

**PENENTUAN BESAR PINJAMAN DI KOPERASI SIMPAN PINJAM
DENGAN ALGORITMA K-NEAREST NEIGHBOR
(Studi Kasus di Koperasi Simpan Pinjam BMT Bina Insani Pringapus)**

Abidah Elcholiqi, Beta Noranita, Indra Waspada

Jurusan Ilmu Komputer/Informatika, FSM, Universitas Diponegoro
Email : abidahel@gmail.com, betta@undip.ac.id, indra.waspada@gmail.com

ABSTRAK

Perkembangan teknologi informasi saat ini memungkinkan penyimpanan data dalam skala yang sangat besar. Melimpahnya data menjadi tantangan dalam dunia teknologi informasi, termasuk perbankan, untuk tidak hanya mendapatkan informasi saja, tetapi juga menemukan pola untuk menghasilkan pengetahuan. *Data mining* merupakan kegiatan menemukan pengetahuan dari sejumlah data yang sangat besar. Dengan memanfaatkan data pinjaman anggota, *data mining* dapat membantu pengambilan keputusan besar pinjaman bagi anggota koperasi simpan pinjam. Algoritma *data mining* yang digunakan untuk membangun aplikasi *data mining* adalah *k-nearest neighbor* (kNN). KNN digunakan untuk mengklasifikasikan besar pinjaman anggota berdasarkan jarak kedekatan atribut. Percobaan menggunakan 25 sampel data dengan nilai $k=8$ dan $k=12$ menghasilkan akurasi sebesar 84%. Hasil proses *mining* dapat digunakan untuk membantu pegawai koperasi simpan pinjam dalam menentukan besar pinjaman bagi anggotanya.

Kata kunci : *data mining*, *k-Nearest Neighbor*, penentuan besar pinjaman

I. PENDAHULUAN

Kemajuan dalam bidang teknologi memberikan manfaat yang luar biasa kepada masyarakat. Kemajuan tersebut memungkinkan penyimpanan data dalam skala yang sangat besar. Melimpahnya data menjadi tantangan dalam dunia teknologi informasi untuk tidak hanya mendapatkan informasi saja, tetapi juga menemukan pola untuk menghasilkan pengetahuan. Proses menganalisis data untuk menemukan pola yang tersembunyi yang dapat dijadikan pengetahuan merupakan pengertian dari *data mining*[7]. *Data mining* dapat digunakan untuk berbagai tujuan, seperti pengelompokan, klasifikasi, prediksi, dan lain sebagainya. Beberapa teknik *data mining* yang umum digunakan diantaranya regresi, *decision tree*, jaringan syaraf tiruan, *k-nearest neighbor*, dan naïve bayes.

Proses *data mining* sudah banyak diterapkan dalam bisnis perusahaan. Hal tersebut ditunjang oleh ketersediaan data. Tidak terkecuali dalam bidang perbankan dan koperasi simpan pinjam, *data mining* sudah digunakan untuk membantu penentuan pemberian kredit, memprediksi penunggakan kredit oleh nasabah dan lain sebagainya.

BMT Bina Insani merupakan koperasi syariah yang mengeluarkan produk pinjaman. Sesuai dengan badan hukumnya sebagai koperasi, BMT Bina Insani mengelola dana dari, oleh dan untuk anggota. Dalam memberikan pinjaman kepada anggotanya, saat ini keputusan besarnya pinjaman ditentukan oleh kepala bagian pembiayaan di kantor pusat. Besarnya pinjaman yang diberikan tersebut didasarkan pada berbagai pertimbangan, misalnya nilai taksiran harga barang jaminan, penghasilan anggota, dan faktor lainnya.

Seiring dengan kemajuan yang dicapai BMT Bina Insani dalam mengelola dana anggota, maka BMT Bina Insani membuka kantor cabang. Dengan bertambahnya kantor cabang, maka keputusan besar pinjaman harus ditangani masing-masing kantor kas. Sedangkan untuk memutuskan pemberian pinjaman tidaklah mudah, perlu pengalaman dalam menganalisis anggota apakah tepat jika anggota diberikan sejumlah pinjaman berdasarkan data anggota dan jaminan. Apalagi di BMT Bina Insani yang berbentuk koperasi, pengelola harus melakukan reorganisasi setiap periode tertentu sehingga bukan tidak mungkin sering terjadi pergantian kepala bagian pembiayaan. Oleh karena itu, perlu penerapan *data mining* untuk menganalisis

besarnya pinjaman yang sudah diberikan di masa lalu dalam membantu menentukan besar pinjaman bagi anggotanya.

II. TINJAUAN PUSTAKA

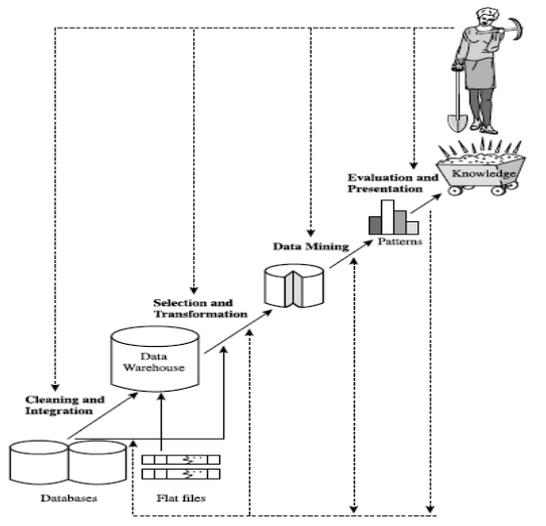
2.1. Pengertian Data Mining

Data mining merupakan kegiatan mengumpulkan dan memakai data historis untuk menemukan keteraturan, pola atau hubungan dalam set data berukuran besar. Keluaran dari data mining dapat dipakai untuk memperbaiki pengambilan keputusan dimasa depan [5]. Secara sederhana, data mining adalah kegiatan menemukan pengetahuan dari sejumlah data yang sangat besar. Data tersimpan dalam bentuk basis data, data warehouse, dan penyimpanan informasi lainnya [1].

Dengan memanfaatkan data mining, pengetahuan menarik atau informasi tingkat tinggi dapat diperoleh dari database. Pengetahuan ini dapat digunakan oleh pemakai untuk mendukung keputusan, proses kontrol, manajemen informasi dan pemrosesan query [1].

2.2. Tahap Data Mining

Data mining merupakan salah satu tahap dalam KDD [1]. Proses KDD digambarkan pada gambar 1.



Gambar 1. Tahapan Proses KDD [1]

Proses KDD terdiri atas 7 tahap [1] yaitu :

- Pembersihan Data (*Data Cleaning*)
Pembersihan data merupakan proses menghilangkan *noise*, data yang tidak lengkap, dan data yang tidak konsisten.
- Integrasi Data (*Data Integration*)

Seringkali data mining menggunakan data dari berbagai sumber data sehingga integrasi data harus dilakukan secara cermat..

- Pemilihan Data (*Data Selection*)
Pemilihan data merupakan proses pemilihan data yang akan digunakan dalam proses data mining.
- Transformasi Data (*Data Transformation*)
Transformasi data merupakan proses perubahan data sebelum digunakan untuk data mining.
- Proses Mining
Proses mining merupakan proses utama dalam menemukan pola yang didapat dengan melaksanakan metode-metode data mining.
- Evaluasi Pola (*Pattern Evaluation*)
Evaluasi pola merupakan proses mengidentifikasi pola yang didapat dari proses mining untuk dievaluasi apakah pola hipotesa yang ada benar-benar tercapai.
- Presentasi Pengetahuan (*Knowledge Presentation*)
Pengetahuan dari pola yang ada divisualisasikan untuk dapat dipahami dan dimanfaatkan oleh pengguna.

2.3. Teknik Data Mining

Metode pembelajaran dibagi ke dalam dua pendekatan yaitu *supervised* dan *unsupervised*. Pendekatan *unsupervised learning* adalah pendekatan yang tidak menggunakan latihan maupun label data. Contoh : *clustering*. Data pada pendekatan ini dikelompokkan ke dalam beberapa kelas yang dikehendaki.

Sedangkan pendekatan *supervised learning* adalah metode belajar dengan adanya pelatih dan latihan, sebagai contoh adalah regresi, analisis diskriminan, jaringan syaraf tiruan, klasifikasi, dan *Support Vector Machine(SVM)*[5].

Data mining dibedakan dalam fungsi minor dan fungsi mayor[6].

- Fungsi **Minor** atau fungsi tambahan, meliputi deskripsi, estimasi, dan prediksi.
- Fungsi **Mayor** atau fungsi utama, meliputi klasifikasi, pengelompokan dan asosiasi.

Fungsi-fungsi dalam data mining adalah sebagai berikut[3] :

- Deskripsi
Deskripsi merupakan analisis sederhana untuk menggambarkan pola dan kecenderungan yang terdapat dalam data.
- Estimasi
Estimasi hampir mirip dengan klasifikasi, hanya saja variabel target biasanya bernilai numerik daripada kategorial.
- Prediksi

- Prediksi hampir mirip dengan klasifikasi dan estimasi, tetapi hasil prediksi bisa berbeda dengan kenyataan di masa datang.
- d. Klasifikasi
Dalam klasifikasi, data didistribusikan ke dalam beberapa kelas sebagai target variabel kategori.
 - e. Pengelompokan (*clustering*)
Pengelompokan *record*, pengamatan, atau memperhatikan dan membentuk kelas objek – objek yang memiliki kemiripan
 - f. Asosiasi
Tugas asosiasi dalam *data mining* adalah menemukan atribut yang muncul dalam satu waktu.

2.4. Algoritma k-Nearest Neighbor

K-nearest neighbor adalah salah satu metode pengklasifikasian dengan cara mencari jarak terdekat antara data baru dengan k tetangga terdekatnya dalam data pelatihan [2]. Setelah sejumlah k tetangga terdekat diperoleh, kemudian dilakukan *voting* untuk menentukan label data baru. *Output* dari sejumlah k yang mempunyai frekuensi paling banyak ditetapkan sebagai label untuk data baru.

Algoritma *k-nearest neighbor* memiliki ketangguhan terhadap data pelatihan yang memiliki banyak *noise* dan efektif jika datanya sangat besar. Namun algoritma ini membutuhkan biaya komputasi yang cukup tinggi karena diperlukan perhitungan jarak data uji ke semua data pelatihan [1].

Algoritma *k-nearest neighbor* dijelaskan sebagai berikut [5]:

- a. Persiapan data pelatihan dan data uji
- b. Menentukan nilai k
- c. Penghitungan jarak data uji ke setiap data pelatihan.
Jarak kedekatan dihitung menggunakan rumus penghitungan jarak *Euclidean*[5] sebagai berikut.
$$d(x,y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$
- d. Tentukan sejumlah k data pelatihan yang memiliki jarak terdekat dengan data uji dan tentukan sejumlah k data pelatihan terdekat.
- e. Periksa label dari k data pelatihan terdekat
- f. Tentukan label yang frekuensinya paling banyak.
- g. Masukkan data uji ke dalam kelas dengan frekuensi paling banyak.
- h. Kondisi berhenti.

III. PEMBAHASAN

3.1. Analisis dan Desain Aplikasi Data Mining

Aplikasi yang dikembangkan digunakan untuk membantu menentukan besar pinjaman bagi anggota.

Hasil yang diperoleh dari aplikasi *data mining* ini berupa *range* nilai pinjaman.

Atribut yang digunakan dalam menentukan nilai pinjaman di BMT Bina Insani adalah taksiran harga barang jaminan, jumlah penghasilan anggota, jumlah tanggungan anggota, jangka waktu dan karakter anggota dilihat dari pernah atau tidaknya mendapatkan pinjaman di BMT Bina Insani. Data yang digunakan adalah data pinjaman BMT Bina Insani cabang Bergas periode 2010 hingga Januari 2012. Data yang digunakan untuk proses *mining* dibatasi pada akad pinjaman yang menggunakan jaminan hanya berupa kendaraan bermotor saja dan atas nama pribadi.

Berikut tahap persiapan data yang dilakukan dalam aplikasi *data mining* BMT Bina Insani.

- a. Pembersihan Data
Pembersihan dilakukan terhadap data yang tidak memiliki kelengkapan nilai untuk setiap atribut yang digunakan dalam aplikasi ini.
- b. Pemilihan Data
Data pinjaman yang digunakan hanya pinjaman atas nama pribadi saja dan hanya menggunakan jaminan berupa kendaraan bermotor.
- c. Transformasi Data
Semua atribut yang digunakan dinormalisasi kedalam skala 0-1, sedangkan nilai pinjaman dibagi kedalam 7 kelas seperti terlihat pada tabel 1.

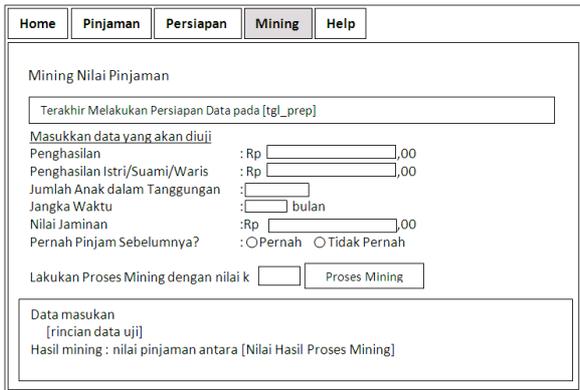
Tabel 1 Kategori Kelas Pembiayaan

Kelas	Keterangan
A	Nilai pinjaman kurang dari 3 juta
B	Nilai pinjaman 3-6 juta
C	Nilai pinjaman 6-9 juta
D	Nilai pinjaman 9-12 juta
E	Nilai pinjaman 12-15 juta
F	Nilai pinjaman 15-25 juta
G	Nilai pinjaman lebih dari 25 juta

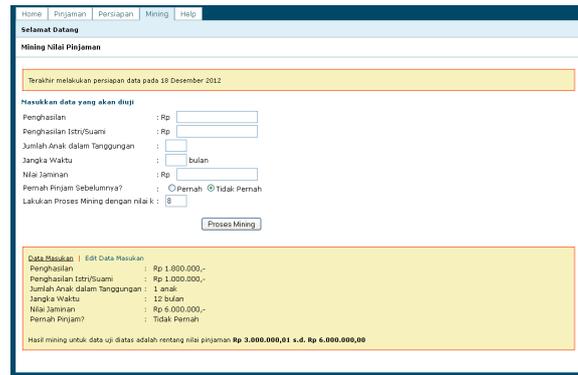
Spesifikasi kebutuhan fungsional pada aplikasi ini merujuk pada kebutuhan akan perancangan *data mining* sebagai berikut.

- a. Sistem dapat menampilkan data pinjaman (*datasource*)
- b. Sistem dapat melayani praproses data
- c. Sistem dapat melayani pemrosesan *data mining*
- d. Sistem dapat menampilkan tanggal terakhir melakukan praproses data

Desain tampilan proses *mining* pada aplikasi *data mining* BMT Bina Insani ini tampak pada gambar 2.



Gambar 2 Desain tampilan proses mining Aplikasi Data Mining



Gambar 3 Tampilan proses mining Aplikasi Data Mining

3.2. Implementasi

Implementasi perancangan antarmuka merupakan transformasi perancangan antarmuka pada aplikasi data mining BMT Bina Insani. Tampilan proses mining dapat dilihat dari gambar 3. Pengguna harus memasukkan nilai data uji kemudian aplikasi akan memproses mining dan menampilkan hasilnya. Hasil dari aplikasi data mining berupa rentang nilai pinjaman seperti terlihat pada tabel 1.

3.3. Pengujian

Pengujian dengan metode *black-box* untuk menemukan kesalahan-kesalahan dan memastikan bahwa input yang dibatasi akan memberikan hasil aktual yang sesuai dengan hasil yang dibutuhkan. Sebuah program yang diuji dengan metode *black-box* dikatakan diterima jika fungsi-fungsinya memenuhi kriteria evaluasi hasil yang diharapkan.

Pengujian terhadap performa aplikasi dalam melakukan mining dilakukan dengan mengambil secara acak sebanyak 25 data sebagai data uji dan sisanya digunakan sebagai data pelatihan. Nilai k yang diujikan adalah k=7 sampai k=13.

Tabel 1. Hasil data mining berupa rentang nilai pinjaman

No	Nilai atribut					Nilai Pembiayaan	Kelas	Klasifikasi dengan sejumlah nilai k						
	Ph	Jk	Tg	Pp	Jm			7	8	9	10	11	12	13
1	0.2	1	0.2	0.85	0.3	6.000.000	B	B	B	B	B	B	B	
2	0	0.277778	0.2	0.85	0.4	1.500.000	A	A	A	A	A	A	A	
3	0	0.666667	0.2	0.85	0.5	5.000.000	B	B	B	B	B	B	B	
4	0.2	0.333333	0.6	0.85	0.2	4.000.000	B	A	B	B	B	B	B	
5	0	0.166667	0	0.85	0.4	2.000.000	A	A	A	A	A	A	A	
6	0.2	0.333333	0.4	0.85	0.2	4.500.000	B	B	B	B	B	B	B	
7	0	0.5	0	0.85	0.4	6.000.000	B	B	B	B	B	B	B	
8	0.4	0.5	0.6	0.85	0.7	15.000.000	E	E	E	E	C	C	C	
9	0.2	0.333333	0.2	0.85	0.2	5.000.000	B	B	B	B	B	B	B	
10	0.2	0.333333	0	1	0.1	4.000.000	B	B	B	B	B	B	B	
11	0	0.666667	0.2	0.85	0.6	10.000.000	D	D	D	D	D	D	B	
12	0	0.333333	0.4	0.85	0.2	2.500.000	A	A	A	A	A	A	A	
13	0.2	0.666667	0.6	0.85	0.4	6.500.000	C	B	B	B	B	B	B	
14	0	0.666667	0.4	0.85	0.1	3.000.000	A	A	A	A	B	B	B	
15	0.2	0.333333	0.2	0.85	0.1	4.000.000	B	B	B	B	B	B	B	
16	0	0.666667	0.2	0.85	0	2.000.000	A	B	B	B	B	B	B	
17	0.6	0.083333	0.6	1	1	35.000.000	G	G	G	G	G	G	G	
18	0	0.5	0.4	0.85	0.1	1.500.000	A	A	A	A	A	A	A	

No	Nilai atribut					Nilai Pembiayaan	Kel as	Klasifikasi dengan sejumlah nilai k						
	Ph	Jk	Tg	Pp	Jm			7	8	9	10	11	12	13
19	0	0.666667	0.4	1	0.1	3.000.000	A	B	B	B	A	A	A	A
20	0.2	0.333333	0.4	0.85	0.2	4.000.000	B	B	B	B	B	B	B	B
21	0	0.666667	0.2	0.85	0.7	10.000.000	D	D	D	D	D	D	D	D
22	0	0.333333	0	1	0.2	3.500.000	B	A	A	A	B	B	B	B
23	0.4	0.666667	0.2	0.85	0.6	14.000.000	E	E	E	C	C	C	E	E
24	0	0.333333	0.4	0.85	0.6	2.000.000	A	A	A	A	A	A	A	A
25	0.4	0.333333	0.6	0.85	0.4	7.500.000	C	C	B	B	B	B	C	B
Jumlah klasifikasi yang sama dg kelas nilai pinjamannya								20	21	19	19	19	21	19

Keterangan :

Ph : Jumlah penghasilan

Jk : Jangka waktu pinjaman

Tg : Jumlah anak dalam tanggungan

Pp : Status pernah pinjam sebelumnya

Jm : Taksiran nilai jaminan

3.4. Analisis Hasil

Analisis pengujian hasil aplikasi *data mining* di atas dapat diambil beberapa informasi sebagai berikut:

- Pengujian data sampel dengan nilai k=9, k=10, k=11, dan k=13 menunjukkan sebanyak 19 data (76% dari total data sampel) memiliki hasil *mining* yang sama dengan kelas nilai pinjamannya.
- Pengujian data sampel dengan nilai k=7 menunjukkan sebanyak 20 data (80% dari total data sampel) memiliki hasil *mining* yang sama dengan kelas nilai pinjamannya.
- Pengujian data sampel dengan nilai k=8 dan k=12 menunjukkan sebanyak 21 data (84% dari total data sampel) memiliki hasil *mining* yang sama dengan kelas nilai pinjamannya.
- Pengujian data sampel dengan nilai k=8 dan k=12 menunjukkan jumlah hasil *mining* paling banyak sesuai dengan kelas nilai pinjamannya.

IV. KESIMPULAN

Kesimpulan yang dapat diambil adalah sebagai berikut:

- Aplikasi yang dikembangkan dapat digunakan untuk membantu menentukan nilai pinjaman bagi anggota koperasi simpan pinjam BMT Bina Insani.

- Dari hasil *mining* menggunakan 25 sampel data uji dengan nilai k=7 sampai dengan k=14 didapatkan nilai k=8 dan k=12 memiliki jumlah kesesuaian antara hasil *mining* dengan kelas nilai pinjaman paling banyak yaitu 84% dari total data sampel sehingga k=8 dan k=12 dapat digunakan sebagai k *default* untuk aplikasi ini. Meskipun demikian, pengguna dapat mengubah nilai k dalam proses mining.

DAFTAR PUSTAKA

- [1] Han J., dan Kamber, M., 2006, "*Data Mining : Concept and Technique, Second Edition*", Morgan Kaufmann Publisher, San Fransisco.
- [2] Hermaduanti, N dan Sri Kusumadewi, 2008, "Sistem Pendukung Keputusan Berbasis SMS untuk Menentukan Status Gizi dengan Metode K-Nearest Neighbor", dalam Seminar Nasional Aplikasi Teknologi Informasi 2008 (SNATI 2008), Yogyakarta
- [3] Larose, Daniel T., 2005, "*Discovering knowledge in Data : An Introduction to Data Mining*", John Willey & Sons Inc.
- [4] MacLennan, J, dkk, 2009, "*Data Mining with Microsoft SQL Server 2008*", Willey Publishing, USA.
- [5] Santosa, Budi, 2007, "*Data Mining : Teknik Pemanfaatan Data untuk Keperluan Bisnis*", Graha Ilmu, Yogyakarta.
- [6] Susanto, S., dan Dedy Suryadi, 2010, "*Pengantar Data Mining : Menggali Pengetahuan dari Bongkahan Data*", ANDI Offset, Yogyakarta.

