

## Aplikasi Data Mining Dengan Menggunakan Metode *Decision Tree* Untuk Prediksi Penentuan Resiko Kredit

Suryadi Hozeng\*<sup>1</sup>, <sup>2</sup> Sitti Aisa

STMIK Dipanegara

Jln. Perintis Kemerdekaan KM.9 Makassar, Telp. (0411) 587194 – Fax. (0411) 588284

e-mail : \*<sup>1</sup>[suryadi\\_hozeng@hotmail.com](mailto:suryadi_hozeng@hotmail.com), <sup>2</sup> [sittiaisa.11@gmail.com](mailto:sittiaisa.11@gmail.com)

### Abstrak

Kredit merupakan salah satu mekanisme pembayaran yang sangat umum di masyarakat. Fungsi pokok kredit yaitu memenuhi pelayanan terhadap kebutuhan masyarakat dalam rangka memperlancar perdagangan, produksi dan jasa-jasa bahkan konsumsi yang kesemuanya itu ditujukan untuk meningkatkan kesejahteraan manusia. PT. Bank BRI merupakan salah satu pelayanan keuangan yang bergerak di bidang perkreditan di kota Makassar dan telah memberikan kemudahan pembayaran kepada pelanggannya dengan menyediakan fasilitas kredit. Ada risiko yang harus dihadapi oleh perusahaan yaitu tidak tepat waktunya pembayaran atau bahkan kegagalan pembayaran dari kredit yang disalurkan. Kredit macet ini sangat berpengaruh terhadap aliran kas perusahaan. Usaha untuk mengurangi jumlah kredit yang macet harus selalu dilakukan. Banyak faktor yang dapat menyebabkan kredit macet, salah satunya adalah kesalahan dalam pengambilan keputusan penerimaan pemohon menjadi debitur perusahaan pada tahap evaluasi kredit. Disamping kesalahan dengan menerima pemohon yang buruk, evaluasi kredit juga memungkinkan membuat kesalahan dengan menolak pemohon yang baik. Oleh karena itu setiap perusahaan yang memberikan kredit harus mampu mengevaluasi pemohon kredit dengan objektif, akurat, dan konsisten. Tool yang dapat membantu dalam mengevaluasi kredit dengan objektif, akurat, dan konsisten adalah *credit scoring*. Penelitian ini berfokus pada pembangunan model *credit scoring* kredit bank BRI dengan teknik *decision tree*. Teknik *decision tree* dapat membangun model dengan objektif, menghasilkan model yang mudah dipahami, dan tingkat akurasi yang tinggi. Model yang dibangun melibatkan variabel target yaitu risiko kredit dan variabel prediktor yaitu penghasilan, cicilan, uang muka, jumlah periode pinjaman, rekening tabungan, umur pemohon, tagihan telepon, dan tagihan listrik. Dari hasil pengembangan model diperoleh bahwa variabel penghasilan merupakan variabel yang paling penting dalam memprediksi risiko kredit dan tingkat akurasi model sebesar 79.57 persen.

**Kata kunci:** *credit scoring; data mining; decision tree; kredit; variabel prediktor; variabel target*

### Abstract

*Credit is a payment mechanism which is very common in the community. A major function of credit that meets the needs of service to the community in order to facilitate trade, production and services and even consumption, all of which was intended to improve human welfare. PT. Bank BRI is one of the financial services engaged in credit in the city of Makassar and has provided payment convenience to its customers by providing credit facilities. There is a risk that must be faced by the company is not timely payments or even failure of payment of the distributed credit. Bad credit is very influential on the company's cash flow. Efforts to reduce the amount of credit must always be done jammed. Many factors can lead to bad debts, one of which is a mistake in the decision admission the applicant becomes a debtor company in evaluating the credit. Besides the mistake of accepting the applicant's bad credit evaluation also allows making a mistake by refusing the applicant's good. Therefore, any company that provides credit loan applicant must be able to evaluate objectively, accurately, and consistently. Tool that can assist in evaluating the credit with an objective, accurate, and consistent is credit scoring. This study focuses on the development of credit scoring models with the BRI bank credit decision tree technique. Techniques can build a decision tree model with the objective to produce a model that is easy to understand, and a high degree of accuracy. Models are constructed involve target variables such as credit risk and predictor variables, namely income, mortgage, down payment, the amount of the loan period, the savings account, the age of the applicant, telephone bills and electricity bills. From the development of the model is obtained that the variable income is the most important variable in predicting credit risk and model accuracy rate of 79.57 percent.*

**Keywords:** *credit scoring; data mining; decision tree; credit; predictor variables; the target variable*

## 1. PENDAHULUAN

Dewasa ini permintaan kredit melalui Bank sudah berkembang dengan sangat pesat. Kredit bukan hanya digunakan bagi masyarakat golongan menengah ke bawah saja melainkan oleh semua lapisan masyarakat untuk memenuhi kebutuhan hidup mereka. Bank Rakyat Indonesia (BRI) adalah salah satu Bank yang dipercaya oleh pemerintah untuk memberikan fasilitas kredit kepada masyarakat. Semakin tingginya minat masyarakat untuk mendapatkan kredit, membuat pihak Bank kesulitan dalam menentukan siapa yang layak menerima kredit atau tidak. Selain itu, proses penentuan siapa yang layak menerima kredit masih dilakukan secara manual dengan menganalisa berkas-berkas pemohon, sehingga kurang efisien dalam pelaksanaannya. Oleh karena itu, penulis berinisiatif untuk merancang suatu sistem yang dapat membantu pihak Bank BRI dalam menentukan siapa yang layak menerima kredit, sehingga dapat lebih efisien dalam pelaksanaannya.

Bank Rakyat Indonesia (BRI) menyediakan fasilitas kredit bagi para nasabahnya diantaranya kredit usaha rakyat, kredit modal kerja, kredit investasi, kredit agunan kas, KUR TKI BRI dan sebagainya. Pengajuan kredit bisa diajukan oleh para pemohon dengan memenuhi syarat-syarat yang ditentukan oleh Bank Rakyat Indonesia (BRI). Pemberian kredit menjadi masalah yang krusial pada tahap evaluasi, karena beresiko menyebabkan kredit macet.

Pada penelitian ini akan mengembangkan *data mining* untuk pembuatan *scoring credit* pada tahap evaluasi yang dapat digunakan sebagai alat penyaringan untuk memisahkan pemohon yang baik dan yang buruk. Dengan menggunakan *data mining* diharapkan membantu proses evaluasi kredit yang dilakukan oleh Bank Rakyat Indonesia.

Tujuan penelitian ini adalah Melakukan *preprocessing data* agar sesuai dengan kebutuhan *format data mining*, pembuatan model untuk *Application Scoring* Bank Rakyat Indonesia untuk menentukan pemohon kartu kredit ke dalam kelas *application status approve* (pemohon dianggap layak karena dianggap tidak berpotensi melakukan kredit macet) atau *reject* (pemohon dianggap tidak layak menerima kredit karena dianggap dapat berpotensi melakukan kredit macet), serta Pembuatan *user interface* dari hasil pemodelan *data mining*.

Data Mining adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan di dalam *database*. Data mining<sup>1</sup> adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terakut dari berbagai *database* besar. Data mining juga di artikan sebagai suatu proses menemukan hubungan yang berarti, pola dan kecenderungan dengan memeriksa dalam sekumpulan besar data yang tersimpan dalam penyimpanan, dengan menggunakan teknik pengenalan pola seperti teknik statistik dan matematika.[5]

Data mining dibagi menjadi beberapa kelompok berdasarkan tugas yang dapat dilakukan, yaitu :

1. Deskripsi
 

Menggambarkan sekumpulan data secara ringkas. Data yang digambarkan berupa:

  - a. Deskripsi grafis : diagram titik, histogram
  - b. Deskripsi lokasi : *mean* (rata-rata), *median* (nilai tengah), modus, kuartil, persentil
  - c. Deskripsi keberagaman : *range* (rentang), varians dan standar deviasi
2. Estimasi
 

Memperkirakan suatu hal dari sejumlah *sample* yang kita miliki (yang tidak kita ketahui). Estimasi hampir sama dengan klasifikasi, kecuali *variable* target. Estimasi lebih kearah *numeric* dari pada kearah kategori.
3. Prediksi
 

Prediksi hampir sama dengan klasifikasi dan estimasi, kecuali bahwa dalam prediksi nilai dari hasil akan ada dimasa datang (memperkirakan hal yang belum terjadi). Kita bisa menunggu hingga hal itu terjadi untuk membuktikan seberapa tepat prediksi kita.
4. Klasifikasi
 

Kegiatan menggolongkan, dengan menggunakan data historis (sebagai data yang digunakan untuk latihan dan sebagai pengalaman). Dalam klasifikasi terdapat *variabel* prediktor dan target *variable*.
5. Pengklusteran
 

Pengklusteran merupakan pengelompokan *record*, pengamatan atau memperhatikan dan membentuk kelas objek-objek yang memiliki kemiripan. Kluster adalah kumpulan *record* yang memiliki kemiripan satu dengan yang lainnya dan memiliki ketidak miripan dengan *record-record* dalam *cluster*.

## 6. Asosiasi

Tugas asosiasi dalam data mining adalah menemukan *atribut* yang muncul dalam satu waktu. Dalam dunia bisnis lebih umum disebut analisis keranjang biasa.[5]

*Decision tree* merupakan salah satu metode klasifikasi yang menggunakan representasi struktur pohon (*tree*) di mana setiap *node* merepresentasikan *atribut*, cabangnya merepresentasikan nilai dari *atribut*, dan daun merepresentasikan kelas. *Node* yang paling atas dari *decision tree* disebut sebagai *root*.

*Decision tree* merupakan metode klasifikasi yang paling populer digunakan. Selain karena pembangunannya relatif cepat, hasil dari model yang dibangun mudah untuk dipahami. Pada *decision tree* terdapat 3 jenis *node*, yaitu: [5]

1. *Root Node*, merupakan *node* paling atas, pada *node* ini tidak ada *input* dan bisa tidak mempunyai *output* atau mempunyai *output* lebih dari satu.
2. *Internal Node*, merupakan *node* percabangan, pada *node* ini hanya terdapat satu *input* dan mempunyai *output* minimal dua.
3. *Leaf node* atau *terminal node*, merupakan *node* akhir, pada *node* ini hanya terdapat satu *input* dan tidak mempunyai *output*.

## 2. Metode Penelitian

### 2.1 Alat dan Bahan

Pada penelitian ini penulis menggunakan alat bantu dalam menganalisis dan mempelajari sistem yang ada dan sistem yang akan dirancang.

Alat desain penelitian, terdiri atas :

- a. Kamus Data
- b. *Flowchart*
- c. *Flowgraph*

Adapun perangkat keras yang digunakan yaitu sebuah laptop dengan spesifikasi :

- a. Processor : Intel® Core 2 Duo 2.6 GHz
- b. Memory (RAM) : 2 GB.
- c. Harddisk : 500 GB.

Perangkat lunak yang digunakan yaitu :

- a. Sistem Operasi Microsoft Windows 7
- b. Visual Basic
- c. Crystal Report 8.5

### 2.2 Metode Pengujian

Metode pengujian yang digunakan dalam penelitian ini adalah metode pengujian *white box* dilakukan untuk menguji prosedur-prosedur yang ada. Lintasan logik yang dilalui oleh setiap bagian prosedur diuji dengan memberikan kondisi/loop spesifik. Pengujian *white box* menjamin pengujian terhadap semua lintasan yang tidak bergantung minimal satu kali, mencoba semua keputusan logik dari sisi 'true' dan 'false', eksekusi semua *loop* dalam batasan kondisi dan batasan operasionalnya dan pengujian validasi data internal.

1. Merupakan bagian dari pengujian *white-box* dalam hal pengujian prosedur-prosedur;
2. Mempergunakan notasi aliran *graph* (*node*, *link* untuk merepresentasikan *sequence*, *if*, *while*, *until* dan lain-lain).
3. Konsep kompleksitas *cyclomatic* antara lain cara perhitungan daerah tertutup pada *graph* planar dimana dapat menghubungkan batas atas jumlah pengujian yang harus direncanakan dan dieksekusi untuk menjamin pengujian seluruh *statement* program.
4. Memunculkan kasus-kasus yang akan diuji dengan membuat daftar lintasan kasus pengujian berdasarkan kompleksitas dan *cyclomatic* yang didapat.
5. Membuat alat bantu matrik *graph* yang membantu pengawasan pengujian.
7. Menghitung *region* (daerah) pengujian sesuai dengan rumus :  
$$V(G) = (E - N) + 2$$
 atau 
$$V(G) = P + 1$$
, dimana;  
E : Edge  
N : Node

P : simpul predikat  
 V(G) : kompleksitas siklomatis

**2.3 Tahap – Tahap Penelitian**

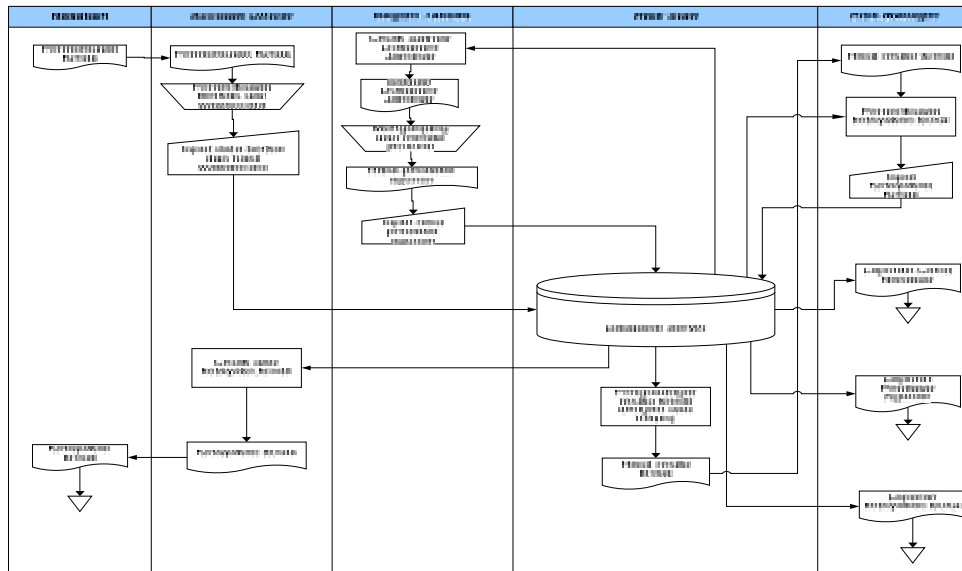
Langkah-langkah yang dilakukan dalam penelitian adalah:

1. Observasi Pendahuluan  
 Melakukan pengamatan secara langsung terhadap kegiatan Divisi Manajemen Resiko di Bank BRI Makassar untuk mengetahui bagaimana proses penetapan resiko kredit dan kelayakan kredit.
2. Pengumpulan Data  
 Mengumpulkan data-data yang dibutuhkan dalam pengembangan sistem. Baik secara langsung maupun tidak langsung.
3. Analisis Data  
 Melakukan analisis terhadap data-data yang telah dikumpulkan untuk digunakan dalam perancangan sistem.
4. Desain Sistem  
 Mendesain perangkat lunak aplikasi berdasarkan hasil analisis data.
5. Pengujian Sistem  
 Melakukan pengujian terhadap perangkat lunak aplikasi yang telah dibuat untuk menemukan kesalahan-kesalahan serta menentukan kualitas perangkat lunak.
6. Implementasi  
 Mengimplementasikan perangkat lunak aplikasi yang telah dibuat di instansi yang terkait

**3. Hasil dan Analisis**

**3.1 Analisis Sistem yang Di Usulkan**

Adapun sistem yang diusulkan dalam bentuk bagan alir dokumen yang menunjukkan arus data dari laporan dan formulir-formulir pada sistem yang diusulkan :[2]

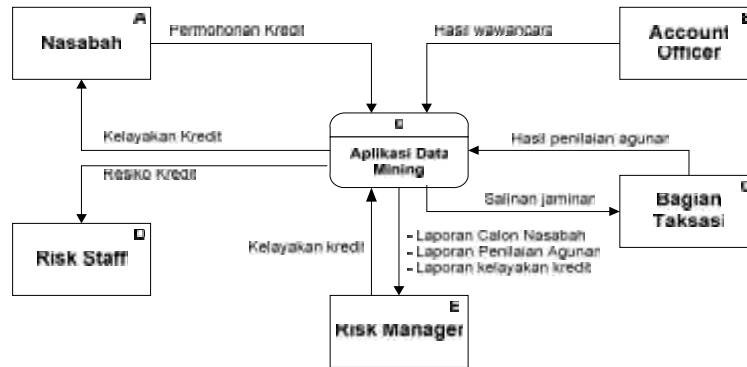


Gambar 3.1. 1 : Bagan Alir Dokumen yang di Usulkan

**3.2 Diagram Arus Data**

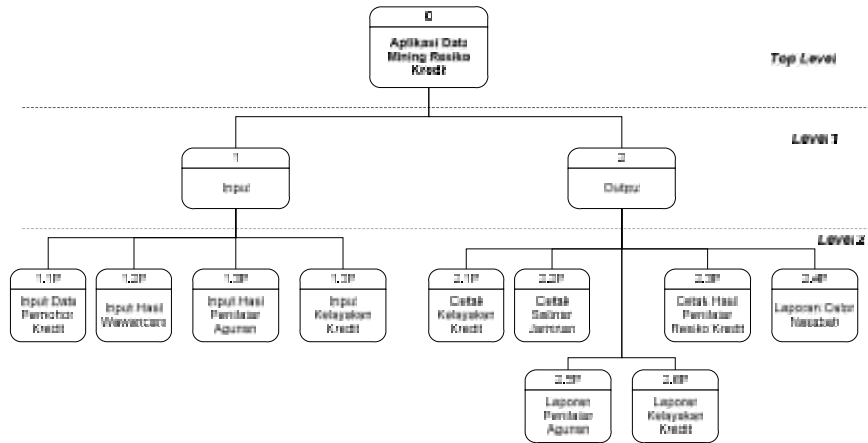
Diagram arus data digunakan untuk menggambarkan suatu sistem yang telah ada atau sistem baru yang akan dikembangkan secara logika tanpa pertimbangan lingkungan fisik dimana data tersebut mengalir atau lingkungan fisik dimana data tersebut tersimpan. Arus data merupakan salah satu simbol yang digunakan dalam diagram arus data.[2]

a. Diagram Konteks



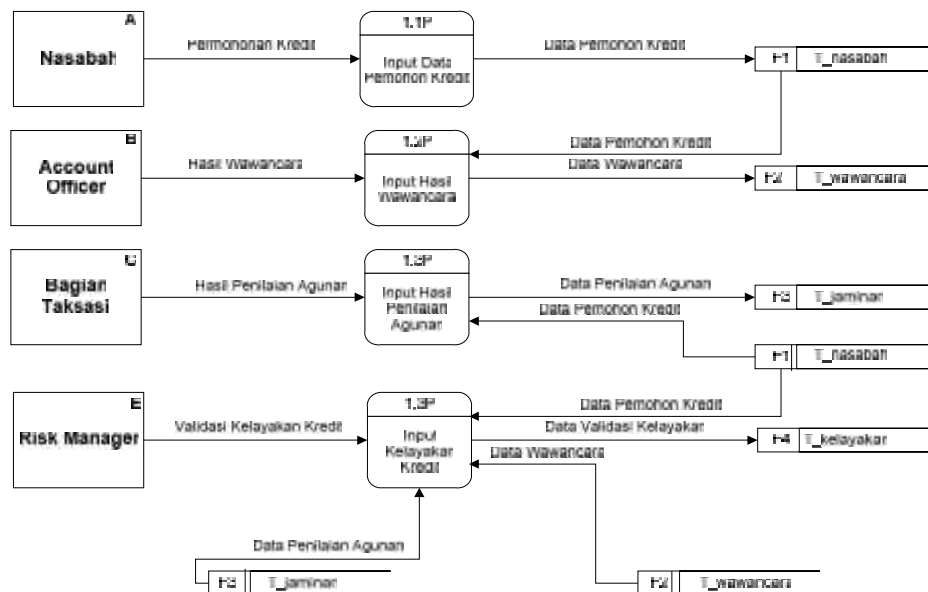
Gambar 3.2.1 : Diagram Konteks

b. Diagram Berjenjang



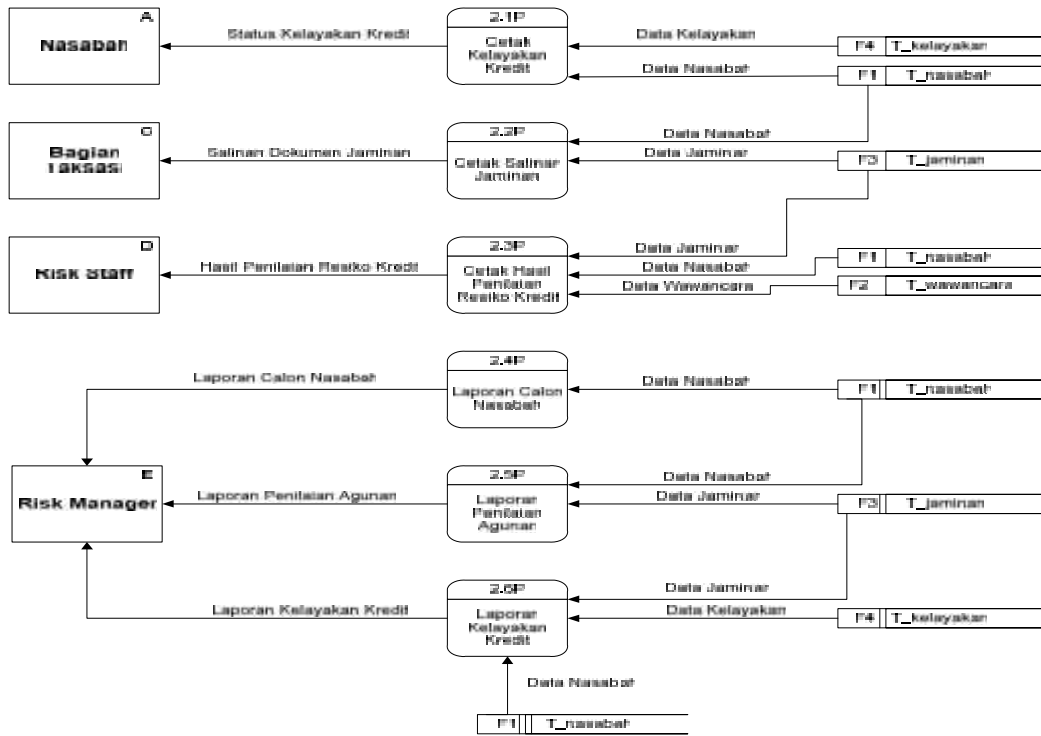
Gambar 3.2.2 : Diagram Berjenjang

c. Diagram Arus Data Level 2 ( Input )



Gambar 3.2.3 : Diagram Arus Data Level 2 ( Input )

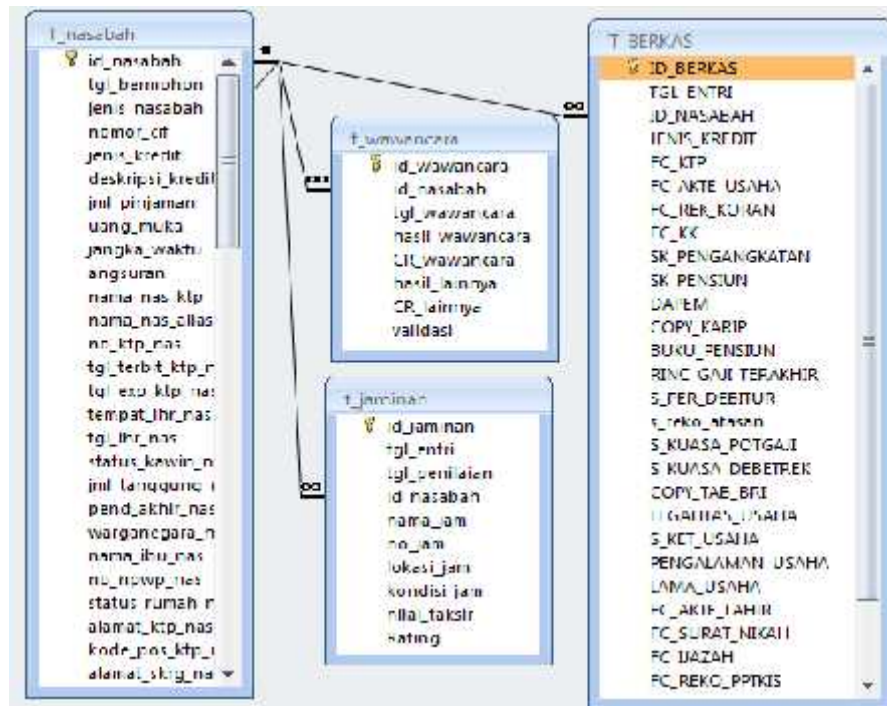
d. Diagram Arus Data Level 2 ( Output )



Gambar 3.2.4 : Diagram Arus Data Level 2 ( Output )

3.3 Perancangan Basis Data

Gambar 3.3.1 adalah perancangan basis data pada aplikasi ini. [1]



Gambar 3.3. 1 : Rancangan Basis Data



### 3.5 Rancangan Output Secara Terinci

#### a. Rancangan Output Hasil Scoring Kredit Data Mining

Gambar 3.5.1 : Rancangan Output Hasil Scoring Kredit Data Mining

#### b. Rancangan Output Hasil Perhitungan Resiko Kredit

Gambar 3.5.3 : Rancangan Output Hasil Perhitungan Resiko Kredit

#### c. Rancangan Output Status Kelayakan Kredit Calon Nasabah

Gambar 3.5.4 : Rancangan Output Status Kelayakan Kredit Calon Nasabah



#### 4. Kesimpulan

Berdasarkan hasil perancangan Aplikasi enkripsi citra dengan algoritma RC4 berbasis PHP maka dapat ditarik kesimpulan adalah Aplikasi *Data Mining* ini dapat digunakan untuk menampilkan informasi resiko kredit pada Bank BRI Makassar. Informasi yang ditampilkan berupa nilai *scoring credit* hubungan antara tingkat resiko kredit macet dan kelayakan kredit dengan data induk nasabah. Semakin tinggi nilai *scoring credit* maka semakin kuat nilai hubungan antar atribut. Data induk nasabah yang diproses *mining* meliputi data formulir nasabah, data wawancara, data penilaian jaminan, dan data berkas. Hasil dari proses *data mining* ini dapat digunakan sebagai pertimbangan dalam mengambil keputusan lebih lanjut tentang faktor yang mempengaruhi tingkat resiko kredit.

#### DAFTAR PUSTAKA

- [1] Fatansyah., 2012, “*Manajemen Basis Data*”, Informatika, Bandung.
- [2] Jogianto.H.M, 2011, “*Analisa dan Perancangan Sistem*”, Penerbit Andi, Yogyakarta.
- [3] Mico Pardosi., 2012, ” *Panduan Menggunakan - Microsoft Visual Basic 6.0*”, Informatika, Bandung.
- [4] Ramlan Ginting , Materi Presentasi diskusi umum Aspek Perbankan Judul “Pengaturan Pemberian Kredit Bank Umum “, Bandung, 2005.
- [5] Santoso, Budi, 2010, “*DATA MINING : Teknik Pemanfaatan Data Untuk Keperluan Bisnis* ”, Graha Ilmu, Yogyakarta.



