

Sistem Pendukung Keputusan Persetujuan Kredit dengan Pemanfaatan Information Gain untuk Pembobotan Atribut Klasifikasi K-Nearest Neighbour

Ivandari, Tria Titiani Chasanah, M. Adib Al Karomi*
STMIK Widya Pratama Pekalongan
E-mail: ivandarialkaromi@gmail.com

RINGKASAN

Kredit merupakan salah satu perilaku ekonomi modern. Dalam prakteknya kredit dapat berupa peminjaman sejumlah uang atau pembelian barang dengan proses pembayaran secara bertahap dan dalam jangka waktu yang telah disepakati. Kondisi perekonomian yang kurang mendukung serta kebutuhan masyarakat yang tinggi membuat masyarakat memilih membeli barang dengan proses kredit ini. Sayangnya kebutuhan yang tinggi ada yang tidak diimbangi dengan kemampuan melakukan pembayaran sesuai dengan perjanjian awal. Kondisi ini yang memungkinkan proses pembayaran menjadi tidak lancar atau disebut juga dengan istilah kredit macet. Penelitian ini menggunakan data publik yaitu dataset kartu kredit dari UCI repository serta data private yaitu dataset persetujuan kredit dari perbankan lokal. Algoritma information gain digunakan untuk menghitung bobot dari masing-masing atribut yang ada. Dari hasil perhitungan diketahui bahwa semua atribut memiliki bobot yang berbeda. Penelitian ini menghasilkan kesimpulan bahwa tidak semua atribut data mempengaruhi hasil klasifikasi. Misalkan atribut A1 pada dataset UCI serta atribut type pinjaman pada dataset lokal yang memiliki bobot information gain 0 (nol). Hasil klasifikasi menggunakan algoritma K-Nearest Neighbour menunjukkan bahwa terjadi peningkatan sebesar 7,53% untuk dataset UCI dan 3,26% untuk dataset lokal setelah dilakukan seleksi fitur pada kedua dataset. Dari peningkatan tersebut maka Sistem Pendukung Keputusan yang dibuat hanya menggunakan atribut yang terpilih untuk meningkatkan akurasi. SPK yang tercipta dapat memberikan rekomendasi kepada manajemen berdasarkan perhitungan dan kedekatan data testing dengan keseluruhan data training yang telah ada sebelumnya.

Kata Kunci : Akurasi KNN, Seleksi Fitur, Rekomendasi Persetujuan Kredit

1. PENDAHULUAN

1.1 Latar Belakang Masalah

Perkembangan perekonomian Indonesia pada saat ini tergolong cepat. Hal ini dibuktikan dengan berkembangnya badan ekonomi dari tingkat lokal maupun nasional. Koperasi dan Bank merupakan salah satu contoh badan ekonomi yang penyebarannya tergolong luas dan cepat. Salah satu produk unggulan dari Bank dan Koperasi adalah pinjaman berupa dana yang dapat dibayarkan secara berkala atau biasa dikenal dengan istilah kredit. Kredit banyak diminati masyarakat kalangan menengah untuk membeli barang dengan proses pembayaran secara berkala dengan syarat dan ketentuan yang berlaku. Dari data yang ada menunjukkan presentase nasabah yang tergolong dalam status kredit macet jauh lebih banyak dari nasabah

dengan status kredit lancar. Nasabah dengan status kredit macet adalah mereka yang dalam pembayarannya mengalami kondisi yang tidak sesuai dengan aturan yang berlaku, jatuh tempo yang telah berlalu dengan angsuran yang belum terselesaikan ataupun angsuran yang tidak pernah terbayarkan.

Seiring perkembangan teknologi, kemungkinan calon nasabah yang tergolong status kredit macet dapat diprediksi menggunakan sebuah sistem. Sistem Pendukung Keputusan (SPK) dapat menjadi salah satu alternatif atau acuan sebuah perusahaan memberikan persetujuan kredit pada calon nasabahnya. Salah satu contohnya adalah SPK pada PT.BPR Artamanunggal Abadi Mranggen (Suhari, Sukur, and Eniyati 2009). SPK serupa juga pernah dibuat oleh mahasiswa STMIK AMIKOM Yogyakarta dengan studi

kasus di Bank Muamalat cabang Yogyakarta (Dwi Cahyani 2013).

Beberapa SPK juga telah menggunakan metode data mining untuk memprediksi kriteria nasabah kredit (Ginanjar Mabur and Lubis 2012). Salah satu metode data mining yang pernah digunakan untuk SPK persetujuan kredit adalah metode bayes (Zahid 2013). Selain itu metode data mining juga digunakan untuk proses analisis data pembayaran kredit pada nasabah bank (Melissa and Oetama 2013). Metode Analytical Hierarchy Process (AHP) juga pernah diimplementasikan untuk SPK pemberian kredit pada nasabah bank (Azwany 2010). Selain untuk memprediksi kriteria nasabah perbankan banyak juga perusahaan pembiayaan seperti halnya pembiayaan motor ataupun mobil yang telah menerapkan metode sejenis. SPK dengan menggunakan metode Simple Additive Weighting (Wahyu Oktaputra and Noersasongko 2014) pernah dibuat untuk nasabah kredit motor. Sedangkan dengan menggunakan metode Scoring Sistem (Wasana 2010) juga pernah tercipta sebuah SPK sejenis.

Banyak algoritma yang dapat digunakan dalam proses klasifikasi data mining. Salah satu algoritma klasifikasi terbaik dan banyak digunakan adalah K-Nearest Neighbour (Wu et al. 2007). Salah satu kelemahan algoritma KNN adalah dalam penentuan parameter k agar tingkat akurasi klasifikasi naik masih digunakan secara manual (Al Karomi 2015). Optimasi parameter k pernah dibahas dalam seminar nasional teknologi dan informatika (Indrayanti, Sugianti, and Al Karomi 2017).

Performa suatu algoritma dapat dipengaruhi oleh dataset dan tipe data yang digunakan (Amancio et al. 2013). Beberapa model algoritma kuat pada tipe data tertentu dan lemah pada tipe data yang lain (Ragab et al. 2014; Patel, Vala, and Pandya 2014; Ashari, Paryudi, and Tjoa 2013). Banyaknya atribut yang relevan yang dipakai dalam klasifikasi dapat mempengaruhi hasil akurasi dari algoritma tersebut (Han and Kamber 2006). Semakin banyak atribut yang relevan digunakan dalam proses klasifikasi akan meningkatkan akurasi dari sebuah algoritma (Maimoon and Rokach 2010; Alpaydin 2010). Atribut yang tidak relevan dalam dataset dapat membebani kinerja algoritma dan menurunkan performa algoritma klasifikasi (Karegowda, Manjunath, and Jayaram 2010).

Seleksi fitur merupakan salah satu tahapan pre processing klasifikasi dengan cara menghilangkan fitur yang tidak relevan dalam data. Proses ini juga dapat mengurangi dimensi data serta meningkatkan akurasi klasifikasi. Salah satu algoritma seleksi fitur yang banyak dipakai dan populer adalah Information gain (Azhagusundari and Thanamani 2013) (Alkaromi 2014). Kelebihan dari information gain adalah baik digunakan dalam memilih atribut khususnya dalam menangani data dengan dimensi tinggi (Koprinska 2010). Atribut yang terlalu banyak dapat membebani proses klasifikasi dan pada akhirnya akan membuat biaya komputasi semakin mahal.

Hasil penelitian telah membuktikan bahwa penerapan algoritma information gain untuk seleksi fitur mempengaruhi performa algoritma KNN dalam klasifikasi persetujuan kredit. Performa algoritma KNN naik hingga lebih dari 3% dengan menggunakan seleksi fitur dibandingkan dengan algoritma KNN tanpa seleksi fitur (Ivandari et al. 2017).

Dari beberapa penelitian terkait yang telah terpapar sebelumnya diketahui bahwa SPK mampu melakukan deteksi sejak dini nasabah pemohon kredit (Maulana and Al Karomi 2016). SPK tersebut nantinya dapat memberikan rekomendasi bagi pihak pemberi kredit untuk melakukan tindakan selanjutnya memberikan kredit atau menolak aplikasi kredit yang diajukan. Dalam penelitian ini akan dibuat sebuah SPK persetujuan kredit menggunakan algoritma seleksi fitur information gain untuk pembobotan atribut klasifikasi KNN. SPK ini nantinya dapat digunakan sebagai salah satu rekomendasi manajemen untuk menentukan persetujuan kredit nasabah baru.

1.2 Pengembangan Sistem

Dalam penelitian ini dilakukan dalam beberapa tahapan. Untuk pencapaian utama dilakukan pengumpulan data, berikutnya dilakukan seleksi fitur menggunakan algoritma *information gain*. Berikutnya atribut terpilih digunakan untuk membangun sebuah sistem pendukung keputusan. Tahapan berikut secara lebih terinci akan dijelaskan dalam sub bab berikut.

1.2.1 Pengumpulan data

Dalam dataset yang ada terdapat keseluruhan 766 record nasabah penerima pinjaman kredit. 210 dari nasabah tersebut tergolong dalam nasabah dengan kriteria kredit lancar dan sisanya 556 nasabah tergolong dalam kredit macet. Dalam dataset terdapat 16 atribut data, satu diantaranya merupakan atribut id yaitu nama nasabah serta satu atribut lain menjadi atribut label atau sering

juga disebut dengan atribut tujuan yaitu status kredit. Sejumlah 14 atribut lain merupakan atribut reguler yang akan dilakukan proses pembobotan menggunakan information gain untuk proses klasifikasi dengan menggunakan algoritma KNN. Tabel 1 merupakan metadata dari dataset yang akan digunakan untuk persetujuan kredit.

Tabel 1. Metadata Data Persetujuan Kredit

No	Atribut	Type	Range
1	Nama	Polynomial (id)	~
2	Jenis Kelamin	Binominal	L, P
3	Umur	numerik	~
4	Jumlah Pinjaman	numerik	~
5	Jangka Waktu	Polynomial	~
6	Angsuran Perbulan	numerik	~
7	Type Pinjaman	nominal	100
8	Jenis Pinjaman	Polynomial	301, 302, 303, 304, 305
9	BI Sektor Ekonomi	Polynomial	6000, 8000, 9990
10	col	Binominal	1, 2
11	BI Golongan Debitur	Polynomial	834, 874, 876
12	BI Golongan Penjamin	Polynomial	000, 800, 835, 874, 875
13	Saldo Nominatif	numerik	~
14	Tunggakan Pokok	numerik	~
15	Tunggakan Bunga	numerik	~
16	Status Kredit	Binominal (label)	Lancar, Macet

1.2.2 Seleksi fitur *information gain*

Setelah tahap pengumpulan data diselesaikan, tahap berikutnya adalah seleksi fitur. Dalam tahap ini akan dilakukan pembobotan semua atribut yang ada dalam dataset sejumlah 14 atribut reguler. Setelah bobot dari keseluruhan atribut diketahui kemudian dilakukan pemilihan atribut sesuai dengan bobot yang ada. Atribut dengan nilai information gain kurang dari nilai treshold atau nilai batas akan dibuang dan tidak akan digunakan dalam klasifikasi selanjutnya. Atribut dengan nilai information gain sama dengan dan lebih besar dari nilai batas yang akan digunakan dalam proses berikutnya. Dataset baru adalah dataset yang telah direvisi dengan hanya

menggunakan atribut terpilih dengan nilai bobot sama dengan atau lebih besar dari nilai treshold.

1.2.3 Pembuatan sistem

Tahap berikutnya setelah desain sistem selesai adalah tahap pembuatan atau coding. Dalam tahap ini akan dibuat sebuah SPK persetujuan kredit menggunakan Microsoft Excell. Ms. Excell digunakan karena dalam penggunaannya mudah diimplementasikan dan menggunakan rumus atau formula yang ada. Proses ini melakukan perhitungan terhadap keseluruhan data training yang ada terhadap data testing. Perhitungan ini untuk mencari nilai kedekatan dengan memanfaatkan bobot yang telah dihitung sebelumnya menggunakan aplikasi rapid miner.

1.3 Pengujian sistem

Tahap terakhir dalam pembuatan SPK adalah tahap pengujian. Tahap ini dibagi menjadi dua yaitu pengujian sistem dan pengujian user. Pengujian sistem dilakukan untuk menguji code program yang ada. Dengan kata lain pengujian sistem adalah pengujian untuk sistem yang terbuat. Lalu pengujian user dilakukan sebagai salah satu indikator keberhasilan program dilihat dari sisi pengguna atau user. Pengujian ini dilakukan dengan wawancara kepada pengguna sistem baru atau menyebar kuisioner.

2. HASIL DAN PEMBAHASAN

2.1 Hasil seleksi fitur *information gain*

Dari proses perhitungan yang telah dilakukan didapatkan hasil seleksi fitur menggunakan *information gain* adalah seperti pada tabel 2 berikut. Keseluruhan atribut yaitu 14 memperoleh masing masing bobot gain yang berbeda.

Tabel 2. Bobot atribut data

Atribut	Weight
type_pinjaman	0.0
bi_golongan_debitur	0.0036749721875092
bi_sektor_ekonomi	0.0055560008895870
jenis_kelamin	0.0066767845564446
jenis_pinjaman	0.0123852930711034
umur	0.0239071875542023
jkw	0.0256974344524127
saldo_nominatif	0.0540959633669987
jml_angsuran_per_bulan	0.1401409412988560
jml_pinjaman	0.2043747467291951
col	0.2241429158475250
bi_gol_penjamin	0.2396212592889049
tunggakan_bunga	0.2842707596587644
tunggakan_pokok	1.0

2.2 Hasil Sistem

Sistem yang tercipta menggunakan basis aplikasi Microsoft Excell. Sistem tersebut juga dapat diakses serta diunduh dari: <http://repository.stmik-wp.ac.id/index.php/download/file/UkVQTzE1MDg5MTAwMjMjMTFtMDgxMS0wMDAxI3BlawVWtMWVhZzc1> Gambar 1 merupakan tampilan utama aplikasi yang dapat digunakan.



Gambar 1 Tampilan Aplikasi

Aplikasi tersebut dapat digunakan dengan hanya merubah atribut data *testing* dengan kondisi calon nasabah. Lalu dari hasil perhitungan apabila dilakukan sorting maka hasil akan tampil. Hasil klasifikasi antara LANCAR atau MACET. Hasil tersebut memiliki tingkat kedekatan dengan data *training* yang dapat dilihat pada teks warna putih di bagian bawah aplikasi.

2.3 Pembahasan

Sistem yang tercipta masih memiliki beberapa kelemahan seperti misalnya ada atribut yang masih belum dapat dipahami yaitu atribut col yang berisi 1 atau 2, serta atribut bi golongan penjamin dengan 5 varian yaitu 000, 800, 835, 874 serta 875. Tabel 3 merupakan hasil analisa dari atribut bi golongan penjamin. Sedangkan tabel 4 merupakan analisa dari atribut col.

Tabel 3. Analisa atribut bi golongan penjamin

Atribut bi golongan penjamin	Lancar	Macet
0	197	322
800	0	8
835	0	1
874	0	9
875	13	216

Tabel 4. Analisa Atribut col

Atribut col	Lancar	Macet
1	208	392
2	2	164

Kedua atribut tersebut masih belum dapat digunakan sebagaimana mestinya untuk dataset lokal. Atau untuk calon nasabah baru pada Baitul maal waa Tamwil (BMT).

3. KESIMPULAN

Dari sistem yang tercipta dengan memanfaatkan Microsoft Excell sudah dapat melakukan perhitungan kedekatan diantara data training dengan data testing. Sistem tersebut dapat melakukan klasifikasi dengan memanfaatkan algoritma knn. Hasil klasifikasi yang diambil adalah data training yang paling mendekati data testing. Dalam hal ini kedekatan mutlak adalah 1 artinya nilai kedekatan yang paling mendekati 1 yang diambil sebagai penunjang keputusan. Aplikasi yang tercipta masih harus menggunakan tombol bantu. Tombol bantu tersebut dibuat dengan menggunakan sorting atau pengurutan berdasarkan nilai yang paling tinggi. Apabila ada yang memiliki nilai kedekatan yang sama maka sistem ini belum dapat melakukan deteksi dan hanya akan mengambil salah satu dari keduanya.

4. SARAN

Aplikasi yang tercipta dengan memanfaatkan Microsoft Excell dapat melakukan perhitungan kedekatan dengan tepat. Selain itu hasil klasifikasi juga dapat diketahui dengan menggunakan beberapa formula yang telah disediakan. Salah satu kelemahan sistem yang tercipta adalah user harus menekan tombol untuk melakukan pengurutan guna mencari kedekatan yang terbaik. Untuk tahap berikutnya dapat dilakukan dengan memanfaatkan aplikasi yang dapat melakukan perhitungan urutan secara otomatis.

5. UCAPAN TERIMA KASIH

Penelitian ini terlaksana atas kerjasama yang baik dari pihak Pusat Peneliti dan Pengabdian kepada Masyarakat STMIK Widya Pratama. Penelitian ini didanai oleh Direktorat Riset dan Pengembangan Kementerian Riset Teknologi dan Perguruan Tinggi

6. DAFTAR PUSTAKA

Alkaromi, M Adib. 2014. "Information Gain Untuk Pemilihan Fitur Pada Klasifikasi Heregistrasi Calon Mahasiswa Dengan Menggunakan K-NN."

Alpaydin, Ethem. 2010. *Introduction to Machine Learning Second Edition*. London: The MIT Press.

Amancio, D. R., C. H. Comin, D. Casanova, G. Travieso, O. M. Bruno, F. a. Rodrigues, and L. Da F. Costa. 2013. "A Systematic Comparison of Supervised Classifiers," October. <http://arxiv.org/abs/1311.0202v1>.

Ashari, Ahmad, Iman Paryudi, and A Min Tjoa. 2013. "Performance Comparison between Naïve Bayes , Decision Tree and K-Nearest Neighbor in Searching Alternative Design in an Energy Simulation Tool" 4 (11): 33–39.

Azhagusundari, B, and Antony Selvadoss Thanamani. 2013. "Feature Selection Based on Information Gain," no. 2: 18–21.

Azwany, Faraby. 2010. "Sistem Pendukung Keputusan Pemberian Kredit Usaha Rakyat Pada Bank Syariah Mandiri Cabang Medan Menggunakan Metode AHP." *Program Studi Ilmu Komputer Departemen Ilmu Komputer Fakultas Matematika Dan Ilmu Pengetahuan Alam Universitas Sumatera Utara Medan*.

Dwi Cahyani, Bangun. 2013. "Sistem Pendukung Keputusan Persetujuan Permohonan Kredit Pinjaman Pada Bank Muamalat Indonesia Cabang Yogyakarta." *Sekolah Tinggi Manajemen Informatika Dan Komputer AMIKOM Yogyakarta*.

Ginanjari Mabrur, Angga, and Riani Lubis. 2012. "Penerapan Data Mining Untuk Memprediksi Kriteria Nasabah Kredit." *Jurnal Komputer Dan Informatika (KOMPUTA)* 1.

Han, Jiawei, and Micheline Kamber. 2006. "Data Mining: Concepts and Techniques Second Edition" 40 (6). Elsevier: 9823. doi:10.1002/1521-3773(20010316)40:6<9823::AID-ANIE9823>3.3.CO;2-C.

Indrayanti, Devi Sugianti, and M Adib Al Karomi. 2017. "Optimasi Parameter K Pada Algoritma K-Nearest Neighbour Untuk Klasifikasi Penyakit Diabetes Mellitus." *Prosiding SNATIF Buku 3 (2017)*: 823–29.

Ivandari, Tria Titiani Chasanah, Satriedi Wahyu Binabar, and M Adib Al Karomi. 2017. "Data Attribute Selection with Information Gain to Improve Credit Approval Classification Performance Using K-Nearest Neighbor Algorithm." *IJIBEC I*:

15–24.

- Karegowda, Asha Gowda, A S Manjunath, and M. A Jayaram. 2010. “Comparative Study of Attribute Selection Using Gain Ratio and Correlation Based Feature Selection.” *International Journal of Information Technology and Knowledge Management* 2 (2): 271–77.
- Karomi, M Adib Al. 2015. “Optimasi Parameter K Pada Algoritma KNN Untuk Klasifikasi Heregistrasi Mahasiswa Program Studi Teknik Informatika STMIK Widya Pratama Jl . Patriot 25 Pekalongan Email : Adib.comp@gmail.com.” *IC-TECH X* (285): 5.
- Koprinska, Irena. 2010. “Feature Selection for Brain-Computer Interfaces,” 100–111.
- Maimoon, Oded, and Lior Rokach. 2010. *Data Mining and Knowledge Discovery Handbook*. Vol. 40. Springer. doi:10.1002/1521-3773(20010316)40:6<9823::AID-ANIE9823>3.3.CO;2-C.
- Maulana, Much. Rifqi, and M Adib Al Karomi. 2016. “Sistem Pendukung Keputusan Persetujuan Kredit Menggunakan Algoritma C4.5.” *Jurnal IC-Tech* Vol. XI No (1): 29–38. <http://jurnal.stmik-wp.ac.id/gdl.php?mod=browse&op=read&id=ictech--muchrifqim-80>.
- Melissa, Ira, and Raymond S Oetama. 2013. “Analisis Data Pembayaran Kredit Nasabah Bank Menggunakan Metode Data Mining” *IV* (1): 18–27.
- Patel, Kanu, Jay Vala, and Jaymit Pandya. 2014. “Comparison of Various Classification Algorithms on Iris Datasets Using WEKA” *1* (1): 1–7.
- Ragab, Abdul Hamid M., Amin Y. Noaman, Abdullah S. Al-Ghamdi, and Ayman I. Madbouly. 2014. “A Comparative Analysis of Classification Algorithms for Students College Enrollment Approval Using Data Mining.” *Proceedings of the 2014 Workshop on Interaction Design in Educational Environments - IDEE '14*. New York, New York, USA: ACM Press, 106–13. doi:10.1145/2643604.2643631.
- Suhari, Yohanes, Muji Sukur, and Sri Eniyati. 2009. “SISTEM PENDUKUNG KEPUTUSAN PEMBERIAN KREDIT PADA PT . BPR ARTAMANUNGGAL ABADI MRANGGEN.” *Dinamika Informatika Fakultas Teknologi Informasi Universitas Stikubank Semarang* 1 (1).
- Wahyu Oktaputra, Alif, and Edi Noersasongko. 2014. “Sistem Pendukung Keputusan Kelayakan Pemberian Kredit Motor Menggunakan Metode Simple Additive Weighting Pada Perusahaan Leasing HD Finance.” *Program Studi Sistem Informasi - S1, Fakultas Ilmu Komputer Universitas Dian Nuswantoro, Semarang*, 1–9.
- Wasana, Emil. 2010. “Sistem Pendukung Keputusan Kelayakan Pengajuan Kredit Motor Menggunakan Metode Scoring System.” *Jurusan Siste Informasi, Sekolah Tinggi Manajemen Informatika & Teknik Komputer Surabaya*, 1–10.
- Wu, Xindong, Vipin Kumar, J. Ross Quinlan, Joydeep Ghosh, Qiang Yang, Hiroshi Motoda, Geoffrey J. McLachlan, et al. 2007. *Top 10 Algorithms in Data Mining. Knowledge and Information Systems*. Vol. 14. doi:10.1007/s10115-007-0114-2.
- Zahid, Ahmad. 2013. “Sistem Pendukung Keputusan Persetujuan Penerimaan Pinjaman Di PD.BPR BKK Lasem Dengan Menggunakan Metode Bayes.” *Sekolah Tinggi Manajemen Informatika Dan Komputer AMIKOM Yogyakarta*.