PENERAPAN ALGORITMA K-NEAREST NEIGHBOR UNTUK PENENTUAN RESIKO KREDIT KEPEMILIKAN KENDARAAN BERMOTOR

MUHAMMAD IQBAL, 1410651080
PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK, UNIVERSITAS MUHAMMADIYAH JEMBER
2020

Analisis kredit adalah kajian yang dilakukan untuk mengetahui kelayakan dari suatu permasalahan kredit. Melalui hasil analisis kreditnya, dapat diketahui apakah usaha nasabah layak (feasible), marketable (hasil usaha dapat dipasarkan), profitable (menguntungkan), serta dapat dilunasi tepat waktu.

K-nearest neighbor (kNN) termasuk kelompok instance-based learning. Algoritma ini juga merupakan salah satu teknik lazy learning. kNN dilakukan dengan mencari kelompok k objek dalam data training yang paling dekat (mirip) dengan objek pada data baru atau data testing.

Metode KNN dapat digunakan untuk melakukan klasifikasi kelayakan pemberian kredit. Dari hasil penelitian pada bab sebelumnya maka didapatkan nilai rata-rata precision senilai 68% dan recall senilai 51%.

Kata Kunci : kredit, K-nearest neighbor, klasifikasi, precision, recall.

1. Pendahuluan

Pada masa sekarang ini, peramalan merupakan hal yang lazim digunakan pada dunia perindustrian dan juga lembaga pemerintahan. Seperti contohnya adalah peramalan permintaan produksi kain, peramalan pemakaian internet pada suatu daerah, dan peramalan penjualan produk. Hal ini juga berlaku oleh lembaga pemerintahan, seperti peramalan penerimaan pajak suatu daerah, peramalan kecelakaan lalu lintas, peramalan kepadatan arus mudik, peramalan penyebaran suatu penyakit dan prediksi dalam konteks ekonomi merupakan salah satu kegiatan matematis tertua dalam bisnis. Prediksi ini telah dilakukan bertahun-tahun sebelum penguasaan komputer, yaitu dengan kalkulator meja.

Menurut Rivai (2006), kredit didefiniskan sebagai berikut. Kredit adalah penyerahan barang, jasa, atau uang dari satu pihak (kreditor/pemberi pinjaman) atas dasar kepercayaan kepada pihak lain (nasabah atau pengutang/*borrower*) dengan janji membayar dari penerima kredit kepada pemberi kredit pada tanggal yang telah disepakati kedua belah pihak. Selain itu juga perlu analisis kredit agar kredit yang diberikan mencapai sasaran dan aman, maka analisis kredit perlu dilakukan.

Analisis kredit adalah kajian yang dilakukan untuk mengetahui kelayakan dari suatu permasalahan kredit. Melalui hasil analisis kreditnya, dapat diketahui apakah usaha nasabah layak (*feasible*), *marketable* (hasil usaha dapat dipasarkan), *profitable* (menguntungkan), serta dapat dilunasi tepat waktu (Rivai, 2006). Untuk mewujudkan hal tersebut, perlu dilakukan persiapan kredit, yaitu dengan mengumpulkan informasi dan data untuk bahan analisis. Kualitas hasil analisis tergantung pada kualitas SDM, data yang diperoleh, dan Teknik analisis.

Perusahaan kredit atau leasing memberikan jasa membantu masyarakat untuk mendapatkan kendaraan transportasi seperti sepeda motor, namun terdapat kendala-kendala yang diterima oleh perusahaaan seperti terjadinya macet dalam pembayaran kredit yang mengakibatkan kerugian yang berdampak bagi perusahaan leasing pada FIF Jember yang menimbulkan kerugian dan resiko.

Konsep penelitian dengan metode klasifikasi telah dilakukan oleh peneliti–peneliti sebelumnya diantaranya yaitu oleh

Ahmad Nouvel pada tahun 2015 mengenai penentuan klasifikasi kendaraan roda empat dengan menggunakan metode *K-Nearest Neighbor* diperoleh kesimpulan bahwa metode *K-Nearest Neighbor* dapat digunakan untuk menentukan kelayakan mobil menurut parameter kondisi fisik dari mobil tersebut, untuk data training yang berjumlah 1728 data dengan k=11 didapat nilai akurasi 95.78% dan semakin besar jumlah data training sistem akan semakin akurat. Sementara pada penelitian lain oleh P.Thamiselvan dan J.G.R. Sathiaseelan pada tahun 2016 menggunakan metode *K-Nearest Neighbor* untuk mengklasifikasikan ganas atau tidaknya kanker paru-paru melalui gambar MRI, akurasi klasifikasi yang dicapai adalah 97%. Maka ditawarkan solusi yaitu menggunakan data mining dalam proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik metode, atau algoritma dalam data mining sangat bervariasi.

Pada pengambilan keputusan selama ini, perbankan biasanya meminta pemohon (calon nasabah) mengisi formulir berupa daftar pertanyaan dan melengkapi permohonan kredit dengan berkas-berkas yang diperlukan oleh perbankan, untuk kemudian dilakukan penilaian permohonan kredit tersebut. Jika *decision maker* tepat dalam mengambil keputusan, maka pihak bank akan mendapatkan nasabah yang menyokong kesehatan dan keberlangsungan usaha perbankan, dan sebaliknya jika salah membuat keputusan maka akan menjatuhkan kelangsungan usaha perbankan.

Dari penelitian-penelitian yang pernah dilakukan, evaluasi resiko kredit merupakan masalah yang menarik dalam analisa keuangan. Penelitian mengenai analisis kelayakan pemberian kredit untuk konsumen khususnya kredit kepemilikan barang dengan metode klasifikasi data mining telah banyak dilakukan. Maka penulis dalam penelitian ini menggunakan data mining dalam penerapan algoritma *k-nearest neighbor* untuk penentuan resiko kredit kepemilikan kendaraan bermotor.

2. Kajian Pustaka Pustaka

Kredit adalah penyerahan barang, jasa, atau uang dari satu pihak (kreditor/pemberi pinjaman) atas dasar kepercayaan kepada pihak lain (nasabah atau pengutang/borrower) dengan janji membayar dari penerima kredit kepada pemberi kredit pada tanggal yang telah disepakati kedua belah pihak (Rivai, etc., 2006).

Istilah kredit berasal dari bahasa Yunani "*Credere*" yang berarti kepercayaan, atau "*Credo*" yang berarti saya percaya, karena itu dasar dari kata kredit adalah kepercayaan bahwa seseorang atau penerima kredit akan memenuhi segala sesuatu yang telah diperjanjikan terlebih dahulu pada masa yang akan datang. Pengertian kredit dalam arti ekonomi adalah suatu penundaan pembayaran, yaitu uang atau barang (prestasi) yang diterima sekarang akan dikembalikan pada masa yang akan datang berikut tambahan suatu kontra prestasi (Supriyono, 2012).

Federal International Finance (FIF GROUP) FIF GROUP adalah perusahaan pembiayaan yang didirikan pada tahun 1989 oleh Astra yang menyediakan fasilitas pembiayaan konvensional dan syariah bagi konsumen yang ingin membeli sepeda motor Honda. FIF GROUP terus berkembang dengan menambahkan rangkaian produk pembiayaan yang luas untuk memenuhi berbagai keperluan pelanggan, yang dilakukan dengan menerapkan manajemen risiko yang terintegrasi.

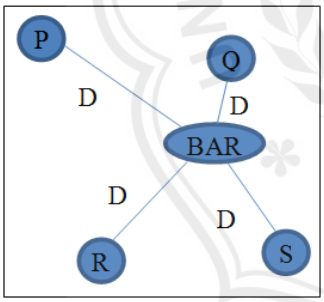
Pada akhir tahun 2017, FIF GROUP mengelola jaringan penjualan dengan 196 kantor cabang dan 416 titik layanan untuk menghadirkan standar layanan terbaik dan jangkauan pelanggan yang luas di seluruh penjuru Indonesia. Akses distribusi juga dioptimalkan melalui pengembangan jalinan bisnis kemitraan, antara lain dengan dealer/jaringan ritel/ATPM, seperti Lotte Mart yang menyediakan fasilitas instalasi

kios lotte, serta kerja sama dalam skema pelanggan grup dengan korporasi, termasuk Glaxo Smith Kline, untuk pengadaan armada sepeda motor yang memberikan kontribusi nilai pembiayaan dalam jumlah besar dengan tingkat risiko yang terkendali.

K-Nearest Neighbor (kNN) termasuk kelompok *instance-based learning*. Algoritma ini juga merupakan salah satu teknik lazy learning. kNN dilakukan dengan mencari kelompok k objek dalam data training yang paling dekat (mirip) dengan objek pada data baru atau data testing (Wu, Cai, & Ao, 2012).

Contoh kasus, misal diinginkan untuk mencari solusi terhadap masalah seorang pasien baru dengan menggunakan solusi dari pasien lama. Untuk mencari solusi dari pasien baru tersebut digunakan kedekatan dengan kasus pasien lama, solusi dari kasus lama yang memiliki kedekatan dengan kasus baru digunakan sebagai solusinya.

Terdapat pasien baru dan 4 pasien lama, yaitu P, Q, R, dan S (Gambar 2). Ketika ada pasien baru maka yang diambil solusi adalah solusi dari kasus pasien lama yang memikili kedekatan terbesar.



Gambar 2.2 ilustrasi kasus algoritma kNN

Misal D1 adalah jarak antara pasien baru dengan pasien P, D2 adalah jarak antara pasien baru dengan pasien Q, D3 adalah jarak antara pasien baru dengan pasien R, D4 adalah jarak antara pasien baru dengan pasien S. Dari ilustrasi gambar terlihat bahwa D2 yang paling terdekat dengan kasus baru. Dengan demikian maka solusi dari kasus pasien Q yang akan digunakan sebagai solusi dari pasien baru tersebut.

Ada banyak cara untuk mengukur jarak kedekatan antara data baru dengan data lama (data training), diantaranya *euclidean distance* dan manhattan *distance* (*city block distance*), yang paling sering digunakan adalah *euclidean distance* (Bramer,2007), yaitu:

$$d = \sqrt{(a_1 + b_1)^2 + (a_2 + b_2)^2 + \cdots + (a_n + b_n)^2}$$

Dimana a = a_1, a_2, …, an, dan b = b_1, b_2, …, b_n mewakili n nilai atribut dari dua record.

Untuk atribut dengan nilai kategori, pengukuran dengan *euclidean distance* tidak cocok. Sebagai penggantinya, digunakan fungsi sebagai berikut (Larose, 2006):

$$different(a_i, b_i) = \begin{cases} 0; \ jika\ a_i = b_i \\ 1; \ selainnya \end{cases}$$

Dimana ai dan bi adalah nilai kategori. Jika nilai atribut antara dua *record* yang dibandingkan sama maka nilai jaraknya 0, artinya mirip, sebaliknya, jika berbeda maka nilai kedekatannya 1, artinya tidak mirip sama sekali. Misalkan atribut warna dengan nilai merah dan merah, maka nilai kedekatannya 0, jika merah dan biru maka nilai kedekatannya

Untuk mengukur jarak dari atribut yang mempunyai nilai besar, seperti atribut pendapatan, maka dilakukan normalisasi. Normalisasi bisa dilakukan dengan *min-max normalization* atau *Z-score standardization* (Larose, 2006). Jika data training terdiri dari atribut campuran antara numerik dan kategori, lebih baik gunakan *min-max normalization* (Larose, 2006).

Untuk menghitung kemiripan kasus, digunakan rumus (Kusrini & Lutfi, 2009):

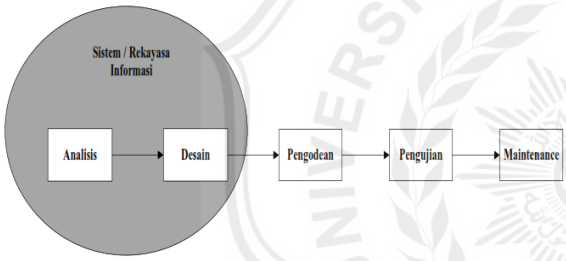
$$Similarity(p, q) = \frac{\sum_{t=1}^{n} f(p_i, q_i) * w_i}{w_i}$$

Keterangan :
p = Kasus baru
q = Kasus yang ada dalam penyimpanan
n = Jumlah atribut dalam tiap kasus

i = Atribut individu antara 1 sampai dengan n
f = Fungsi similarity atribut i antara kasus p dan kasus q
w = Bobot yang diberikan pada atribut ke-i

3. Metodologi Penelitian

Metode yang digunakan dalam penelitian ini yaitu metode *waterfall*. Model SDLC air terjun (*waterfall*) sering juga disebut model sekuensial linier (sequential linear) atau alur hidup klasik (*classic life cycle*). Model air terjun (*waterfall*) menyediakan pendekatan alur hidup perangkat lunak secara sekuensial atau terurut dimulai analisis, desain, pengodean, pengujian, dan tahap pendukung (*support*). Berikut adalah gambar model air terjun (*waterfall*) (Sommerville, 2011):



Gambar 3.1. Ilustrasi model *waterfall*

a. Analisis kebutuhan perangkat lunak
Proses pengumpulan kebutuhan dilakukan secara intensif untuk menspesifikasikan kebutuhan perangkat lunak agar dapat dipahami perangkat lunak seperti apa yang dibutuhkan user. Spesifikasi kebutuhan perangkat lunak pada tahap ini perlu untuk didokumentasikan.
b. Desain (perancangan)
Desain perangkat lunak adalah proses multi langkah yang fokus pada desain pembuatanprogram perangkat lunak termasuk struktur data, arsitektur perangkat lunak, representasi antar muka, dan prosedur pengodean. Tahap ini mentranslasi kebutuhan perangkat lunak dari tahap analisis kebutuhan ke representasi desain agar dapat diimplementasikan menjadi program pada tahap selanjutnya. Desain perangkat lunak yang dihasilkan pada tahap ini juga perlu didokumentasikan.
c. Pembuatan kode program
Desain harus ditranslasikan kedalam program perangkat lunak. Hasil dari tahap ini adalah program komputer sesuai dengan desain yang telah dibuat pada tahap desain.
d. Pengujian
Pengujian fokus pada perangkat lunak secara dari segi lojik dan fungsional dan memastikan bahwa semua bagian sudah diuji. Hal ini dilakukan untuk meminimalisir kesalahan (*error*) dan memastikan keluaran yang dihasilkan sesuai dengan yang diinginkan.
e. Pendukung (*support*) atau pemeliharaan (*maintenance*)
Tidak menutup kemungkinan sebuah perangkat lunak mengalami perubahan ketika sudah dikirimkan ke user. Perubahan bisa terjadi karena adanya kesalahan yang muncul dan tidak terdeteksi saat pengujian atau perangkat lunak harus beradaptasi dengan lingkungan yang baru. Tahap pendukung atau pemeliharaan dapat mengulangi proses pengembangan mulai dari analisis spesifikasi untuk perubahan perangkat lunak yang sudah ada, tapi tidak untuk membuat perangkat lunak baru.

Model air terjun (*waterfall*) sangat cocok digunakan untuk kebutuhan pelanggan yang sudah sangat dipahami dan kemungkinan terjadinya perubahan kebutuhan selama pengembangan perangkat lunak kecil. Hal positif dari model air terjun (*waterfall*) adalah struktur tahap pengembangan sistem jelas, dokumentasi dihasilkan di setiap tahap pengembangan, dan sebuah tahap dijalankan setelah tahap sebelumnya selesai dijalankan.

Metode ini digunakan karena merupakan suatu metode yang praktis dan cukup menghemat biaya karena semua

parameter-parameter yang dibutuhkan serta hasil yang diinginkan dapat langsung dimodelkan dan disimulasikan dengan menggunakan suatu program komputer (Personal Komputer) dalam bentuk perangkat lunak berbasis sistem pakar.

4. Hasil dan Pembahasan

Pengujian sistem merupakan serangkaian tahapan untuk menguji ketangguhan dari sistem yang telah dibangun, dengan tujuan untuk menemukan kesalahan-kesalahan atau kekurangan-kekurangan yang tidak ditemukan sebelumnya pada sistem, dan apakah sistem yang dibangun telah sesuai dengan perencanaan awal. Pada tahap pengujian sistem terdiri dari beberapa bagian, yaitu skenario pengujian, hasil pengujian, dan kesimpulan pengujian.

Kriteria	**Alternatif**	**Layak**		**Tidak Layak**	
		Kelas	**Responden**	**Kelas**	**Responden**
Perkawinan	Janda/ Duda	0,391	0,709	0,552	1
Jumlah Tanggungan	1 Orang	0,174	0,504	0,138	0,4
Umur	21-55 Tahun	0,232	0,448	0,207	0,4
Kepemilikan Rumah	Dinas	0,058	1	0,103	0,56
Lama Tinggal	< 1 Tahun	0,493	0,14	0,483	0,143
Pekerjaan	Jaksa	0,232	0,438	0,207	0,49
Penghasilan	>3x Angsuran	0,319	0,324	0,103	1

Berdasarkan bobot diatas maka jika diimplementasikan ke *Euclidean distance* sebagai berikut:

$$Perkawinan\ (Layak|Janda/Duda) = \sqrt{(0{,}391 - 0{,}709)^2} = 0{,}318$$

$$Jumlah\ Tanggungan\ (Layak|1\ Orang) = \sqrt{(0{,}174 - 0{,}504)^2} = 0{,}33$$

$$Umur(Layak|\ 21 - \ 55\ \text{tahun}) = \sqrt{(0{,}232 - 0{,}448)^2} = 0{,}216$$

$$\text{Kepemilikan Rumah}(Layak|\text{Dinas}) = \sqrt{(0{,}058 - 1)^2} = 0{,}942$$

$$\text{Lama Tinggal}(Layak| < 1\text{Tahun}) = \sqrt{(0{,}493 - 0{,}14)^2} = 0{,}353$$

$$\text{Pekerjaan}(Layak|\text{Jaksa}) = \sqrt{(0{,}232 - 0{,}438)^2} = 0{,}206$$

$$\text{Penghasilan}(Layak| > 3\text{x Angsuran}) = \sqrt{(0{,}319 - 0{,}324)^2} = 0{,}005$$

$$\mathbf{1.}\boldsymbol{Layak} = \mathbf{0,318 + 0,33 + 0,216 + 0,942 + 0,353 + 0,206 + 0,005 = 2,37}$$

$$Perkawinan\ (T.Layak|Janda/Duda) = \sqrt{(0{,}552 - 1)^2} = 0{,}448$$

$$Jumlah\ Tanggungan\ (T.Layak|1\ Orang) = \sqrt{(0{,}138 - 0{,}4)^2} = 0{,}262$$

$$Umur(T.Layak|21 - \ 55\ \text{tahun}) = \sqrt{(0{,}207 - 0{,}4)^2} = 0{,}193$$

$$\text{Kepemilikan Rumah}(T.Layak|\text{Dinas}) = \sqrt{(0{,}103 - 0{,}56)^2} = 0{,}457$$

$$\text{Lama Tinggal}(T.Layak| < 1\ \text{Tahun}) = \sqrt{(0{,}483 - 0{,}143)^2} = 0{,}34$$

$$\text{Pekerjaan}(T.Layak|\text{Jaksa}) = \sqrt{(0{,}207 - 0{,}49)^2} = 0{,}283$$

$$\text{Penghasilan}(T.Layak| > 3\text{x Angsuran}) = \sqrt{(0{,}103 - 1)^2} = 0{,}897$$

$$\mathbf{2.}\boldsymbol{Tidak\ Layak} = \mathbf{0,448 + 0,262 + 0,193 + 0,457 + 0,34 + 0,283 + 0,}897 = \mathbf{2,88}$$

Dari pengajuan 4 diatas, maka dapat di simpulkan bahwa pengajuan 4 Layak untuk pemberian kredit. Alasannya, karena score K-NN jarak terdekat yaitu Layak (2,37) < Tidak Layak (2,88)

5. Kesimpulan

Dari hasil penelitian pada bab sebelum maka dapat di simpulkan hasil penelitian sebagai berikut:

1. Metode KNN dapat digunakan untuk melakukan klasifikasi kelayakan pemberian kredit.
2. Dari hasil penelitian pada bab sebelumnya maka didapatkan nilai rata-rata precision senilai 68% dan recall senilai 51%.

DAFTAR PUSTAKA

Bramer, Max. 2007. Principles of Data Mining. London : Springer.

Han, J, Kamber, M, & Pei, J. 2006. Data Mining: Concept and Techniques, Second Edition. Waltham: Morgan Kaufmann Publishers.

Junaedi, Fajar. 2005. Panduan Lengkap Pemrograman PHP untuk Membuat WEB Dinamis. Yogyakarta:PD. Anindya

Kustiyahningsih, Y, Anamisa, R,D, 2011.Pemograman Basis Data Berbasis Web Menggunakan PHP & MySQL.Graha Ilmu : Yogyakarta.

Kusrini & Luthfi,E.T. 2009. Algoritma Data Mining. Yogyakarta : Andi Publishing.

Larose, D. T. 2005. Discovering Knowledge in Data. New Jersey : John Willey & Sons, Inc.

Murdianingsih, Y, Lukmana, A. 2017, Sistem Pendukung Keputusan Analisis Pengajuan Kredit Motor Menggunakan Metode KNN (K Nearest Neighbor) (Study Kasus pada PT. Federal International Finence (FIF) Cabang Subang), Subang, Program Studi Teknik Informatika, STMIK Subang

Render, B., Heizer, J., 2001, Prinsip-prinsip Manajemen Operasi, PT. Salemba Emban Patria, Jakarta.

Rivai, Veithzal., &Veithzal, Andria Permata. 2006. Credit Management Handbook. Jakarta : Raja GrafindoPersada.

Roger S. Pressman, 2002. Rekayasa Perangkat Lunak Pendekatan Praktisi (Buku Satu), ANDI Yogyakarta.

Sumathi, & S., Sivanandam, S.N. 2006. Introduction to Data Mining and its Applications. Berlin Heidelberg New York: Springer

Supriyono, M, 2011, Buku Pintar Perbankan, Yogyakarta, Penerbit Andi.

Supardi, Y., 2000, Pascal dan Flowchart, Dinastindo, Jakarta.

Sibero, Alexander F.K. 2012. “Kitab Suci Web Programing”. Jakarta: Mediakom.

Sommerville, Ian. 2011. Software Engineering (Rekayasa Perangkat Lunak). Jakarta: Erlangga.

Witten, I. H., Frank, E., & Hall, M. A. 2011. Data Mining: Practical Machine Learning and Tools. Burlington : Morgan Kaufmann Publisher.

Witten, etc., 2011. Data Mining Practical Machine Learning Tools and Techniques (3rd ed). USA: Elsevier.

Wu, Cai, & Ao, 2012, *Hybrid dynamic k-nearest-neighbour and distance and attribute weighted method for classification*, Wuhan,School of Computer Science, China University of Geosciences.