

PREDIKSI KELANCARAN PEMBAYARAN CICILAN CALON DEBITUR DENGAN METODE K-NEAREST NEIGHBOR

Sri Ayu Rizky¹, Rolly Yesputra^{2*}, Santoso³

¹Sistem Informasi, STMIK Royal

²Sistem Komputer, STMIK Royal

³Manajemen Informatika, STMIK Royal

email: *rollyesputra1@gmail.com

Abstract: In this research, a prediction system has been successfully developed to predict whether or not a prospective money borrower will run smoothly. Prospective borrowers who will borrow, some of the data that meet the criteria will be inputted by the office clerk into a prediction application system interface to be processed using the Data Mining method, namely the K-Nearest Neighbor Algorithm with the Codeigniter programming language 3. The results of the Euclidean calculation process are based on predetermined criteria Between training data (training) to testing data (test) will be displayed with a table that has been sorted from smallest to largest containing 9 closest neighbors according to the K value that has been determined, namely 9. The nine neighbors will be taken the dominant category. This dominant category can be used as a guideline that makes it easier for the leader to make a decision on the next borrower.

Keywords: Data Mining; Euclidean; K-Nearest Neighbor; Prospective Borrowers;

Abstrak: Dalam penelitian ini telah berhasil dibuat sebuah sistem prediksi untuk memprediksi lancar atau tidak lancarnya seorang calon peminjam uang. Calon peminjam uang yang akan meminjam, sebagian datanya yang memenuhi kriteria akan diinputkan petugas kantor ke dalam sebuah *interface* sistem aplikasi prediksi untuk diolah menggunakan metode *Data Mining* yaitu Algoritma K-Nearest Neighbor dengan bahasa pemrograman Codeigniter 3. Hasil proses perhitungan Euclidean berdasarkan kriteria yang sudah ditentukan antara data training (latih) ke data testing (uji) tersebut akan ditampilkan dengan sebuah tabel yang sudah diurutkan dari yang terkecil ke terbesar berisi 9 tetangga terdekat sesuai dengan nilai K yang sudah ditentukan yaitu 9. Sembilan tetangga tersebut akan diambil kategori yang dominan. Kategori yang dominan tersebut bisa dijadikan suatu pedoman yang memudahkan pimpinan dalam mengambil sebuah keputusan kepada calon peminjam selanjutnya.

Kata kunci: Debitur; Data Mining; Euclidean; K-Nearest Neighbor

PENDAHULUAN

Pandemi Covid-19 menyebabkan banyak orang-orang yang kehilangan

pekerjaan, sehingga perlu mencari solusi untuk tetap bisa menghasilkan demi memenuhi kebutuhan sehari-hari. Sulitnya dalam mendapatkan pekerjaan

menyebabkan orang ingin berwiraswasta, akan tetapi banyak yang terkendala dengan keterbatasan modal. Modal merupakan sesuatu yang sangat utama dan penting dalam bisnis atau perusahaan. Tanpa itu bisnis tidak dapat berlangsung selayaknya. Banyak yang menyediakan solusi peminjaman dana modal usaha yaitu bank, pegadaian, *leasing*, situs penyedia modal dan koperasi simpan pinjam. Salah satu diantaranya yang sering digunakan masyarakat adalah koperasi simpan pinjam. Koperasi simpan pinjam Menurut UU No. 17 tahun 2012 Pasal 1 adalah koperasi yang menjalankan usaha simpan pinjam sebagai satu-satunya usaha.

Banyak faktor yang harus dipertimbangkan, untuk meminimalisir kerugian. Usaha pengkreditan memiliki untung yang besar tetapi juga memiliki resiko yang besar pula. Untuk menghindari resiko dapat dilakukan dengan melakukan pengolahan data dengan menggunakan data mining metode K-NN dengan menghitung kedekatan antara data lama dan data baru[1].

Data mining merupakan proses ekstraksi data untuk mendapat pengetahuan. Metode K-NN merupakan proses pengelompokan dari hasil ekstraksi data dari pembelajaran data yang sudah di kelompokkan sebelumnya[2]. K-NN juga disebut sebagai metode klasifikasi untuk menentukan nilai K atau tetangga yang terdekat ketika menentukan kelompok pada data yang baru[3][4].

Metode *K-Nearest Neighbor* yang mempunyai kemampuan untuk melakukan diagnosa dengan cara memperhatikan data awal dalam kasus sebelumnya dan membandingkan dengan data baru, proses ini bertujuan agar menghasilkan diagnosa yang lebih baik

serta akurat[5]. Kelebihan metode K-NN yaitu memiliki kekuatan dalam mengolah data *training* yang mempunyai banyak *noise* serta sangat efektif jika data training yang digunakan memiliki jumlah yang besar[6]. Semakin banyak data training maka hasil prediksi akan menjadi lebih baik.

Sistem prediksi dengan menggunakan metode K-NN ini diimplementasi dengan menggunakan Framework PHP. Framework yang digunakan adalah CodeIgniter. CodeIgniter yaitu sebuah framework yang berbasis PHP dikembangkan dengan menerapkan paradigma pemrograman berorientasi objek dengan arsitektur model-view-controller (MVC)[7]. MVC akan memisahkan pengembangan komponen utama sistem antara pengolahan data (model) dengan antar muka (view) serta dengan bagian kontrol web (controller)[8].

Sehubungan dengan perkembangan teknologi komputer dibidang sistem informasi masalah tersebut bisa diminimalisir menggunakan suatu aplikasi untuk memprediksi peminjam uang lancar atau tidak lancar selanjutnya pada KSP Riama.

METODE

K-Nearest Neighbor (K-NN) merupakan sebuah metode untuk mengklasifikasikan objek berdasarkan data terdahulu dengan jarak terdekat dari objek tersebut. Secara umum untuk mendefinisikan jarak antara dua objek x dan y , digunakan rumus jarak yang disebut *Euclidean* berikut rumus (1)[9]:

$$d_{xy} = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (1)$$

Keterangan :

d_{xy} : Jarak *Euclidean*

x_i : *record* ke- i

x_j : record ke- j

i, j : 1,2,3,...n

Untuk menggunakan metode ini perlu menentukan banyaknya nilai K tetangga terdekat dalam melakukan klasifikasi data baru[10]. Selanjutnya perlu mengikuti secara bertahap dalam penerapan algoritma K-NN, sebagai berikut: (1) Menentukan parameter K. (2) Menghitung jarak antara data yang akan diuji dengan semua data latih. (3) Mengurutkan jarak tersebut dari terkecil ke terbesar. (4) Menentukan jarak yang

paling dekat sesuai nilai K. (5) Cari dan simpulkan hasil yang dominan kedekatannya dengan K.

HASIL DAN PEMBAHASAN

Total data yang menjadi sampel ialah sebanyak 164 data. Data tersebut terdiri dari 7 atribut, di mana 6 atribut merupakan prediktor dan 1 atribut merupakan label, data yang digunakan proses prediksi dengan menggunakan metode K-NN dapat dilihat tabel 1-7.

Tabel 1. Atribut dan Transformasi Data Prediksi Calon Peminjam Uang

Atribut	Keterangan			
	Tipe	Sub Atribut	Bobot	Variabel
Nomor Pinjaman	Polinomial	I	1	X1
		II	2	
		III	3	
		IV	4	
		V	5	
		VI	6	
		VII	7	
		VIII	8	
		IX	9	
		X	10	
		XI	11	
Jenis Kelamin	Binomial	Laki-laki	1	X2
		Perempuan	2	
Status Pernikahan	Polinomial	Janda	1	X3
		Duda	2	
		Menikah	3	
Jenis Usaha	Polinomial	Ibu Rumah Tangga	1	X4
		Nelayan	2	
		Wiraswasta	3	
		Petani	4	
		Pedagang	5	
Jumlah Pinjaman	Polinomial	200.000 – 499.000	1	X5
		500.000 – 999.999	2	
		1.000.000 – 1.999.999	3	
		2.000.000 – 2.999.999	4	
		3.000.000 – 3.999.999	5	
		4.000.000 – 4.999.999	6	
		5.000.000 – 5.999.999	7	
		6.000.000 – 6.999.999	8	
		7.000.000 – 7.999.999	9	
		8.000.000 – 8.999.999	10	
		9.000.000 – 10.000.000	11	
Jenis Pinjaman	Binomial	Harian	1	X6
		Mingguan	2	
Kategori	Label	Lancar	L	Y
		Tidak Lancar	TL	

Tabel 2. Sampel Data *Training*

Nama	No Pinj.	Jenis Kel.	Status Pernikahan	Jenis Usaha	Jlh Pinj.	Jen. Pinj.	Kat
Debitur 1	IV	P	Menikah	Wiraswasta	600000	Mingguan	L
Debitur 2	V	L	Menikah	Wiraswasta	400000	Mingguan	L
Debitur 3	I	P	Menikah	Wiraswasta	500000	Mingguan	L
Debitur 4	I	P	Menikah	Wiraswasta	500000	Mingguan	L
Debitur 5	V	P	Menikah	Wiraswasta	500000	Mingguan	L
Debitur 6	I	P	Menikah	Wiraswasta	800000	Mingguan	L
Debitur 7	VIII	P	Menikah	Pedagang	300000	Harian	TL
Debitur 8	I	L	Menikah	Wiraswasta	500000	Mingguan	L
Debitur 9	III	P	Menikah	Wiraswasta	600000	Mingguan	TL
Debitur 10	I	P	Menikah	Pedagang	500000	Mingguan	TL
Debitur 11	I	P	Menikah	Wiraswasta	500000	Harian	TL
Debitur 12	I	P	Menikah	Wiraswasta	500000	Mingguan	L
Debitur 13	I	P	Menikah	Wiraswasta	500000	Mingguan	L
Debitur 14	I	L	Menikah	Pedagang	500000	Mingguan	L
Debitur 15	IX	P	Menikah	Wiraswasta	500000	Mingguan	L

Tabel 3. Sampel Data *Testing*

Nama	No Pinj.	Jenis Kel.	Status Pernikahan	Jenis Usaha	Jlh Pinj.	Jen. Pinj.	Kat
Debitur 16	I	L	Menikah	Kp. Subur	Wiraswasta	1000000	?
Debitur 17	I	P	Menikah	P. Tanjung	Wiraswasta	200000	?

Tabel 4. Transformasi Data *Training*

Nama	No Pinj.	Jenis Kel.	Status Pernikahan	Jenis Usaha	Jlh Pinj.	Jen. Pinj.	Kat
Debitur 1	4	2	3	3	2	2	L
Debitur 2	5	1	3	3	1	2	L
Debitur 3	1	2	3	3	2	2	L
Debitur 4	1	2	3	3	2	2	L
Debitur 5	5	2	3	3	2	2	L
Debitur 6	1	2	3	3	2	2	L
Debitur 7	8	2	3	5	1	1	TL
Debitur 8	1	1	3	3	2	2	L
Debitur 9	3	2	3	3	2	2	TL
Debitur 10	1	2	3	5	2	2	TL
Debitur 11	1	2	3	3	2	1	TL
Debitur 12	1	2	3	3	2	2	L
Debitur 13	1	2	3	3	2	2	L
Debitur 14	1	1	3	5	2	2	L
Debitur 15	9	2	3	3	2	2	L

Tabel 5. Transformasi Data *Testing*

Nama	No Pinj.	Jenis Kel.	Status Pernikahan	Jenis Usaha	Jlh Pinj.	Jen. Pinj.	Kat
Debitur 16	1	1	3	3	3	2	?
Debitur 17	1	2	3	3	1	1	?

Tabel 6. Hasil Perhitungan Jarak

Nama	Jarak	Kategori
Debitur 1	3,316624	L
Debitur 2	4,472135	L
Debitur 3	1,414213	L
Debitur 4	1,414213	L
Debitur 5	4,242640	L
Debitur 6	1,414213	L
Debitur 7	7,681145	TL
Debitur 8	1	L
Debitur 9	2,449489	TL
Debitur 10	2,449489	TL
Debitur 11	1,732050	TL
Debitur 12	1,414213	L
Debitur 13	1,414213	L
Debitur 14	2,236067	L
Debitur 15	8,124038	L

Tabel 7. Perhitungan Jarak Terurut

Nama	Jarak	K = 5	Kategori i
Debitur 8	1	Ya	L
Debitur 3	1,414213	Ya	L
Debitur 4	1,414213	Ya	L
Debitur 6	1,414213	Ya	L
Debitur 12	1,414213	Ya	L
Debitur 13	1,414213	Ya	L
Debitur 11	1,732050	Ya	TL
Debitur 14	2,236067	Ya	L
Debitur 9	2,449489	Ya	TL
Debitur 10	2,449489	Tidak	TL
Debitur 1	3,316624	Tidak	L
Debitur 5	4,242640	Tidak	L
Debitur 2	4,472135	Tidak	L
Debitur 7	7,681145	Tidak	TL
Debitur 15	8,124038	Tidak	L

Rumus *Euclidean* haruslah angka maka dari itu diperlukan perubahan data dari alfabetis menjadi numeric sesuai dengan bobot atau nilai yang sudah

ditentukan pada Tabel 1. Pada Tabel 4 adalah hasil transformasi data latih atau *training* dari Tabel 2. Untuk data *testing* juga dilakukan transformasi dengan nilai yang sama seperti pada Tabel 5.

Langkah pertama ialah menentukan parameter *K*. Nilai *K* yang digunakan pada penelitian ini adalah 9. Langkah kedua ialah menghitung jarak menggunakan rumus *Euclidean*. Objek atau keseluruhan data *training* dihitung jaraknya masing-masing dengan data *testing* pertama:

$$\begin{aligned}
 d1 &= \sqrt{(4-1)^2 + (2-1)^2 + (3-3)^2 + (3-3)^2} \\
 &= \sqrt{(2-3)^2 + (2-2)^2} \\
 &= 3,316624
 \end{aligned}$$

Selanjutnya langkah ketiga yaitu mengurutkan jarak seperti pada Tabel 7 sesuai nilai parameter *K*. Kemudian langkah keempat pada kolom ketiga Tabel 7 untuk memperjelas dari lima belas sampel data *training* tersebut yang termasuk ke sembilan tetangga terdekat. Sedangkan data dengan melebihi nilai *K* tidak termasuk sebagai tetangga terdekat. Untuk sampel data *testing* kedua dilakukan langkah yang sama dengan data *testing* pertama.

Langkah terakhir menyimpulkan hasil mayoritas kategori yang di dapat dari perhitungan *K-Nearest Neighbor*, karena nilai *K=9* berarti data yang menunjukkan Lancar(L)=7 dan data yang menunjukkan Tidak Lancar(TL)=2. Merujuk hal diatas maka disimpulkan bahwa kategori data *testing* pertama

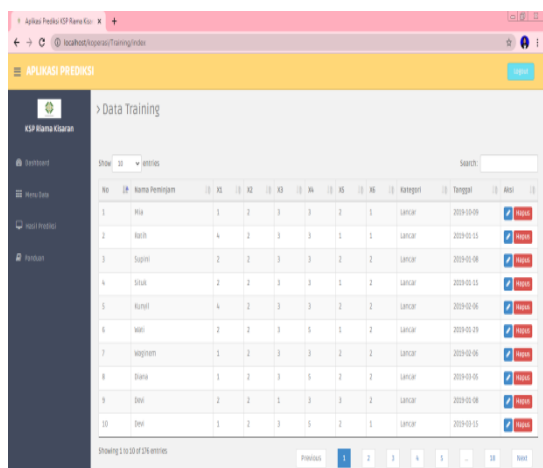
dengan nama calon peminjam Bambang di prediksi Lancar(L).

Hasil prediksi terhadap data testing yang ada pada tabel 3 adalah:

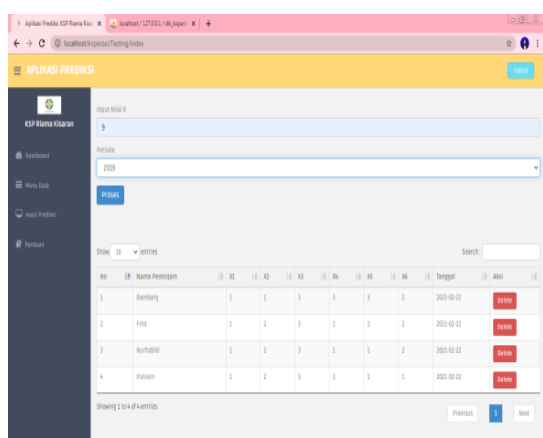
Tabel 8. Hasil Prediksi

No	Nama Nasabah	Kategori
1	Debitur 16	Lancar
2	Debitur 17	Tidak Lancar

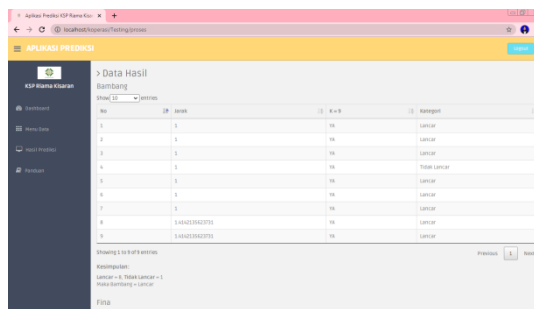
Aplikasi prediksi kelancaran calon peminjam uang di KSP Riama Kisan dimodelkan dengan menggunakan pemodelan *Unified Modeling Language* (UML). Setelah dibuatkan model maka dilakukan proses konstruksi(pengkodean) dan hasil implementasi sistemnya.



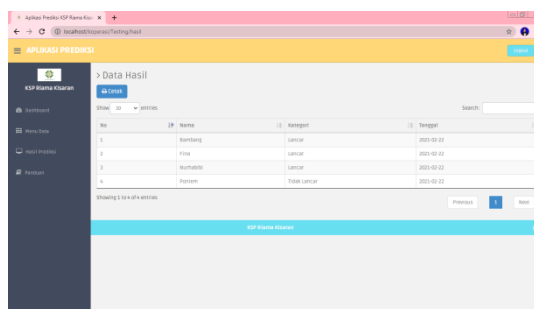
Gambar 1. Data Training



Gambar 2. Data Testing



Gambar 3. Proses Perhitungan



Gambar 4. Hasil Prediksi

Setelah dilakukan proses prediksi secara manual dan aplikasi yang dibuat, berikutnya adalah mengukur kinerja dengan menghitung accuracy yang tersaji pada Gambar 5.

accuracy: 73.37%					
	true Kategori	true L	true TL	true ?	class precision
pred. Kategori	0	0	0	0	0.00%
pred. L	1	106	34	3	73.61%
pred. TL	0	6	18	1	72.00%
pred. ?	0	0	0	0	0.00%
class recall	0.00%	94.64%	34.62%	0.00%	

Gambar 5. Accuracy

Pengujian akurasi data menggunakan metode K-NN pada penelitian ini dengan aplikasi Rapid Miner. Pada pengujian yang telah dilakukan didapatkan hasil accuracy memperlihatkan bahwa kinerja prediksi metode K-NN dengan data yang ada, memiliki akurasi sebesar 73,37%.

Pengujian User Interface dari aplikasi Prediksi ini menggunakan metode *Focus Group Discussion* dengan pihak KSP Riama.

Tabel 9. FGD Penggunaan Sistem

No Komponen	Tanggapan
1 Dashboard	Setuju
2 Tata Letak Menu	Setuju
3 Hasil Prediksi	Setuju dengan catatan
4 Format Form Input	Setuju
5 Format Laporan	Setuju

SIMPULAN

Berdasarkan hasil penelitian metode K-NN dapat digunakan untuk melakukan prediksi terhadap kelancaran calon peminjam uang. Dengan menggunakan 164 data training dan 4 data testing. Hasil perhitungan manual, *Rapid Miner* dan aplikasi prediksi yang dibuat menunjukkan hasil yang sama. Nilai akurasi yang terbentuk ketika diuji menggunakan *Rapid Miner* ialah 73,37% dengan nilai $K=9$. Maka dari hasil penelitian ini banyaknya data training maupun data testing akan sangat mempengaruhi kualitas hasil dan akurasi prediksi. Hal ini juga harus sebanding dengan klasifikasi serta kriteria yang digunakan.

Dari hasil yang ada, penerapan algoritma k-nearest neighbor pada sebuah sistem berbasis web dapat digunakan dengan baik untuk memprediksi calon peminjam uang apakah akan lancar atau tidak lancar berdasarkan kedekatan dengan data training.

DAFTAR PUSTAKA

- [1] E. P. W. Mandala, "Data Mining Algoritma Nearest Neighbor Untuk Memprediksi Tingkat Resiko Pinjaman Dana Di Bank Perkreditan Rakyat," *JIK J. Ilmu Komputer.*, vol. 1, no. 2, hal. 91–103, 2016.
- [2] I. W. Supriana, M. A. Raharja, dan P. W. Gunawan, "Sistem Informasi Prediksi Penilaian Kredit Perbankan Menggunakan Algoritma K-Nearest Neighbor Classification," *JST (Jurnal Sains dan Teknol.*, vol. 8, no. 1, hal. 44, 2019, doi: 10.23887/jst-undiksha.v8i1.16470.
- [3] M. Reza Noviansyah, T. Rismawan, dan D. Marisa Midyanti, "Penerapan Data Mining Menggunakan Metode K-Nearest Neighbor Untuk Klasifikasi Indeks Cuaca Kebakaran Berdasarkan Data Aws (Automatic Weather Station) (Studi Kasus: Kabupaten Kubu Raya)," *J. Coding, Sist. Komput. Untan*, vol. 06, no. 2, hal. 48–56, 2018.
- [4] H. Risman, D. Nugroho, dan Y. Retno, "Penerapan Metode K-Nearest Neighbor Pada Aplikasi Penentu Penerima Beasiswa Mahasiswa Di Stmik Sinar Nusantara Surakarta," *TIKOMSiN*, hal. 19–25, 2012
- [5] P. S. Ramadhan, "Penerapan K-Nearest Neighbor dalam Pendeteksian Abcessus," *InfoTekJar (Jurnal Nas. Inform. dan Teknol. Jaringan)*, vol. 3, no. 2, hal. 61–70, 2019, doi: 10.30743/infotekjar.v3i2.1003.
- [6] I. A. D. P. Sari, B. Hidayat, dan U. Sunarya, "Pengenalan Aksara Bali Dengan Metode Local Binary Pattern," *e-Proceeding Eng.*, vol. 2, no. 2, hal. 2697–2704, 2015.
- [7] M. Destiningrum dan Q. J. Adrian, "Sistem Informasi Penjadwalan Dokter Berbasis Web Dengan Menggunakan Framework Codeigniter (Studi Kasus: Rumah

- Sakit Yukum Medical Centre),” *J. Teknoinfo*, vol. 11, no. 2, hal. 30, 2017.
- [8] Yesputra, Rolly, Marpaung Nasrun, “Penerapan Arsitektur Model View Contoller (Mvc) Pada Sistem Informasi E-Skripsi Stmik Royal,” *j. Inform. Sains dan Teknol.*, vol. 3, no. 2, hal. 281–290, 2018.
- [9] Y. I. Kurniawan dan T. I. Barokah, “Klasifikasi Penentuan Pengajuan Kartu Kredit Menggunakan K-Nearest Neighbor,” *J. Ilm. Matrik*, vol. 22, no. 1, hal. 73–82, 2020, doi: 10.33557/jurnalmatrik.v22i1.843.
- [10] E. Laksono, A. Basuki, F. Bachtiar, “Optimasi Nilai K pada Algoritma KNN untuk Klasifikasi Spam dan Ham Email,” vol. 1, no. 10, hal. 377–383, 2021.