



**IMPLEMENTASI K-MEANS CLUSTERING DALAM
ANALISIS RISIKO KREDIT NASABAH BERMASALAH
DI PT. BPR SUPRADANAMAS**

SKRIPSI

NADHIRA JASMINE NURRAHMA

2110512138

**PROGRAM STUDI S1 SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN" JAKARTA
JAKARTA
2025**



**IMPLEMENTASI K-MEANS CLUSTERING DALAM
ANALISIS RISIKO KREDIT NASABAH BERMASALAH
DI PT. BPR SUPRADANAMAS**

SKRIPSI

**Diajukan Sebagai Salah Satu Syarat untuk Memperoleh Gelar Sarjana
Komputer**

NADHIRA JASMINE NURRAHMA

2110512138

**PROGRAM STUDI S1 SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN" JAKARTA
JAKARTA
2025**

LEMBAR PERSETUJUAN

LEMBAR PERSETUJUAN

Yang bertanda tangan dibawah ini:

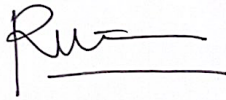
Nama : Nadhira Jasmine Nurrahma
NIM : 2110512138
Program Studi : Sistem Informasi Program Sarjana
Judul Tugas Akhir : Implementasi K-Means Clustering Dalam Analisis Risiko Kredit
Nasabah Bermasalah Di PT.BPR Supradanamas

Dinyatakan telah memenuhi syarat dan menyetujui untuk mengikuti ujian sidang Tugas Akhir.

Jakarta, 17 April 2025

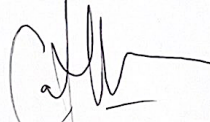
Menyetujui,

Dosen Pembimbing I,



Ruth Mariana Bunga Wadu, S.Kom.,
M.M.Si
NIDN. 0429038801

Dosen Pembimbing II,



Catur Nugrahaeni Ruspita Dewi, SKom.,
MKom.
NIDN. 0023088701

Mengetahui,

Koordinator Program Studi,



Anita Muliawati, S.Kom., MTI.
NIDN. 0321057001

PERNYATAAN HAK CIPTA

PERNYATAAN MENGENAI SKRIPSI DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa skripsi dengan judul “Implementasi K-Means Clustering Dalam Analisis Risiko Kredit Nasabah Bermasalah Di PT. BPR Supradanamas” adalah karya saya dengan arahan dari dosen pembimbing dan belum diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka di bagian akhir skripsi ini.

Dengan ini saya melimpahkan hak cipta dari karya tulis saya kepada Universitas Pembangunan Nasional “Veteran” Jakarta.

Jakarta, 25 April 2025



Nadhira Jasmine Nurrahma
2110512138

PERNYATAAN ORISINALITAS

PERNYATAAN ORISINALITAS

Skripsi ini merupakan hasil karya sendiri serta semua sumber referensi yang dikutip maupun yang dirujuk telah saya nyatakan benar.

Nama : Nadhira Jasmine Nurrahma

NIM : 2110512138

Tanggal : 04 Juli 2025

Bilamana di kemudian hari ditemukan ketidaksesuaian dengan pernyataan ini, maka saya bersedia dituntut dan diproses sesuai dengan ketentuan berlaku.

Jakarta, 04 Juli 2025

Yang Menyatakan,



Nadhira Jasmine Nurrahma

PERNYATAAN PERSETUJUAN PUBLIKASI

PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK KEPENTINGAN AKADEMIS

Sebagai civitas akademik Universitas Pembangunan Nasional Veteran Jakarta, saya yang bertanda tangan di bawah ini:

Nama : Nadhira Jasmine Nurrahma

NIM : 2110512138

Fakultas : Ilmu Komputer

Program Studi : S1 Sistem Informasi

Demi pembangunan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Pembangunan Nasional Veteran Jakarta Hak Bebas Royalti Eksklusif (*Non-exclusive Royalty Free Right*) atas karya ilmiah saya yang berjudul:

Implementasi K-Means Clustering Dalam Analisis Risiko Kredit Nasabah Bermasalah Di PT. BPR Supradanamas

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti ini Universitas Pembangunan Nasional Veteran Jakarta berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan Skripsi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sesungguhnya.

Dibuat di : Jakarta

Pada Tanggal : 04 Juli 2025

Yang Menyatakan,



Nadhira Jasmine Nurrahma

LEMBAR PENGESAHAN

LEMBAR PENGESAHAN

Judul : Implementasi K-Means Clustering Dalam Analisis Risiko Kredit
Nasabah Bermasalah Di PT.BPR Supradanamas
Nama : Nadhira Jasmine Nurrahma
NIM : 2110512138
Program Studi : SI Sistem Informasi

Disetujui oleh:

Penguji 1:
Ika Nurlaili Isnainiyah, S.Kom., M.Sc.

Penguji 2:
Bambang Tri Wahyono, S.Kom., M.Si.

Pembimbing 1:
Ruth Mariana Bunga Wadu, S.Kom., MMSI.

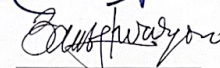
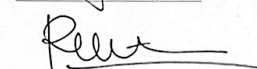
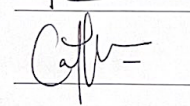
Pembimbing 2:
Catur Nugrahaeni Puspita Dewi, SKom., MKom.

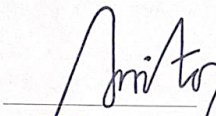
Diketahui oleh:

Koordinator Program Studi:
Anita Muliawati, S.Kom., MTI.
NIP. 197005212021212002

Dekan Fakultas Ilmu Komputer:
Prof. Dr. Ir. Supriyanto, S.T., M.Sc., IPM.
NIP. 197605082003121002

Tanggal Ujian Tugas Akhir :
12 Juni 2025




IMPLEMENTASI K-MEANS CLUSTERING DALAM ANALISIS RISIKO KREDIT NASABAH BERMASALAH DI PT. BPR SUPRADANAMAS

Nadhira Jasmine Nurrahma

ABSTRAK

Kredit bermasalah menjadi tantangan utama yang dihadapi PT. BPR Supradanamas dalam mengelola risiko kredit. Penelitian ini bertujuan untuk mengelompokkan nasabah PT. BPR Supradanamas berdasarkan perilaku pembayaran kredit menggunakan algoritma K-Means. Pengelompokan dilakukan untuk membedakan nasabah ke dalam tiga kategori, yaitu risiko kredit minimal, terkendali, dan tidak terkendali. Proses penelitian berisi preprocessing data, normalisasi, serta penentuan jumlah kluster optimal menggunakan metode *Elbow* dan *Silhouette Score*. Hasil kluster menunjukkan terdapat perbedaan karakteristik antar nasabah di setiap kluster, dari segi nilai tunggakan, durasi keterlambatan pembayaran, hingga keberhasilan pembayaran dan menunjukan nilai K yang optimal 3 berdasarkan *Elbow Method*, terjadi penurunan pada nilai *Sum of Squared Errors (SSE)* sehingga jumlah *cluster* 3. Hasil visualisasi dan rekomendasi dengan *framework* Streamlit menampilkan *clustering* risiko dari tahun 2022-2024, diharapkan dapat membantu perusahaan dalam memahami kluster risiko nasabah. Dengan klasterisasi, perusahaan dapat meminimalkan risiko kredit bermasalah, menentukan nasabah yang layak mendapat kredit baru dan mengoptimalkan strategi pengelolaan kredit di masa mendatang.

Kata Kunci: K-Means, Risiko Kredit, Perilaku Pembayaran, Kredit Bermasalah.

IMPLEMENTATION OF K-MEANS CLUSTERING FOR CUSTOMER NON PERFORMING LOANS RISK ANALYSIS AT PT. BPR SUPRADANAMAS

Nadhira Jasmine Nurrahma

ABSTRACT

Problematic loans are a major challenge for PT. BPR Supradanamas in managing credit risk. This study aims to segment customers of PT. BPR Supradanamas based on their credit payment behavior using the K-Means algorithm. The clustering is conducted to categorize customers into three risk levels: minimal risk, controlled risk, and uncontrolled risk. The process includes data preprocessing, normalization, and determining the optimal clusters using the Elbow Method and Silhouette Score. Clustering results indicate differences customer characteristics across cluster in terms of outstanding amounts, delay duration, and payment success rates. The optimal number of clusters was determined to be three based on the Elbow Method, indicated by a significant drop in the Sum of Squared Errors (SSE). Visualization and recommendation system developed using the Streamlit framework display the clustering of customer credit risk from 2022–2024. The study help company understand customer risk clusters. With clustering, company can minimize problematic credit risks, identify eligible customers for new credit, and optimize future credit strategies.

Keywords: *K-Means, Credit Risk, Payment Behavior, Non-Performing Loans.*

KATA PENGANTAR

Puji dan syukur penulis panjatkan kepada Allah SWT, yang telah memberikan rahmat serta karunianya, berkat hidayah dan kehendak-Nya, penulis dapat menyelesaikan Skripsi ini dengan judul "Implementasi K-Means *Clustering* Dalam Analisis Risiko Kredit Nasabah Bermasalah Di PT. BPR Supradanamas". Penulis menyadari dalam penyelesaian skripsi ini banyak pihak yang telah membimbing dan memberikan dukungan. Oleh karena itu, pada kesempatan ini penulis mengucapkan terima kasih yang sebesar besarnya kepada:

1. Kedua Orang Tua penulis, yaitu Mama dan Papa yang selalu memberi doa, dukungan, semangat dan motivasi tanpa henti dalam setiap langkah perjalanan penulis hingga saat ini.
2. Bapak Prof. Dr. Ir. Supriyanto, S.T., M.Sc., selaku Dekan Fakultas Ilmu Komputer UPN Veteran Jakarta.
3. Ibu Anita Muliawati, S.Kom., M.TI, selaku Ketua Program Studi S1 Sistem Informasi UPN Veteran Jakarta.
4. Ibu Ruth Mariana Bunga Wadu, S.Kom., M.M.Si, selaku Dosen Pembimbing skripsi 1 dan Ibu Catur Nugrahaeni Puspita Dewi, SKom., MKom. selaku dosen pembimbing 2, yang telah membimbing, memberikan masukan, bersedia meluangkan waktu dan memberikan arahan kepada penulis dalam proses pengerjaan skripsi ini.
5. Bapak Dr. Bambang Saras Yulistiawan, S.T., M.Kom, selaku dosen pembimbing akademik yang telah memberikan arahan dan ilmu sejak awal perkuliahan.
6. Seluruh dosen di Fakultas Ilmu Komputer yang telah memberikan ilmu serta arahan selama masa perkuliahan.
7. Teman-teman perkuliahan penulis yaitu Najwaa, Risa, Rasya, Kania, Aulia serta teman perkuliahan lainnya yang selalu memberikan dukungan, semangat dan saling mendukung dari awal perkuliahan hingga proses penyusunan skripsi ini.

Penulis menyadari bahwa dalam penyusunan skripsi ini masih jauh dari sempurna. Oleh karena itu, penulis mengharapkan kritik dan saran yang membangun untuk perbaikan di masa yang akan datang. Penulis berharap semoga skripsi ini dapat memberikan manfaat bagi semua pihak, serta dapat menjadi referensi untuk pengembangan ilmu pengetahuan di bidang Sistem Informasi, khususnya dalam analisis risiko kredit menggunakan metode K-Means Clustering.

Jakarta, 25 April 2025



Nadhira Jasmine Nurrahma

DAFTAR ISI

LEMBAR PERSETUJUAN	iii
PERNYATAAN HAK CIPTA	iv
PERNYATAAN ORISINALITAS	v
PERNYATAAN PERSETUJUAN PUBLIKASI	vi
LEMBAR PENGESAHAN	vii
ABSTRAK	viii
<i>ABSTRACT</i>	ix
KATA PENGANTAR	x
DAFTAR ISI	xii
DAFTAR GAMBAR	xv
DAFTAR TABEL	xvi
DAFTAR RUMUS	xvii
DAFTAR SIMBOL	xviii
DAFTAR LAMPIRAN	xix
BAB I PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	2
1.3. Batasan Masalah	3
1.4. Tujuan Penelitian	3
1.5. Manfaat Penelitian	4
1.6. Luaran yang Diharapkan	4
1.7. Sistematika Penulisan	5
BAB II TINJAUAN PUSTAKA	7
2.1. Kredit Bermasalah	7
2.2. Tunggakan Pembayaran	8
2.3. Risiko Kredit	8
2.4. Data Mining	9
2.5. Clustering	11
2.6. Algoritma K-Means Clustering	11
2.7. Python	13
2.8. Metode <i>Elbow</i>	13
2.9. Metode <i>Silhouette Coefficient</i>	14
2.10. Peneliti Terdahulu	15

BAB III METODE PENELITIAN	18
3.1. Alur Penelitian	18
3.1.1. Identifikasi Masalah.....	19
3.1.2. Studi Literatur	19
3.1.3. Pengumpulan Data	19
3.1.4. Data Selection	19
3.1.5. Data Preprocessing	20
3.1.6. Data Transformation	20
3.1.7. Data Mining	21
3.1.8. Interpretation/Evaluation	21
3.1.9. Analisis Hasil dan Visualisasi Rekomendasi	21
3.2. Alat Pendukung Penelitian	22
3.2.1. Perangkat Keras (<i>Hardware</i>)	22
3.2.2. Perangkat Lunak (<i>Software</i>)	22
3.3. Jadwal Penelitian	22
BAB IV HASIL DAN PEMBAHASAN	24
4.1. Profil Perusahaan	24
4.2. Data Selection	24
4.3. Data Preprocessing	27
4.4. Data Transformation	29
4.5. Data Mining	30
4.5.1. Menentukan Jumlah <i>Cluster</i> Optimal dengan Metode <i>Elbow</i> ...	30
4.5.2. Inisialisasi Centroid dan Penghitungan Jarak	32
4.5.3. Pengelompokan Data	33
4.5.4. Pembaruan Centroid	34
4.5.5. Proses Iterasi	34
4.6. Data Interpretation and Evaluation	34
4.6.1. Hasil Clustering	35
4.7. Analisis Hasil dan Visualisasi Rekomendasi	37
4.7.1. Distribusi Cluster Berdasarkan Data Nasabah	37
4.7.2. Rata-rata Total Tunggakan dan Durasi Tunggakan per Cluster	38
4.7.3. Rata-rata Plafon dan Tenor per Cluster	40
4.7.4. Rata-rata Coll Rate per Cluster	41
4.7.5. Heatmap Rata-rata Variabel Risiko per Cluster	42
4.7.6. Boxplot Risiko Nasabah Per Cluster	43

4.7.7.	Dashboard Visualisasi Streamlit.....	44
4.7.8.	Fitur Form Rekomendasi Kredit Baru Nasabah	49
BAB V PENUTUP		53
5.1.	Kesimpulan	53
5.2.	Saran	54
DAFTAR PUSTAKA		55
RIWAYAT HIDUP		58
LAMPIRAN		59

DAFTAR GAMBAR

Gambar 2.1. Tahapan Knowledge Discovery in Databases (Haris Kurniawan et al., 2020)	10
Gambar 3.1. Alur Penelitian	18
Gambar 4.1. Data Kredit Nasabah PT. BPR Supradanamas	24
Gambar 4.2. Source Code Data Selection.....	25
Gambar 4.3. Data setelah diproses Data Selection	25
Gambar 4.4. Source Code Data Preprocessing	27
Gambar 4.5. Source Code Imputasi Numerik dan Kategorikal	28
Gambar 4.6. Hasil Data Preprocessing	28
Gambar 4.7. Source Code dan Hasil Data Transformation	30
Gambar 4.8. Source Code Elbow Method	30
Gambar 4.9. Elbow Jumlah Cluster	31
Gambar 4.10. Silhoutte Score	31
Gambar 4.11. Inisialisasi Centroid.....	32
Gambar 4.12. Source Code Pembaruan Centroid	34
Gambar 4.13. Proses Iterasi	34
Gambar 4.14. Hasil Clustering	35
Gambar 4.15. Data Hasil Clustering.....	36
Gambar 4.16. Visualisasi Distribusi Nasabah Berdasarkan Cluster.....	37
Gambar 4.17. Visualisasi Rata-rata Total Tunggakan dan Durasi Tunggakan	38
Gambar 4.18. Visualisasi Rata-rata Total Tunggakan dan Durasi Tunggakan	39
Gambar 4.19. Visualisasi Plafon dan Tenor per Cluster.....	40
Gambar 4.20. Visualisasi Coll Rate per Cluster	41
Gambar 4.21. Visualisasi Heatmap Risiko per Cluster	42
Gambar 4.22. Visualisasi Boxplot Risiko Nasabah per Cluster	43
Gambar 4.23. Visualisasi Dashboard Tabel per Cluster.....	44
Gambar 4.24. Visualisasi Dashboard Distribusi Nasabah per Cluster	45
Gambar 4.25. Visualisasi Dashboard Total Tunggakan Tertinggi	45
Gambar 4.26. Visualisasi Scatter Plot Durasi dan Total Tunggakan.....	46
Gambar 4.27. Visualisasi Line Plot Rata Total dan Durasi Tunggakan.....	46
Gambar 4.28. Visualisasi Bar Chart Total dan Durasi Tunggakan	47
Gambar 4.29. Visualisasi Rata Coll Rate Per Cluster.....	47
Gambar 4.30. Visualisasi Rata Plafon dan Tenor per Cluster	48
Gambar 4.31. Distribusi Variabel Numerik Boxplot per Cluster	48
Gambar 4.32. Source Code Rules Sederhana Prediksi Kredit.....	49
Gambar 4.33. Output Form Rekomendasi Kredit Baru Nasabah Tidak Layak.....	51
Gambar 4.34. Output Form Rekomendasi Kredit Baru Perlu Pertimbangan	51
Gambar 4.35. Output Form Rekomendasi Kredit Baru Nasabah Layak	52

DAFTAR TABEL

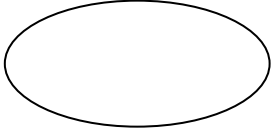
Tabel 2.1. Interpretasi Nilai Silhouette Coefficient	14
Tabel 2.2. Penelitian Terdahulu	15
Tabel 3.1. Jadwal Penelitian	23
Tabel 4.1. Rincian Atribut Dataset	25
Tabel 4.2. Kolom Normalisasi	29
Tabel 4.3. Hasil Silhoutte Score	32

DAFTAR RUMUS

Nomor	Judul Rumus	Halaman
2.1.	Normalisasi <i>Min Max Scaler</i>	10
2.2.	<i>Euclidean Distance</i>	12
2.3.	<i>Silhouette Coefficient</i>	14

DAFTAR SIMBOL

1. Simbol *Flowchart*

Simbol	Nama	Keterangan
	<i>Process</i>	Simbol yang menunjukkan suatu aktivitas atau proses yang dilakukan.
	<i>Terminator</i>	Simbol yang menyatakan titik awal (<i>start</i>) atau akhir (<i>end</i>) dari <i>flowchart</i> .
	<i>Flow</i>	Simbol yang menunjukkan arah proses dari satu langkah ke langkah berikutnya.

DAFTAR LAMPIRAN

Lampiran 1. Surat Permohonan Riset	59
Lampiran 2. Website Perusahaan.....	60
Lampiran 3. Source Code Python	61
Lampiran 4. Source Code Streamlit.....	68
Lampiran 5. Hasil Turnitin	73