KLASTERISASI DAN KLASIFIKASI STATUS KOLEKTABILITAS DATA PINJAMAN ANGGOTA MENGGUNAKAN ALGORITMA *K-MEANS* DAN *RANDOM FOREST*

(Studi Kasus Pada Unit Simpan Pinjam Koperasi Konsumen Bank BJB ZIEBAR Kota Bandung)

Disusun Oleh :

Silvi Nurinsan

2C2220006

SKRIPSI

Diajukan Untuk Memenuhi Salah Satu Syarat Memperoleh Gelar

Sarjana Sains Data

Dosen Pembimbing :

Drs. Udin Hidayat, S.E., M.Ti.

Muhammad Haris Fadhillah, SE., MM.

PROGRAM STUDI SARJANA SAINS DATA

FAKULTAS SAINS DAN TEKNOLOGI

UNIVERSITAS KOPERASI INDONESIA

2026

PERNYATAAN MENGENAI LAPORAN SKRIPSI DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa laporan skripsi dengan judul "**Klasterisasi Dan Klasifikasi Status Kolektabilitas Data Pinjaman Anggota Menggunakan Algoritma *K-Means* Dan *Random Forest***" adalah karya saya dengan arahan dari dosen pembimbing dan belum diajukan dalam bentuk apapun kepada perguruan tinggi manapun. 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 laporan akhir skripsi ini.

Dengan ini saya melimpahkan hak cipta dari karya tulis saya kepada Universitas Koperasi Indonesia.

Jatinangor, 14 September 2026

Silvi Nurinsan

2C2220006

IKOPIN University

LEMBAR PENGESAHAN

Judul	:	KLASTERISASI DAN KLASIFIKASI STATUS KOLEKTABILITAS DATA PINJAMAN ANGGOTA MENGGUNAKAN ALGORITMA *K-MEANS* DAN *RANDOM FOREST* (Studi Kasus Pada Unit Simpan Pinjam Koperasi Konsumen Bank BJB ZIEBAR Kota Bandung)
Nama	:	Silvi Nurinsan
NRP	:	2C2220006
Fakultas	:	Sains dan Teknologi
Program Studi	:	Sains Data
Konsentrasi	:	Digital Business Analytic

Menyetujui dan mengesahkan

Pembimbing 1	Pembimbing 2
Drs. Udin Hidayat, S.E., M.Ti.	M. Haris Fadhillah, S.E., M.M.
Pjs. Dekan Fakultas Sains dan Teknologi	Pjs. Ketua Program Studi Sains Data
Ir. H. Dady Nurpadi, M.P., CIIQA.	Ir. H. Dady Nurpadi, M.P., CIIQA.

RIWAYAT HIDUP



Silvi Nurinsan, lahir di Sumedang pada tanggal 11 September 2003. Penulis merupakan anak ke-1 dari 4 bersaudara, dari pasangan Bapak Andi Nurdin dan Ibu Santi Maryati. Pada tahun 2022 penulis diterima sebagai mahasiswa Program Studi Sarjana Sains Data di Universitas Koperasi Indonesia.

Jenjang pendidikan formal yang telah diselesaikan penulis adalah sebagai berikut :

1. Tahun 2010 menyelesaikan pendidikan di TK Miftahuttaufiq
2. Tahun 2016 menyelesaikan pendidikan Sekolah Dasar di SDN Margaluyu
3. Tahun 2019 menyelesaikan pendidikan di SMPN 1 Sukasari
4. Tahun 2022 menyelesaikan pendidikan di SMK Yadika Tanjungsari Sumedang

Pada tahun 2022, penulis terdaftar sebagai mahasiswi aktif Universitas Koperasi Indonesia pada Program Studi Sains Data. Adapun pengalaman organisasi yang diikuti selama perkuliahan yaitu :

1. Himpunan Mahasiswa Sains Data sebagai Koordinator Sumber Daya Manusia (SDM) Periode 2023-2024 kemudian menjadi Anggota Sumber Daya Manusia (SDM) Periode 2024-2025.
2. Unit Kegiatan Mahasiswa Ikopin Badminton Club sebagai anggota Periode 2023-2024
3. Staff Magang BEM KM Universitas Koperasi Indonesia sebagai anggota Kementerian Luar Negeri periode 2021-2022

Selama menempuh pendidikan di perguruan tinggi, penulis aktif terlibat dalam kegiatan akademik dengan menjadi Asisten laboratorium mata kuliah Big Data Analyst di Lab Sains Data pada tahun 2025. Dalam peran tersebut, penulis bertugas membantu dosen dalam membimbing mahasiswa selama kegiatan praktikum.

IKOPIN University

ABSTRACT

Silvi Nurinsan, 2026. *Clustering and Classification of Member Loan Data Collectability Status Using K-Means and Random Forest Algorithms (A Case Study at the Savings and Loan Unit of Koperasi Konsumen Bank BJB ZIEBAR, Bandung City), under the guidance of Udin Hidayat and Muhammad Haris Fadhillah.*

Membership of Koperasi Konsumen Bank BJB (ZIEBAR) declined from 2,923 in 2024 to 2,685 in 2025, driven partly by problematic loan repayments. This phenomenon is reflected in a Non-Performing Loan (NPL) rate of 2.98%, while manual identification of loan characteristics and collectability status has become less effective as data volume grows. This study segments members by loan characteristics using K-Means and classifies loan collectability status using Random Forest, supporting more objective, data-driven credit risk decisions.

This research used a quantitative descriptive and implementative approach with 2,712 member loan records, of which 1,567 met the selection criteria after preprocessing. Clustering used six variables: savings, loan ceiling, outstanding balance, loan term, total installment, and total arrears. The optimal cluster count was determined using the Elbow Method, Silhouette Score, and Davies-Bouldin Index. The Random Forest model was built using seven predictor features-the six variables above plus the K-Means cluster label-with collectability status as the target variable, evaluated using accuracy, precision, recall, and F1-score.

Results show K-Means produced four optimal clusters (k = 4): small-scale loans, large long-term loans, high-risk loans with significant arrears, and the healthiest loan profiles. Based on the seven predictor features, Random Forest achieved an accuracy of 0.9904, weighted precision of 0.9835, weighted recall of 0.9904, and weighted F1-score of 0.9865 in classifying collectability status. Feature importance results show total arrears, total installment, and outstanding balance as the three most influential features. However, substantially lower macro-average scores indicate that performance was affected by class imbalance, limiting the model's ability to accurately classify minority collectability categories.

Keywords: ***K-Means Clustering, Random Forest Classification,*** **L*oan Collectability* Status**

IKOPIN University

ABSTRAK

Silvi Nurinsan, 2026. Klasterisasi dan Klasifikasi Status Kolektabilitas Data Pinjaman Anggota Menggunakan Algoritma *K-Means* dan *Random Forest* (Studi Kasus Pada Unit Simpan Pinjam Koperasi Konsumen Bank BJB ZIEBAR Kota Bandung), di bawah bimbingan Udin Hidayat dan Muhammad Haris Fadhillah.

Jumlah anggota Koperasi Konsumen Bank BJB (ZIEBAR) menurun dari 2.923 orang pada 2024 menjadi 2.685 orang pada 2025, antara lain dipicu kondisi pembayaran pinjaman bermasalah. Fenomena ini ditunjukkan oleh tingkat *Non-Performing Loan* (NPL) sebesar 2,98%, sementara identifikasi karakteristik pinjaman dan penentuan status kolektabilitas manual menjadi kurang efektif seiring bertambahnya volume data. Penelitian ini bertujuan mengelompokkan anggota berdasarkan karakteristik pinjaman menggunakan *K-Means* serta mengklasifikasikan status kolektabilitas menggunakan *Random Forest*, guna mendukung pengelolaan risiko kredit yang lebih objektif dan berbasis data.

Penelitian menggunakan pendekatan kuantitatif deskriptif dan implementatif dengan 2.712 data pinjaman anggota, di mana 1.567 data memenuhi kriteria setelah seleksi dan praproses. Klasterisasi menggunakan enam variabel, yaitu simpanan, plafon pinjaman, baki debet, jangka waktu pinjaman, total angsuran, dan total tunggakan. Jumlah klaster optimal ditentukan menggunakan *Elbow Method*, *Silhouette Score*, dan *Davies-Bouldin Index*. Model *Random Forest* dibangun menggunakan tujuh fitur prediktor, yaitu keenam variabel tersebut ditambah label klaster hasil *K-Means*, dengan status kolektabilitas sebagai variabel target, dievaluasi menggunakan *accuracy*, *precision*, *recall*, dan *F1-score*.

Hasil penelitian menunjukkan *K-Means* menghasilkan empat klaster optimal (k = 4): pinjaman berskala kecil, pinjaman besar berjangka panjang, kelompok berisiko tinggi dengan tunggakan besar, serta kelompok berprofil paling sehat. Berdasarkan ketujuh fitur prediktor, *Random Forest* memperoleh *accuracy* 0,9904, *precision (weighted)* 0,9835, *recall (weighted)* 0,9904, dan *F1-score (weighted)* 0,9865 dalam mengklasifikasikan status kolektabilitas. Evaluasi *feature importance* menunjukkan total tunggakan, total angsuran, dan baki debet sebagai tiga fitur paling berpengaruh. Namun, nilai *macro average* yang jauh lebih rendah menunjukkan performa model dipengaruhi ketidakseimbangan kelas, sehingga kemampuan model mengklasifikasikan kelas minoritas status kolektabilitas masih terbatas.

Kata Kunci: Klasterisasi K-Means, Klasifikasi Random Forest, Status Kolektabilitas Pinjaman

KATA PENGANTAR

Puji dan syukur penulis panjatkan ke hadirat Tuhan Yang Maha Esa atas segala rahmat dan karunia-Nya, sehingga penulis dapat menyelesaikan skripsi yang berjudul "Klasterisasi dan Klasifikasi Status Kolektabilitas Data Pinjaman Anggota Menggunakan Algoritma *K-Means* dan *Random Forest* (Studi Kasus Pada Unit Simpan Pinjam Koperasi Konsumen Bank BJB ZIEBAR Kota Bandung)". Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana Sains Data pada Program Studi Sarjana Sains Data, Fakultas Sains dan Teknologi, Universitas Koperasi Indonesia.

Penyusunan skripsi ini bertujuan menganalisis karakteristik data pinjaman melalui klasterisasi *K-Means* dan mengklasifikasikan kolektabilitas pinjaman menggunakan *Random Forest* sebagai informasi pendukung pengelolaan risiko kredit secara lebih objektif dan berbasis data pada USP Koperasi Konsumen Bank BJB (ZIEBAR) Kota Bandung.

Penulis menyadari bahwa penyelesaian skripsi ini tidak terlepas dari bantuan, bimbingan, dan dukungan dari berbagai pihak. Oleh karena itu, pada kesempatan ini penulis ingin menyampaikan ucapan terima kasih yang sebesar-besarnya kepada:

1. Kedua orang tua tercinta, Bapak Andi Nurdin dan Ibu Santi Maryati, yang senantiasa memberikan doa, kasih sayang, dukungan, serta motivasi kepada penulis selama menjalani pendidikan hingga penyelesaian skripsi ini.

2. Yang terhormat, Bapak Udin Hidayat, SE., M.Ti. dan Bapak Muhammad Haris Fadhillah, SE., MM., MOS. selaku dosen pembimbing yang telah meluangkan waktu, tenaga, dan pikiran untuk memberikan bimbingan, arahan, serta masukan yang sangat berharga selama proses penyusunan skripsi ini.
3. Yang terhormat, Bapak Agus Nugraha, S.T, M.Kom selaku penguji konsentrasi yang telah memberikan banyak masukan dan arahan.
4. Yang terhormat, Bapak Ir. H. Dady Nurpadi., M.P. selaku penguji koperasi dan Wakil Dekan sekaligus Kaprodi Sains Data, yang telah memberikan banyak masukan dan arahan juga dukungan dan fasilitas selama penulis menempuh pendidikan di Universitas Koperasi Indonesia.
5. Bapak/Ibu Dosen Pengajar Universitas Koperasi Indonesia yang telah memberikan ilmu pengetahuan serta nilai-nilai kehidupan selama kegiatan perkuliahan.
6. Pihak pengurus dan pengelola Unit Simpan Pinjam Koperasi Konsumen Bank BJB (ZIEBAR) Kota Bandung, yang telah memberikan izin dan kesempatan kepada penulis untuk melakukan penelitian serta memperoleh data yang diperlukan.
7. Kakak dari Ibu penulis Arip Yuniardi, yang dengan penuh ketulusan telah membiayai pendidikan penulis sejak awal perkuliahan hingga terselesaikannya skripsi ini. Tanpa pengorbanan, dukungan, dan kepercayaan yang beliau berikan, penulis tidak akan dapat mencapai titik

ini. Semoga segala kebaikan dan jasa beliau senantiasa mendapat balasan yang berlimpah dari Tuhan Yang Maha Esa.

8. Nenek tercinta, Siti Aisyah dan kakek tercinta, Alm. Zaenal Arifin yang telah merawat, membesarkan, dan mendampingi penulis sejak kecil hingga saat ini. Terima kasih atas kasih sayang, perhatian, doa, pengorbanan, dan ketulusan yang telah diberikan kepada penulis. Kehadiran dan dukungan Nenek menjadi salah satu sumber kekuatan bagi penulis dalam menjalani setiap proses hingga dapat menyelesaikan pendidikan dan skripsi ini.
9. Serta Seluruh keluarga yang senantiasa memberikan doa, perhatian, dan dukungan kepada penulis.
10. Teman-teman seperjuangan Program Studi Sains Data, khususnya kepada Mulyati Eka Saputri, Teni Deinarosa Hermalia, Indri Nur Sukmawati, Cahya Rahmawati, Ashafa Multazam, Sagara Andi L. Yang telah memberikan bantuan, semangat, dan kebersamaan selama masa perkuliahan.
11. Kepada seluruh pengurus HIMA Sains Data periode 2023–2025, penulis mengucapkan terima kasih atas segala kebersamaan, dukungan, pengalaman, dan cerita yang telah dilalui selama dua periode kepengurusan. Terima kasih telah menjadi bagian dari perjalanan perkuliahan yang penuh dengan pembelajaran, tantangan, suka maupun duka. Setiap proses yang telah dilalui bersama akan menjadi kenangan berharga yang tidak akan penulis lupakan.

12. Kepada pihak yang tidak dapat penulis sebutkan satu per satu, yang telah hadir dan menemani penulis sejak awal hingga akhir masa perkuliahan, serta senantiasa memberikan dukungan, semangat, bantuan, dan motivasi, khususnya selama proses penyusunan skripsi ini.
13. Terakhir, penulis ingin mengucapkan terima kasih kepada diri sendiri yang telah berusaha, bertahan, dan terus melangkah hingga tahap ini. Terima kasih karena telah memilih untuk tidak menyerah meskipun dalam proses penyusunan skripsi ini terdapat berbagai kesulitan, keraguan, kelelahan, dan tantangan yang harus dihadapi. Terima kasih telah tetap berusaha menyelesaikan setiap proses dengan sebaik mungkin, belajar dari setiap kesalahan, serta percaya bahwa setiap langkah kecil tetap memiliki arti.

Penulis menyadari bahwa skripsi ini masih jauh dari kesempurnaan, baik dari segi isi maupun penyajiannya. Oleh karena itu, penulis sangat mengharapkan kritik dan saran yang membangun demi perbaikan di masa mendatang. Semoga skripsi ini dapat memberikan manfaat, khususnya bagi pengembangan ilmu Sains Data serta bagi pihak-pihak yang membutuhkan.

Jatinangor, September 2026

Silvi Nurinsan

IKOPIN University

DAFTAR ISI

RIWAYAT HIDUP ..... **i**
ABSTRACT ..... **iii**
ABSTRAK ..... **iv**
KATA PENGANTAR ..... **v**
DAFTAR ISI ..... **ix**
DAFTAR TABEL ..... **xiii**
DAFTAR GAMBAR ..... **xiv**
DAFTAR LAMPIRAN ..... **xv**

BAB I PENDAHULUAN **1**
1.1 Latar Belakang Penelitian **1**
1.2 Identifikasi Masalah **11**
1.3 Maksud dan Tujuan Penelitian **11**
1.3.1 Maksud Penelitian 11
1.3.2 Tujuan Penelitian 12
1.4 Kegunaan Penelitian **13**
1.4.1 Kegunaan Teoritis 13
1.4.2 Kegunaan Praktis 14

BAB II PENDEKATAN MASALAH, KERANGKA PIKIR, DAN METODE PENELITIAN **15**
2.1. Pendekatan Masalah, Kerangka Pikir **15**
2.1.1. Pendekatan Masalah 15
2.1.1.1 Pendekatan Perkoperasian 16
2.1.1.2 Pendekatan Variabel Data Pinjaman 33
2.1.1.3 Pendekatan Sains Data dan Knowledge Discovery in Databases (KDD) 37
2.1.1.4 Pendekatan Machine Learning 40
2.1.1.5 Pendekatan Klasterisasi 41
2.1.1.6 Pendekatan Klasifikasi 49
2.1.1.7 Perangkat Penelitian 56

2.1.2 Kerangka Pikir 60
2.2 Metode Penelitian 61
2.2.1 Metode Penelitian yang Digunakan 61
2.2.2 Data yang Diperlukan (Operasionalisasi Variabel)..... 61
2.2.3 Sumber Data dan Cara Menentukan Sumber Data 65
2.2.3.1 Data Primer 65
2.2.3.2 Data Sekunder 65
2.2.3.3 Populasi Penelitian 66
2.2.4 Teknik Pengumpulan Data 66
2.2.4.1 Dokumentasi 66
2.2.4.2 Observasi 67
2.2.4.3 Wawancara 67
2.2.5 Analisis Data 68
2.2.5.1 Diagram Alur Penelitian 69
2.2.5.2 Kondisi Pemanfaatan Data 69
2.2.5.3 Selection (Seleksi Data) 70
2.2.5.4 Preprocessing (Praproses Data)..... 71
2.2.5.5 Exploratory data analysis (EDA) 73
2.2.5.6 Transformation (Transformasi Data) 74
2.2.5.7 Penentuan Jumlah Cluster Menggunakan Elbow Method 75
2.2.5.8 K-Means Clustering 76
2.2.5.9 Evaluasi Klaster 77
2.2.5.10 Analisis Karakteristik Cluster 78
2.2.5.11 Feature Engineering 79
2.2.5.12 Klasifikasi Menggunakan Random Forest 79
2.2.5.13 Evaluasi Model Klasifikasi 82
2.2.5.14 Interpretasi Hasil dan Pemanfaatan Hasil 84
2.2.5.15 Knowledge 84
2.2.5.16 Pengendalian Kebocoran Data dan Validasi 85
2.2.5.17 Studi Ablasi Fitur Klaster..... 85
2.2.5.18 Keterulangan dan Etika Data 85
2.2.6 Tempat/Lokasi Penelitian..... 86

2.2.7 Jadwal Penelitian..... 87

BAB III KEADAAN UMUM TEMPAT PENELITIAN 88
3.1 Keadaan Umum Organisasi 88
3.1.1 Sejarah Terbentuknya Koperasi 88
3.1.2 Jenis Koperasi 89
3.1.3 Struktur Organisasi..... 91
3.2 Keanggotaan Koperasi 95
3.3 Kegiatan Usaha dan Permodalan Koperasi 96
3.3.1 Kegiatan Usaha Koperasi Konsumen Bank BJB ZIEBAR..... 96
3.3.2 Permodalan Koperasi 98
3.4 Kinerja Keuangan..... 99
3.4.1 Pertumbuhan dan Kemandirian..... 100
3.4.2 Efisiensi..... 102
3.4.3 Likuiditas..... 102

BAB IV HASIL DAN PEMBAHASAN..... 106
4.1 Identitas Data Penelitian dan Informan 106
4.1.1 Identitas Data Penelitian 106
4.1.2 Informan..... 107
4.2 Kondisi Pemanfaatan Data Pinjaman Anggota 108
4.3 Penerapan *K-Means*..... 109
4.3.1 Penentuan Jumlah Klaster..... 109
4.3.2 Hasil Klasterisasi..... 113
4.3.3 Karakteristik *Cluster* 114
4.3.4 Hasil Klasterisasi Anggota..... 115
4.4 Penerapan *Random Forest* 116
4.4.1 Persiapan Data untuk Pembentukan Model *Random Forest*..... 116
4.4.2 Hasil Evaluasi Model *Random Forest* 118
4.4.3 Analisis *Feature Importance*..... 121
4.4.4 Hasil Klasifikasi..... 122
4.5 Pemanfaatan Hasil Penelitian..... 125

BAB V SIMPULAN DAN SARAN 128
5.1 Simpulan 128
5.2 Saran 130
5.2.1 Saran Teoritis 130
5.2.2 Saran Praktis 131

DAFTAR PUSTAKA 133
LAMPIRAN 139

IKOPIN University

DAFTAR TABEL

Tabel 1.1 Perkembangan Jumlah Anggota Koperasi Konsumen Bank BJB (ZIEBAR) Kota Bandung Tahun 2021–2025 3

Tabel 2.1 Tahapan KDD dan Penerapannya dalam penelitian 39

Tabel 2.2 Kriteria Jumlah Klaster 47

Tabel 2.3 Ringkasan Penelitian Terdahulu yang Mendukung Pendekatan 55

Tabel 2.4 Perangkat Lunak dan Library yang digunakan Dalam Penelitian..... 59

Tabel 3 1 Pertumbuhan SHU Koperasi Konsumen Bank BJB ZIEBAR Tahun 2021-2025..... 97

Tabel 3 2 Perkembangan Rasio Likuiditas Koperasi Konsumen Bank BJB ZIEBA 103

Tabel 4.1 Nilai Silhoutte Score 110

Tabel 4.2 Nilai Davies-Bouldin Indeks..... 111

Tabel 4.3 Perbandingan Hasil Evaluasi Penentuan Jumlah Klaster..... 112

Tabel 4.4 Hasil Klasterisasi..... 113

Tabel 4.5 Karakteristik Setiap Cluster 114

Tabel 4.6 Distribusi Status Kolektabilitas Setelah Proses Label Encoding 117

Tabel 4.7 Pemanfaatan Hasil Klasterisasi dan Klasifikasi Status Kolektabilitas 126

IKOPIN University

DAFTAR GAMBAR

Gambar 2.1 Diagram Kerangka Pikir..... 60

Gambar 2.2 Diagram Alur Penelitian..... 69

Gambar 3.1 Struktur Organisasi Koperasi Konsumen Bank BJB ZIEBAR 91

Gambar 3.2 Struktur Organisasi Koperasi berdasarkan UU No.25 Tahun 1992 Tentang Perkoperasian..... 95

Gambar 4.1 Grafik Elbow Method..... 110

Gambar 4.2 Hasil Evaluasi metrik Confusion Matrix..... 118

Gambar 4.3 Hasil Classification Report..... 119

Gambar 4.4 Hasil Feature Importance 121

IKOPIN University

DAFTAR LAMPIRAN

Lampiran 1. Pedoman Wawancara 139

Lampiran 2. Dokumentasi Bersama Ketua SDM USP Koperasi ZIEBAR 140

IKOPIN University