

## Sistem Pendukung Keputusan Kredit Usaha Rakyat PT. Bank Rakyat Indonesia Unit Kaliangkrik Magelang

Agung Nugroho\*<sup>1</sup>, Kusrini<sup>2</sup>, M. Rudyanto Arief<sup>3</sup>

<sup>1</sup>STMIK AMIKOM Yogyakarta

<sup>2,3</sup>Magister Teknik Informatika STMIK AMIKOM Yogyakarta

E-mail: \*[agung.nugroho.888@gmail.com](mailto:agung.nugroho.888@gmail.com), [kusrini@amikom.ac.id](mailto:kusrini@amikom.ac.id), [rudy@amikom.ac.id](mailto:rudy@amikom.ac.id)

### Abstrak

Banyak faktor dan variabel yang mempengaruhi risiko kredit dalam pengambilan keputusan pada permasalahan Kredit Usaha Rakyat (KUR). Faktor-faktor yang digunakan sebagai dasar penilaian Kredit Usaha Rakyat pada PT. Bank Rakyat Indonesia Unit Kaliangkrik menggunakan prinsip dasar yang dikenal dengan prinsip "5 of Credit" yaitu Character, Capacity, Capital, Condition dan Collateral. Dari factor-faktor yang digunakan sebagai dasar penilaian kredit, digunakan metode Mining Classification Rule dalam membuat Sistem Pendukung Keputusan pemberian KUR. Terdapat beberapa algoritma yang dapat digunakan dalam data mining untuk metode klasifikasi salah satunya adalah algoritma k-nearest neighbor. Konsep sistem pendukung keputusan pemberian KUR ini dirancang dapat melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut dan memberikan solusi nasabah yang layak menerima KUR berdasarkan masukan dari user dengan menggunakan metode k-nearest neighbors (knn). Data-data transaksi pembayaran nasabah lama akan dijadikan sebagai data training dimana sebelumnya akan ditentukan kelasnya terlebih dahulu. Penentuan kelas dilakukan dengan proses klasifikasi data berdasarkan kategori status nasabah sesuai jumlah tunggakan pembayaran kreditnya. Dari hasil perhitungan kemiripan kasus antara data calon nasabah baru dengan nasabah lama atau data training menggunakan algoritma K-Nearest Neighbor, hasil dengan nilai tertinggi akan dijadikan acuan seorang decision maker dalam mengambil keputusan.

**Kata Kunci** — sistem, objek, kredit, keputusan, klasifikasi

### Abstract

Many factors and variables that affect credit risk in decision-making on issues People's Business Credit (KUR). The factors are used as the basis of assessment of the People's Business Credit Unit at PT Bank Rakyat Indonesia Kaliangkrik using basic principle known as the principle of "5 of Credit" ie Character, Capacity, Capital, Collateral Condition and. Of the factors that are used as a basis for credit assessment, Classification Rule Mining method used in making the administration of KUR Decision Support Systems. There are several algorithms that can be used in data mining for classification methods one of which is the k-nearest algorithm neighbor. The concept of the provision of decision support system is designed KUR can perform the classification of objects based on distance learning data that is closest to the object and provide a viable solution customers receive KUR based on input from the user by using the k-nearest neighbors (KNN). Payment transaction data will be used as a customer long training data which will be determined prior to first class. Grading is done with the data classification process based on customer status categories according to the amount of credit outstanding payments. From the calculation of the similarity between the case of data with prospective new customers or old customers training data using the K-Nearest Neighbor algorithm, the results with the highest scores will be used as a reference to a decision maker in making decisions.

**Keywords** — systems, object, credit, decision, classification

## 1. PENDAHULUAN

PT. Bank Rakyat Indonesia adalah salah satu lembaga keuangan yang memfasilitasi pinjaman atau kredit untuk masyarakat. Pemberian kredit merupakan kegiatan usaha yang mengandung risiko tinggi dan berpengaruh terhadap kesehatan dan keberlangsungan usaha perbankan. Didalam kegiatan perkreditan sering terjadi masalah kredit macet yang disebabkan oleh gagalnya pengembalian sebagian pinjaman yang diberikan kepada para peminjam. Pada kasus permohonan kredit oleh nasabah, seorang *decision maker* pada suatu perbankan harus mampu mengambil keputusan yang tepat untuk menerima atau menolak permohonan kredit tersebut. Masalah ini dapat diatasi, salah satunya dengan mengidentifikasi dan memprediksi nasabah dengan baik sebelum memberikan pinjaman dengan cara memperhatikan data riwayat pinjaman nasabah lama dan membandingkan dengan data pemohon kredit. Dalam dunia Teknologi Informasi (TI) ada beberapa teknik yang bisa digunakan untuk mengatasi hal tersebut, salah satunya adalah teknik *Data Mining*. *Data Mining* adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan didalam *database*. *Data Mining* adalah proses menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar [8]. Salah satu metode *Data Mining* yang bisa digunakan adalah *Mining Classification Rule* atau metode Klasifikasi Data Mining. *Mining Classification Rule* merupakan proses menentukan kelas (label) dari suatu objek yang tidak memiliki label. Pelabelan objek dilakukan berdasarkan kesamaan karakteristik antara sekumpulan objek (*training set*) dengan objek baru tersebut [1]. Salah satu metode yang digunakan untuk membangun model klasifikasi dalam mengidentifikasi debitur ke dalam kategori baik atau buruk adalah *k-nearest neighbor (KNN)*. *K-Nearest Neighbor* adalah pendekatan untuk mencari kasus dengan menghitung kedekatan antara kasus baru dengan kasus lama, yaitu berdasarkan pada pencocokan bobot dari sejumlah fitur yang ada [2].

Penelitian terkait tentang sistem pendukung keputusan menggunakan data mining adalah penelitian Emerensye.S dalam jurnal yang ditulis pada tahun 2012, menjelaskan teknik data mining dapat digunakan untuk menilai kemampuan nasabah berdasarkan data-data masa lalu. Data *debitur* yang telah melalui tahapan data mining selanjutnya diproses menggunakan algoritma data mining yaitu *k-NN* untuk memprediksi pengajuan kredit berdasarkan tingkat kemiripan sejumlah nilai variabel *k*. Kemudian data diuji tingkat *error*-nya menggunakan teknik *cross validation* dengan 20 *fold* dan 10 nilai *k*. Hasil pengujian menunjukkan persentase tingkat *error* data pada angka kurang dari 3.7% dan mencapai kestabilan data pada nilai  $k=3$  sampai  $k=11$  sehingga ditentukan nilai *k* dengan tingkat *error* terendah yaitu pada  $k=7$  sebagai nilai *k* terbaik *k-NN* untuk prediksi kategori kredit. [5]. Penelitian terkait lainnya yaitu penelitian Leidiyana dalam jurnal yang ditulis pada tahun 2013, membahas algoritma *k-Nearest Neighbor (kNN)* yang diterapkan pada data konsumen yang menggunakan jasa keuangan kredit kendaraan bermotor. Hasil testing untuk mengukur performa algoritma ini menggunakan metode *Cross Validation*, *Confusion Matrix* dan kurva ROC dan menghasilkan akurasi dan nilai AUC berturut-turut 81,46 % dan 0,984. Karena nilai AUC berada dalam rentang 0,9 sampai 1,0 maka metode tersebut masuk dalam kategori sangat baik (excellent). [3].

Penelitian ini mencoba memanfaatkan data riwayat pembayaran kredit debitur lama yang ada di obyek penelitian serta menggunakan *Decision Support System (DSS)* dengan metode *Mining Classification Rule* menggunakan algoritma *k-nearest neighbor* untuk membentuk model prediksi pembayaran kredit calon nasabah. Variable inputan berasal dari data nasabah berdasarkan parameter prinsip dasar pada bank yang dikenal dalam "5 of Credit" yaitu *Character* (Keadaan Watak), *Capacity* (Kemampuan), *Capital* (Modal), *Condition* (Kondisi Sosial Ekonomi) dan *Collateral* (Barang yang diserahkan) kreditur yang bersangkutan. Variabel *output* berupa nilai kedekatan nasabah baru dengan nasabah lama berdasarkan dari hasil hitung kemiripan variabel kasus baru dengan kasus nasabah lama. Nilai kedekatan kasus terbesar akan dijadikan acuan oleh seorang *decision maker* dalam mengambil keputusan.

Metode-metode yang dipilih dalam penelitian ini berdasarkan pertimbangan analisa dan kebutuhan data yang dibutuhkan dalam konsep sistem pendukung keputusan yang dibuat.

### 1.1. Konsep Sistem Pendukung Keputusan

Sebuah sistem keputusan merupakan model dari sistem yang terdiri atas keputusan yang bersifat tertutup atau terbuka. Sistem keputusan tertutup menganggap bahwa keputusan dipisah dari masukan yang tidak diketahui lingkungannya. sistem keputusan terbuka memandang keputusan sebagian berada dalam suatu lingkungan yang rumit dan sebagian tak diketahui. Keputusan dipengaruhi oleh lingkungan dan pada gilirannya proses keputusan kemudian mempengaruhi lingkungan[7].

Dalam keputusan model tertutup, komputer bertindak sebagai sebuah alat penghitung untuk bisa menghitung hasil optimum. Dalam keputusan model terbuka, komputer bertindak sebagai pembantu bagi pengambilan keputusan dalam menghitung, menyimpan, mencari kembali, menganalisis data dan sebagainya. Perancangan tersebut memungkinkan manusia mengambil keputusan mengalokasikan tugas bagi dirinya atau pada computer[5].

### 1.2. Data Mining

Sebuah sistem keputusan merupakan model dari sistem yang terdiri atas keputusan yang bersifat tertutup atau terbuka. Sistem keputusan tertutup menganggap bahwa keputusan dipisah dari masukan yang tidak diketahui lingkungannya. sistem keputusan terbuka memandang keputusan sebagian berada dalam suatu lingkungan yang rumit dan sebagian tak diketahui. Keputusan dipengaruhi oleh lingkungan dan pada gilirannya proses keputusan kemudian mempengaruhi lingkungan[8].

Dalam keputusan model tertutup, komputer bertindak sebagai sebuah alat penghitung untuk bisa menghitung hasil optimum. Dalam keputusan model terbuka, komputer bertindak sebagai pembantu bagi pengambilan keputusan dalam menghitung, menyimpan, mencari kembali, menganalisis data dan sebagainya. Perancangan tersebut memungkinkan manusia mengambil keputusan mengalokasikan tugas bagi dirinya atau pada computer[5].

### 1.3. Teknik Klasifikasi (Mining Classification Rule)

Klasifikasi merupakan proses untuk menempatkan suatu objek ke dalam suatu kategori/kelas yang sudah didefinisikan sebelumnya berdasarkan model tertentu[5]. Data mining merupakan penjelasan tentang masa lalu dan prediksi masa depan berdasarkan analisa pada sekelompok data. Secara umum, proses klasifikasi dimulai dengan diberikannya sejumlah data yang menjadi acuan untuk membuat aturan klasifikasi data. Data-data ini biasa disebut dengan training sets. Dari *training sets* tersebut kemudian dibuat suatu model untuk mengklasifikasikan data. Model tersebut kemudian digunakan sebagai acuan untuk mengklasifikasikan data-data yang belum diketahui kelasnya yang biasa disebut dengan *datatest set*.

### 1.4. Algoritma *k*-Nearest neighbour (*k*-NN)

Algoritma *k*-Nearest Neighbour (*k*-NN) adalah algoritma pengklasifikasian data sederhana dimana penghitungan jarak terpendek dijadikan ukuran untuk mengklasifikasikan suatu kasus baru berdasarkan ukuran kemiripan. Algoritma *k*-NN tergolong dalam algoritma supervised yaitu proses pembentukan algoritma diperoleh melalui proses pembelajaran (*learning*) pada *record-record* lama yang sudah terklasifikasi dan hasil pembelajaran tersebut dipakai untuk mengklasifikasikan *record* baru dengan output yang belum diketahui[3].

---

## 2. METODE PENELITIAN

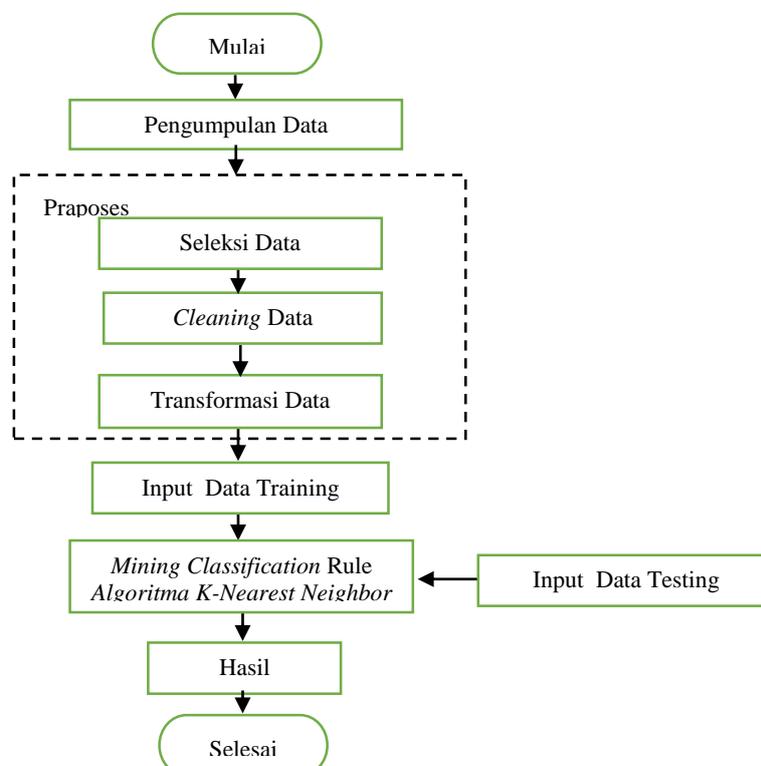
### 2.1. Bahan Penelitian

Sumber data yang digunakan diperoleh dari data riwayat pembayaran angsuran debitur lama yang telah selesai, data informasi umum seperti formulir pengajuan kredit calon kreditur, dan data informasi pola dan proses pengajuan kredit yang diperoleh dari proses wawancara kepada pihak bank. Dalam sistem pendukung keputusan pemberian kredit usaha rakyat di PT.BRI Unit Kaliangkrik dibutuhkan *data training* dan *data testing*. *Data training* sebagai dasar perbandingan dengan data *testing* yang variabelnya sudah ditentukan melalui beberapa tahap yaitu tahap seleksi, *cleaning*, dan transformasi untuk menentukan kelas yang sudah ditentukan.

Data training diambil dari variabel *independen* (*predictor*) atau variabel bebas yang mempengaruhi atau menjadi sebab perubahan dan timbulnya variabel terikat atau variabel *dependen* (kelas). Variabel *dependen* merupakan variabel yang dipengaruhi atau menjadi akibat dari adanya variabel *independen*. Variabel *dependen* dalam penelitian ini adalah tingkat pengembalian KUR Mikro. Variabel *independen* diperoleh dari data riwayat pembayaran nasabah lama dan *form* pengajuan kredit. Kemudian setelah ditentukan variabel *dependen* dan *independen* maka dihasilkan satu *set* data yang berisi nilai dari kedua komponen di atas yang digunakan untuk menentukan kelas yang cocok berdasarkan *predictor*. Berisi data baru atau nasabah baru yang akan dicari kedekatan variabelnya dengan data training dan diklasifikasi berdasarkan kelas yang telah ditentukan.

### 2.2. Analisa Model

Analisis model bertujuan untuk menentukan model yang digunakan untuk memecahkan masalah menjadi sebuah solusi. Urutan langkah-langkah pemecahan masalah dalam penelitian ini dapat digambarkan pada Gambar.1.



Gambar 1. Alur Model Analisis Data Sistem Pendukung Keputusan

Penjelasan dari gambaran alur model analisis proses data sistem pendukung keputusan diatas sebagai berikut:

- 1) Pengumpulan Data, teknik pengumpulan data dilakukan melalui teknik *observasi* dengan cara mengumpulkan informasi secara langsung ke lokasi penelitian untuk mengamati bagaimana syarat dan proses pengajuan kredit usaha rakyat yang dilakukan oleh pihak BRI Unit Kaliangkrik Cabang Magelang dengan cara pengamatan dan pencatatan dengan peninjauan langsung dan wawancara dengan melakukan wawancara terstruktur dengan menyediakan daftar pertanyaan untuk mewawancarai kepala unit dan *Acount Officer* atau mantri selaku pelaku penindak putusan pemberian kredit usaha rakyat (KUR). Teknik pengumpulan data juga dilakukan dengan cara dokumentasi untuk mengetahui informasi yang dibutuhkan, dengan melakukan pengumpulan data arsip transaksi pembayaran debitur lama, data calon debitur, dan data karakteristik debitur lama pada *form* pengajuan kredit.
- 2) Seleksi Data, dari data yang dikumpulkan kemudian dilakukan penyeleksian dengan memilih dan memisahkan data berdasarkan kriteria-kriteria yang ditentukan.
- 3) *Cleaning Data*, dalam proses ini dilakukan pembuangan data pada data karakteristik debitur lama yang tidak akan memberikan pengaruh terhadap hasil prediksi. Proses *clening data* bertujuan untuk membentuk data training yang berisi data *predictor* dan kelas.
  - a. *Predictor*  
Variabel independen yang direpresentasikan oleh karakteristik (atribut) data.
  - b. Kelas  
Variabel dependen yang berupa kategorikal yang merepresentasikan ‘label’ yang terdapat pada objek.
  - c. *Training Dataset*  
Satu set data yang berisi nilai dari kedua komponen di atas yang digunakan untuk menentukan kelas yang cocok berdasarkan predictor.
- 4) Transformasi Data, adalah proses pembobotan atribut *predictor*. Pada tahap ini dilakukan pembobotan pada tiap-tiap atribut pada data *training*.
- 5) *Mining Classification Rule* Algoritma *K-Nearest Neighbor*, tahap ini dipusatkan untuk ekstraksi data training dan data testing dengan menggunakan teknik Mining Classification Rule dengan menerapkan algoritma K-nearest neighbor.
- 6) Hasil, setelah dicari nilai kededekatan antara data *testing* terhadap data *training* menggunakan algoritma *K-nearest neighbour* maka data *testing* akan diklasifikasi berdasarkan kelas yang telah ditentukan.

### 2.3. Proses Minning Klasifikasi

#### 2.3.1. Seleksi Data

Dari data yang dikumpulkan, dilakukan penyeleksian dengan memilih dan memisahkan data berdasarkan kriteria-kriteria yang ditentukan yang nantinya akan digunakan sebagai data *sample*. Untuk menentukan data *sample* nasabah digunakan metode *probability sampling*. *Probability sampling* adalah teknik sampling yang memberikan peluang yang sama bagi setiap unsur (anggota) populasi untuk dipilih menjadi anggota *sampel* [5]. Jumlah sampel didapat melalui perhitungan *Slovin*, persamaan perhitungan *Slovin* dirumuskan seperti persamaan (1) sebagai berikut:

$$n = \frac{N}{N.e^2+1} \quad (1)$$

Dimana:  $n$  = jumlah contoh (ukuran sampel)  
 $N$  = jumlah populasi  
 $e$  = tingkat kesalahan yang masih bisa ditolerir (10%)

### 2.3.2. Cleaning Data

Dilakukan pembuangan data pada data karakteristik debitur lama yang tidak akan memberikan pengaruh terhadap hasil prediksi. Atribut yang tidak dibutuhkan akan dibuang [4]. Atribut dipilih berdasar faktor-faktor yang diturunkan dari prinsip-prinsip yang digunakan dalam pertimbangan pengajuan kredit menggunakan prinsip penilaian “5C” yaitu *character, capital, capacity, conditions of economy, dan collateral*.

### 2.3.3. Pembagian Kelas

Setelah dilakukan proses *cleaning* selanjutnya dilakukan pembagian kelas dengan membagi satu *set* data *training* atau data nasabah lama yang berisi nilai dari komponen pada tabel data *predictor* yang digunakan untuk menentukan kelas yang cocok berdasarkan *predictor*-nya.

### 2.3.4. Pembobotan Nilai Atribut

Pada tahap ini yang pertama kali dilakukan adalah mendefinisikan bobot tiap-tiap atribut pada data training. Metode pemberian bobot atribut menggunakan metode pembobotan ROC (*Rank Order Centroid*). Teknik pembobotan ROC adalah teknik memberikan bobot pada setiap kriteria sesuai dengan *ranking* yang dinilai berdasarkan tingkat prioritas kriterianya [9]. Biasanya dibentuk dengan pernyataan “Kriteria 1 lebih penting dari kriteria 2, yang lebih penting dari kriteria 3” dan seterusnya hingga kriteria ke  $n$ , jika ditulis menjadi:

$$Cr1 \geq Cr2 \geq Cr3 \dots \geq Crn.$$

Untuk menentukan bobotnya, diberikan aturan yang sama yaitu:

$$W1 \geq W2 \geq W3 \dots \geq Wn,$$

Dimana  $W1$  merupakan bobot untuk kriteria  $Cr1$ . Secara umum pembobotan ROC dapat dirumuskan seperti persamaan (2) sebagai berikut:

$$Wk = 1/k \sum_{n=1}^k \binom{1}{i} \quad (2)$$

Dimana:  $Wk$  : bobot kriteria  $n$   
 $k$  : banyaknya kriteria

### 2.3.5. Algoritma K-Nearest Neighbor

Setelah mendefinisikan bobot semua atribut, selanjutnya menerapkan algoritma *K-Nearest Neighbor* (K-NN) untuk memprediksi potensi calon debitur. Secara umum untuk mendefinisikan jarak antara dua objek  $x$  dan  $y$ , digunakan rumus jarak *Euclidean*, seperti terlihat pada persamaan (3)

$$Dxy = \sqrt{\sum_{i=1}^n (xi - yi)^2} \quad (3)$$

Dimana :  $d$  : jarak kedekatan  
 $X$  : data *training*  
 $Y$  : data *testing*  
 $n$  : jumlah atribut individu antara 1 s.d.  $n$   
 $f$  : fungsi *similitary* atribut  $i$  antara kasus  $X$  dan kasus  $Y$   
 $Wi$  : bobot yang diberikan pada atribut ke- $i$

Jarak antara objek  $x$  dan  $y$  didefinisikan sebagai  $Dxy$ , dimana  $xi$  merupakan *record* yang akan diprediksi dan  $yi$  merupakan *record* data pola sedangkan nilai  $n$  didefinisikan sebagai jumlah atribut dan nilai  $i$  merujuk kepada *record* ke- $i$  [6].

### 3. HASIL DAN PEMBAHASAN

Dari hasil proses cleaning diperoleh data akhir yang nantinya akan dijadikan data *predictor*. Berikut data *predictor* yang digunakan:

Tabel 1. Tabel Data Atribut dan Nilainya

No	Atribut	Nilai Atribut
1	Penghasilan Perbulan	<ul style="list-style-type: none"><li>• &gt;3x angsuran</li><li>• 3x angsuran</li><li>• 2x angsuran</li><li>• 1x angsuran</li></ul>
2	Nilai Jaminan	<ul style="list-style-type: none"><li>• &gt; Jumlah Pinjaman</li><li>• = Jumlah Pinjaman</li><li>• &lt; Jumlah Pinjaman</li></ul>
3	Jumlah tanggungan	<ul style="list-style-type: none"><li>• Tidak ada</li><li>• 1 orang</li><li>• 2-3 orang</li><li>• &gt; 3 orang</li></ul>
4	Pendidikan terakhir	<ul style="list-style-type: none"><li>• Universitas</li><li>• SLTA</li><li>• SLTP</li><li>• SD</li></ul>
5	Lama Usaha	<ul style="list-style-type: none"><li>• &gt;5 tahun</li><li>• 3 - 5 tahun</li><li>• 1 - 2 tahun</li><li>• &lt; 1 tahun</li></ul>
6	Status Perusahaan	<ul style="list-style-type: none"><li>• Swasta besar</li><li>• Swasta menengah</li><li>• Swasta kecil</li><li>• Perorangan</li></ul>
7	Usia	<ul style="list-style-type: none"><li>• &lt; 20 tahun / &gt; 60 tahun</li><li>• 20 - 50 tahun</li><li>• 51 - 60 tahun</li></ul>
8	Status	<ul style="list-style-type: none"><li>• Menikah</li><li>• Belum menikah</li><li>• Janda/Duda</li></ul>
9	Kelas	<ul style="list-style-type: none"><li>• Bad</li><li>• Normal</li><li>• Good</li></ul>

Proses akhir dari tahap ini menghasilkan data *training* seperti pada tabel 2.

Tabel 2. Data Training Hasil Proses *Cleaning*

No	KDT	Nama	Alamat	JK	A	B	C	D	E	F	G	H	T	S
1	001	xxxx	xxxx	L	2x	=	2	SMA	1	Kecil	20-50	Menikah	4	Baik
2	002	xxxx	xxxx	L	2x	=	3	UNI	1-2	Menengah	20-50	Menikah	3	Baik
3	003	xxxx	xxxx	P	1x	=	2	SMA	1	Kecil	51-60	Menikah	2	Baik
4	004	xxxx	xxxx	L	1x	>	1	SMA	1	Kecil	51-60	Menikah	1	Baik
5	005	xxxx	xxxx	L	2x	>	1	SMA	3-5	Menengah	20-50	Menikah	0	Baik
6	006	xxxx	xxxx	L	2x	<	1	UNI	1-2	Menengah	20-50	Menikah	5	Normal
7	007	xxxx	xxxx	L	3x	<	2	UNI	>5	Menengah	<20 >60	Menikah	6	Normal
8	008	xxxx	xxxx	P	2x	>	1	SMA	1-2	Menengah	20-50	Menikah	7	Normal
...	...	...	.....	.....	...	...	...	.....	...	.....	.....	.....	...	.....
...	...	...	.....	.....	...	...	...	.....	...	.....	.....	.....	...	.....
42	042	xxxx	xxxx	L	2x	=	1	SLTA	3-5	Menengah	<20 >60	Menikah	4	Buruk
43	043	xxxx	xxxx	L	3x	=	>3	SLTA	3-5	Menengah	20-50	Menikah	5	Baik

Keterangan:

No : Nomor Urut

KDT : Kode Data Training

JK : Jenis Kelamin

A : Pendapatan Perbulan

B : Nilai Jaminan

C : Jumlah Tanggungan

D : Pendidikan Terakhir

E : Lama Usaha

F : Kondisi Usaha

G : Usia

H : Status Pernikahan

T : Jumlah Tunggakan

S : Status Debitur

Setelah diperoleh data *training*, kemudian dilakukan pendefinisian bobot pada setiap variable yang telah ditentukan. Pendefinisian bobot didasarkan pada atribut yang paling berpengaruh untuk pemberian kredit di BRI Kaliangkrik Bagian KUR.

Tabel 3. Tabel Bobot Atribut

Nama Atribut	Tingkat Kepentingan	Bobot
Penghasilan Perbulan	1	0,340
Nilai Jaminan	2	0,215
Jumlah Tanggungan Keluarga	3	0,152
Pendidikan Terakhir	4	0,111
Lama Usaha	5	0,079
Kondisi Usaha	6	0,054
Usia	7	0,034
Status Pernikahan	8	0,016

Setelah bobot tiap-tiap atribut didefinisikan, kedekatan antara nilai-nilai dalam atribut juga didefinisikan.

Perbandingan nilai sub-atribut penghasilan perbulan

Tabel 4. Tabel Perbandingan Nilai Sub-Atribut Penghasilan Perbulan

	3x angsuran	2x angsuran	1x angsuran
3x angsuran	1	0,45	0,18
2x angsuran	0,45	1	0,39
1x angsuran	0,18	0,39	1

Tabel 5. Tabel Kedekatan Nilai Atribut Penghasilan Perbulan

No	Sub-Atribut	Nama Atribut	Bobot
1	3x Angsuran	3x Angsuran	1
2	3x Angsuran	2x Angsuran	0,45
3	3x Angsuran	1x Angsuran	0,18
4	2x Angsuran	2x Angsuran	1
5	2x Angsuran	1x Angsuran	0,39
6	2x Angsuran	3x Angsuran	0,45
7	1x Angsuran	1x Angsuran	1
8	1x Angsuran	3x Angsuran	0,18
9	1x Angsuran	2x Angsuran	0,39

Berikut adalah contoh mengitung kedekatan data *testing* dengan data *training*:

Tabel 6. Tabel Contoh Data *Training*

Kode Training	A	B	C	D	E	F	G	H	Status
TRXX	2x angsuran	= Nilai Jaminan	1	SLTA	3-5	Menengah	<20 / >60	Menikah	Buruk
TRXY	3x angsuran	= Nilai Jaminan	>3	SLTA	3-5	Menengah	20-50	Menikah	Baik

Tabel 7. Tabel Contoh Data *Testing*

Kode Testing	A	B	C	D	E	F	G	H	Status
TSXX	2x angsuran	= Jumlah Pinjaman	Tidak Ada	SLTP	>5	Kecil	<20 / >60	Belum Menikah	Buruk

Nilai kedekatan kasus baru TSXX dengan kasus lama TRXX:

Tabel 8. Tabel Nilai Kedekatan Data TSXX Dengan Data TRXX

Nama Atribut	Bobot Atribut	Kedekatan Bobot Atribut		
		TR042	TS001	Nilai
Pendapatan Perbulan (A)	(Ai) 0,340	2x Angsuran	2x Angsuran	1
Nilai Jaminan (B)	(Bi) 0,215	=Jumlah Pinjaman	=Jumlah Pinjaman	1
Jumlah Tanggungan (C)	(Ci) 0,152	1 Orang	Tidak Ada	0,65
Pendidikan Terakhir (D)	(Di) 0,111	SLTA	SLTP	0,5
Lama Usaha (E)	(Ei) 0,079	3-5	>5	0,75
Kondisi Usaha (F)	(Fi) 0,054	Menengah	Kecil	0,35
Usia (G)	(Gi) 0,034	<20/>60	<20/>60	1
Status Pernikahan (H)	(Hi) 0,016	Menikah	Belum Menikah	0,5
Hasil Similarity k-nn			0,83	

Nilai kedekatan kasus baru TSXX dengan kasus lama TRXY:

Tabel 9. Tabel Nilai Kedekatan Data TSXX Dengan Data TRXY

Nama Atribut	Bobot	Kedekatan Bobot Atribut		Nilai
		TR002	TS001	
Status Pernikahan (A)	(Ai) 0,016	Menikah	Belum Menikah	0,5
Jumlah Tanggungan (B)	(Bi) 0,152	>3 Orang	Tidak Ada	0,15
Pendidikan Terakhir (C)	(Ci) 0,111	SLTA	SLTP	0,5
Usia (D)	(Di) 0,034	20-50	<20/>60	0,5
Lama Usaha (E)	(Ei) 0,079	3-5	>5	0,75
Kondisi Usaha (F)	(Fi) 0,054	Menengah	Kecil	0,35
Pendapatan Perbulan (G)	(Gi) 0,340	3x Angsuran	2x Angsuran	0,5
Nilai Jaminan (H)	(Hi) 0,215	=Jumlah Pinjaman	=Jumlah Pinjaman	1
Hasil Similarity k-nn			0,62	

Dari langkah diatas dapat diketahui bahwa nilai tertinggi adalah kedekatan kasus TSXX dengan TRXX. Berdasarkan hasil pada langkah diatas, maka klasifikasi dari kasus TRXX yang akan digunakan untuk memprediksi kasus baru, kemungkinan nasabah baru akan bermasalah, karena data TRXX mempunyai catatan status “Buruk” pada data history dan sama seperti pada catatan status data TSXX. Data TSXX adalah data sampling yang mempunyai catatan status “Buruk”. Oleh karena itu dari perhitungan ini membuktikan hasil yang sama.

### 3.1. User Interface

Menu data Manajemen Atribut, petugas dapat memanipulasi data Atribut. Dapat dilihat pada gambar 2.

Manajemen Atribut		
Nama Atribut	Bobot	Pilihan
jumlah_tanggungan	0.152	<a href="#">Edit</a>
kondisi_usaha	0.054	<a href="#">Edit</a>
lama_usaha	0.079	<a href="#">Edit</a>
nilai_jaminan	0.215	<a href="#">Edit</a>
pendidikan_terakhir	0.111	<a href="#">Edit</a>
penghasilan_perbulan	0.340	<a href="#">Edit</a>
status_pernikahan	0.160	<a href="#">Edit</a>
usia	0.034	<a href="#">Edit</a>

Showing 1 to 8 of 8 entries

← Previous 1 Next →

Gambar 2. Tampilan Manajemen Atribut

Menu Perbandingan digunakan untuk memasukan nilai atau bobot nilai untuk masing-masing atribut. Dapat dilihat pada gambar 3.

**Perbandingan Atribut**

Nama Atribut	Pilihan
jumlah_tanggungan	<a href="#">Q Detail</a>
kondisi_usaha	<a href="#">Q Detail</a>
lama_usaha	<a href="#">Q Detail</a>
nilai_jaminan	<a href="#">Q Detail</a>
pendidikan_terakhir	<a href="#">Q Detail</a>
penghasilan_perbulan	<a href="#">Q Detail</a>
status_pernikahan	<a href="#">Q Detail</a>
usia	<a href="#">Q Detail</a>

Showing 1 to 8 of 8 entries

← Previous 1 Next →

Gambar 3. Tampilan Form Perbandingan Atribut

Menu Kasus Lama berfungsi sebagai menampilkan, menambah, mengedit, dan menghapus data nasabah lama.

**Kasus Lama**

[+ Tambah Kasus](#)

DTR	Nama	Omset	Jaminan	Tanggungan	Pendidikan	Lama	Kondisi	Umur	Status	Remark	Pilihan
TR001	xxxx	2x angsuran	= jumlah pinjaman	2-3 orang	SLTA	1-2 tahun	kecil	20-50	menikah	baik	<a href="#">Edit</a> <a href="#">Hapus</a>
TR002	xxxx	2x angsuran	= jumlah pinjaman	2-3 orang	Universitas	1-2 tahun	menengah	20-50	menikah	baik	<a href="#">Edit</a> <a href="#">Hapus</a>
TR003	xxxx	1x angsuran	= jumlah pinjaman	1-2 orang	SLTA	1-2 tahun	kecil	51-60	menikah	baik	<a href="#">Edit</a> <a href="#">Hapus</a>
TR004	xxxx	1x angsuran	> jumlah pinjaman	1 orang	SLTA	1-2 tahun	kecil	51-60	menikah	baik	<a href="#">Edit</a> <a href="#">Hapus</a>
TR005	xxxx	2x angsuran	> jumlah pinjaman	1 orang	SLTA	3-5 tahun	menengah	20-50	menikah	baik	<a href="#">Edit</a> <a href="#">Hapus</a>
TR006	xxxx	2x angsuran	< jumlah pinjaman	1 orang	Universitas	1-2 tahun	menengah	20-50	menikah	normal	<a href="#">Edit</a> <a href="#">Hapus</a>
TR007	xxxx	3x angsuran	< jumlah pinjaman	2-3 orang	Universitas	>5 tahun	menengah	<20 / >60	menikah	normal	<a href="#">Edit</a> <a href="#">Hapus</a>
TR008	xxxx	2x angsuran	> jumlah pinjamn	1 orang	SLTA	1-2 tahun	menengah	20-50	menikah	normal	<a href="#">Edit</a> <a href="#">Hapus</a>

Gambar 4. Tampilan Form Nasabah Lama

Menu Kasus Baru berfungsi untuk menginputkan data nasabah baru yang akan diproses dan dianalisis hasilnya dengan nasabah lama atau kasus lama

**Kasus Baru**

[+ Tambah Kasus](#)

DTS	Nama	Alamat	JK	Penghasilan	Jaminan	J.Tanggung	Pendidikan	Lama	Kondisi	Umur	Status	Pilihan
1	Agungr	Magelang	L	3x angsuran	> jumlah pinjaman	tidak ada	universitas	>5 tahun	besar	20-50	menikah	<a href="#">Edit</a> <a href="#">Hapus</a>
2	dian	gg	L	1x angsuran	< jumlah pinjaman	> 3 orang	SD	<1 tahun	kecil	<20 / >60	janda / duda	<a href="#">Edit</a> <a href="#">Hapus</a>
4	nugroho	xxxxxxx	L	2x angsuran	= jumlah pinjaman	1 orang	SLTA	1-2 tahun	besar	20-50	belum menikah	<a href="#">Edit</a> <a href="#">Hapus</a>
5	Agung nugroho	magelang	L	2x angsuran	= jumlah pinjaman	tidak ada	universitas	>5 tahun	besar	20-50	belum menikah	<a href="#">Edit</a> <a href="#">Hapus</a>

Showing 1 to 4 of 4 entries [← Previous](#) [1](#) [Next →](#)

Gambar 5. Tampilan Form Input Nasabah Baru (*Data Testing*)

Menu Analisis Kasus berfungsi untuk memproses data kasus baru dengan kasus lama, dan akan menampilkan hasil perbandingan kasus baru atau nasabah baru berdasarkan kedekatannya dengan nasabah lama atau kasus lama.

**Analisi Kasus**

Nama:

[Tampil](#)

DTR	Nama	Omset	Jaminan	Tanggung	Pendidikan	Lama	Kondisi	Umur	Status	Remark
TR001	xxxx	2x angsuran	= jumlah pinjaman	2-3 orang	SLTA	1-2 tahun	kecil	20-50	menikah	baik
TR002	xxxx	2x angsuran	= jumlah pinjaman	2-3 orang	Universitas	1-2 tahun	menengah	20-50	menikah	baik
TR003	xxxx	1x angsuran	= jumlah pinjaman	1-2 orang	SLTA	1-2 tahun	kecil	51-60	menikah	baik
TR004	xxxx	1x angsuran	> jumlah pinjaman	1 orang	SLTA	1-2 tahun	kecil	51-60	menikah	baik
TR005	xxxx	2x angsuran	> jumlah pinjaman	1 orang	SLTA	3-5 tahun	menengah	20-50	menikah	baik
TR006	xxxx	2x angsuran	< jumlah pinjaman	1 orang	Universitas	1-2 tahun	menengah	20-50	menikah	normal
TR007	xxxx	3x angsuran	< jumlah pinjaman	2-3 orang	Universitas	>5 tahun	menengah	<20 / >60	menikah	normal
TR008	xxxx	2x angsuran	> jumlah pinjamn	1 orang	SLTA	1-2 tahun	menengah	20-50	menikah	normal
TR009	xxxx	1x angsuran	< jumlah pinjaman	tidak ada	SMP	<1 tahun	menengah	51-60	belum menikah	buruk
TR010	xxxx	1x angsuran	= jumlah pinjaman	2-3 orang	SD	<1 tahun	perorangan	51-60	menikah	buruk

Showing 1 to 10 of 43 entries [← Previous](#) [1](#) [2](#) [3](#) [4](#) [5](#) [Next →](#)

Gambar 6. Tampilan Menu Analisis Kasus

### 3.2. Pengujian Algoritma

Pengujian algoritma *K-nearest neighbor* menggunakan *n-fold Cross Validation*. *n-fold cross validation* atau *k-fold cross validation* merupakan salah satu metode yang digunakan untuk mengetahui rata-rata keberhasilan dari suatu sistem dengan cara melakukan perulangan dengan mengacak atribut masukan sehingga sistem tersebut teruji untuk beberapa atribut input yang acak. *n-fold cross validation* diawali dengan membagi data sejumlah *n-fold* yang diinginkan [5].

Dalam iterasi ke-*i* partisi *Di* akan menjadi data *testing* dan sisanya akan menjadi data *training*. Untuk penggunaan jumlah *fold* terbaik untuk uji validitas, dianjurkan menggunakan 10-

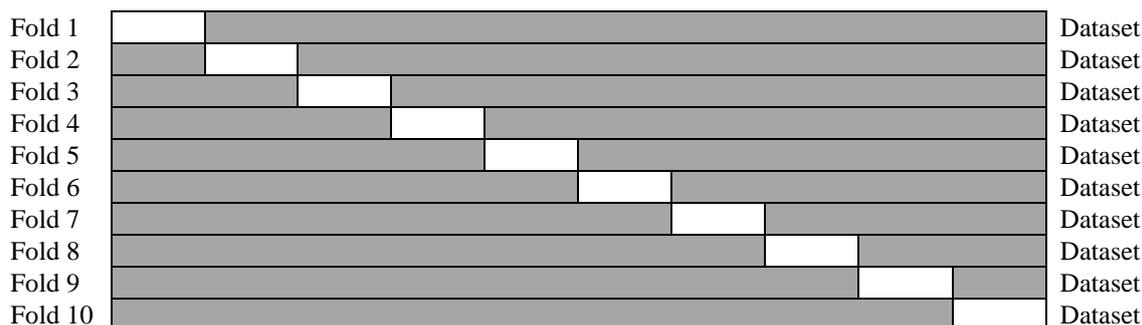
*fold cross validation* dalam model [6]. Pembagian *dataset* dalam proses *10-fold cross validation* seperti terlihat pada gambar berikut :



Gambar 7. Pembagian *Dataset 10-fold Cross Validatioan*

Pembagian *Dataset* dibagi menjadi 10 partisi sebagai berikut:

Dalam proses *cross validation* data akan dibagi dalam  $n$  buah partisi dengan ukuran yang sama  $D1, D2, D3 \dots Dn$  selanjutnya proses *testing* dan *training* dilakukan sebanyak  $n$  kali.



Gambar 8. Proses *Cross Validatioan*

Dari hasil evaluasi *n-fold cross validation* terhadap perhitungan algoritma *k-nearest neighbor* terhadap perhitungan kedekatan kasus lama pada data training dengan kasus baru pada data testing. Diketahui ada 43 dataset, dibagi menjadi 3 kelas yaitu:

1. Kelas Baik
  - Ada 23 data dengan status atau dikategorikan ke dalam kelas “Baik”, hasil uji validasi:
    - 21 data diprediksi baik dan sesuai
    - 2 diprediksi baik ternyata normal
2. Kelas Normal
  - Ada 5 data dengan status atau dikategorikan ke dalam kelas “Normal”, dari hasil uji validasi diperoleh hasil sebagai berikut :
    - 4 data diprediksi normal dan sesuai
    - 1 data diprediksi normal ternyata buruk
3. Kelas Buruk
  - Ada 15 data dengan status atau dikategorikan ke dalam kelas “Buruk”, dari hasil uji validasi diperoleh hasil sebagai berikut:
    - 15 buruk
    - 15 data diprediksi buruk dan sesuai

Dari hasil evaluasi diatas didapat:

- Data yang diprediksi sesuai (akurat) = 40
- Data yang tidak sesuai dengan prediksi (*error*) = 3

Jadi dapat disimpulkan persentase tingkat akurasi penerapan algoritma *kNN* adalah tingkat *error rate* sebesar 6,98 % , dan tingkat akurasi kebenarannya 93,023%.

#### 4. KESIMPULAN

Dari data-data informasi yang diperoleh dari PT. Bank Rakyat Indonesia Unit Kaliangkrik khususnya dibidang Kredit Usaha Rakyat (KUR) Sebenarnya diperoleh data-data transaksi yang banyak tersimpan dan dapat dimanfaatkan melalui *Mining Classification Rule* dengan menggunakan algoritma *K-Nearest Neighbor* yang menghasilkan sebuah kecerdasan bisnis seperti memprediksi calon debitur yang memiliki potensi baik, normal, dan buruk berdasarkan data history dari data transaksi pembayaran debitur-debitur lama. Kecerdasan bisnis yang telah dibuat dapat dijadikan acuan untuk menentukan langkah apa yang harus dilakukan oleh PT. BRI Unit Kaliangkrik Cabang Magelang dalam memberikan Kredit Usaha Rakyat (KUR) kepada calon debitur yang sudah diprediksi potensinya.

Dari hasil penelitian ini dihasilkan beberapa point kesimpulan sebagai berikut:

1. Sistem ini hanya dapat digunakan untuk membantu menentukan baik, buruk, atau normal calon debitur Kredit Usaha Rakyat di BRI Kaliangkrik dan hasil dari sistem bukan sebagai penentu akhir dalam proses seleksi, keputusan akhir tertinggi tetap berada dalam pihak pengambil keputusan.
2. Dalam penelitian ini dilakukan penerapan algoritma *k-nn* pada data debitur lama yang mendapat pembiayaan kredit usaha rakyat. Agar didapat data yang berkualitas, dilakukan preprosesing sebelum diterapkan ke dalam algoritma. Tahap pra proses dalam penerapan data mining klasifikasi menggunakan metode nearest neighbor memakan waktu lebih banyak, karena perlunya penambahan proses penentuan bobot variabel dan penentuan bobot perbandingan antar nilai variabel. Kedekatan antara kasus baru dengan kasus lama dilakukan untuk menentukan termasuk kelas mana kasus baru tersebut. Untuk mengukur kinerja algoritma tersebut digunakan metode Cross Validation dengan hasil tingkat akurasi penerapan algoritma *k-nn* diperoleh tingkat *error rate* sebesar 6,98 %, dan 93,023% akurat.
3. Pada metode *k-nearest neighbor* semua kasus akan terklasifikasi meskipun nilai kedekatannya belum tentu 100%. Namun hasil klasifikasi tidak dapat dijamin kebenarannya karena banyak kasus yang memiliki kedekatan yang sama dengan beberapa klasifikasi yang berbeda.

#### 5. SARAN

Adapun saran untuk penelitian ini adalah sebagai berikut:

1. Pada PT. Bank BRI Unit Kaliangkrik terdapat beberapa jenis kredit. Akan lebih baik jika data training yang digunakan untuk penelitian selanjutnya sebaiknya di bagi sesuai dengan jenis kredit yang diambil.
  2. Sebaiknya aplikasi yang dihasilkan bisa mendukung file dengan format xlsx (Microsoft Office Excel 2010) agar aplikasi bisa menyesuaikan dengan perkembangan dunia teknologi informasi (TI),
  3. Sebaiknya data-data yang digunakan bisa digunakan sebagai laporan keuangan atau pencatatan pembayaran debitur, sehingga data training nantinya akan up-to-date,
  4. Hasil penelitian ini sebaiknya dapat diterapkan guna memudahkan kepada pihak Kepala Unit dan Mantri (*Account Officer*) PT. BRI Unit Kaliangkrik Cabang Magelang untuk memberikan kredit.
  5. Dalam memasukkan data dilakukan dengan teliti, agar dalam pengambilan keputusan tidak terjadi kesalahan.
  6. Sebaiknya setiap beberapa periode dilakukan evaluasi terhadap system yang digunakan, sehingga selalu dapat dilakukan penyesuaian system.
  7. Penentuan bobot variabel dan bobot jarak perbandingan nilai variabel yang dimasukkan secara manual memungkinkan ada nilai-nilai yang berbeda-beda sesuai dengan persepsi pengguna, padahal besarnya nilai bobot variabel dan bobot jarak nilai variabel sangat mempengaruhi besarnya kedekatan antara satu kasus dengan kasus yang lain.
-

Karena jumlah variable dan/atau jumlah nilai variable cukup banyak, akan lebih membantu jika memasukkan data bobot dari tiap-tiap variable dan bobot jarak antara nilai-nilai dalam variable bisa dilakukan secara otomatis dan dinamis.

## 6. UCAPAN TERIMA KASIH

Terimakasih kepada pihak PT. Bank BRI yang telah mengizinkan dilakukan penelitian ini. Terimakasih kepada Drs. Kusriani dan M. Rudiyanto Arief, M.Kom selaku dosen pembimbing atas bimbingan dan dukungannya sehingga penelitian ini berjalan dengan lancar. Terimakasih kepada pihak STMIK AMIKOM Yogyakarta atas beasiswa studi S2 yang telah dipercayakan kepada penulis.

## DAFTAR PUSTAKA

- [1] Abidin, Taufik. 2006. Eksplorasi Data (Data Mining), <http://www.keudekupi.com/content/view/26/1/>, diakses 24 Juli 2014.
  - [2] Kusriani, Hartati, S., Wardoyo, R., Harjoko, A., 2009, Perbandingan Metode Nearest Neighbor dan Algoritma C4.5 untuk Menganalisis Kemungkinan Pengunduran Diri Calon Mahasiswa di STMIK AMIKOM Yogyakarta, *Jurnal Dasi*, Vol. 10, No. 1.
  - [3] Leidiyana, H., 2013, Penerapan Algoritma K-Nearest Neighbor untuk Penentuan Resiko Kredit Kepemilikan Kendaraan Bermotor, *Jurnal Penelitian Ilmu Komputer, System Embedded & Logic*, Vol 1, No 1, Hal 65-76.
  - [4] Sholichah, A., 2009. Data Mining untuk pembiayaan murabahah menggunakan Association Rule, <http://lib.uin-malang.ac.id>, diakses 26 September 2012.
  - [5] Pandie, E.S.Y., 2012, Sistem Informasi Pengambilan Keputusan Pengajuan Kredit dengan Algoritma K-Nearest Neighbour (Studi Kasus: Koperasi Simpan Pinjam), *Tesis*, Magister Sistem Informasi, Universitas Diponegoro, Semarang.
  - [6] Prasetya, E., 2006, Case Based Reasoning untuk Mengidentifikasi Kerusakan Bangunan, *Tesis*, Program Pasca Sarjana Ilmu Komputer, Univ. Gadjah Mada, Yogyakarta.
  - [7] Kusriani, 2007, *Konsep dan Aplikasi Sistem Pendukung Keputusan*, Andi Offset, Yogyakarta.
  - [8] Kusriani, Luthfi, E.T., 2009, *Algoritma Data Mining*, Andi Offset, Yogyakarta.
  - [9] Rahmah, A., 2013, Sistem Pendukung Keputusan Seleksi Masuk Mahasiswa Menggunakan Metode Smarter, *Skripsi*, Jurusan Pendidikan Matematika, Universitas Pendidikan Indonesia, Bandung.
-