

IMPLEMENTASI K-NEAREST NEIGHBOR (K-NN) UNTUK PREDIKSI PENERIMA PINJAMAN (Study Kasus: KSP. Sukri Jaya Mandiri)

Sidik Rahmatullah^{1*}, Irwandi², Naili Khoiriyah⁴

^{1,2,3}. Institut Teknologi Bisnis dan Bahasa Dian Cipta Cendikia, ⁴Penulis Koresponden
e-mail : ¹sidik@dcc.ac.id, ²irwandits@yahoo.co.id, ³naili@gmail.com.,

ABSTRACT

KSP Sukri Jaya Mandiri Central Lampung is one of the savings and loan cooperative financial institutions that has succeeded in attracting many customers to become members each year, as evidenced by the increasing number of new members each year. The problem that often occurs is the pattern of behavior or customer characteristics that are not good. At Ksp. Sukri Jaya Mandiri Central Lampung classifying whether or not customers who get loans are appropriate is not accurate or not right. For classification, it still uses untested manual methods and uses untested Microsoft Excel into programming systems such as Rapidminer. In this research, the method of collecting data is literature study, and interviews. The customer data used is data from the last 3 years with a total of 200 customers. Several attributes are used including Name, Age, Occupation, Salary, Type of Residence, Condition of the House, Marital Status, Number of Dependents, Type of Loan, Collateral, membership status, and Other Loans. Data mining predictions using the K-Nearest Neighbor (K-NN) algorithm have been successfully implemented with the help of RapidMiner software. In this experiment, the k value parameter that produces the best level of accuracy is k=5, with an accuracy level of 82.50%. Based on the results of the analysis, a pattern was obtained that allows the classification of new customers into worthy or unworthy categories with a high level of confidence. This pattern is expected to help make more effective decisions in customer management.

Keywords: Data Mining , K-Nearest Neighbor, Rapidminer

INTISARI

KSP. Sukri Jaya Mandiri Lampung Tengah merupakan salah satu lembaga keuangan koperasi simpan pinjam yang sukses menarik banyak nasabah menjadi anggota tiap tahunnya, terbukti dengan meningkatnya jumlah anggota baru tiap tahun. masalah yang sering terjadi yaitu pola perilaku atau karakteristik nasabah yang kurang baik. Pada Ksp. Sukri Jaya Mandiri Lampung Tengah pengklasifikasian layak atau tidaknya nasabah yang mendapat pinjaman belum akurat atau belum tepat. Untuk pengklasifikasiannya masih menggunakan manual yang belum teruji dengan metode dan menggunakan *Microsoft excel* yang belum teruji ke dalam sistem pemrograman seperti Rapidminer. Pada penelitian ini digunakan metode pengumpulan data studi pustaka, dan wawancara. Data nasabah yang digunakan yaitu data dalam waktu 3 tahun terakhir dengan jumlah nasabah sebanyak 200 data. Digunakan beberapa atribut diantaranya Nama, Umur, Pekerjaan, Gaji, Jenis Tinggal, Kondisi Rumah, Status Pernikahan, Jumlah Tanggungan, Jenis Pinjaman, Jaminan, status keanggotaan, dan Pinjaman Lain. Prediksi data mining menggunakan algoritma K-Nearest Neighbor (K-NN) telah berhasil diimplementasikan dengan bantuan perangkat lunak RapidMiner. Dalam eksperimen ini, parameter nilai k yang menghasilkan tingkat akurasi terbaik adalah k=5, dengan tingkat akurasi mencapai 82,50%. Berdasarkan hasil analisis, diperoleh suatu pola yang memungkinkan klasifikasi nasabah baru ke dalam kategori layak atau tidak layak dengan tingkat kepercayaan yang tinggi. Pola ini diharapkan dapat membantu pengambilan keputusan yang lebih efektif dalam pengelolaan nasabah.

Kata kunci: Data Mining , K-Nearest Neighbor, Rapidminer

1. PENDAHULUAN

Di era modern, layanan pemodalan dan peminjaman dana semakin diminati oleh masyarakat. KSP Sukri Jaya Mandiri Lampung Tengah merupakan salah satu lembaga keuangan koperasi simpan pinjam yang sukses menarik minat nasabah baru setiap tahunnya. Terbukti dengan meningkatnya jumlah anggota baru, koperasi ini secara konsisten menyediakan layanan peminjaman dana yang bertujuan untuk meningkatkan kesejahteraan anggotanya dan masyarakat secara umum (Munandar, 2020).

Namun, seiring perkembangan zaman dan teknologi, diprediksi akan semakin banyak pengajuan pinjaman, dan hal ini menghadirkan tantangan tersendiri dalam hal evaluasi kelayakan nasabah. KSP Sukri Jaya Mandiri menghadapi masalah terkait dengan pengklasifikasian nasabah yang layak mendapatkan pinjaman, karena metode penilaian saat ini masih dilakukan secara manual menggunakan Microsoft Excel, yang belum teruji dalam sistem

pemrograman yang lebih canggih seperti RapidMiner. Akibatnya, hasil pengklasifikasian belum cukup akurat untuk meminimalkan risiko pinjaman macet.

Algoritma *K-Nearest Neighbor* (K-NN) dapat menjadi solusi dalam konteks ini. Algoritma ini bekerja berdasarkan teori bahwa data dengan karakteristik serupa cenderung memiliki hasil yang serupa pula. Dalam prediksi kelayakan penerima pinjaman, K-NN dapat digunakan untuk mengelompokkan pemohon berdasarkan karakteristik dan riwayat kredit mereka, sehingga memudahkan pengambilan keputusan (Ahmed et al., 2023). Dengan memanfaatkan data historis, K-NN dapat membantu menghasilkan prediksi yang akurat mengenai kemampuan calon nasabah dalam memenuhi kewajiban pinjaman mereka.

KSP Sukri Jaya Mandiri membutuhkan prediksi akurat terhadap data calon nasabah penerima pinjaman untuk meminimalkan risiko keterlambatan pembayaran pinjaman. Dalam penelitian ini, data dari nasabah KSP Sukri Jaya Mandiri akan diolah menggunakan *data mining* untuk mengidentifikasi parameter pinjaman yang relevan. Prediksi ini akan membantu dalam memberikan penilaian kelayakan pinjaman, sehingga dapat menjadi alternatif solusi untuk mengurangi risiko pinjaman macet (Hasibuan, 2024). Kriteria yang digunakan mencakup nama, umur, pekerjaan, gaji, jenis tempat tinggal, kondisi rumah, status pernikahan, jumlah tanggungan, jenis pinjaman, jaminan, status keanggotaan, dan pinjaman lain.

Penelitian ini akan menggunakan metode K-NN untuk mengklasifikasikan data berdasarkan kedekatan objek dalam dataset yang relevan (Irawan et al., 2023). Data yang diperoleh kemudian akan diimplementasikan menggunakan perangkat lunak RapidMiner. Oleh karena itu, penelitian ini berjudul "Implementasi K-Nearest Neighbor (K-NN) untuk Prediksi Penerima Pinjaman (Studi Kasus: KSP Sukri Jaya Mandiri Lampung Tengah)". Penelitian ini bertujuan untuk mengeksplorasi dan mengimplementasikan algoritma K-NN dalam prediksi penerima pinjaman. Melalui dataset yang relevan, penelitian ini menganalisis efektivitas K-NN dalam mengidentifikasi pemohon yang berpotensi mengalami gagal bayar. Diharapkan hasil penelitian ini dapat memberikan kontribusi signifikan dalam praktik penyaluran pinjaman serta membantu lembaga keuangan dalam meminimalkan risiko dan meningkatkan efisiensi dalam pengambilan (Windy Mardiyah et al., 2024).

2. METODE PENELITIAN

2.1 Data Mining

Data mining adalah proses untuk mengekstrak atau menggali pola, pengetahuan, atau informasi berharga dari kumpulan data yang besar. Proses ini melibatkan penggunaan teknik statistik, matematika, dan kecerdasan buatan untuk mengidentifikasi pola atau hubungan dalam data yang sebelumnya tidak terlihat, sehingga dapat mendukung pengambilan keputusan (Afandi et al., 2022).

Dalam data mining, metode seperti klasifikasi, clustering, regresi, dan asosiasi sering digunakan untuk tujuan analisis. Kegunaan data mining bervariasi, mulai dari bidang bisnis, keuangan, kesehatan, hingga pemasaran, dan berperan penting dalam menganalisis perilaku pelanggan, prediksi penjualan, pengelolaan risiko, dan banyak aplikasi lainnya (Waspah et al., 2022).

2.2 K-Nearest Neighbor (K-NN)

K-Nearest Neighbor (K-NN) adalah algoritma pembelajaran mesin berbasis *instance-based learning*, yang digunakan dalam masalah klasifikasi dan regresi. Algoritma ini bekerja berdasarkan prinsip bahwa objek-objek yang memiliki kemiripan karakteristik atau fitur akan cenderung berada dalam kelas yang sama atau memiliki hasil yang serupa (Nyoman & Smrti, 2015). Dalam konteks ini, "tetangga" merujuk pada objek data yang berada pada jarak terdekat di ruang fitur. Berikut adalah cara kerja dasar K-NN (Zhang, 2022):

$$\sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + \dots + (a_n - b_n)^2} \dots\dots\dots (1)$$

keterangan :

dimana $a = a_1, a_2, \dots, a_n$, dan $b = b_1, b_2, \dots, b_n$ 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):

$$\text{different}(a_i, b_i) = \begin{cases} 0 & \text{jika } a_i = b_i \\ 1 & \text{jika } a_i \neq b_i \end{cases}$$

Dimana a_i dan b_i 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 1 [6].

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

$$\text{Similarity (p, q)} = \frac{\sum_{i=1}^n f(p_i, q_i) \times w_i}{w_i} \dots\dots\dots (2)$$

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

Tahapan Langkah Algoritma K-NN sebagai berikut:

1. Menentukan parameter k (jumlah tetangga paling dekat).
2. Menghitung kuadrat jarak eucliden objek terhadap data training yang diberikan.
3. Mengurutkan hasil no 2 secara *ascending* (berurutan dari nilai tinggi ke rendah)
4. Mengumpulkan kategori Y (Klasifikasi *nearest neighbor* berdasarkan nilai k)
5. Dengan menggunakan kategori *nearest neighbor* yang paling mayoritas maka dapat dipredisikan kategori objek.

2.3 Rapid Miner

RapidMiner merupakan salah satu solusi data mining yang paling banyak digunakan diseluruh dunia. RapidMiner tidak hanya digunakan oleh peneliti di bidang *computer science* dan matematika saja, tetapi banyak juga digunakan untuk bisnis dan industri. Didalam dunia *computer science* dan matematika, rapid miner digunakan untuk pengembangan metode *data minuing*, *machine learning* dan metode statistik. RapidMiner dapat digunakan sebagai alat, karena menyediakan berbagai macam alat berbagai metode dari evaluasi statistik sederhana seperti analisis korelasi hingga regresi, klasifikasi dan prosedur pengelompokan serta pengurangan dimensi dan optimasi parameter (Junaedi et al., 2022).

2.4. Teknik Pengumpulan Data

Adapun Teknik pengumpulan data yang digunakan oleh penulis dalam penelitian ini adalah:

a. Observasi

Proses pengambilan data dilakukan secara langsung di KSP Sukri Jaya Mandiri, Lampung Tengah, untuk memperoleh data nasabah yang diperlukan dalam penelitian.

b. Studi Kepustakaan (*Library Research*)

Kepustakaan merupakan metode pengumpulan data dimana penulis mengumpulkan data dengan membaca buku-buku, jurnal atau catatan yang ada hubungannya judul penelitian.

c. Wawancara (*interview*)

Wawancara langsung dilakukan dengan pegawai KSP Sukri Jaya Mandiri pada saat berkunjung ke KSP Sukri Jaya Mandiri. Tujuan wawancara ini adalah untuk memperoleh informasi yang tepat dan jelas yang dibutuhkan dalam penyusunan penelitian.

2.5 Tahapan Penelitian

Pada penelitian metode yang diusulkan untuk proses data mining pada kasus ini menggunakan metode prediksi dengan algoritma *K-Neares Neighbor (K-NN)*. Tahapan Langkah Algoritma K-NN sebagai berikut :



Gambar 1. Tahapan Penelitian

Dari Gambar 1 tahapan penelitian dapat dijelaskan sebagai berikut :

1. Mulai

Tahap awal penelitian dimulai dengan merumuskan masalah dan tujuan penelitian. Pada tahap ini, juga ditentukan metode yang akan digunakan, misalnya *K-Nearest Neighbor (K-NN)*, serta identifikasi kebutuhan data.

2. Mengumpulkan Data

Data yang relevan dengan penelitian dikumpulkan dari sumber terpercaya, seperti dataset yang tersedia secara publik, eksperimen, atau survei.

3. Seleksi Data

Tahapan ini melibatkan pembersihan dan seleksi fitur data. Data yang tidak relevan, duplikat, atau memiliki nilai hilang diolah agar kualitas data menjadi baik. Pada tahap ini, juga dilakukan normalisasi data jika diperlukan.

4. Implementasi K-NN

Setelah data siap, algoritma K-NN diterapkan. Peneliti menentukan parameter penting, seperti jumlah tetangga (k) yang digunakan, dan mengimplementasikan algoritma menggunakan perangkat lunak yaitu rapidminer

5. Validasi dan Analisis Hasil

Model divalidasi menggunakan teknik seperti *cross-validation* atau membagi data menjadi *training* dan *test set*. Pada tahap ini, metrik evaluasi seperti akurasi, presisi, atau *recall* dianalisis untuk mengukur performa model.

6. Selesai

Penelitian disimpulkan dengan menyusun laporan yang memuat hasil, analisis, dan rekomendasi untuk penelitian lanjutan. Jika memungkinkan, hasil penelitian diimplementasikan pada aplikasi nyata.

2.5 Tahapan Pengujian K-Nearest Neighbor Rapid Minner

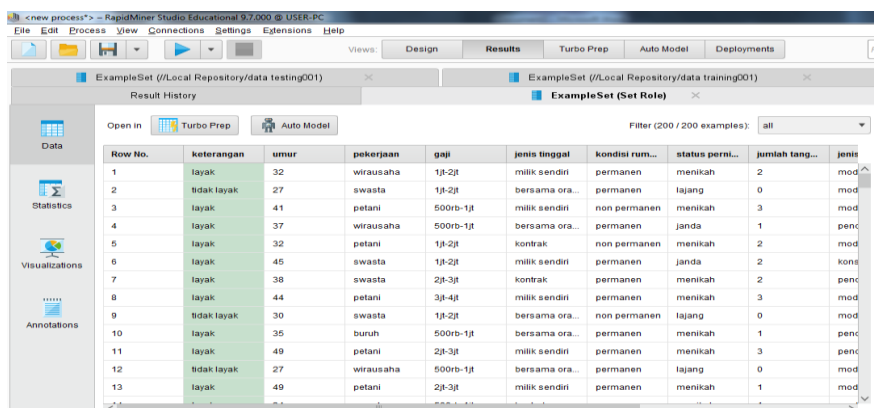
Adapun langkah untuk melakukan perhitungan algoritma *K-Nearest Neighbor* menggunakan rapidminer adalah

1. Setelah melakukan *import* data langsung ke proses implementasi, drag Data Training kedalam lembar kerja. Kemudian kita pilih dan drag *set role* pada *operators* ke lembar kerja, lalu hubungkan keduanya seperti gambar dibawah ini. *Set role* digunakan untuk menetapkan atribut mana yang akan dijadikan label pada *parameters*.



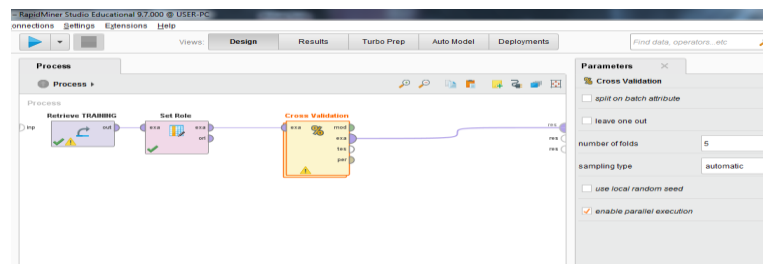
Gambar 2. Tampilan Data Training dan Set Role

2. Jika atribut yang sudah diubah menjadi label maka kolom pada atribut tersebut berubah warna menjadi hijau.

The screenshot shows the 'Results' view in RapidMiner Studio. It displays a data table with 13 rows and 10 columns. The columns are: 'Row No.', 'keterangan', 'umur', 'pekerjaan', 'gaji', 'jenis tinggal', 'kondisi rum...', 'status perni...', 'jumlah lang...', and 'jenis'. The 'keterangan' column is highlighted in green, indicating it is the target variable. The data rows contain various attributes such as 'layak', 'tidak layak', 'wirausaha', 'petani', 'buruh', 'gaji', 'jenis tinggal', 'kondisi rum...', 'status perni...', 'jumlah lang...', and 'jenis'.

Gambar 3. Tampilan Data Set

3. Selanjutnya kita akan memasukan *cross validation* kedalam lembar kerja kemudian hubungkan *set role* dengan *cross validation*. Karena kita akan menggunakan nilai k adalah 5, maka pada *number of folds* ubah menjadi 5.



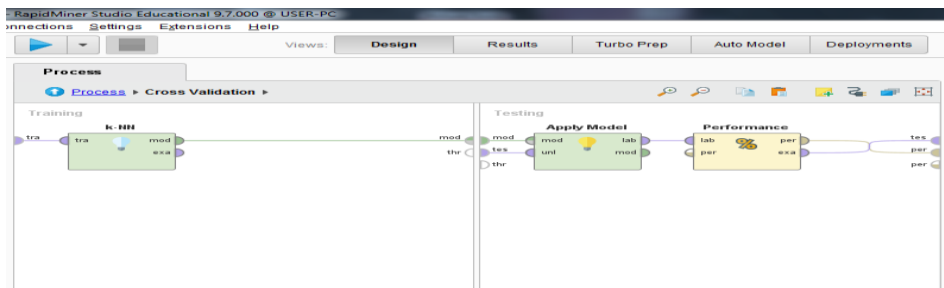
Gambar 4. Tampilan Pengaplikasian Cross Validation

4. Kemudian klik 2x pada *cross validation*, maka muncul lembar seperti gambar dibawah ini. Jadi sisi kiri merupakan training dan sisi kanan merupakan testing.



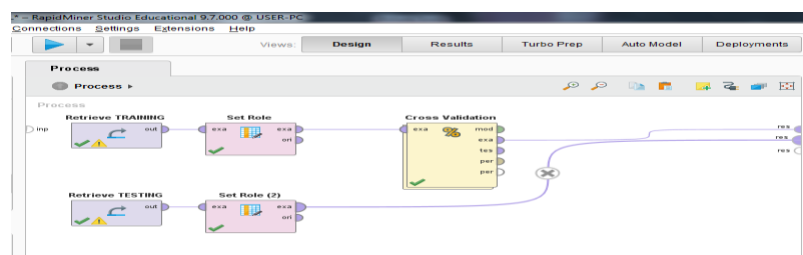
Gambar 5. Tampilan Cross Validation

5. Setelah itu masukan K-NN dengan nilai k=5 pada sisi training dan *apply models* serta *performance* pada sisi testing, lalu hubungkan ketiganya.



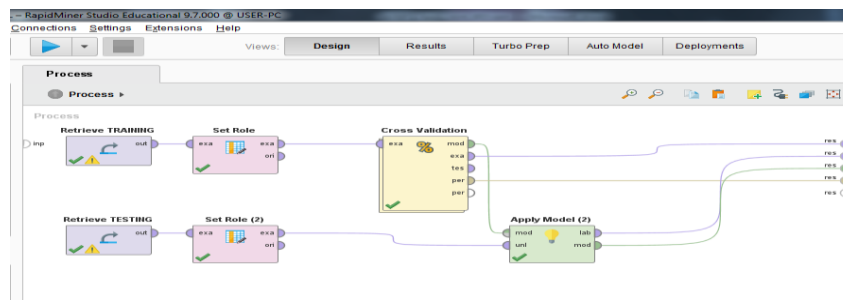
Gambar 6. Proses Implementasi K-NN

6. Lalu kembali ke *Process*, masukan data testing dan *Set Role* dan tentukan atribut keterangan menjadi label. Setelah itu coba *Run* dan lihat jika atribut keterangan berwarna hijau maka sudah menjadi label.



Gambar 7. Tampilan Proses Data Testing

7. Lalu gunakan *Apply Model* untuk melihat hasil akhirnya. Hubungkan *Cross Validation* dan *Set Role* data testing pada *Apply Model* seperti gambar dibawah ini, kemudian *Run* untuk melihat hasilnya.



Gambar 8. Tampilan Proses Akhir

3. HASIL DAN PEMBAHASAN

3.1 Perhitungan Algoritma *K-Nearest Neighbor*

Menghitung Kedekatan Atribut Kasus Baru (Data Testing)

Untuk mengetahui hasil analisa dari metode *K-Nearest Neighbor* maka dilakukan perhitungan manual antara data training (kasus lama) dan data testing (kasus baru) seperti pada tabel 1

Tabel 1. Menghitung Kedekatan Atribut Kasus Baru

No	nama	umur	pekerjaan	gaji	jenis tinggal	kondisi rumah	status pernikahan	jumlah tanggungan	jenis pinjaman	jaminan	status keanggotaan	pinjaman lain
1	aan sulpaeni	41	buruh	500rb-1jt	milik sendiri	non permanen	menikah	2	konsumsi	tidak ada	baru	tidak ada
2	sri juwanti	34	wirusaha	1jt-2jt	milik sendiri	permanen	menikah	2	modal usaha	tidak ada	baru	ada
3	nuraida	29	swasta	2jt-3jt	bersama orang tua	permanen	menikah	0	konsumsi	tidak ada	baru	tidak ada
4	handika yuliono	45	buruh	500rb-1jt	milik sendiri	permanen	menikah	3	konsumsi	ada	baru	ada
5	ahmad yamin	50	pns	>4jt	milik sendiri	permanen	menikah	0	modal usaha	ada	lama	ada
6	rustam	48	buruh	500rb-1jt	milik sendiri	non permanen	menikah	4	pendidikan	tidak ada	baru	tidak ada
7	dwi widyastuti	32	buruh	1jt-2jt	milik sendiri	permanen	menikah	1	konsumsi	tidak ada	lama	tidak ada
8	widya setya	35	swasta	1jt-2jt	bersama orang tua	permanen	menikah	0	modal usaha	ada	baru	tidak ada
9	zalman	46	swasta	3jt-4jt	milik sendiri	permanen	menikah	3	pendidikan	ada	baru	tidak ada
10	siswanto	47	buruh	1jt-2jt	kontrak	non permanen	menikah	4	konsumsi	tidak ada	baru	ada
11	nusirwan	39	petani	3-4jt	milik sendiri	permanen	menikah	3	modal usaha	tidak ada	lama	ada
17	muhajir	43	petani	1jt-2jt	milik sendiri	non permanen	menikah	2	pendidikan	ada	baru	tidak ada
.
18	leni marlina	32	swasta	3jt-4jt	kontak	permanen	menikah	1	konsumsi	tidak ada	lama	ada
19	tuti herawati	46	pns	>4jt	milik sendiri	permanen	menikah	4	pendidikan	ada	lama	ada

Dari 20 data testing diatas maka hanya 5 data testing yang akan digunakan dalam perhitungan menggunakan metode *K-Nearest Neighbor*. Untuk menghitung kedekatan kasus antara data training dengan data testing diatas berikut adaah penyelesaiannya. Menghitung nilai kedekatan atribut antara data testing 1 atas nama “Aan Sulpaeni” dengan data training 1 sampai dengan 200.

Tabel 2. Menghitung Nilai Kedekatan Kasus

Nama	Atribut1	Atribut2	Atribut 3	Atribut 4	Atribut 5	Atribut 6	Atribut 7	Atribut 8	Atribut 9	Atribut 10	Atribut 11
Aan Sulpaeni	41	Buruh	500rb-1jt	Milik sendiri	Non Permanen	Menikah	2	Konsumsi	Tidak ada	Baru	Tidak ada
Arif	32	Wirusaha	1jt-2jt	Milik Sendiri	Permanen	Menikah	2	Modal usaha	Tidak ada	Baru	Tidak ada
Nilai kedekatan	1	0.6	0.8	1	0.6	1	1	0.6	1	1	1
Nilai atribut	a	c	e	G	i	k	m	o	q	s	u

Perhitungan :

Jarak

$$= \frac{(a * b) + (c * d) + (e * f) + (g * h) + (i * j) + (k * l) + (m * n) + (o * p) + (q * r) + (s * t) + (u * v)}{b + d + f + h + j + l + n + p + r + t + v}$$

$$= \frac{(1 * 0.7) + (0.6 * 0.8) + (0.8 + 1) + (1 * 1) + (0.6 * 0.8) + (1 * 0.8) + (1 + 1) + (0.6 * 0.9) + (1 * 1) + (1 * 1) + (1 * 1)}{0.7 + 0.8 + 1 + 1 + 0.8 + 0.8 + 1 + 0.9 + 1 + 1 + 1}$$

$$= \frac{0.7 + 0.48 + 0.8 + 1 + 0.48 + 0.8 + 1 + 0.54 + 1 + 1 + 1}{10}$$

$$= \frac{8.8}{10} = 0.88$$

Setelah dilakukan perhitungan nilai kedekatan kasus baru dengan kasus lama dengan Algoritma *K-Nearest Neighbor*, maka berikut adalah hasil rekapitulasi nilai kedekatan kasus baru dengan kasus lama :

Tabel 3. Rekapitulasi Nilai Kedekatan

No	Kasus	Data Uji 1	Data Uji 2	Data Uji 3	Data Uji 4	Data Uji 5
1	Kasus 1	0.88	0.97	0.807	0.852	0.793
2	Kasus 2	0.743	0.801	0.912	0.715	0.751
3	Kasus 3	0.842	0.856	0.685	0.87	0.851
4	Kasus 4	0.768	0.782	0.785	0.74	0.715
5	Kasus 5	0.812	0.806	0.715	0.78	0.781
6	Kasus 6	0.854	0.84	0.813	0.826	0.755
7	Kasus 7	0.78	0.842	0.799	0.872	0.777
8	Kasus 8	0.79	0.888	0.717	0.882	0.883
9	Kasus 9	0.78	0.774	0.815	0.748	0.749
10	Kasus 10	0.832	0.782	0.817	0.804	0.747
11	Kasus 11	0.755	0.817	0.737	0.847	0.882
12	Kasus 12	0.703	0.813	0.82	0.735	0.811
13	Kasus 13	0.725	0.823	0.737	0.817	0.918
14	Kasus 14	0.772	0.786	0.717	0.8	0.811
15	Kasus 15	0.785	0.823	0.827	0.757	0.888
16	Kasus 16	0.837	0.787	0.747	0.869	0.852
17	Kasus 17	0.869	0.755	0.745	0.777	0.79
18	Kasus 18	0.833	0.747	0.888	0.805	0.747
19	Kasus 19	0.78	0.838	0.847	0.752	0.813
⋮	⋮					
197	Kasus 197	0.785	0.879	0.747	0.877	0.872
198	Kasus 198	0.872	0.918	0.715	0.9	0.821
199	Kasus 199	0.846	0.832	0.905	0.818	0.747
200	Kasus 200	0.82	0.842	0.809	0.852	0.787

Kasus baru 5 Berdasarkan tabel diatas merupakan hasil dari perhitungan nilai kedekatan data uji/kasus baru 1, 2, 3, 4 dan 5. Kemudian akan di cari nilai tetangga terdekat. Nilai k yang digunakan adalah k=5. Di dapat hasil sebagai berikut :

1. Kasus baru 1

Dari pencarian nilai k5 (jumlah tetangga paling dekat), di ketahui bahwa kasus baru 1 memiliki kedekatan dengan kasus 58 sebanyak 0.938 dengan keterangan “layak”, kasus 86 sebanyak 0.938 dengan keterangan “layak”, kasus 108 sebanyak 0.95 dengan keterangan “tidak layak”, kasus 170 sebanyak 0.938 dengan keterangan “layak”, dan kasus 91 sebanyak 0.94 dengan keterangan “tidak layak”. Jadi di dapat hasil kasus baru 1 “LAYAK” mendapat pinjaman.

2. Kasus baru 2

Dari pencarian nilai k5 (jumlah tetangga paling dekat), di ketahui bahwa kasus baru 2 memiliki kedekatan dengan kasus 1 sebanyak 0.97 dengan keterangan “layak”, kasus 37 sebanyak 0.95 dengan keterangan “layak”, kasus 89 sebanyak 0.95 dengan keterangan “layak”, kasus 132 sebanyak 0.96 dengan keterangan “layak”, dan kasus 180 sebanyak 0.97 dengan keterangan “layak”. Jadi di dapat hasil kasus baru 2 “LAYAK” mendapat pinjaman.

3. Kasus baru 3

Dari pencarian nilai k5 (jumlah tetangga paling dekat), di ketahui bahwa kasus baru 3 memiliki kedekatan dengan kasus 78 sebanyak 0.94 dengan keterangan “tidak layak”, kasus 129 sebanyak 0.948 dengan keterangan “tidak layak”, kasus 146 sebanyak 0.948 dengan keterangan “tidak layak”, kasus 159 sebanyak 0.948 dengan keterangan “tidak layak”, dan kasus 191 sebanyak 0.92 dengan keterangan “layak”. Jadi di dapat hasil kasus baru 3 atas nama “Nuraida” “TIDAK LAYAK” mendapat pinjaman.

4. Kasus baru 4

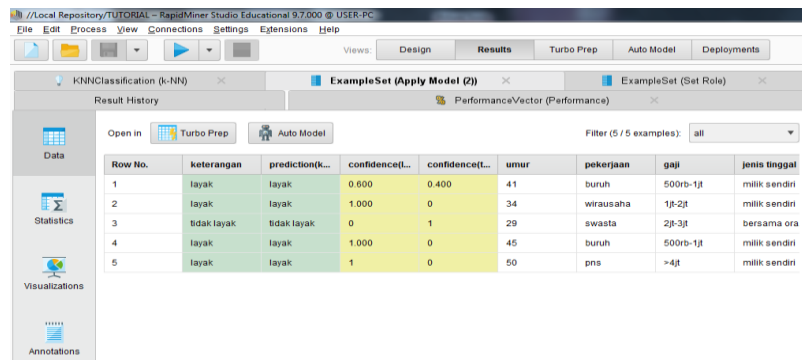
Dari pencarian nilai k5 (jumlah tetangga paling dekat), di ketahui bahwa kasus baru 4 memiliki kedekatan dengan kasus 45 sebanyak 0.94 dengan keterangan “layak”, kasus 54 sebanyak 0.97 dengan keterangan “layak”, kasus 106 sebanyak 0.918 dengan keterangan “tidak layak”, kasus 147 sebanyak 0.932 dengan keterangan “layak”, dan kasus 183 sebanyak 0.938 dengan keterangan “tidak layak”. Jadi di dapat hasil kasus baru 4 atas nama “Handika Yuliono” “LAYAK” mendapat pinjaman.

5. Kasus baru 5

Dari pencarian nilai k5 (jumlah tetangga paling dekat), di ketahui bahwa kasus baru 5 memiliki kedekatan dengan kasus 13 sebanyak 0.918 dengan keterangan “layak”, kasus 40 sebanyak 0.95 dengan keterangan “layak”, kasus 60 sebanyak 0.918 dengan keterangan “layak”, kasus 72 sebanyak 0.918 dengan keterangan “layak”, dan kasus 182 sebanyak 0.948 dengan keterangan “layak” Jadi di dapat hasil kasus baru 5 atas nama “Ahmad Yamin” “LAYAK” mendapat pinjaman.

3.2 Hasil Pengujian Rapidminer

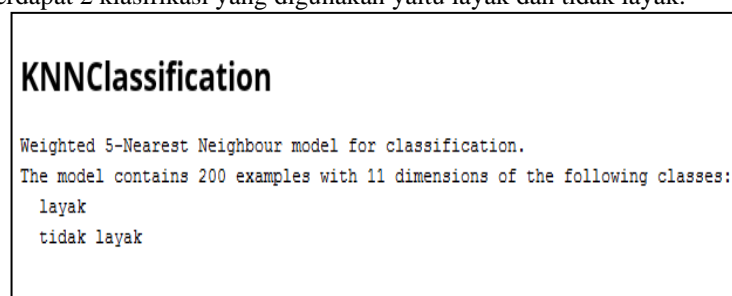
Setelah melakukan *Running* maka akan di dapatkan hasil seperti gambar di bawah ini. Setelah itu lihat hasil prediksinya apakah klasifikasi yang telah dibuat dengan perhitungan manual maupun dengan *Microsoft excel* sama / *balance* dengan hasil impelentasi menggunakan Rapidminer.



Row No.	keterangan	prediction(k...	confidence(L...	confidence(R...	umur	pekerjaan	gaji	jenis tinggal
1	layak	layak	0.000	0.400	41	buruh	500rb-1jt	milik sendiri
2	layak	layak	1.000	0	34	wirausaha	1jt-2jt	milik sendiri
3	tidak layak	tidak layak	0	1	29	swasta	2jt-3jt	bersama ora
4	layak	layak	1.000	0	45	buruh	500rb-1jt	milik sendiri
5	layak	layak	1	0	50	pns	>4jt	milik sendiri

Gambar 9. Tampilan Example Set

8. Pada *K-NN Clasification* berisi deskripsi dimana data set yang digunakan sebanyak 200 data dengan 11 dimensi/atribut dan terdapat 2 klasifikasi yang digunakan yaitu layak dan tidak layak.



Gambar 10. Tampilan K-NN Classification

9. Untuk melihat akurasi dari data yang telah kita gunakan terdapat pada *Performance Vector*. Akurasi yang di dapat yaitu sebanyak 82.50%.

accuracy: 82.50% +/- 9.19% (micro average: 82.50%)

	true layak	true tidak layak	class precision
pred. layak	113	21	84.33%
pred. tidak layak	14	52	78.79%
class recall	88.98%	71.23%	

Gambar 11. Tingkat Akurasi Performance

Berdasarkan penelitian yang telah dilakukan dengan menggunakan data training nasabah pada KSP. Sukri Jaya Mandiri tahun 2018-2020 dengan jumlah data set sebanyak 200 data dan data training sebanyak 20 data, yang telah dilakukan perhitungan manual sebanyak 5 data testing menggunakan algoritma *K-Nearest Neighbor* untuk mengetahui prediksi nasabah yang layak atau tidak layak mendapat pinjaman. Hasil yang didapat yaitu prediksi yang menggunakan Rapidminer memiliki hasil yang sama dengan perhitungan manual algoritma *K-Neares Neighbor*. Kemudian pada gambar *Performance Vector* seperti yang dihasilkan dengan menggunakan Algoritma *K-Neares Neighbor*, perhitungan kasus lama pada data training dengan kasus baru yang ada di data testing tingkat akurasi yang didapat sebesar 82.50%.

Row No.	keterangan	prediction(L)	confidence(L)	umur	pekerjaan	gaji
1	layak	layak	0.600	41	buruh	5000-1jt
2	layak	layak	1.000	34	wirusaha	1jt-2jt
3	tidak layak	tidak layak	0	29	swasta	2jt-3jt
4	layak	layak	1.000	45	buruh	5000-1jt
5	layak	layak	1	50	pns	>4jt

Gambar 12. Hasil Prediksi Rapidminer

accuracy: 82.50% +/- 9.19% (micro average: 82.50%)

	true layak	true tidak layak	class precision
pred. layak	113	21	84.33%
pred. tidak layak	14	52	78.79%
class recall	88.98%	71.23%	

Gambar 13. Tampilan Persentase

4. KESIMPULAN

Berdasarkan penelitian yang telah dilakukan dengan menggunakan data training nasabah pada KSP. Sukri Jaya Mandiri tahun 2018-2020 dengan jumlah data set sebanyak 200 data dan data training sebanyak 20 data, yang telah dilakukan perhitungan manual sebanyak 5 data testing menggunakan algoritma *K-Nearest Neighbor* untuk mengetahui prediksi nasabah yang layak atau tidak layak mendapat pinjaman. Hasil yang didapat yaitu prediksi yang menggunakan Rapidminer memiliki hasil yang sama dengan perhitungan manual algoritma *K-Neares Neighbor*. Kemudian pada gambar *Performance Vector* seperti yang dihasilkan dengan menggunakan Algoritma *K-Neares Neighbor*, perhitungan kasus lama pada data training dengan kasus baru yang ada di data testing tingkat akurasi yang didapat sebesar 82.50%. Penerapan data mining menggunakan algoritma *K-Nearest Neighbor* untuk prediksi kelayakan penerima pinjaman dapat menghasilkan prediksi yang akurat, sehingga membantu dan memudahkan koperasi dalam menentukan apakah seorang nasabah layak atau tidak layak menerima pinjaman.

DAFTAR PUSTAKA

- Afandi, A., Noviana, N., & Nurdianah, D. (2022). Naive Bayes Method and C4.5 in Classification of Birth Data. *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, 16(4), 435. <https://doi.org/10.22146/ijccs.78198>
- Ahmed, R., Bibi, M., & Syed, S. (2023). Improving Heart Disease Prediction Accuracy Using a Hybrid Machine Learning Approach: A Comparative study of SVM and KNN Algorithms. *International Journal of Computations, Information and Manufacturing (IJCIM)*, 3(1), 49–54. <https://doi.org/10.54489/ijcim.v3i1.223>
- Hasibuan, A. M. (2024). *Classification of Customer Loan Eligibility Using The K-Nearest Neighbor (K-NN) Algorithm Klasifikasi Kelayakan Peminjaman Nasabah Menggunakan Algoritma K-Nearest Neighbor (K-NN)*. 4(10), 1525–1532.
- Irawan, D., Riswanto, P., Sistem Informasi, P., Teknologi Bisnis Dan Bahasa Dian Cipta Cendikia, I., Teknologi Komputer, P., Negara Candimas, J., & Lampung Utara, K. (2023). Penerapan Algoritma K-Nearest Neighbor (KNN) Untuk Mengklasifikasi Jenis Penerimaan Bantuan Studi Kasus Desa Madukoro Lampung Utara. *Jurnal Teknologi Komputer Dan Sistem Informasi*, 6(2), 204–207. <https://jurnal.ftikomibn.ac.id/>
- Junaedi, E., Siregar, A. M., & Nurlaelasari, E. (2022). Implementasi C4.5 Dan Algoritma K-Nearest Neighbor Untuk Prediksi Kelayakan Pemberian Kredit Menggunakan RapidMiner Studio. *Scientific Student Journal for Information, Technology and Science*, 3(1), 83–90.
- Munandar, K. (2020). *Analisis Sistem Peminjaman Untuk Nasabah Pada Koperasi Artha Sejati Cabang Kroya Menggunakan Data Mining Dan Algoritma C4. 5* <http://eprints.uty.ac.id/4827/>
- Nyoman, N., & Smrti, E. (2015). Otomatisasi Klasifikasi Buku Perpustakaan Dengan Menggabungkan Metode K-Nn Dengan K-Medoids. *Lontar Komputer : Jurnal Ilmiah Teknologi Informasi*, 4(1), 201–214.
- Waspah, A. I., Afandi, A., Efendi, D. M., Sartika, D., Waspah, A. I., Afandi, A., Efendi, D. M., Sartika, D., Informasi, S., Mining, D., & Algorithm, E. (2022). Expectation Maximization Algorithm Memprediksi Penjualan Susu Murni Pada Pt . Sewu Primatama Indonesia Lampung. *JUTIM (Jurnal Teknik Informatika Musirawas)*, 7(1), 27–38.
- Windy Mardiyah, N., Rahaningsih, N., & Ali, I. (2024). Penerapan Data Mining Menggunakan Algoritma K-Nearest Neighbor Pada Prediksi Pemberian Kredit Di Sektor Finansial. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(2), 1491–1499. <https://doi.org/10.36040/jati.v8i2.9010>
- Zhang, S. (2022). Challenges in KNN Classification. *IEEE Transactions on Knowledge and Data Engineering*, 34(10), 4663–4675. <https://doi.org/10.1109/TKDE.2021.3049250>