

Penerapan Metode K-Nearest Neighbour Dalam Menentukan Kelayakan Calon Nasabah Yang Layak Untuk Kredit Mobil (Studi Kasus : Pt. Astra International,Tbk-Toyota)

Tito Cakra Pratama

Program Studi Teknik Informatika STMIK Budi Darma, Medan, Indonesia
Jl. Sisingamangaraja No. 338 Simpang Limun, Medan, Indonesia

Abstrak

Sejalan dengan pertumbuhan bisnis, kredit merupakan masalah yang menarik untuk diteliti. Beberapa riset bidang komputer untuk mengurangi kredit yang telah banyak dilakukan dalam rangka menghindari kehancuran suatu perusahaan pembiayaan dan mengurangi kesalahan dalam proses penentuan kelayakan calon pembeli untuk mendapatkan kredit dalam pembelian mobil. Salah satu faktor kegagalan dalam usaha perkreditan yaitu kurang akuratnya penilaian terhadap kemampuan debitor, sehingga mengakibatkan kesalahan dalam keputusan pemberian kredit yang berujung pada kemacetan kredit.

Kata Kunci: K-Nearest Neighbor, Kredit, Kredit Mobil

Abstract

In line with business growth, credit is an interesting issue to research. Some computer research to reduce the credit that has been done in order to avoid the collapse of a finance company and reduce errors in the process of determining the eligibility of prospective buyers to get credit in the purchase of cars. One of the failure factors in the credit business is the lack of accurate assessment of the ability of the debtor, resulting in mistakes in credit decisions that led to credit congestion.

Keywords: K-Nearest Neighbor, Credit, Car Loan

1. PENDAHULUAN

Data mining adalah kegiatan yang meliputi pengumpulan, pemakaian data historis yang menemukan keteraturan, pola dan hubungan dalam set data berukuran besar. Maksud dari pengertian ini yaitu proses pencarian informasi yang tidak diketahui sebelumnya dari sekumpulan data besar.

Data dengan skala besar biasanya sering terdapat pada perusahaan maupun instansi-instansi pendidikan seperti perguruan tinggi. Dari beberapa data yang ada pada suatu perusahaan salah satunya pada PT. Astra International TBK-Toyota yaitu data kelayakan pelanggan untuk mendapatkan kredit mobil. Sistem yang berjalan saat ini di PT. Astra International TBK-Toyota saat ini belum semua berjalan dengan lancar dan profesional karena ada beberapa masalah pada proses pengelolaan datanya, diantaranya dalam penentuan keputusan sering tidak sesuainya nasabah yang mendapatkan kredit dengan nilai kriteria dan nilai bobot yang ditentukan oleh perusahaan, lamanya penyerahan laporan keputusan kepada kepala perusahaan dan informasi kepada calon pembeli membuat lambatnya proses penentuan keputusan calon pembeli untuk membeli mobil secara kredit.

Pada data pelanggan yang mendapatkan kelayakan kredit dalam pembelian mobil terdapat berbagai macam atribut yang ada pada data diri pelanggan yang dikumpulkan saat mendata pelanggan tersebut. Data yang dikumpulkan diantaranya adalah data kepribadian dari pelanggan itu sendiri seperti pekerjaan, umur, dan jumlah tanggungan. Selain itu terdapat juga berkas-berkas yang harus dilengkapi seperti KTP dan KK yang digunakan sebagai kelengkapan data untuk menentukan layak atau tidaknya pelanggan dalam mendapatkan kelayakan kredit mobil. Terdapat beberapa penelitian sebelumnya yang telah dilakukan, yaitu menggunakan proses data mining dalam menentukan kelayakan kredit baik itu mobil dan pinjaman dari bank seperti yang dilakukan oleh Nandang Iriadi yang berjudul kajian penerapan metode klasifikasi data mining Algoritma C4.5 untuk prediksi kelayakan kredit pada bank Mayapada Jakarta [1].

Selain itu ada juga penelitian yang berjudul Evaluasi Pemohon Kredit Mobil Di PT. Karya Motor dengan menggunakan teknik *Data Mining Decision Tree* yang dilakukan oleh Yogi Yusuf, pokok pembicaraan jurnal tersebut membahas bagaimana menangani pembayaran yang sering macet dan tertunda. [2], dan terakhir Penerapan Algoritma *K-Means* Untuk Penentuan Resiko Kredit Kepemilikan Kendaraan Bermotor Yang Dilakukan Oleh Henny Leidiyana Pada Tahun 2013 [3].

Berdasarkan penelitian-penelitian yang sudah dipaparkan diatas bahwa konsep data mining dapat dimanfaatkan untuk proses penentuan kelayakan untuk mencari dan mengidentifikasi apakah data pelanggan yang ingin mengajukan kredit apakah layak atau tidak untuk mendapatkan kredit mobil. Teknik yang digunakan adalah dengan memanfaatkan data pelanggan lama yang sudah pernah mengajukan kredit nantinya akan dijadikan sebagai data training yang digunakan untuk menguji data baru apakah data baru atau calon pelanggan baru berhak mendapatkan kredit pembelian mobil berdasarkan kemiripan atribut yang ada di data lama dengan data baru. [2]

Metode K-Nearest Neighbour bertujuan untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut. Data pembelajaran diproyeksikan ke ruang berdimensi banyak, dimana masing – masing dimensi merepresentasikan fitur dari data [1]. Pada penelitian akan

dianalisis data pelanggan baru yang akan mendapatkan kredit mobil berdasarkan karakter, pendidikan, pekerjaan, jumlah tanggungan, dan berkas pelanggan dengan data lama yang sudah ada dengan menggunakan Algoritma K-Nearest Neighbour.

2. TEORITIS

2.1 Data Mining

Data mining adalah kegiatan yang meliputi pengumpulan, pemakaian data historis yang menemukan keteraturan, pola dan hubungan dalam set data berukuran besar. Maksud dari pengertian ini yaitu proses pencarian informasi yang tidak diketahui sebelumnya dari sekumpulan data besar.

2.2 Algoritma K-Nearest Neighbor (K-NN)

Algoritma K-Nearest Neighbor (K-NN) adalah sebuah metode melakukan klasifikasi terhadap objek berdasarkan data pembelajaran jaraknya paling dekat dengan objek tersebut. K-NN termasuk algoritma supervised learning dimana hasil dari query instance yang baru diklasifikasikan berdasarkan mayoritas dari kategori pada K-NN, kemudian kelas yang paling banyak muncul yang akan menjadi kelas hasil klasifikasi.

Tujuan dari algoritma ini adalah mengklasifikasikan obyek berdasarkan atribut dan training sample. *Klasifier* tidak menggunakan apapun untuk dicocokkan dan hanya berdasarkan pada memori. Diberikan titik *query*, akan ditemukan sejumlah k obyek atau (titik training) yang paling dekat dengan titik *query*. Klasifikasi menggunakan voting terbanyak diantara klasifikasi dari k obyek. *Algoritma K-Nearest Neighbor (K-NN)* menggunakan klasifikasi ketetanggaan sebagai nilai prediksi dari *query instance* yang baru.

Algoritma metode K-Nearest Neighbor (K-NN) sangat sederhana, bekerja berdasarkan jarak terpendek dari *query instance* ke training sample untuk menentukan K-NN. Training sample diproyeksikan ke ruang berdimensi banyak, dimana masing-masing dimensi merepresentasikan fitur dari data. Ruang ini dibagi menjadi bagian-bagian berdasarkan klasifikasi training sample. Sebuah titik pada ruang ini ditandai jika merupakan klasifikasi yang paling banyak ditemukan pada k buah tetangga terdekat dari titik tersebut. Metode pencarian jarak, ada dua jenis yaitu *metode Cosine Similarity* atau *Euclidean Distance* yaitu perhitungan jarak terdekat. Perhitungan jarak terdekat dibutuhkan untuk menentukan jumlah kemiripan yang dihitung dari kemiripan kemunculan teks yang dimiliki suatu paragraf. Setelah itu kemunculan teks yang sedang diujikan dibandingkan terhadap masing-masing sample data asli. Ada dua metode untuk menghitung jarak antar tetangga yaitu metode Euclidean Distance dan Cosine Similarity. Jika vector $a = x_1i + y_2i$ dan $b = x_2i + y_2j$, maka jarak euclidan a dan b adalah sebagai berikut :

$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \dots\dots\dots(1)$$

dimana *matriks D(a,b)* adalah jarak skalar dari kedua vektor a dan b dari matriks dengan ukuran d dimensi. Satuan jarak yang digunakan Euclidian, jenis dari metode ini, jika dilihat dari nilai Nnya ada dua macam yaitu

1. 1-NN
Pengklasifikasian dilakukan pada 1 label data terdekat, algoritmanya sebagai berikut :
 - a. Menghitung jarak antara data baru ke setiap pelabelan data
 - b. Menentukan 1 pelabelan data yang mempunyai jarak paling minimal
 - c. Klasifikasi data baru ke dalam pelabelan data tersebut
2. k-NN
Pengklasifikasian dilakukan dengan menentukan nilai pada k label data terdekat, dengan syarat nilai $k > 1$, algoritmanya sebagai berikut :
 - a. Menghitung jarak antara data baru ke setiap pelabelan data
 - b. Menentukan k pelabelan data yang mempunyai jarak paling minimal
 - c. Klasifikasi data baru ke dalam pelabelan data yang mayoritas

Berikut ini adalah langkah-langkah menghitung *K-Nearest Neighbors* pada dokumen :

1. Menentukan parameter k (jumlah tetangga paling dekat).
2. Menghitung kuadrat jarak euclidean (euclidean distance) masing-masing obyek terhadap data sampel yang diberikan.

$$d_i = \sqrt{\sum_{i=1}^p (x_{2i} - x_{1i})^2} \dots\dots\dots(2)$$

Keterangan:

x = Sampel Data

2 x = Data Uji / Testing

i = Variabel Data

d = Jarak

p = Dimensi Data

3. Mengurutkan objek-objek tersebut ke dalam kelompok yang mempunyai jarak euclidean terkecil
4. Mengumpulkan kategori y (klasifikasi nearest neighbor)

Kelebihan dari *metode K-Nearest Neighbor* ini adalah sebagai berikut :

1. Lebih efektif di data training yang besar
2. Dapat menghasilkan data yang lebih akurat
3. Metode yang baik dalam hal ruang pencarian, misalnya, kelas tidak harus linear dipisahkan.
4. Sangat cocok terhadap training data yang *noise*

Kekurangan dari *metode K-Nearest Neighbor* ini adalah perlu ditentukan nilai k yang paling optimal yang menyatakan jumlah tetangga terdekat [5]

2.3 Kredit

Kata kredit berasal dari bahasa latin yaitu “*credere*”, yang artinya percaya atau *to believe* atau *to trust*. Oleh karena itu, dasar pemikiran persetujuan pemberian kredit oleh bank pada seseorang atau badan usaha adalah kepercayaan. Bila dikaitkan dengan kegiatan usaha, kredit berarti suatu kegiatan member nilai ekonomi (*econimi value*) kepada seseorang atau badan usaha yang berlandaskan kepercayaan saat itu, bahwa nilai ekonomi yang sama akan dikembalikan pada kreditur (bank) setelah jangka waktu sesuai dengan kesepakatan yang sudah disetujui antara kreditur dan debitur.

Menurut pasal 1 butir (11) UU No. 10 Tahun 1998, “kredit adalah penyediaan uang atau tagihan yang dapat dipersamakan dengan itu, berdasarkan persetujuan atau kesepakatan antara bank dengan pihak lain yang mewajibkan pihak peminjaman untuk melunasi hutangnya setelah angka waktu tertentu dengan pemberian bunga”[6].

3. ANALISA DAN PEMBAHASAN

Dalam menentukan seseorang layak atau tidak menerima kredit sepeda motor ada beberapa kriteria yang digunakan pihak PT. Astra International TBK-Toyota. Tidak hanya kriteria yang bersifat objektif, tetapi kriteria yang bersifat subjektif juga menjadi penilaian. Berdasarkan hasil wawancara dengan pihak PT. Astra International TBK-Toyota khususnya pada bagian perkreditan, diperoleh bentuk penilaian pemohon kredit berupa pertanyaan mengenai pemohon kredit dimana jawaban dari jawaban tersebut memiliki nilai yang berbeda-beda dan sebagian jawaban diperoleh dari penilaian hasil survey tim surveyor. Berikut merupakan kriteria yang digunakan sebagai penilaian pemohon kredit mobil :

1. Pekerjaan

Variabel Pekerjaan dikonversikan dengan sebuah nilai, dapat dilihat pada tabel di bawah ini

Tabel 1. Kriteria Pekerjaan

No	Pekerjaan	Nilai
1	Wiraswasta Produktivitas Rendah, Lain-lain	1
2	Wiraswasta Produktivitas Sedang	2
3	Wiraswasta Produktivitas Tinggi	3
4	Profesi	4
5	PNS/BUMN	5

2. Jumlah Tanggungan

Variabel Jumlah Tanggungan dikonversikan dengan sebuah nilai, dapat dilihat pada tabel di bawah ini

Tabel 2. Kriteria Jumlah Tanggungan

No	Jumlah tanggungan	Nilai
1	> 6 Orang	1
2	5 Orang	2
3	3-4 Orang	3
4	1-2 Orang	4
5	0	5

3. Status Rumah

Variabel status rumah dikonversikan dengan sebuah nilai, dapat dilihat pada tabel di bawah ini

Tabel 3. Kriteria Status Rumah

No	Status Rumah	Nilai
1	Kost/Kontrakan	1
2	KPR	2
3	Milik Instansi	3
4	Milik Orang tua	4
5	Milik Sendiri	5

4. Pendapatan

Variabel Pendapatan dikonversikan dengan sebuah nilai, dapat dilihat pada tabel di bawah ini

Tabel 4. Kriteria Pendapatan

No	Pendapatan	Nilai
1	< 1.000.000	1
2	1.000.000 s/d 1.500.000	2
3	1.500.000 s/d 2.500.000	3
4	2.500.000 s/d 3.500.000	4
5	> 3.500.000	5

5. Berkas Nasabah

Variabel Berkas Nasabah dikonversikan dengan sebuah nilai, dapat dilihat pada tabel di bawah ini

Tabel 5. Kriteria Berkas Nasabah

No	Berkas Nasabah	Nilai
1	Sangat Kurang	1
2	Kurang	2
3	Cukup	3
4	Baik	4
5	Sangat Baik	5

Di bawah ini adalah rekam data yang sudah ada sebelumnya, dimana data ini nantinya akan dijadikan data training dalam proses perhitungan untuk data baru.

1. Data Training

- a. Id_Nasabah : ADLAN AZHAR
Pekerjaan : Wiraswasta Produktivitas Tinggi (Bobot 3)
Jumlah Tanggungan : 2 (Bobot 3)
Status Rumah : Milik Sendiri (Bobot 5)
Pendapatan : 1,5 Juta /bulan (Bobot 2)
Berkas : Sangat Kurang (Bobot 1)
- b. Id_Nasabah : YULIANTO
Pekerjaan : PNS (Bobot 5)
Jumlah Tanggungan : 4 (Bobot 2)
Status Rumah : Milik Orang tua (Bobot 4)
Pendapatan : 2 Juta /bulan (Bobot 3)
Berkas : Cukup (Bobot 3)
- c. Id_Nasabah : SELVIANI
Pendidikan : Diploma (Bobot 4)
Pekerjaan : Akuntan (Bobot 4)
Jumlah Tanggungan : 5 (Bobot 2)
Status Rumah : KPR (Bobot 2)
Pendapatan : >3.5 Juta (Bobot 5)
Berkas : Cukup (Bobot 3)
- d. Id_Nasabah : ANDIKA
Pekerjaan : Pedagang Asongan (Bobot 1)
Jumlah Tanggungan : 3 (Bobot 3)
Status Rumah : Kontrak (Bobot 1)
Pendapatan : 1,5 Juta /bulan (Bobot 2)
Berkas : Sangat Baik (Bobot 5)

Tabel 6. Alternatif dan Kriteria

Kriteria/Alternatif	A	B	C	D
Pekerjaan	3	5	4	4
Jumlah Tanggungan	3	2	2	2
Status Rumah	5	4	2	2
Pendapatan	2	3	5	5
Berkas	1	3	3	3
Hasil	Tidak Memenuhi	Memenuhi	Memenuhi	Tidak Memenuhi

Dari data training di atas dapat di buat tabel seperti di bawah ini

Setelah mendapatkan data untuk dijadikan sebuah data training, kemudian akan dibuatkan data test atau data baru dan ditentukan apakah data baru tersebut memenuhi atau tidak memenuhi untuk mendapatkan kredit mobil

2. Data Testing

Id_Nasabah : TINA SIBURIAN

Pekerjaan : Perancang Busana (Bobot 4)

Jumlah Tanggungan : 6 (Bobot 1)

Status Rumah : Milik Sendiri (Bobot 5)

Pendapatan : 3 Juta /bulan (Bobot 4)

Berkas : Baik (Bobot 4)

Dari data diatas dapat kita buat tabel untuk data testing seperti di bawah ini

Tabel 7. Data Testing

Kriteria/Alternatif	E
Pekerjaan	4
Jumlah Tanggungan	1
Status Rumah	5
Pendapatan	4
Berkas	4
Hasil	dicari dengan Algoritma K-NN

1. Dari data training dan data testing akan digabung menjadi 1 seperti tabel di bawah ini

Tabel 8. Gabungan data training dan data Testing

No	ID Konsumen	Kriteria				
		K1	K2	K3	K4	k5
1	K001	3	3	5	2	1
2	K002	5	2	4	3	3
3	K003	4	2	2	5	3
4	K004	4	2	2	5	3
5	K005	4	1	5	4	4

2. Menghitung Euclidean Distance

Menghitung kuadrat jarak euclid (*euclidean distance*) masing-masing obyek terhadap data sampel yang diberikan

$$d_{1,d5} = \sqrt{(d_{53} - d_{13})^2 + (d_{54} - d_{14})^2 + (d_{55} - d_{15})^2 + (d_{56} - d_{16})^2 + (d_{57} - d_{17})^2}$$

$$d_{1,d5} = \sqrt{(4 - 3)^2 + (1 - 3)^2 + (5 - 5)^2 + (4 - 2)^2 + (4 - 1)^2}$$

$$d_{1,d5} = \sqrt{18}$$

$$d1,d5 = 4.242$$

$$d2,d5 = \sqrt{(d53 - d23)^2 + (d54 - d24)^2 + (d55 - d25)^2 + (d56 - d26)^2 + (d57 - d27)^2}$$

$$d2,d5 = \sqrt{(4 - 5)^2 + (1 - 2)^2 + (5 - 4)^2 + (4 - 3)^2 + (4 - 3)^2}$$

$$d2,d5 = \sqrt{5}$$

$$d2,d5 = 2.236$$

$$d3,d5 = \sqrt{(d51 - d31)^2 + (d52 - d32)^2 + (d53 - d33)^2 + (d54 - d34)^2 + (d55 - d35)^2 + (d56 - d36)^2 + (d57 - d37)^2}$$

$$d3,d5 = \sqrt{(4 - 4)^2 + (1 - 2)^2 + (5 - 2)^2 + (4 - 5)^2 + (4 - 3)^2}$$

$$d3,d5 = \sqrt{12}$$

$$d3,d5 = 3.464$$

$$d4,d5 = \sqrt{(d53 - d43)^2 + (d54 - d44)^2 + (d55 - d45)^2 + (d56 - d46)^2 + (d57 - d47)^2}$$

$$d4,d5 = \sqrt{(4 - 4)^2 + (1 - 2)^2 + (5 - 2)^2 + (4 - 5)^2 + (4 - 3)^2}$$

$$d4,d5 = \sqrt{12}$$

$$d4,d5 = 3.464$$

3. Mengurutkan objek-objek tersebut ke dalam kelompok yang mempunyai jarak euclidian terkecil Untuk mengurutkan kita hanya perlu membuat urutan dari data yang mempunyai jarak terkecil ke terbesar. Berikut hasil urutan berupa rangking

Tabel 9. Jarak euclidian terkecil

No	Nama	Distance	Ranking
1	d1,d5	4.242	4
2	d2,d5	2.236	1
3	d3,d5	3.464	2
4	d4,d5	3.464	3

4. Mengumpulkan kategori Y (klasifikasi nearest neighbor), Pada tahap ini kita hanya mengambil data sesuai dengan jumlah k yang kita tentukan di langkah 1, Pada langkah 1, k yang kita tentukan adalah k=2, jadi kita memilih 2 data terbaik saja. Hasilnya sebagai berikut.

Tabel 10. Klasifikasi nearest neighbor

No	Nama	Distance	Ranking
1	d1,d5	4.242	4
2	d2,d5	2.236	1
3	d3,d5	3.464	2
4	d4,d5	3.464	3

Jadi dari data 3 terbaik adalah data nomor 2 dan 3

2	d2,d5	2.236	1
3	d3,d5	3.464	2

5. Dengan menggunakan kategori mayoritas, maka dapat hasil klasifikasi Dari ketiga data terbaik tersebut kita harus melihat labelnya berikut data ketika dengan label sesuai dengan dataset awal

Tabel 11. Hasil Klasifikasi

No	Nama	Distance	Ranking	Hasil
2	d2,d5	2.236	1	Memenuhi
3	d3,d5	3.464	2	Memenuhi

Data terbaik adalah data ke 2 dan 3 dari data tersebut terdapat 2 Memenuhi dan 0 tidak memenuhi Sehingga mayoritas Memenuhi. Jadi hasil klasifikasi untuk Konsumen K001 adalah **Memenuhi**

4. KESIMPULAN

Berdasarkan pembahasan, implementasi dan pengujian dari sistem, diperoleh beberapa kesimpulan:

1. Sistem data mining yang dibangun dapat membantu pihak toyota untuk proses pengambilan keputusan dalam pemberian kredit mobil dengan cepat dan akurat
2. Sistem data mining ini menampilkan hasil kesimpulan dari pemohon kredit sebagai bahan pertimbangan dan alat bantu dalam pengambilan keputusan perusahaan dalam memberikan kredit mobil
3. Proses penentuan kesimpulan dari pemohon kredit yang dilakukan dengan algoritma K-NN, dimulai dengan menyiapkan data training yang sudah ada. Dari data training yang ada akan dijadikan perbandingan dengan data pemohon baru dengan menggunakan teknik euclidean distance sehingga menghasilkan nilai kedekatan antara data pemohon baru dan pemohon yang sudah ada atau data training

REFERENCES

- [1] Kusriani and Taufik Luthfi Empa, Algoritma Data Mining, Theresia Ari Prabawati, Ed. Yogyakarta, Indonesia: PENERBIT ANDI, 2009.
- [2] budi sentosa, Data Mining Terapan Dengan Matlab, 1st ed. Yogyakarta, Indonesia: Graha Ilmu, 2007.
- [3] D. Nofriansyah, Konsep Data Mining Vs Sistem Pendukung Keputusan, 2015.
- [4] A.P. Windarto, Penerapan Data Mining Pada Ekspor Buah-Buahan Menurut Negara Tujuan Menggunakan K-Means Clustering, Techno.COM. 16 (2017) 348–357.
- [5] M. Maharani, N.A. Hasibuan, N. Silalahi, S.D. Nasution, M. Mesran, S. Suginam, D.U. Sutiksno, H. Nurdianto, E. Buulolo, Y. Yuhandri, IMPLEMENTASI DATA MINING UNTUK PENGATURAN LAYOUT MINIMARKET DENGAN MENERAPKAN ASSOCIATION RULE, J. Ris. Komput. 4 (2017) 6–11. <https://www.researchgate.net/publication/312495968>.
- [6] Iko Pramudiono, Jakarta, Indonesia, 2010.
- [7] Ashok dan Mehran Sahami Srivastava, Jakarta, Indonesia, 2009.
- [8] B. Nugroho, Jakarta, Indonesia, 2004.
- [9] veithzal dan andriana permata vethzal Rivai, credit manajemen handbook. Jakarta, Indonesia, 2006.
- [10] Prastuti Sulistyorini, Algoritma. Jakarta, Indonesia, 2009.
- [11] MINING, D. (12). Junjie Wu. JAKARTA, INDONESIA.