentu, masyarakat yang berada di kalangan menengah ataupun menengah ke atas hampir semua mempunyai kartu kredit untuk melakukan pembayaran yang mereka lakukan sehari-hari. Salah satu bank di Indonesia yang menunjang adanya kartu kredit adalah Bank Mega.

Bank Mega berawal dari sebuah usaha milik keluarga bernama PT. Bank Karman yang didirikan pada tahun 1969 dan berkedudukan di Surabaya, selanjutnya pada tahun 1992 berubah nama menjadi PT. Bank Mega dan melakukan relokasi kantor pusat ke Jakarta. Seiring dengan perkembangan PT. Bank Mega pada tahun 1996 diambil alih oleh Para Grup (PT. Para Global Investindo dan PT. Para Rekan Investama), sebuah holding company milik pengusaha nasional Chairul Tanjung, selanjutnya Para Grup berubah nama menjadi CT Corpora.

Kredit itu sendiri adalah cara menjual barang atau pinjaman uang dengan pembayaran secara tidak tunai dimana pembayaran ditangguhkan atau diangsur dengan pinjaman sampai batas tertentu yang diizinkan oleh bank atau badan tertentu. Salah satu tugas utama dari sebuah lembaga keuangan adalah mengembangkan beberapa set model dan teknik untuk memungkinkan mereka untuk memprediksi kelayakan kredit.

Pengertian kartu kredit sendiri adalah alat pembayaran pengganti uang tunai yang dapat digunakan oleh konsumen untuk ditukarkan dengan barang dan jasa yang diinginkannya di tempat-tempat yang dapat menerima pembayaran dengan menggunakan kartu kredit (merchant). Kartu kredit juga dapat diartikan sebagai salah satu fasilitas dari perbankan yang memudahkan transaksi nasabah. Anda tinggal menggesek kartu kredit dan kita tinggal membayarnya saat tagihan tiba. Baik tagihan lembaran fisik yang dikirimkan ke rumah ataupun e-statement yang dikirimkan via email.

Sebagai bank yang memberikan pelayanan kartu kredit Bank Mega punya beberapa jenis kartu kredit yang memudahkan pelanggan yaitu kartu platinum, gold, metro, carrefour, tsm, fifa, mega first, barca itu adalah jenis-jenis dari kartu kredit yang dimiliki Bank Mega, dan dari beberapa jenis kartu kredit tersebut kebanyakan masyarakat mengajukan aplikasi kartu kredit yaitu platinum, gold atau carefour dikarenakan ketiga jenis kartu itu adalah kartu yang bisa digunakan oleh masyarakat umum dikarenakan ada beberapa jenis kartu yang hanya diperuntukan untuk masyarakat kalangan atas yaitu kartu mega first.

Masyarakat sekarang ini lebih memilih menggunakan pembayaran dengan cara kartu kredit, karena lebih mudah dan tidak perlu membawa uang yang berlebihan tetapi sekarang ini semua masyarakat lebih mengenal kartu kredit dan banyak yang mengajukan kartu kredit tanpa mengerti syarat yang harus dipatuhi oleh seorang calon nasabah. Maka dari itu banyak sekali kejadian penipuan bank oleh sejumlah nasabah pemegang kartu kredit karena sebelumnya calon nasabah pemegang kartu kredit tidak mengerti syarat yang diajukan oleh pihak bank atau sengaja melakukan hal penipuan untuk kepentingan nasabah sendiri. Melalui data aplikasi calon nasabah pemegang kartu kredit, dapat dimanfaatkan untuk dilakukan segmentasi kelayakan calon pemegang kartu kredit untuk mengambil sebuah keputusan layak atau tidaknya calon nasabah pemegang kartu kredit.

Pada penelitian sebelumnya Timbo Faritcan Parlaungan Siallagan (2015) telah melakukan penelitian tentang calon nasabah kredit dengan metode algoritma C4.5 untuk menilai kelayakan calon nasabah kredit dengan menggunakan rapid miner sebagai program pembantu menilai kelayakan calon nasabah kredit. Timbo Faritcan Parlaungan Siallagan dengan penelitiannya tersebut juga menggunakan prinsip 6 C's analisis yaitu character, capital, capacity, collateral dan constraint untuk lebih memudahkan menilai calon nasabah yang layak diberikan kredit atau tidak selain menggunakan rapid miner sebagai penghitungan hasil akhir.

Pada penelitian kali ini dengan menggunakan data mining algoritma C4.5 untuk melakukan kelayakan calon pemegang kartu kredit. Maka akhirnya penyusun mengambil judul untuk penelitian ini "Penerapan Sistem Metode Algoritma C4.5 untuk Menentukan Kelayakan Calon Nasabah pemegang Kartu Kredit Bank Mega Card Center Kuningan".

2. Rumusan Masalah

Saat ini proses untuk menentukan kelayak calon nasabah pemegang kartu kredit di Bank Mega Card center kuningan masih kurang efektif. Pemilihan calon nasabah pemegang kartu kredit masih menggunakan cara manual walaupun memang ada metode yang digunakan Bank Mega Card Center masih kurang tepat untuk menentukan calon nasabah pemegang kartu kredit. Karyawan Card Center pun masih mengacu pada metode yang lama untuk mengerjakan aplikasi calon nasabah, kriteria-kriteria pemegang calon nasabah harus memenuhi syarat yang di ajukin oleh Bank Mega Contoh nya harus mempunyai SID (Sistem Informasi debitur) apabila mempunyai kartu kredit atau leasing di bank lain karena data calon nasabah terdapat dalam database Bank Indonesia harus bersih atau dengan kata lain tidak ada tunggakan hutang, lalu usia calon nasabah harus mencukupi, AKKI (Asosiasi Kartu Kredit Indonesia) dari calon nasabah pun harus bersih tidak ada catatan buruk. Tetapi dari persyaratan yang di berikan oleh Bank Mega pihak *Analyst* yang menyetujui calon nasabah layak atau tidak nya mempunyai kartu kredit masih saja kesulitan untuk meyetujui calon nasabah layak atau tidak karena beberapa pihak dari Bank Mega masih bersifat pemaksaan atau subjektif dalam menentukan kelayakan calon nasabah, dari contoh yang bisa dilihat yaitu calon yang kurang memiliki kriteria pendapatan atau pendapatan kurang mencukupi dan SID bagus masih bisa diterima atau layak ataupun pendapatan lebih dan SID tidak bagus masih diberikan kelayakan oleh pihak bank. Apabila terjadi hal seperti itu calon nasabah tidak berhak diberikan kelayakan untuk mempunyai atau memegang kartu kredit. Maka dari itu dilihat dari contoh diatas masih adanya sifat subjektif. Dari contoh diatas penulis ingin menjabarkan kelayakan calon nasabah yang layak dan tidak layak dari cara sistematis dengan menggunakan Algoritma C4.5 untuk lebih memudahkan menentukan layak atau tidaknya calon nasabah pemegang kartu kredit untuk meminimaliskan kesalahan penyetujuan. Contoh data dapat dilihat ditabel 1.

Tabel 1. Tabel contoh permasalahan kelayakan Kartu Kredit

No.	No. Aplikasi	Nama	Nama Perusahaan	Pendapatan	BI Checking	Penyetujuan
1.	77.7.153 .9001.0	Mr AB	Sinode Gereja Kristen Jakarta	13.000.000	Kolek 5 atau mempunyai masalah dengan bank lain	Kartu disetujui
2.	182.7153 .9003.0	Mr CD	KNS Group	18.500.000	Kolek 4 bermasalah dengan kredit Bank lain	Kartu disetujui
3.	003.7153 .9001	Mr EF	Pitagiri Hotel	3.461.538	SID bersih	Kartu tidak disetujui karena gaji tidak sesuai kriteria
4.	252.7153 .9001	Mr OCD	KAP DRS Kartoyo dan rekan	3.384.000	Hasil pencarian SID bagus tidak ada tunggakan di Bank lain	Kartu tidak disetujui karena gaji tidak sesuai kriteria
5.	315.7156 .9011.0	Mrs COD	PT indo surya inti finance	16.250.000	Kolek 5 atau mempunyai masalah dengan bank lain	Kartu disetujui

Berdasarkan penjelasan dari latar belakang diatas, maka dapat diidentifikasi masalah sebagai berikut:

1. Sulit menentukan aturan untuk menentukan kalayakan calon nasabah pemegang kartu kredit.
2. Sulitnya memilih calon pemegang kartu kredit.

Pernyataan masalah:

Sulitnya untuk menentukan calon pemegang kartu kredit.

Pertanyaan masalah:

Bagaimana penerapan algoritma C4.5 dalam proses menentukan calon pemegang kartu kredit.

3. Tujuan

Maksud penelitian ini adalah untuk menerapkan metode algoritma C4.5 untuk menentukan kelayakan calon pemegang kartu kredit.

Adapun tujuan dari penelitian ini adalah:

1. Dapat memudahkan dalam memilih calon pemegang kartu kredit yang tepat.
2. Mendapatkan proses penyeleksian pemegang kartu kredit secara tepat.

4. Tinjauan Pustaka

a. Pengertian *Data Mining*

Data Mining adalah langkah analisis terhadap proses penemuan pengetahuan didalam basis data atau *knowledge discovery in databases* yang disingkat KDD (Kusrina dan Emha Taufiq, 2009) . pengetahuan bisa berupa pola data atau relasi antar data yang *valid* (yang tidak diketahui sebelumnya).

Secara umum, kegunaan data *mining* dapat dibagi menjadi dua: deskriptif dan prediktif. Deskriptif berarti data *mining* digunakan untuk mencari pola-pola yang dapat dipahami manusia yang menjelaskan karakteristik data. Sedangkan prediktif berarti data *mining* digunakan untuk membentuk sebuah model pengetahuan bisa yang akan digunakan untuk melakukan prediksi. Berdasarkan fungsionalitasnya, tugas-tugas *data mining* bisa dikelompokkan kedalam enam kelompok berikut ini (Kusrina dan Emha Taufiq, 2009):

- a. Klasifikasi (*Classification*): men-generasikan struktur struktur yang diketahui untuk diaplikasikan pada data-data baru. Misalkan, Klasifikasi penyakit kedalam sejumlah jenis, klasifikasi email kedalam spam atau bukan.
- b. Klasterisasi (*Clustering*): mengelompokan data, yang tidak diketahui label kelasnya kedalam sejumlah kelompok tertentu sesuai dengan ukuran kemiripannya.
- c. Regresi (*regression*): menemukan suatu fungsi yang memodelkan data dengan galat (kesalahan prediksi) seminimal mungkin.
- d. Deteksi anomali (*anomali detection*): mengidentifikasi data yang tidak umum, bisa berupa *outliner* (pencilan), perubahan atau devisiasi yang mungkin sangat penting dan perlu investigasi lebih lanjut.
- e. Pembelajaran aturan asosiasi (association rule learning) atau pemodelan kebergantungan (*dependecy modeling*): mencari relasi variabel .
- f. Perangkuman (summarization): menyediakan representasi data yang lebih sederhana, meliputi visualisasi dan pembuat laporan.

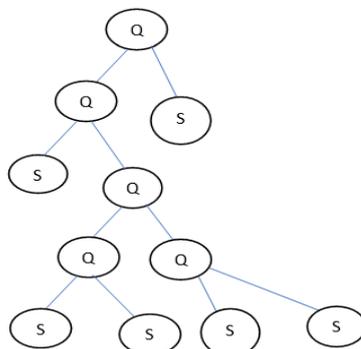
b. Pohon Keputusan

Pohon Keputusan merupakan metode klasifikasi dan prediksi yang sangat sulit dan terkenal. Metode Pohon Keputusan mengubah fakta yang sangat besar. Menjadi pohon keputusan yang mempresentasikan aturan. Aturan dapat dengan mudah dipahami dengan bahasa alami. Dan mereka juga dapat diekspresikan dalam bentuk bahasa basis data seperti *Structured Query Language* untuk mencari *record* pada kategori tertentu. Pohon keputusan juga berguna untuk mengeksplorasi data menentukan hubungan tersembunyi antara jumlah calon variabel *input* dengan sebuah variabel target.

Karena pohon keputusan memadukan antara eksplorasi data dan pemodelan, pohon keputusan sangat bagus sebagai langkah awal dalam proses pemodelan bahkan ketika dijadikan sebagai model akhir dari beberapa tehnik lain. Sebuah pohon keputusan adalah sebuah struktur yang dapat digunakan untuk membagi kumpulan data yang besar menjadi himpunan-himpunan record yang lebih kecil dengan menerapkan serangkaian keputusan. Dengan masing-masing rangkaian pembagi, anggota himpunan hasil menjadi mirip satu dengan yang lain (Berry dan Linoff, 2004).

Pada sebuah pohon keputusan setiap node merepresentasikan atribut yang telah diuji, setiap cabang merupakan suatu pembagian hasil uji, dan node daun (*leaf*) merepresentasikan kelompok tertentu. Level node teratas dari sebuah pohon keputusan adalah node akar (*root*) yang biasanya berupa atribut yang memiliki pengaruh terbesar pada suatu kelas tertentu. Pada umumnya pohon keputusan melakukan strategi pencarian secara *top-down* untuk

solusinya, Pada proses mengklasifikasi data yang tidak diketahui, nilai atribut akan diuji dengan cara melacak jalur dari node akar (*root*) sampai node akhir (daun) dan kemudian akan diprediksi kelas yang memiliki oleh suatu data baru tertentu. Proses dalam pohon keputusan (*tree*) kemudian mengubah model pohon tersebut menjadi aturan (*rule*). Model pohon keputusan ditunjukkan pada Gambar 1.



Gambar 1. Pohon Keputusan

Pohon keputusan merupakan himpunan aturan *IFTHEN*. Tahap awal dilakukan pengujian simpul akar yang ditandai dengan simbol Q (*Question*), jika pada pengujian simpul akar atau memenuhi syarat memilih yang dilambangkan dengan Y (*Yes*) dan N (*No*) lalu menghasilkan sesuatu maka proses pengujian juga dilakukan pada setiap cabang berdasarkan hasil dari pengujian. Hal ini berlaku juga untuk simpul internal dimana suatu kondisi pengujian baru akan diterapkan pada simpul daun berikutnya.

Untuk membangun pohon keputusan digunakan *algoritma*, ada beberapa teknik *algoritma* yang biasa digunakan seperti: *Algoritma Hunt*, CART (*Classification And Regression Tree*), CHAID (*Chisquared Automatic Interaction Detector*), QUEST (*Quick Unbiased Efficient Statistical Tree*), *Commercial version 4.5 (C4.5)* and *Interactive Dichotomizer version 3 (ID3)*. Masing-masing metode memiliki keunggulan dan kelemahan tersendiri tergantung penggunaan pada jenis masalah data klasifikasi (Ture *et al.*, 2009).

c. Algoritma C4.5

Algoritma C4.5 merupakan algoritma yang digunakan untuk membentuk pohon keputusan. Secara umum Algoritma C4.5 untuk membangun keputusan (*Iterative dichotomiser 3*) yang diciptakan oleh J.Ross Quinlan pada 1993. Menurut Witten Ian H, dk (2001,108) adalah sebagai berikut:

- a. Pilih atribut yang akan digunakan sebagai akar
- b. Buatlah sebuah cabang untuk setiap nilai
- c. Ulangi proses untuk setiap cabang sampai semua kasus pada cabang memiliki kelas yang sama.

Ada beberapa tahapan dalam membuat sebuah pohon keputusan dengan algoritma C4.5 yaitu:

1. Menyiapkan data training. Data Training biasanya diambil dari data history yang pernah terjadi sebelumnya dan sudah dikelompokkan kedalam kelas-kelas tertentu.
2. Menentukan akar dari pohon. Akar akan diambil dari masing atribut, nilai gain yang paling tinggi akan menjadi akar pertama. Sebelum menghitung nilai gain dari atribut, hitung dahulu nilai entropy yaitu:

$$\text{Entropy}(s) = - \sum_{i=1}^n p_i \log_2 p_i \quad (1)$$

Keterangan :

S : Keterangan kasus

N : Jumlah Partial S

Pi : Proporsi dari Si terhadap s

Entropy menyatakan Impurity suatu kumpulan objek dan digunakan untuk memilih nilai optimal untuk memecahkan node berdasarkan maksimilisasi informasi. Jika

semua objek memiliki label kelas yang sama maka entropinya adalah 0 dan akan meningkat nilai entropy hingga maksimum ketika semua kelas sama sama didistribusikan.

3. Kemudian hitung nilai gain dengan metode information gain:

$$gain(S, A) = Entropy(s) - \sum_{i=1}^n \frac{|S_i|}{|S|}$$

Keterangan:

S : Himpunan Kasus

A : Atribut

n : Jumlah Partisi Atribut A

|S_i| : Jumlah Kasus Pada Partisi ke-1

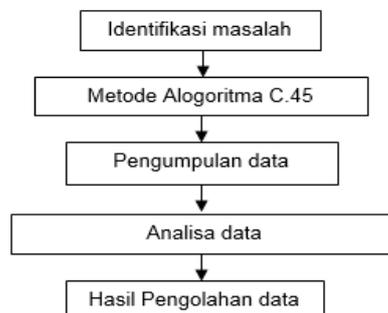
|S| : Jumlah kasus dalam s

4. Ulangi langkah ke-2 hingga semua tupel terpartisi

5. Proses Partisi pohon keputusan akan berhenti saat:
 - a. Semua tupel dalam mode N mendapat kelas yang sama
 - b. Tidak ada lagi atribut di dalam tupel yang dipartisi lagi.
 - c. Tidak ada tupel di dalam cabang kosong.

d. Kerangka Pemikiran

Pada proses pemilihan kelayakan calon pemegang kartu kredit yang baik untuk dapat dijadikan acuan dalam menentukan pemegang kartu kredit. Tetap dalam kelayakan calon pemegang kartu kredit yang layak seperti menerapkan suatu metode atau upaya untuk memenuhi syarat kelayakan calon pemegang kartu kredit. Sehingga umpan balik yang dihasilkan dapat digunakan sebagai dasar untuk pengambilan keputusan yang tepat berdasarkan kriteria yang ada di Bank Mega Card Center Kuningan yang akan di tetapkan. Metode yang digunakan untuk menganalisa dalam pemilahan kelayakan calon pemegang kartu kredit berdasarkan syarat adalah menggunakan metode Data Mining Algoritma C4.5. Metode ini pertama dengan mendefinisikan permasalahan dalam Bank Mega Center Kuningan untuk dicarikan solusi mengklasifikasikan *class*/kriteria dan atribut yang telah ditentukan untuk mencari nilai gain yang tertinggi dari setiap class/kriteria tersebut untuk membuat pohon keputusan diharapkan dengan menggunakan Algoritma C4.5 ini dapat menentukan calon pemegang kartu kredit yang layak. Kerangka dapat dilihat pada gambar 2.



Gambar 2. Kerangka Pemikiran

Gambar 2 menjelaskan tentang kerangka pemikiran dimana didalamnya terdapat.

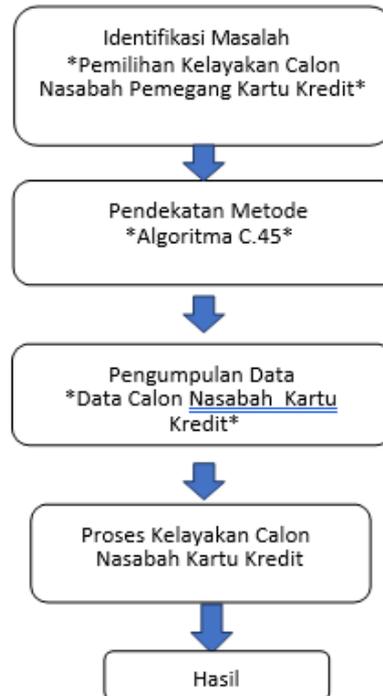
1. Identifikasi masalah, yang dilakukan dalam tahapan ini untuk menetapkan calon pemegang kartu kredit pada Bank Mega Card Center Kuningan.
2. Selanjutnya dilakukan tahap pendekatan dengan menggunakan metode Algoritma C4.5.
3. Kemudian dilakukan tahapan pengumpulan data, dimana dalam tahapan ini peneliti melakukan observasi, dokumentasi.
4. Setelah itu terdapat proses analisa data yang telah dikumpulkan akan dilakukan uji coba menggunakan teori yang telah ditentukan yaitu dengan metode Algoritma C4.5.
5. Setelah itu adalah proses terakhir yaitu hasil dimana calon nasabah sudah layak menjadi nasabah kartu kredit Bank Mega.

B. METODE

1. Desain Penelitian

Penelitian yang di lakukan pada Bank Mega Card Center Kuningan Jakarta Selatan untuk menentukan kelayakan calon nasabah pemegang kartu kredit. Dimana penelitian ini menentukan calon nasabah yang layak untuk jadi pemegang kartu kredit, yang semula hanya calon nasabah berdasarkan syarat yang dilampirkan dalam aplikasi kartu kredit.

Adapun tahapan-tahapan yang dilakukan dalam proses pemecahan suatu masalah yang dimulai dari mengidentifikasi permasalahannya terlebih dahulu, selanjutnya menentukan metode yang akan digunakan untuk meneliti, pengumpulan data, menganalisis data dengan menggunakan data yang telah dipilih, kemudian terakhir mendapatkan hasil akhir. Desain penelitian dapat dilihat dari bagian gambar 3.



Gambar 3. Desain Penelitian

Dari gambar diatas dapat dijelaskan tahapan dari desain penelitian sebagai berikut:

1. Identifikasi Masalah

Tahapan dimana melakukan identifikasi terhadap masalah yang ada, dan permasalahan yang telah ditentukan akan di carikan metode yang cocok untuk diterapkan, kemudian dilakukan penelitian atau uji coba untuk membuktikan metode tersebut.

2. Penerapan Metode

Pada penelitian ini, metode yang akan diterapkan menggunakan Algoritma C4.5.

3. Pengumpulan Data

Pada tahap ini, ada beberapa teknik yang digunakan pada pengumpulan data, salah satunya dengan dokumentasi. Guna memperoleh informasi data yang akurat dalam dalam menunjang penelitian yang sedang dilakukan dengan mendapatkan hasil yang diharapkan.

4. Proses Kelayakan Calon Nasabah

Kelayakan calon nasabah pemegang kartu kredit, dimana tahapan ini menilai kelayakan calon nasabah dari persyaratan yang diajukan hingga mengecek BI *Checking* calon nasabah.

5. Hasil Akhir

Hasil dari proses analisis dengan menggunakan metode Algoritma C4.5 adalah menentukan kelayakan calon nasabah pemegang lartu kredit menjadi nasabah yang layak menjadi pemegang kartu kredit.

2. Variable Penelitian

Dasar penetapan variabel dalam penyetujuan kredit, sudah ditetapkan dalam jurnal “ *Credit Information System*. Bank Mega, 2013”. Adalah sebagai berikut:

1. Variable dari data yang telah di tentukan yaitu SID (Sistem Informasi Debitur), Umur, Status Pekerjaan, Gaji.
2. Atribut dari Variable
 - a. SID (Sistem Informasi Debitur)
 - 1) Apakah calon pemegang mempunyai leasing di kantor Leasing
 - 2) Apakah calon pemegang mempunyai kartu kredit selain Bank Mega
 - b. Umur
 - 1) 21 s/d 65 tahun untuk kartu basic atau kartu utama
 - 2) 17 s/d 70 tahun untuk kartu supplement atau kartu tambahan
 - c. Mulai Dari Rp. 3.500.000 -,
 - d. Status Pekerjaan
 - 1) Pekerja Tetap
 - 2) Pekerja Kontrak

3. Sumber Data dan pengumpulan Data

Data yang digunakan sebagai bahan untuk menganalisis data berasal data aplikasi calon nasabah pemegang kartu kredit pada bagian analis bulan juli 2016 aplikasi dari Bank Mega Card center Kuningan Jakarta Selatan.

4. Instrumen penelitian

Menurut Suharsimi Arikunto, instrumen penelitian merupakan alat bantu yang dipilih oleh peneliti dalam melakukan kegiatan untuk mengumpulkan data agar kegiatan tersebut menjadi sistematis dan dipermudah.

- a. Data

Penggunaan data pada penelitian ini adalah data yang bersumber dari PT Bank Mega dalam menentukan kelayakan calon pemegang kartu kredit.
- b. Tools

Dalam penelitian ini *tools* yang digunakan adalah *RapidMiner*.

5. Metode Analisis Data

Berdasarkan dari data calon aplikasi calon pemegang kartu kredit yang diperoleh maka metode yang akan digunakan adalah Algoritma C4.5, serta menambahkan Confussion Matrix untuk melakukan perhitungan akurasi pada konsep data mining informasi dalam Confusion Matrix diperlukan untuk menentukan kinerja model klasifikasi ringkasan informasi ini ke dalam sebuah nilai digunakan untuk membandingkan kinerja model-model yang berbeda. Hal ini dapat dilakukan dengan menggunakan Performance Metrica (Brammer, 2007:89).

Tabel 2. Confussion Matrix

Classification	Predcited Class	
	Class = yes	Class = no
Class = Yes	a (true positive- tp)	c (false negative- fn)
Class = No	b (False postive- fp)	d (True negative- fp)

Rumus ini melakukan perhitungan dengan 4 keluaran, yaitu: recall, precision, accuracy dan error rate.

1. *Recall atau Postive rate* (tp) adalah proporsi kasus positif yang diidentifikasi dengan benar.
 Rumus dari *recall atau true postive* (tp)

$$TP = \frac{d}{(C + D)}$$

2. *Precision* (P) adalah proporsi dengan hasil positif yang benar dari *precision*(P)

$$P = \frac{d}{(b + d)}$$

3. *Accuracy (AC)* adalah perbandingan kasus yang diidentifikasi benar dengan jumlah semua kasus. Rumus dari *accuracy (AC)*

$$AC = \frac{(a + c)}{(a + b + c + d)}$$

4. *Error Rate(ER)* adalah kasus yang diidentifikasi salah dengan sejumlah semua kasus. Rumus dari *Error Rate (ER)*

$$ER = \frac{(b + c)}{(a + b + c + d)}$$

C. HASIL DAN PEMBAHASAN

1. HASIL

a. Proses Pengklasifikasian Data Mining Menggunakan Algoritma C4.5

Data dalam penelitian ini berdasarkan data calon pemegang kartu kredit yang dapat dilihat pada table 3.

Tabel 3. Daftar Calon Pemegang kartu Kredit

NO	No Aplikasi	Nama Nasabah	Nama perusahaan	Pedapatan	SID Dan Parameter	Status
1	097.7271.9005.0	Mr AA	ABC	5.000.000,00	1	Lancar
2	291.7271.9002.0	Miss BC	DEF	3.000.000,00	5	Tidak Lancar
3	327.7271.9001.0	Mrs CC	WER	35.000.000,00	1	Lancar
4	206.7271.9001.0	Miss DF	ERT	35.000.000,00	1	Lancar
5	128.7271.9001.0	Mr BB	RTY	4.000.000,00	1	Lancar
6	185.7271.9002.0	Mrs QQ	QWE	4.000.000,00	1	Lancar
7	176.7271.9001.0	Mr XX	AXE	50.000.000,00	1	Lancar
8	074.7271.9008.0	Miss HK	ERT	5.000.000,00	1	Lancar
9	185.7271.9001.0	Mrs RT	BER	5.000.000,00	1	Lancar
10	043.7271.9005.0	Mrs Gh	CFE	50.000.000,00	1	Lancar
11	202.7271.9003.0	Miss YI	MJU	30.000.000,00	1	Lancar
12	300.7271.9009.0	Mrs LL	MIK	35.000.000,00	1	Lancar
13	274.7271.9004.0	Mr w	ASD	5.000.000,00	2	Lancar
14	060.7271.9005.0	Miss A	BGT	5.000.000,00	2	Lancar
15	034.7271.9003.0	Mr W	KIO	5.000.000,00	2	Lancar
16	092.7271.9001.0	Mrs P	LOP	5.000.000,00	2	Lancar
17	146.7271.9001.0	Miss K	XDF	3.000.000,00	4	Tidak Lancar
18	260.7271.9001.0	Miss D	BGV	3.000.000,00	4	Tidak Lancar
19	153.7271.9001.0	Mrs H	ZXC	6.000.000,00	2	Lancar
20
21
22
23
24
200	117.7271.9002.0	Mr U	MNB	6.000.000,00	2	Lancar

Tahap berikutnya adalah dilakukan pengolahan data. Proses pengolahan *data mining* yang dilakukan pada penelitian ini terdiri dari beberapa tahap yaitu, tahap data *selection* dan data *transformation*.

Data yang terkumpul pada tabel 4.1 terdapat beberapa atribut yang tidak terpakai dalam proses pengolahan data. Atribut yang tidak memiliki kontribusi terhadap tujuan penelitian, tidak digunakan dalam penelitian ini. Sehingga perlu adanya *pruning* atau pemotongan atribut yang tidak dipakai dalam pengklasifikasian metode. Proses penentuan seleksi data adalah dimana proses tersebut dilakukan melalui tahap pemecahan nilai-nilai kriteria.

Rumus *IF AND* yang digunakan untuk melakukan pemecahan nilai klasifikasian yang akan dilakukan, adalah sebagai berikut:

Rumus dalam mencari klasifikasi BI Checking atau SID:

=IF(OR(B3=1;B3=2);"Lancar";IF(B3=3;"RaguRagu";IF(OR(B3=4;B3=5);"Tidak Lancar";"Check Your Data"))).

Adapun rumus *IF AND* dalam menentukan nilai hasil layak atau tidak layaknya calon nasabah dari data yang didapat adalah:

=IF(AND(C3="Lancar";E3="Kecil");"Layak";

IF(AND(C3="Lancar";E3="Sedang");"Layak";

IF(AND(C3="Lancar";E3="Besar");"layak";

IF(AND(C3="Ragu Ragu";E3="Kecil");"Tidaklayak";

IF(AND(C3="RaguRagu";E3="Sedang");"Layak";

IF(AND(C3="RaguRagu";E3="Besar");"Layak";

IF(AND(C3="Tidak Lancar";E3="Kecil");"Layak";
 IF(AND(C3="Tidak Lancar";E3="Sedang");"TidakLayak";
 IF(AND(C3="TidakLancar";E3="Besar");"TidakLayak";"CheckYourData"))))))))

Dari proses manual melakukan rumus persyaratan *IF AND* tersebut pada format Excel, maka didapat hasil pengklasifikasian sesuai dengan klasifikasi nilai-nilai kriteria yang diinginkan.

Adapun penjelasan sederhana dari proses rumus tersebut adalah:

- Syarat kondisi 1: Jika SID atau kol 1 lancar pendapatan kecil maka layak.
- Syarat kondisi 2: Jika SID atau kol 2 lancar pendapatan sedang maka layak.
- Syarat kondisi 3: Jika SID atau kol 2 lancar pendapatan besar maka layak.
- Syarat kondisi 4: Jika SID atau kol 3 ragu-ragu pendapatan kecil maka layak.
- Syarat kondisi 5: Jika SID atau kol 3ragu-ragu pendapatan sedang maka layak.
- Syarat kondisi 5: Jika SID atau kol 3ragu-ragu pendapatan besar maka layak.
- Syarat kondisi 6: Jika SID atau kol 3ragu-ragu pendapatan sedang maka layak.
- Syarat kondisi 7: Jika SID atau kol 4tidak lancar pendapatan sedang maka tidak layak.
- Syarat kondisi 8: Jika SID atau kol 4tidak lancar pendapatan besar maka tidak layak.
- Syarat kondisi 9: Jika SID atau kol 5 tidak lancar pendapatan sedang maka tidak layak.

Berdasarkan kriteria diatas dengan menggunakan rumus persyaratan *IF* maka diperoleh hasil penyeleksian data, yang dapat dilihat pada tabel 4.

Tabel 4. Daftar Calon Pemegang kartu Kredit

NO	SID	KLASIFIKASI	Pendapatan	KLASIFIKASI	HASIL
		SID		Pendapatan	
1	1	Lancar	6.000.000,00	Sedang	layak
2	1	Lancar	6.000.000,00	Kecil	Layak
3	1	Lancar	6.000.000,00	Besar	layak
4	1	Lancar	6.000.000,00	Besar	layak
5	1	Lancar	5.000.000,00	Kecil	Layak
6	1	Lancar	35.000.000,00	Kecil	Layak
7	5	Tidak Lancar	3.000.000,00	Besar	Tidak Layak
8	5	Tidak Lancar	3.000.000,00	Sedang	Tidak Layak
9	2	Lancar	35.000.000,00	Sedang	Layak
10	2	Lancar	35.000.000,00	Besar	layak
11	4	Tidak Lancar	3.000.000,00	Besar	Tidak Layak
12	1	Lancar	35.000.000,00	Besar	layak
13	2	Lancar	35.000.000,00	Sedang	Layak
14	2	Lancar	35.000.000,00	Sedang	Layak
15	2	Lancar	35.000.000,00	Sedang	Layak
16	2	Lancar	6.000.000,00	Sedang	Layak
17	2	Lancar	6.000.000,00	Kecil	Layak
18	2	Lancar	6.000.000,00	Kecil	Layak
19	1	Lancar	6.000.000,00	Sedang	Layak
20	1	Lancar	6.000.000,00	Sedang	Layak
21
22
23	1	Lancar	35.000.000,00	Sedang	Layak

2. Penentuan Atribut Dalam Penelitian

Variabel yang digunakan dalam penentuan kelayakan calon nasabah adalah SID, dan pendapatan yang terdapat dari jumlah banyaknya calon nasabah pemegang kartu kredit.

a. Kolektibilitas

Atribut SID dapat dilihat pada tabel 5.

Tabel 5. Atribut SID

Atribut	Kriteria
SID	Lancar
	Ragu -Ragu
	TidakLancar

Pada PT Bank Mega Card Center Kuningan kriteria yang digunakan untuk penelitian hanya 2 kategori, yaitu melalui kriteria lancar dan tidak lancar. Berikut penjelasan dari atribut SID:

- (1) Kol 1 dan 2 adalah SID yang memiliki kelancaran angsuran. Hasil SID tersebut yang bisa menentukan hasil layak.
- (2) Kol 3 adalah SID yang memiliki kelancaran atau ragu-ragu angsuran. Tetapi hasil SID tersebut masih bisa menentukan hasil layak.
- (3) Kol 4 dan 5 adalah SID yang memiliki angsuran tidak lancar. Hasilnya pun sudah diketahui tidak layak.

3. Tabel Rekapitulasi Calon Nasabah Pemegang Kartu Kredit

Dari daftar tabel rekapitulasi Calon Nasabah pemegang Kartu Kredit yang sudah di klasifikasi ada 200 nama calon nasabah yang diambil tahun 2017 pada periode bulan juli dan september, yang nantinya akan diteliti layak atau tidaknya calon nasabah pemegang kartu kredit, untuk perhitungan sendiri ada beberapa cara yaitu dengan menghitung nilai entropy kasus, menghitung nilai entropy dan gain lalu setelah kita mengetahui nilai entropy dan gain di lakukan cara proses pembuatan pohon keputusan dengan *RapidMiner* setelah kita mengetahui pohon keputusan dari *RapidMiner* dilanjutkan dengan menghitung *accuracy* dengan menggunakan *confusion matrix*, berikut ini adalah langkah-langkah cara perhitungannya:

Tabel 6. Tabel Rekapitulasi Calon Pemegang kartu Kredit

NO	SID	KLASIFIKASI	Gaji	KLASIFIKASI	HASIL
		SID		Pendapatan	
1	1	Lancar	5.000.000,00	Sedang	Layak
2	5	Tidak Lancar	3.000.000,00	Kecil	Layak
3	1	Lancar	35.000.000,00	Besar	layak
4	1	Lancar	35.000.000,00	Besar	layak
5	1	Lancar	4.000.000,00	Kecil	Layak
6	1	Lancar	4.000.000,00	Kecil	Layak
7	1	Lancar	50.000.000,00	Besar	layak
8	1	Lancar	5.000.000,00	Sedang	Layak
9	1	Lancar	5.000.000,00	Sedang	Layak
10	1	Lancar	50.000.000,00	Besar	layak
11	1	Lancar	30.000.000,00	Besar	layak
12	1	Lancar	35.000.000,00	Besar	layak
13	2	Lancar	5.000.000,00	Sedang	Layak
14	2	Lancar	5.000.000,00	Sedang	Layak
15	2	Lancar	5.000.000,00	Sedang	Layak
16	2	Lancar	5.000.000,00	Sedang	Layak
17	4	Tidak Lancar	3.000.000,00	Kecil	Layak
18	4	Tidak Lancar	3.000.000,00	Kecil	Layak
19	2	Lancar	6.000.000,00	Sedang	Layak
20	2	Lancar	6.000.000,00	Sedang	Layak
21
22
23
24
200	2	Lancar	6.000.000,00	Sedang	Layak

4. Menghitung Nilai Entorpy

Entropy adalah merupakan jumlahan perkalian antara frekuensi *sampel* dengan *information gain* tiap atribut.

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

Keterangan :

S : Keterangan kasus

N : Jumlah Partial S

Pi : Proporsi dari Si terhadap s

Nilai *Entropy* kasus adalah hasil perhitungan dari jumlah seluruh kasus data nasabah calon pemegang kartu kredit yang mencantumkan data yaitu sejumlah 200 kasus. Tabel 7 adalah hasil perhitungan dengan menggunakan rumus entropy.

Tabel 7. Hasil Perhitungan *Entropy* Kasus

Hitungan Entropy Kasus		
Kasus	Layak	Tidak layak
	P	N
200	180	20

Kasus	p/(p+n)	Log 2	n/(P+n)	Log 2	Entropi
200	180/200	-0,152003093	20/200	-	0,468
	0,9		0,1	3,32193	
(-(0.9)x(-0.152)-(0.1)x(-3.321))					

Kesimpulan dari Tabel 7 Hasil Perhitungan Entropy Kasus yaitu peneliti dapat mengetahui jumlah dari entropy yang bisa dilihat dari tabel bahwa dari 200 kasus Calon Nasabah yang layak dan tidak layak ada 180 dan 20, maka dihitung dengan rumus dan mendapatkan hasil entropy kasus yaitu 0,468.

5. Menghitung Nilai *Entropy* dan *Gain*

Gain adalah merupakan nilai pengurangan dalam *entropy* yang disebabkan karena diketahuinya nilai suatu atribut.

Rumus *Gain*:

$$gain(S, A) = Entropy(s) - \sum_{i=1}^n \frac{|S_i|}{|S|} \log_2 \frac{|S_i|}{|S|}$$

Keterangan:

- S : Himpunan Kasus
- A : Atribut
- n : Jumlah Partisi Atribut A
- |S_i| : Jumlah Kasus Pada Partisi ke-1
- |S| : jumlah kasus dalam s

6. Ulangi langkah ke-2 hingga semua tupel terpartisi.

7. Proses Partisi pohon keputusan akan berhenti saat:

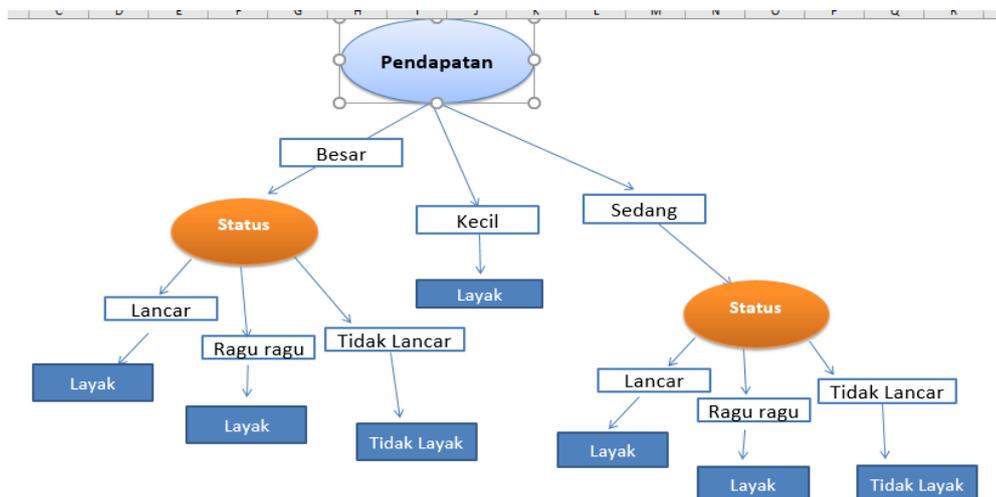
- a. Semua tupel dalam mode N mendapat kelas yang sama
- b. Tidak ada lagi atribut di dalam tupel yang dipartisi lagi.
- c. Tidak ada tupel di dalam cabang kosong.

Tabel 8. Hasil Perhitungan Entropy dan Gain

Cabang			Layak	Tidak	P Layak	P Tidak	P Setuju x log2 P Setuju	P Tidak x log2 P Tidak	Total	P	Total x p	E Total * p	Gain
1	SID	Lancar	148	0	0	0	0	0	0	0,74	0	0,205	0,107
		Ragu-ragu	22	0	0,5	0	0,5	0	0,5	0,11	0,055		
		Tidak Lancar	10	20	0,5	0,5	0,5	0,5	1	0,15	0,15		
2	pendapatan	Kecil	15	0	0,5	0	0,5	0	0,5	0,075	0,0375	0,9625	0,494
		Sedang	109	5	0,5	0,5	0,5	0,5	1	0,57	0,57		
		Besar	56	15	0,5	0,5	0,5	0,5	1	0,355	0,355		

Kesimpulan dari Tabel 8. Hasil Perhitungan Entropy dan Gain yaitu peneliti dapat mengetahui jumlah dari entropy yang bisa dilihat dari tabel bahwa dari 200 kasus Calon Nasabah yang layak dan tidak layak ada 180 dan 20, dan menghasilkan jumlah entropy sebesar 0,468, lalu dilakukan

perhitungan entropy dan gain dari setiap cabang yang ada yaitu cabang SID dan Pendapatan maka dapat dilihat hasilnya dari masing masing cabang yaitu 0,107 dan 0,494. Berikut ini adalah Pohon Keputusan yang diambil dari perhitungan entropy kasus, entropy dan gain.



Gambar 4. Pohon Keputusan 1

Berikut ini penjelasan Pohon keputusan pada gambar 4:

1. Jika Pendapatan nya besarterlebih dahulu dilihat status SID nya apabila lancar maka nasabah layak untuk mendapatkan kartu kredit.
2. Jika Pendapatan nya sedang terlebih dahulu dilihat status SID nya apabila lancar maka nasabah layak untuk mendapatkan kartu kredit.
3. Jika Pendapatan nya kecil maka status nya lancar dan calon nasabah layak.
4. Jika Pendapatan nya besar terlebih dahulu dilihat status SID nya apabila ragu-ragu maka nasabah layak untuk mendapatkan kartu kredit.
5. Jika Pendapatan nya sedang terlebih dahulu dilihat status SID nya apabila ragu-ragu maka nasabah layak untuk mendapatkan kartu kredit.
6. Jika Pendapatan nya besar terlebih dahulu dilihat status SID nya apabila tidak lancar maka nasabah tidak layak untuk mendapatkan kartu kredit.
7. Jika Pendapatan nya sedang terlebih dahulu dilihat status SID nya apabila tidak lancar maka nasabah tidak layak untuk mendapatkan kartu kredit.

Kesimpulan yang ditampilkan dari pohon keputusan secara manual didapat 7 buah keputusan untuk menentukan layak atau tidak layak calon nasabah.

8. Proses Pembuatan Pohon Menggunakan *RapidMiner 5*

Untuk pembuatan Pohon Decision Tree C4.5 menggunakan *software RapidMiner 5*. *RapidMiner 5* adalah salah satu software untuk pengolahan *data mining*. Pekerjaan yang dilakukan oleh *RapidMiner* text mining adalah berkisar dengan analisis teks, mengekstrak pola-pola dari data set yang besar dan mengkombinasikannya dengan metode statistika, kecerdasan buatan, dan database. Tujuan dari analisis teks ini adalah untuk mendapatkan informasi bermutu tertinggi dari teks yang diolah.

RapidMiner menyediakan prosedur data mining dan machine learning, di dalamnya termasuk: *ETL (extraction, transformation, loading)*, data preprocessing, visualisasi, modelling dan evaluasi. Proses data mining tersusun atas operator-operator yang nestable, dideskripsikan dengan XML, dan dibuat dengan GUI. Penyajiannya dituliskan dalam bahasa pemrograman Java. Dan ini adalah beberapa langkah yang harus diselesaikan untuk mengetahui pohon keputusan menggunakan *RapidMiner 5*:

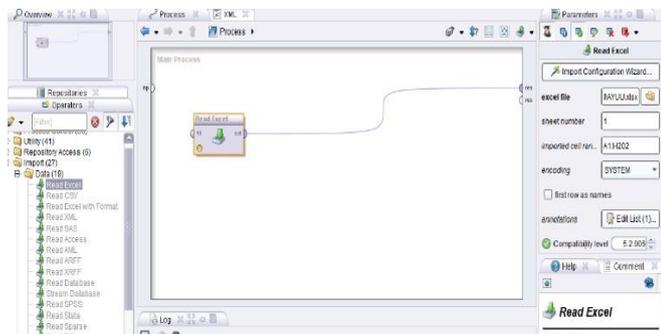
a. Menambah Repositories baru

Repositories baru dibuat dari data awal dari data excel yang berisi data nasabah, tiap atribut memiliki tipe data untuk software RapidMiner 5. Atribut pinjaman tipe data yang digunakan adalah polynominal, atribut nomor aplikasi digunakan, atribut nama nasabah polynominal (pernyataan matematika yang melibatkan jumlahan perkalian pangkat dalam satu atau lebih variabel dengan koefisien), atribut nama perusahaan integer, atribut pendapatan integer, atribut SID dan Parameter Integer, atribut status polynominal.

NO	No Aplikasi	Nama Nasa	Nama perus	Pedapatan	SID Dan Par	Status	Hasil
1	097.7271.90	Mr AA	ABC	5000000	1	Lancar	Layak
2	291.7271.90	Miss BC	DEF	3000000	5	Lancar	Tidak Layak
3	327.7271.90	Mrs CC	WER	35000000	1	Lancar	Layak
4	206.7271.90	Miss DF	ERT	35000000	1	Lancar	Layak
5	128.7271.90	Mr BB	RTY	4000000	1	Lancar	Layak
6	185.7271.90	Mrs QQ	QWE	4000000	1	Lancar	Layak
7	176.7271.90	Mr XX	AXE	50000000	1	Lancar	Layak
8	074.7271.90	Miss HK	ERT	5000000	1	Lancar	Layak
9	185.7271.90	Mrs RT	BER	5000000	1	Lancar	Layak
10	043.7271.90	Mrs Gh	CFE	50000000	1	Lancar	Layak
11	202.7271.90	Miss YI	MJU	30000000	1	Lancar	Layak
12	300.7271.90	Mrs LL	MIK	35000000	1	Lancar	Layak
13	274.7271.90	Mr w	ASD	5000000	2	Lancar	Layak
14	060.7271.90	Miss A	BGT	5000000	2	Lancar	Layak
15	034.7271.90	Mr VV	KIO	5000000	2	Lancar	Layak

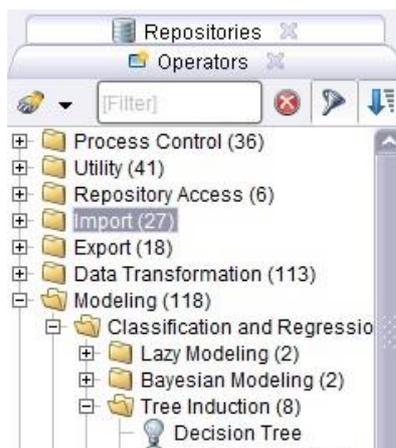
Gambar 5. Import Data

Setelah data yang kita punya dibuat dalam bentuk tabel format xls, selanjutnya lakukan Importing Data kemudian tarik atau drag pada main process kemudian klik Import Configuration Wizard untuk import data.



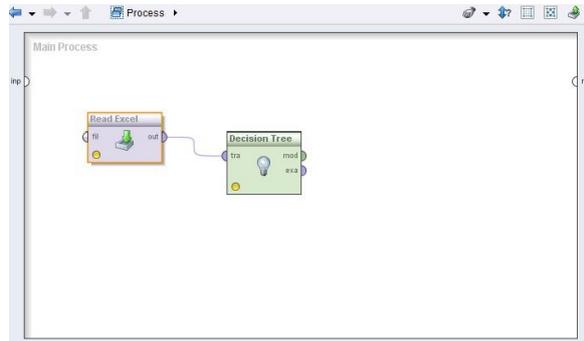
Gambar 6. Main Process

Untuk membuat Decision Tree dengan menggunakan RapidMiner, kita membutuhkan operator Decision Tree, operator ini terdapat pada view Operators. Untuk menggunakannya pilih Modelling pada view Operators, lalu pilih *Classification and regression*, lalu pilih Tree Induction dan pilih Decision Tree.



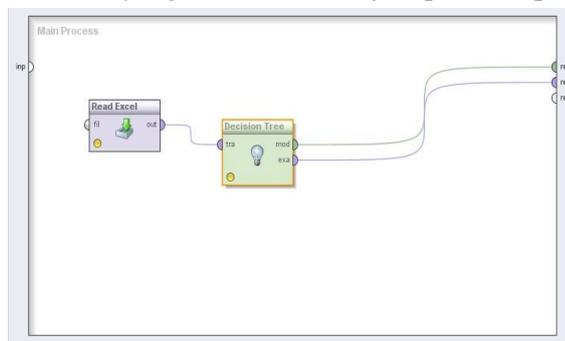
Gambar 7. Daftar Operator pada View Operators

Setelah menemukan operator Decision Tree, seret (drag) operator tersebut lalu letakan (drop) ke dalam *view Process*. Kemudian susun posisinya disamping operator, seperti yang tampak pada Gambar 7.



Gambar 8. Posisi operator Decision Tree

Selanjutnya, hubungkan Operator *Decession Tree* dengan menarik garis dari tabel ke operator *Decession Tree* dan menarik garis lagi dari operator *Decession Tree* ke result sisi kanan, seperti yang tampak pada Gambar 7 Operator *Decession Tree* berguna untuk memprediksikan keputusan dari atribut-atribut yang dimasukkan ke dalam operator, Dengan mengubah tabel (atribut) yang dimasukkan menjadi pohon keputusan.



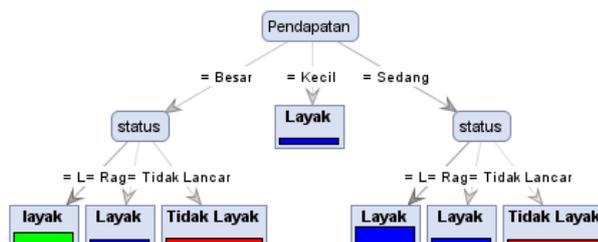
Gambar 9. Menghubungkan Table dengan Operator Decision Tree

Setelah parameter diatur, klik ikon *Run* pada *toolbar*, seperti pada gambar 8. untuk menampilkan hasilnya. Tunggu bebarapa saat komputer membutuhkan waktu menyelesaikan perhitungan.



Gambar 10. Ikon Run

Setelah beberapa detik maka *RapidMiner* akan menampilkan hasil keputusan pada *view Result*. Jika pilih *Graph View*, maka akan ditampilkan hasilnya berbentuk pohon keputusan.



Gambar 11. Hasil Graphic View Decision Tree

Rule pada gambar 11 yang nantinya implementasi program. Berikut ini penjelasan Algoritma C4.5 pada gambar 11.

1. Jika Pendapatan nya besar terlebih dahulu dilihat status SID nya apabila lancar maka nasabah layak untuk mendapatkan kartu kredit.
2. Jika Pendapatan nya sedang terlebih dahulu dilihat status SID nya apabila lancar maka nasabah layak untuk mendapatkan kartu kredit.
3. Jika Pendapatan nya kecil maka status nya lancar dan calon nasabah layak.
4. Jika Pendapatan nya besar terlebih dahulu dilihat status SID nya apabila ragu-ragu maka nasabah layak untuk mendapatkan kartu kredit.
5. Jika Pendapatan nya sedang terlebih dahulu dilihat status SID nya apabila ragu-ragu maka nasabah layak untuk mendapatkan kartu kredit.
6. Jika Pendapatan nya besar terlebih dahulu dilihat status SID nya apabila tidak lancar maka nasabah tidak layak untuk mendapatkan kartu kredit.
7. Jika Pendapatan nya sedang terlebih dahulu dilihat status SID nya apabila tidak lancar maka nasabah tidak layak untuk mendapatkan kartu kredit.

Selain menampilkan hasil decision tree berupa *graphic* atau tampilan pohon keputusan, *RapidMiner* juga menyediakan *tool* untuk menampilkan hasil berupa teks view dengan mengklik button *Text View* seperti yang tampak pada Gambar 12.

```

Tree

Pendapatan = Besar
| status = Lancar: layak {Layak=0, layak=49, Tidak Layak=0}
| status = Ragu Ragu: Layak {Layak=7, layak=0, Tidak Layak=0}
| status = Tidak Lancar: Tidak Layak {Layak=0, layak=0, Tidak Layak=15}
Pendapatan = Kecil: Layak {Layak=15, layak=0, Tidak Layak=0}
Pendapatan = Sedang
| status = Lancar: Layak {Layak=94, layak=0, Tidak Layak=0}
| status = Ragu Ragu: Layak {Layak=15, layak=0, Tidak Layak=0}
| status = Tidak Lancar: Tidak Layak {Layak=0, layak=0, Tidak Layak=5}
    
```

Gambar 12. Hasil *Text View Decision tree*

2. PEMBAHASAN

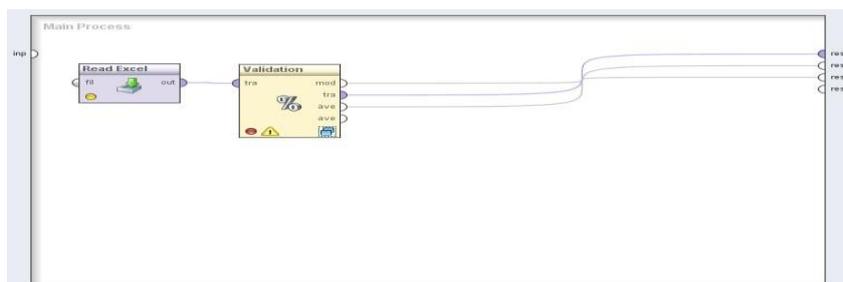
a. Pembahasan pembuatan Akurasi Matrix Pada RapidMiner

1. Pemrosesan Awal

Tahap pertama proses data mining adalah memahami domain aplikasi untuk mengetahui dan menggali pengetahuan awal serta sasaran pengguna. Dalam pemrosesan ini yang menjadi sasarannya adalah data calon pemegang kartu kredit pada bulan juli dan september 2017. Pembuatan model diawali dengan pembacaan file data (*read excel*). Data training dari data testing disimpan dala satu file excel 2016.

2. Validation

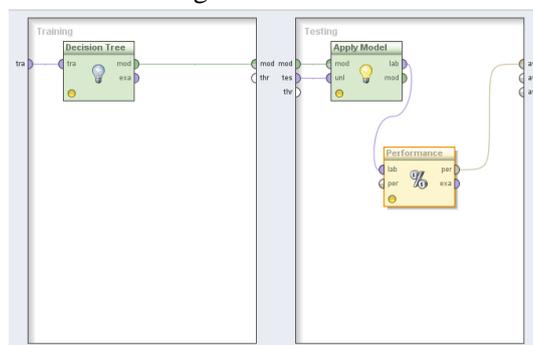
Melakukan validation yaitu melakukan analisi berbagai model dan memilih model dengan kinerja prediksi yang baik. Pada gambar 13 merupakan proses validasi, setelah pembacaan file data, blok *read excel* dihubungkan dengan blok validation. Proses data training dan testing didalam blok proses (validation) dapat ditampilkan dengan melakukan klik 2x menggunakan mouse sehingga muncul proses training dan testing.



Gambar 13. Validation

3. Proses Training dan testing

Proses training yaitu melakukan proses pelatihan data pada model (*decision tree*). Sedangkan proses testing yaitu melakukan pengujian data yang akan menghasilkan grafik atau pola. Pada gambar 14 dijelaskan bahwa proses training digunakan untuk blok model *decision tree*, *decision tree*, dihubungkan dengan garis penghubung pada blok *apply model* dan blok *performance* dengan bagian testing sebagai testing penampilan informasi hasil dari pengujian data, hasil pengujian ini akan menghasilkan arsitektur *decision tree*. Dengan metode *Decision Tree* atau Algoritma C4.5.



Gambar 14. Training dan Testing

4. Pembahasan Hasil Confusion Matrix

Langkah selanjutnya mempresentasikan hasil dari evaluasi dengan menggunakan *confusion matrix*, hasil pengklasifikasian secara benar dengan pengajuan data set terdiri dari dua kelas, kelas pertama dianggap *postive*, dan kelas kedua dianggap *negative*. Evaluasi menggunakan *confusion matrix* menghasilkan nilai *accuracy*, akurasi dalam klasifikasi merupakan presentase ketepatan record data yang diklasifikasi secara benar setelah dilakukan pengujian data pada hasil klasifikasi.

accuracy: 98.33%				
	true Layak	true layak	true Tidak Layak	class precision
pred Layak	36	1	0	97.30%
pred layak	0	18	0	100.00%
pred. Tidak Layak	0	0	5	100.00%
class recall	100.00%	94.74%	100.00%	

Gambar 15. Hasil Accuracy C4.5

Berdasarkan pada gambar 15 terdapat data testing sebanyak 200 data dengan 3 atribut yaitu limit, kolektibilitas dan pendapatan yang dimodelkan dengan Algoritma C4.5 diperoleh hasil sebagai berikut:

1. Nilai *Accuracy* menggunakan Algoritma C4.5 diperoleh sebesar 98,33%, dengan hasil kelayakan nasabah sebanyak 36 nasabah kartu kredit.
2. Nilai *Recall* yang diperoleh sebesar 94,74 % dalam menemukan jumlah data.
3. Nilai *precision* yang diperoleh dari confusion matrix sebesar 97,30%.
4. Nilai *Error rate* yang diperoleh dari confusion matrix adalah sebesar 0 %.

D. KESIMPULAN

Dari hasil klasifikasi yang dilakukan pada Rekapitulasi Calon Nasabah Pemegang Kartu Kredit Bank Mega Kuningan, ini dapat disimpulkan sebagai berikut:

1. Pola data yang di hasilkan untuk menentukan klasifikasi Calon Pemegang Kartu Kredit dari hasil *Decision Tree* diperoleh 7 aturan.
2. Dan dapat mengetahui hasil klasifikasi Calon Pemegang Kartu Kredit dari pohon keputusan, maka dari 200 nasabah calon pemegang kartu kredit 160 yang disetujui dan 40 yang tidak disetujui.

3. Perhitungan nilai akurasi penentuan kelayakan calon pemegang kartu kredit dari data yang diukur dengan menggunakan *Confusion Matrix* di aplikasi *Rapid Miner 5* diperoleh nilai akurasi sebesar 98,33%.

E. DAFTAR PUSTAKA

- [1] Fitriani, Linda M., and Andik Setyono. "Penerapan Algoritma C4.5 Untuk Seleksi Penerimaan Siswa Baru Pada SD Islam Terpadu Permata Bunda Demak Implementation of Decision Tree Algorithm for Selection of New Student Admission on Permata Bunda Integrated Islamic Elementary School." *Creative Communication and Innovative Technology Journal*, vol. 11, no. 2, 2018, pp. 158-170.
- [2] Irmayansyah, Irmayansyah, and Aulia A. Firdaus. "Penerapan Algoritma C4.5 untuk Klasifikasi Penentuan Penerimaan Bantuan Langsung di Desa Ciomas." *Teknois*, vol. 8, no. 1, May. 2018, pp. 17-28, doi:10.36350/jbs.v8i1.18.
- [3] Irmayansyah, I., & Kastrilia, M. (2020). Penerapan Algoritma C4.5 untuk Prediksi Mahasiswa Berpotensi Lulus Tidak Tepat Waktu. *Teknois : Jurnal Ilmiah Teknologi Informasi dan Sains*, 10(2), 9-18. doi:https://doi.org/10.36350/jbs.v10i2.82
- [4] Irmayansyah, I., & Lastrini, E. (2021). Penerapan Metode Algoritma C4.5 Untuk Prediksi Mahasiswa Non Aktif. *Teknois : Jurnal Ilmiah Teknologi Informasi dan Sains*, 11(2), 95-106. doi:https://doi.org/10.36350/jbs.v11i2.119
- [5] Nanda, Adi P., et al. "Menentukan Tingkat Kepuasan Mahasiswa terhadap Pelayanan Akademik Menggunakan Metode Algoritma K-Means." *Explore: Jurnal Sistem Informasi dan Telematika*, vol. 11, no. 1, Jun. 2020, pp. 23-28, doi:10.36448/jsit.v11i1.1455.
- [6] Prabowo, Indra M., and Subiyanto Subiyanto. "Recommendation System for Vocational Major Streaming by C4.5 Algorithm." *Jurnal Kependidikan: Penelitian Inovasi Pembelajaran*, vol. 1, no. 1, Jun. 2017, pp. 139-149, doi:10.21831/jk.v1i1.8964.
- [7] Rahmawan, Hanif. "Penentuan Rekomendasi Pelatihan Pengembangan Diri Bagi Pegawai Negeri Sipil Menggunakan Algoritma C4.5 dengan Principal Component Analysis dan Diskritisasi." *Jurnal Tekno Kompak*, vol. 14, no. 1, 16 Feb. 2020, pp. 5-10, doi:10.33365/jtk.v14i1.531.
- [8] Turnip, Syarif M., and Parasian D. Silitonga. "Analisis Pola Penyebaran Penyakit dengan Menggunakan Algoritma C4.5." *Jurnal Teknik Informatika Unika Santo Thomas*, vol. 3, no. 1, 2018, pp. 1-5, doi:10.17605/jti.v3i1.238.