

PENENTUAN RESIKO KREDIT PADA KOPERASI BAROKAH JAYA MENGGUNAKAN ALGORITMA K-NEAREST NEIGHBOR CLASSIFICATION

Bagus Khikmawan¹⁾, Agung Nilogiri

Program Studi Teknik Informatika, Fakultas Teknik, Universitas Muhammadiyah Jember

Jl. Semeru IV No. E4 Jember Kode Pos 68121

Email: bkhikmawan@gmail.com¹⁾,

ABSTRAK

Koperasi adalah salah satu tempat untuk memberikan pinjaman atau kredit kepada nasabahnya. Banyak nasabah yang mengajukan pinjaman atau kredit dengan kondisi ekonomi yang berbeda, sehingga dapat menimbulkan berbagai macam masalah pada koperasi. Salah satunya yaitu terjadinya kredit yang macet dalam pembayaran angsuran pinjaman. Penelitian ini bertujuan untuk mengetahui tingkat akurasi dan presisi hasil klasifikasi lancar atau tidak lancarnya nasabah menggunakan metode K-nearest neighbor berdasarkan atribut kisaran usia, kisaran angsuran, lama tempo, kisaran pendapatan, bidang pekerjaan, status pernikahan, status tempat tinggal, dimana kemiripan dilihat dari riwayat nasabah sebelumnya. Setelah dilakukan pengujian data sebanyak 105 kali dengan K 3,5,7,9,11 dan k-fold 2,4,5,10 menggunakan Teknik validasi cross validation diperoleh nilai akurasi dan presisi yang berbeda-beda pada setiap pengujian. Terdapat dua akurasi dan presisi yang sama namun K yang berbeda yaitu dengan nilai K (ketetangaan knn) 3 dan 9 pada nilai Kf (K-fold) 10 pengujian 9 nilai akurasi 90% dan presisi 100%, Maka dapat disimpulkan bahwa metode klasifikasi k-nearest neighbor dapat digunakan dalam rekomendasi nasabah selanjutnya.

Kata kunci : Koperasi, *k-nearest neighbor*, *Confusion Matrix*, *Cross Validation*

DETERMINATION OF CREDIT RISK IN BAROKAH JAYA COOPERATION USING K-NEAREST NEIGHBOR CLASSIFICATION ALGORITHM

Bagus Khikmawan¹⁾, Agung Nilogiri

Informatics Engineering Study Program, Faculty of Engineering, Muhammadiyah University, Jember

Jl. Semeru 4 No. E4 Jember Postal Code 68121

Email: bkhikmawan@gmail.com¹⁾,

ABSTRACT

Cooperative is one of place to provide loans or credit to their customers. Many customers who apply for loan or credit with the different economic conditions, so it can cause many kinds of problems in cooperative, which is the occurrence of bad loans in the payment of loan installment. This study aims to determine the level of accuracy and precision of the classification results fluent or not using K-nearest neighbor method based on the age range, installment range, timelines, income range, occupation, marital status, residence status, where the similarity is seen from the previous customer history. After testing the data 105 times with K 3,5,7,9,11 and k-fold 2,4,5,10 using cross validation techniques obtained accuracy and precision values are different in each test. The value in both of accuracy and precision are same, but the value of K is different, which is K values (knn neighboring) 3 and 9 on 10 Kf values (K-fold) and the values in 9 testing is 90% accuracy and 100% precision. Then it can be concluded that the k-nearest neighbor classification method can be used in further customer recommendations.

Keywords : *Cooperative*, *k-nearest neighbor*, *Confusion Matrix*, *Cross Validation*.

PENDAHULUAN

Koperasi adalah salah satu tempat untuk memberikan pinjaman atau kredit kepada nasabahnya. Banyak nasabah yang mengajukan pinjaman atau kredit dengan kondisi ekonomi yang berbeda, sehingga dapat menimbulkan berbagai macam masalah pada koperasi. Salah satunya yaitu terjadinya kredit yang macet dalam pembayaran angsuran pinjaman. Hal tersebut terjadi dikarenakan pihak koperasi kurang memperhatikan beberapa faktor yang menyebabkan kredit macet, antara lain kisanan angsuran, lama tempo, pendapatan, bidang pekerjaan.

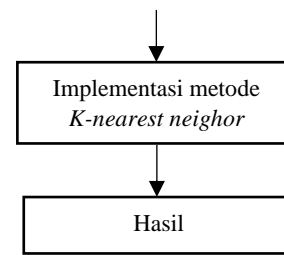
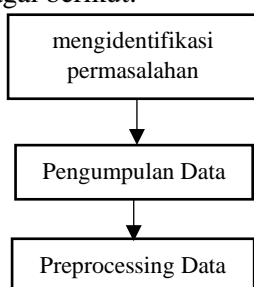
Pada koperasi barokah jaya, analisis klasifikasi nasabah yang berpotensi kredit macet masih menggunakan teknik sederhana, seperti masih menggunakan keputusan personal berdasarkan jaminan saja. Dikarenakan keterbatasan koperasi dalam melibatkan tenaga analis kredit yang membutuhkan biaya cukup mahal dan seringkali memberatkan nasabah, kesederhanaan analisis inilah yang memicu terjadinya kredit macet. Antisipasi yang biasa dilakukan hanyalah dengan melakukan pendekatan-pendekatan personal pada semua nasabah kredit.

Untuk meminimalisir masalah kredit macet, peneliti memilih menggunakan metode K-Nearest Neighbor. Dengan diterapkannya metode K-Nearest Neighbor maka diharapkan Koperasi dapat membuat suatu klasifikasi potensi kredit macet calon nasabah berdasarkan riwayat data transaksi pembayaran kredit sebelumnya serta dapat memanfaatkan data karakteristik nasabah untuk dijadikan sebagai informasi tambahan yang dapat membantu Koperasi dalam membuat analisis kelancaran pinjaman.

METODE PENELITIAN

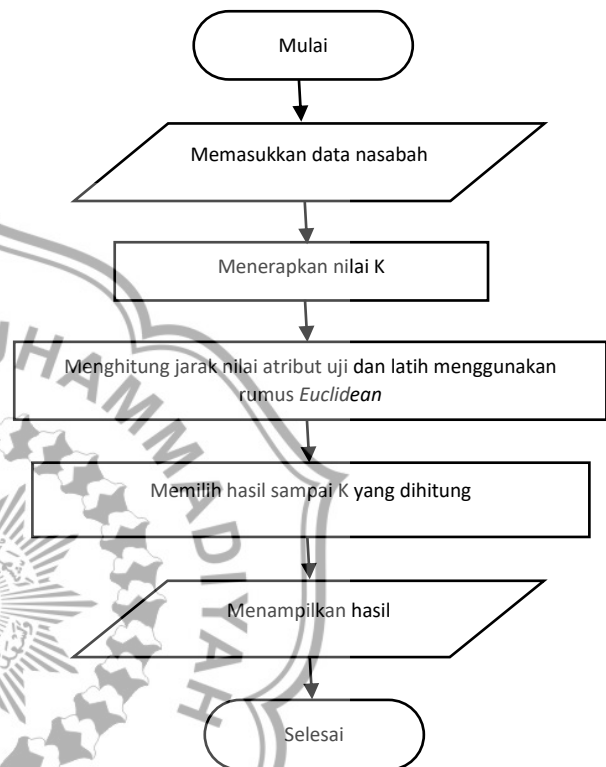
a. Kerangka Penelitian

Penjelasan mengenai tahap-tahapan penelitian tentang penerapan metode K-Nearest Neighbor dalam klasifikasi nasabah koperasi Barokah Jaya untuk menentukan persetujuan peminjaman dana. Metodologi ini melalui beberapa tahapan yaitu analisa masalah, mempelajari literature, mengumpulkan data, preprocessing data, implementasi, hasil, kesimpulan dan saran. Berikut ini adalah diagram metodologi penelitian yang berisi tahapan penelitian yang akan dilakukan sebagai berikut:



Gambar 3.1 Tahapan Penelitian

b. Implementasi Metode K-Nearest Neighbor



Gambar 2.1 Flowchart Algoritma K-nearest neighbor Pada Gambar 2.1 adalah alur implementasi metode knn (Anshori L, 2018).

Proses implementasi pada data menggunakan metode *K-Nearest Neighbor* awalnya adalah. Menentukan parameter k (jumlah tetangga paling dekat). Kemudian menghitung kuadrat jarak *euclidean* objek terhadap data latih yang diberikan., mengurutkan hasil b secara ascending, kategori yang digunakan yaitu L (lancar) dan TL (tidak lancar) (Klasifikasi K-nearest neighbor berdasarkan nilai k). Terakhir, dengan menggunakan kategori K-nearest neighbor yang paling mayoritas maka dapat diprediksikan kategori objek. Untuk menghitung kuadrat jarak euclidean objek terhadap data latih yang diberikan, kita dapat menggunakan rumus:

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

Keterangan:

X_1 : Sampel data atau data latih

X_2 : Data uji atau data testing
 i : Variabel data
 d : Jarak
 p : Dimensi data

Berikut contoh pengerjaannya:

$$d_1 = \sqrt{\frac{((4-4)^2 + (1-1)^2 + (2-2)^2 + (1-2)^2 + (4-1)^2 + (1-1)^2 + (2-1)^2)}{7}} = 2.6457$$

$$d_2 = \sqrt{\frac{((3-4)^2 + (3-1)^2 + (2-2)^2 + (3-2)^2 + (2-1)^2 + (1-1)^2 + (1-1)^2)}{7}} = 2.8284$$

c. Implementasi K-Fold Cross Validation

K-Fold Cross Validation adalah teknik validasi dengan membagi data secara acak kedalam k bagian. Dengan data berjumlah 100 record, data tersebut di bagi menjadi 2 bagian yaitu 80% sebagai data training dan 20% sebagai data testing kemudian data tersebut diproses menggunakan metode K-Fold Cross Validation. Dalam penelitian ini nilai knn yang digunakan sebanyak 3, 5, 7, 9 dan 11 sedangkan untuk nilai K-Fold nya digunakan sebanyak 2, 4, 5, dan 10 bagian karena untuk dipartisi menjadi data training dan data testing. Pemilihan model disesuaikan dengan kebutuhan sesuai dengan tujuan penelitian.

d. Confusion Matrix

Confusion matrix merupakan data set yang hanya memiliki dua kelas, kelas yang satu sebagai positif dan kelas yang lain sebagai negative. Confusion matrix berisi informasi perbandingan label hasil klasifikasi dengan label sebenarnya. Terlihat pada gambar di bawah tabel confusion matrix:

Tabel 2.1 Confusion Matrix Dua Kelas

Classification	Classification Class		
		Class = Yes	Class = No
Real Class	Class = Yes	TP	FN
	Class = No	FP	TN

Keterangan:

- TP (*true positive*) : hasil prediksi positif dengan kelas sebenarnya positif
- FN (*false negative*) : hasil prediksi negatif dengan kelas sebenarnya positif
- FP (*false positive*) : hasil prediksi positif dengan kelas sebenarnya negatif

TN (*true negative*) : hasil prediksi negatif dengan kelas sebenarnya negatif

Untuk menghitung akurasi digunakan persamaan di bawah ini:

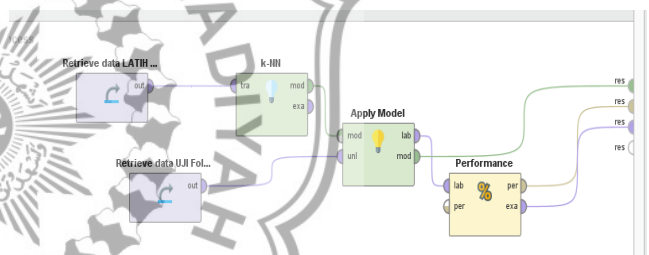
$$\text{Presisi} = \frac{TP}{FP+TP} \times 100\%$$

$$\text{Akurasi} = \frac{TP+TN}{TP+TN+FP+FN} \times 100\%$$

IMPLEMENTASI DAN PENGUJIAN

Implementasi Knn pada Rapidminer

Tahap implementasi tools Rapidminer. RapidMiner merupakan perangkat lunak yang bersifat terbuka (open source). RapidMiner adalah sebuah solusi untuk melakukan analisis terhadap data mining, text mining dan analisis prediksi. RapidMiner menggunakan berbagai teknik deskriptif dan prediksi dalam memberikan wawasan kepada pengguna sehingga dapat membuat keputusan yang paling baik. Maka dilakukan implementasi Rapidminer untuk metode *K-nearest Neighbor*, pada "implementasi metode *K-nearest Neighbor* Dalam Klasifikasi penentuan resiko kredit pada koperasi barokah jaya.



Gambar 4.2 Model KNN pada RapidMiner

Pengujian Data

Pada penelitian ini data yang sudah dikumpulkan akan dilakukan preprocessing data dimana tahap ini memilah atribut yang akan dipakai untuk pengujian, atribut yang dipakai antara kisaran usia, kisaran angsuran, lama tempo, kisaran pendapatan, bidang pekerjaan, status pernikahan, status tempat tinggal. Data yang akan digunakan sebagai pengujian pada tugas akhir ini sebanyak 100 data. Sebelumnya data akan dibagi menjadi 5 kategori pengujian data yaitu pembagian pengujian dengan data latih dan data uji yang berbeda-beda :

1) Hasil Uji K 3 K-Fold 2 4 5 10

Pada hasil pengujian K 3 K-Fold 2 4 5 10 terdapat akurasi dan presisi yang paling tinggi yaitu pada K-fold 10 pengujian 9 skenario 20 nilai akurasi 90% dan nilai presisi 100%.

2) Hasil Uji K 5 K-Fold 2 4 5 10

Pada hasil pengujian K 5 K-Fold 2 4 5 10 terdapat akurasi dan presisi yang paling tinggi yaitu pada K-fold 10 pengujian 9 skenario 42 nilai akurasi 90% dan nilai presisi 85.71%.

3) Hasil Uji K 7 K-Fold 2 4 5 10

Pada hasil pengujian K 7 K-Fold 2 4 5 10 terdapat akurasi dan presisi yang paling tinggi yaitu pada K-fold 10 pengujian 9 skenario 63 nilai akurasi 90% dan nilai presisi 85.71%.

4) Hasil Uji K 9 K-Fold 2 4 5 10

Pada hasil pengujian K 9 K-Fold 2 4 5 10 terdapat akurasi dan presisi yang paling tinggi yaitu pada K-fold 10 pengujian 9 skenario 83 nilai akurasi 90% dan nilai presisi 100%.

5) Hasil Uji K 11 K-Fold 2 4 5 10

Pada hasil pengujian K 11 K-Fold 2 4 5 10 terdapat akurasi dan presisi yang paling tinggi yaitu pada K-fold 10 pengujian 9 skenario 99 nilai akurasi 90% dan nilai presisi 87.50%.

	10	pengujian 4	90.00%	87.50%	99
--	----	-------------	--------	--------	----

KESIMPULAN DAN SARAN

Dari hasil penelitian terhadap Penentuan Resiko Kredit Pada Koperasi Barokah Jaya dapat diambil beberapa kesimpulan dan saran sebagai berikut:

a. Kesimpulan

Komparasi hasil perhitungan excel dan rapidminer menghasilkan nilai akurasi dan presisi yang sama.

Dari perhitungan metode K-Nearest Neighbor menggunakan tool rapidminer dengan nilai K (Ketertanggaan) 3,5,7,9,11 dan skenario uji menggunakan K-Fold Cross Validation dengan nilai 2,4,5,10, untuk Klasifikasi Penentuan Resiko Kredit memperoleh hasil optimal dengan K (ketertanggaan knn) 3 dan 9 pada nilai Kf (K-Fold) 10 pada pengujian 9 dengan nilai akurasi sebesar 90% dan presisi 100%. Sedangkan dari perhitungan metode K-Nearest Neighbor menggunakan Excel dengan K (ketertanggaan knn) 3 dan 9 pada nilai Kf (K-Fold) 10 pada pengujian 9 dengan nilai akurasi sebesar 90% dan presisi 100%.

b. Saran

Beberapa saran untuk pengembangan penelitian dimasa akan datang adalah sebagai berikut:

1. Penelitian ini dapat dikembangkan dengan menambahkan metode optimasi untuk dapat dibandingkan dengan penelitian ini.
2. Pada penelitian ini data yang dipakai kurang seimbang (balance) pada keterangan lancar dan tidak lancar maka sebaiknya menggunakan data yang lebih banyak dan seimbang (balance) agar dapat menghasilkan nilai akurasi lebih baik.

DAFTAR PUSTAKA

- Amin, A. M. (2017). Klasifikasi Kelompok Umur Manusia Berdasarkan Analisis Dimensifraktal Box Counting Dari Citra Wajah Dengan Deteksi Tepi Canny. *MATHunesa*, 2(6).
- Astuti, P. (2016). Komparasi Penerapan Algoritma C45, KNN dan Neural Network Dalam Proses Kelayakan Penerimaan Kredit Kendaraan Bermotor. *Faktor Exacta*, 9(1), 87–101.
- Banjarsari, M. A., Budiman, I., Farmadi, A. (2016). Penerapan K-Optimal Pada Algoritma Knn Untuk Prediksi Kelulusan Tepat Waktu Mahasiswa Program Studi Ilmu Komputer Fmipa Unlam Berdasarkan Ip Sampai Dengan Semester 4. *Klik - K Jurnal Ilmu Komputer*, 2(2), 159–173.
<https://doi.org/10.20527/KLIK.V2I2.26>

Tabel 4.14 Hasil K Terbaik

K	K_f	Pengujian	Akurasi	Presisi	Skenario
3	2	pengujian 2	78.00%	84.83%	2
	4	pengujian 4	80.00%	87.50%	6
	5	pengujian 1	80.00%	87.50%	7
	10	pengujian 9	90.00%	100.00%	20
5	2	pengujian 2	82.00%	83.33%	23
	4	pengujian 4	84.00%	88.24%	27
	5	pengujian 5	85.00%	92.86%	32
	10	pengujian 10	90.00%	85.71%	42
7	2	pengujian 2	84.00%	85.71%	44
	4	pengujian 4	88.88%	88.89%	48
	5	pengujian 1	85.00%	81.82%	49
	10	pengujian 10	90.00%	85.71%	63
9	2	pengujian 2	88.00%	88.57%	65
	4	pengujian 4	88.88%	88.89%	69
	5	pengujian 5	80.00%	92.31%	74
	10	pengujian 9	90.00%	100.00%	83
11	2	pengujian 2	88.00%	88.57%	86
	4	pengujian 4	88.88%	88.89%	90
	5	pengujian 2	85.00%	81.25%	92

- Elholiqi, A., Noranita, B., Waspada, I. (2012). PENENTUAN BESAR PINJAMAN DI KOPERASI SIMPAN PINJAM DENGAN ALGORITMA K-NEAREST NEIGHBOR (Studi Kasus di Koperasi Simpan Pinjam BMT Bina Insani Pringapus). *Jurnal Masyarakat Informatika*, 3(6), 15–20. <https://doi.org/10.14710/jmasif.3.6.15-20>
- Harlina, S. (2018). *Data Mining Pada Penentuan Kelayakan Kredit Menggunakan Algoritma K-Nn Berbasis Forward Selection Data Mining on Credit Feasibility Determination Using K-Nn Algorithm Based on Forward Selection*. 11(2), 236–244.
- Indrayanti, S. D., Al Karomi, M. A. (2017). Optimasi Parameter K Pada Algoritma K-Nearest Neighbour Untuk Klasifikasi Penyakit Diabetes Mellitus. *Prosiding SNATIF Ke-4 2017*, 823–829. <https://doi.org/10.1007/s10115-007-0114-2>
- Leidiyana, H. (2013). Penerapan Algoritma K-Nearest Neighbor Untuk Penentuan Resiko Kredit Kepemilikan Kendaraan Bermotor. *Jurnal Penelitian Ilmu Komputer, System Embedded & Logic*, 1(1), 65–76.
- Menarianti, I. (2015). Klasifikasi data mining dalam menentukan pemberian kredit bagi nasabah koperasi. *Jurnal Ilmiah Teknosains*, 1(1), 1–10. <http://e-jurnal.upgrismg.ac.id/index.php/JITEK/article/view/836>
- Sholeh, M. (2014). Prosiding Seminar Nasional Aplikasi Sains & Teknologi (SNAST) 2014 Yogyakarta, 15 November 2014 ISSN: 1979-911X. *Snast, November*, 211–216.
- Mustakim, & Oktaviani F, G. (2016). *Algoritma K-Nearest Neighbor Classification Sebagai Sistem Prediksi Predikat Prestasi Mahasiswa*. 13(2), 195–202.
- Nugroho, A. (2016). Implementasi Algoritma K-Nearest Neighbor Dalam Memprediksi Potensi Calon Kreditur Di Ksp Galih Manunggal. *Data Manajemen Dan Teknologi Informasi*, 17(2), 1–6.
- Sumarlin, S. (2015). Implementasi Algoritma K-Nearest Neighbor Sebagai Pendukung Keputusan Klasifikasi Penerima Beasiswa PPA dan BBM. *Jurnal Sistem Informasi Bisnis*, 5(1), 52–62. <https://doi.org/10.21456/vol5iss1pp52->
- Supriana, I. W., Astuti, L. G. (2019). Implementasi K-Nearest Neighbor Pada Penentuan Keluarga Miskin Bagi Dinas Sosial Kabupaten Tabanan. *Jurnal Teknologi Informasi Dan Komputer*, 5(1), 120–129. <https://doi.org/10.36002/jutik.v5i1.645>
- Wahyuningsih, S. (2018). Perbandingan Metode K-Nearest Neighbor, Naïve Bayes dan Decision Tree untuk Prediksi Kelayakan Pemberian Kredit. *Konferensi Nasional Sistem Informasi 2018*, 8–9. <http://jurnal.atmaluhur.ac.id/index.php/knsi2018/article/view/424/349>