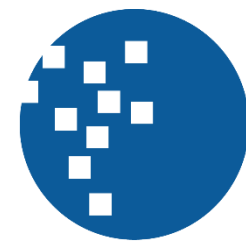


**PERBANDINGAN ALGORITMA *K-MEANS* DAN *K-MEDOIDS*
DALAM SEGMENTASI TINGKAT KESEHATAN BANK
PEREKONOMIAN RAKYAT BERDASARKAN
RASIO RGEC**



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

SKRIPSI

Salsabillah Isnain Zahirah

00000071543

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG**

2025

PERBANDINGAN ALGORITMA *K-MEANS* DAN *K-MEDOIDS*
DALAM SEGMENTASI TINGKAT KESEHATAN BANK
PEREKONOMIAN RAKYAT BERDASARKAN
RASIO RGEK



SKRIPSI

Diajukan sebagai Salah Satu Syarat untuk Memperoleh

Gelar Sarjana Komputer (S.Kom.)

Salsabillah Isnain Zahirah

00000071543

PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA

TANGERANG

2025

HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Salsabillah Isnain Zahirah

Nomor Induk Mahasiswa : 00000071543

Program Studi : Sistem Informasi

Skripsi dengan judul:

PERBANDINGAN ALGORITMA *K-MEANS* DAN *K-MEDOIDS* DALAM
SEGMENTASI TINGKAT KESEHATAN BANK PEREKONOMIAN RAKYAT
BERDASARKAN RASIO RGEC

Merupakan hasil karya saya sendiri bukan plagiat dari laporan karya tulis ilmiah yang ditulis oleh orang lain, dan semua sumber, baik yang dikutip maupun dirujuk, telah saya nyatakan dengan benar serta dicantumkan di Daftar Pustaka.

Jika di kemudian hari terbukti ditemukan kecurangan/penyimpangan, baik dalam pelaksanaan maupun dalam penulisan laporan karya tulis ilmiah, saya bersedia menerima konsekuensi dinyatakan TIDAK LULUS untuk mata kuliah yang telah saya tempuh.

Tangerang, 28 November 2025



Salsabillah Isnain Zahirah

LEMBAR PERNYATAAN PENGGUNAAN BANTUAN KECERDASAN ARTIFISIAL (AI)

Saya yang bertanda tangan di bawah ini:

Nama Lengkap : Salsabillah Isnain Zahirah
NIM : 00000071543
Program Studi : Sistem Informasi
Judul Skripsi : Perbandingan Algoritma *K-Means* Dan *K-Medoids*
Dalam Segmentasi Tingkat Kesehatan Bank Perekonomian Rakyat
Berdasarkan Rasio RGEC

Dengan ini saya menyatakan secara jujur menggunakan bantuan Kecerdasan Artifisial (AI) dalam pengerjaan Tugas Akhir sebagai berikut (beri tanda centang yang sesuai):

- ☐ Menggunakan AI sebagaimana diizinkan untuk membantu dalam menghasilkan ide-ide utama saja
- ☒ Menggunakan AI sebagaimana diizinkan untuk membantu menghasilkan teks pertama saja
- ☒ Menggunakan AI untuk menyempurnakan sintaksis dan tata bahasa untuk pengumpulan tugas
- ☐ Karena tidak diizinkan: Tidak menggunakan bantuan AI dengan cara apa pun dalam pembuatan tugas

Saya juga menyatakan bahwa:

- (1) Menyerahkan secara lengkap dan jujur penggunaan perangkat AI yang diperlukan dalam tugas melalui Formulir Penggunaan Perangkat Kecerdasan Artifisial (AI)
- (2) Saya mengakui bahwa saya telah menggunakan bantuan AI dalam tugas saya baik dalam bentuk kata, paraphrase, penyertaan ide atau fakta penting yang disarankan oleh AI dan saya telah menyantumkan dalam sitasi serta referensi

Terlepas dari pernyataan di atas, tugas akhir ini sepenuhnya merupakan karya saya sendiri

Tangerang, 28 November 2025
Yang Menyatakan,



Salsabillah Isnain Zahirah



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH

Yang bertanda tangan di bawah ini:

Nama : Salsabillah Isnain Zahirah

NIM : 00000071543

Program Studi : Sistem Informasi

Jenjang : S1

Judul Karya Ilmiah : PERBANDINGAN ALGORITMA K-MEANS DAN K-MEDOIDS DALAM SEGMENTASI TINGKAT KESEHATAN BANK PEREKONOMIAN RAKYAT BERDASARKAN RASIO RGEK

Menyatakan dengan sesungguhnya bahwa saya bersedia* (pilih salah satu):

- ☐ Saya bersedia memberikan izin sepenuhnya kepada Universitas Multimedia Nusantara untuk mempublikasikan hasil karya ilmiah saya ke dalam repositori Knowledge Center sehingga dapat diakses oleh Sivitas Akademika UMN/Publik. Saya menyatakan bahwa karya ilmiah yang saya buat tidak mengandung data yang bersifat konfidensial.
- ☐ Saya tidak bersedia mempublikasikan hasil karya ilmiah ini ke dalam repositori Knowledge Center, dikarenakan: dalam proses pengajuan publikasi ke jurnal/konferensi nasional/internasional (dibuktikan dengan *letter of acceptance*) **.
- ☒ Lainnya, pilih salah satu:
 - ☒ Hanya dapat diakses secara internal Universitas Multimedia Nusantara
 - ☐ Embargo publikasi karya ilmiah dalam kurun waktu 3 tahun.

Tangerang, 28 November 2025



Salsabillah Isnain Zahirah

KATA PENGANTAR

Puji syukur atas selesainya penulisan skripsi dengan judul: “Perbandingan Algoritma *K-Means* Dan *K-Medoids* Dalam Segmentasi Tingkat Kesehatan Bank Perekonomian Rakyat Berdasarkan Rasio RGEC” yang dilakukan untuk memenuhi salah satu syarat untuk mencapai gelar Sarjana Komputer pada program studi Sistem Informasi. Saya menyadari bahwa, tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai pada penyusunan skripsi ini, sangatlah sulit bagi saya untuk menyelesaikan skripsi yang telah dilakukan ini. Oleh karena itu, saya mengucapkan terima kasih kepada:

1. Ir. Andrey Andoko, M.Sc., Ph.D. selaku Rektor Universitas Multimedia Nusantara.
2. Dr. Eng. Niki Prastomo, S.T. M.Sc, selaku Dekan Fakultas Teknik dan Informatika Universitas Multimedia Nusantara.
3. Ibu Ririn Ikana Desanti, S.Kom., M.Kom., selaku Ketua Program Studi Sistem Informasi Universitas Multimedia Nusantara.
4. Ibu Monika Evelin Johan, S.Kom., M.M.S.I., sebagai Pembimbing yang telah banyak meluangkan waktu untuk memberikan bimbingan, arahan dan motivasi atas terselesainya Tugas Akhir ini.
5. Keluarga yang telah memberikan bantuan dukungan material dan moral, sehingga dapat menyelesaikan Tugas Akhir ini.
6. Otoritas Jasa Keuangan (OJK), sebagai institusi yang telah memberikan dukungan dan kesempatan dalam memperoleh data serta wawasan yang relevan, sehingga proses penyusunan Tugas Akhir ini dapat berjalan dengan optimal.
7. Tim Data Analitik Otoritas Jasa Keuangan, yang telah memberikan bantuan melalui diskusi serta dukungan dalam pemahaman data selama proses penyusunan Tugas Akhir ini.
8. Teman-teman seperjuangan dalam *Fast Track Career Acceleration* Program, yang telah menjadi rekan diskusi, berbagi semangat, dan saling

mendukung dalam menghadapi berbagai tantangan selama pelaksanaan program.

9. Kepada teman-teman magang atas kebersamaan dan dukungan yang telah memberikan warna tersendiri serta menjaga suasana positif selama proses penyusunan Tugas Akhir ini.

Akhir kata, ucapan terima kasih disampaikan kepada seluruh pihak yang telah memberikan kontribusi dan dukungan selama proses penyusunan Tugas Akhir ini, sejak tahap awal hingga penyelesaiannya.

Tangerang, 28 November 2025



Salsabillah Isnain Zahirah



PERBANDINGAN ALGORITMA *K-MEANS* DAN *K-MEDOIDS* DALAM SEGMENTASI TINGKAT KESEHATAN BANK PEREKOMIAN RAKYAT MENGGUNAKAN RASIO RGEK

Salsabillah Isnain Zahirah

ABSTRAK

Bank Perekonomian Rakyat (BPR) memiliki peran penting dalam mendukung sektor keuangan mikro, Namun, penilaian tingkat kesehatannya oleh Otoritas Jasa Keuangan (OJK) masih dilakukan melalui sistem *Monitoring Tools* yang belum sepenuhnya terotomatisasi, sehingga memerlukan waktu dan tenaga analisis yang cukup besar. Penelitian ini bertujuan untuk membandingkan kinerja algoritma *K-Means* dan *K-Medoids* dalam mengelompokkan tingkat kesehatan BPR berdasarkan rasio keuangan *Risk Profile*, *Good Corporate Governance*, *Earnings*, dan *Capital* (RGEK), serta mengembangkan model prediksi berbasis *Random Forest* untuk melakukan klasifikasi tingkat kesehatan secara otomatis. Data yang digunakan merupakan data *dummy* yang merepresentasikan kondisi keuangan BPR periode Juni 2024 hingga Agustus 2025. Hasil penelitian menunjukkan bahwa algoritma *K-Medoids* memiliki performa lebih baik dengan nilai *Silhouette Score* sebesar 0,8287 dibandingkan *K-Means* sebesar 0,7843. Model *Random Forest* yang dibangun berdasarkan hasil klasterisasi tersebut menghasilkan akurasi 100%, dengan variabel PK_NPL, PK_ROA, dan PK_LDR sebagai faktor paling berpengaruh. Sistem prediksi berbasis *Streamlit* yang dikembangkan mampu memberikan hasil klasifikasi secara interaktif dan efisien, mendukung proses pengawasan kesehatan BPR berbasis data.

Kata kunci: *Bank Perekonomian Rakyat, K-Means, K-Medoids, Random Forest, RGEK*

A COMPARATIVE STUDY OF K-MEANS AND K-MEDOIDS ALGORITHMS FOR SEGMENTING THE FINANCIAL HEALTH OF RURAL BANKS USING RGEC RATIOS

Salsabillah Isnain zahirah

ABSTRACT (English)

Rural Banks or Bank Perekonomian Rakyat play a crucial role in supporting the microfinance sector; However, the assessment of their financial soundness by the Financial Services Authority (OJK) is still conducted through the Monitoring Tools system, which has not yet been fully automated.. This study aims to compare the performance of the K-Means and K-Medoids algorithms in clustering the financial health level of BPR based on the Risk Profile, Good Corporate Governance, Earnings, and Capital (RGEC) ratios, as well as to develop a predictive classification model using the Random Forest algorithm for automated health assessment. The dataset used in this research is a dummy dataset representing BPR financial ratios from June 2024 to August 2025. The results show that the K-Medoids algorithm outperforms K-Means with a higher Silhouette Score of 0.8287 compared to 0.7843. The Random Forest model achieved 100% accuracy, with PK_NPL, PK_ROA, and PK_LDR identified as the most influential variables. Furthermore, a web-based prediction system developed using the Streamlit framework enables efficient and interactive classification, supporting a data-driven supervisory process for BPR.

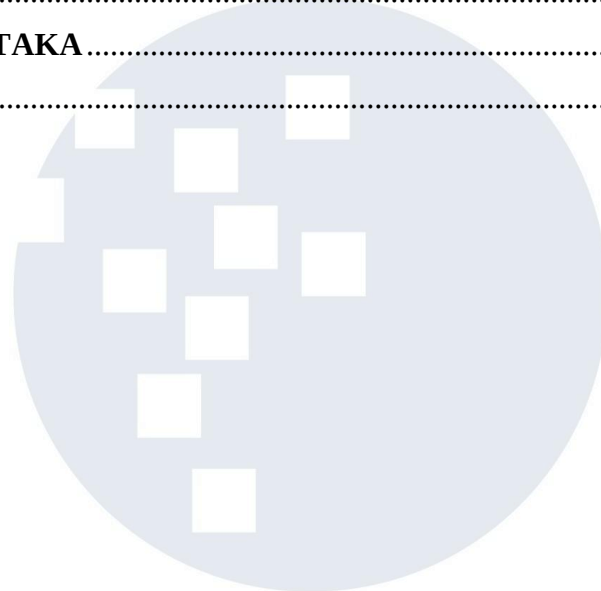
Keywords: K-Means, K-Medoids, Random Forest, RGEC, Rural Banks

DAFTAR ISI

HALAMAN PERNYATAAN TIDAK PLAGIAT	ii
KATA PENGANTAR.....	vi
ABSTRAK	viii
ABSTRACT (<i>English</i>).....	ix
DAFTAR ISI.....	x
DAFTAR TABEL.....	xiii
DAFTAR GAMBAR.....	xiv
DAFTAR RUMUS	xvi
DAFTAR LAMPIRAN	xvii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah.....	6
1.3 Batasan Masalah.....	6
1.4 Tujuan dan Manfaat Penelitian	7
1.4.1 Tujuan Penelitian	7
1.4.2 Manfaat Penelitian	7
1.5 Sistematika Penulisan	8
BAB II LANDASAN TEORI	10
2.1 Penelitian Terdahulu.....	10
2.2 Tinjauan Teori.....	19
2.2.1 Bank Perekonomian Rakyat	19
2.2.2 Penilaian Tingkat Kesehatan Bank	20
2.2.3 <i>Machine Learning</i>	25
2.3 Framework/Algoritma yang digunakan.....	25
2.3.1 CRISP-DM.....	25
2.3.2 Standardisasi Data	27
2.3.3 <i>K-means</i>	27
2.3.4 <i>K-medoids</i>	28
2.3.5 <i>Davies Bouldin Index</i>	29
2.3.6 <i>Silhouette Coefficient</i>	29

2.3.7	<i>Random Forest</i>	29
2.3.8	<i>Evaluasi Random Forest</i>	30
2.4	Tools/software yang digunakan	32
2.4.1	<i>Jupyter Notebook</i>	32
2.4.2	<i>Python</i>	32
2.4.3	<i>Streamlit</i>	33
BAB III METODOLOGI PENELITIAN		34
3.1	Gambaran Umum Objek Penelitian	34
3.2	Metode Penelitian	34
3.2.1	Perbandingan Framework	34
3.2.2	Alur Penelitian	36
3.3	Teknik Pengumpulan Data	40
3.4	Variabel Penelitian	40
3.4.1	Variabel Independen	40
3.4.2	Variabel Dependen	41
3.5	Teknik Analisis Data	41
BAB IV ANALISIS DAN HASIL PENELITIAN		43
4.1	<i>Business Understanding</i>	43
4.2	<i>Data Understanding</i>	45
4.2.1	<i>Inisialisasi Library</i>	45
4.2.2	<i>Read Dataset</i>	46
4.2.3	<i>Eksplorasi Struktur Awal Data</i>	46
4.2.4	<i>Pemeriksaan Data Duplikat</i>	49
4.3	<i>Data Preparation</i>	50
4.3.1	<i>Pembersihan Data (Data Cleaning)</i>	50
4.3.2	<i>Exploratory Data Analysis</i>	52
4.3.3	<i>Transformasi Data</i>	57
4.3.4	<i>Strandarisasi Data</i>	60
4.4	Modelling	62
4.4.1	Algoritma K-means	62
4.4.2	Algoritma K-Medoids	65
4.4.3	Algoritma Random Forest	72

4.5	<i>Evaluation</i>	74
4.6	<i>Deployment</i>	80
4.7	Hasil dan Pembahasan	84
BAB V SIMPULAN DAN SARAN		90
5.1	Simpulan	90
5.2	Saran	91
DAFTAR PUSTAKA		93
LAMPIRAN		99



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

DAFTAR TABEL

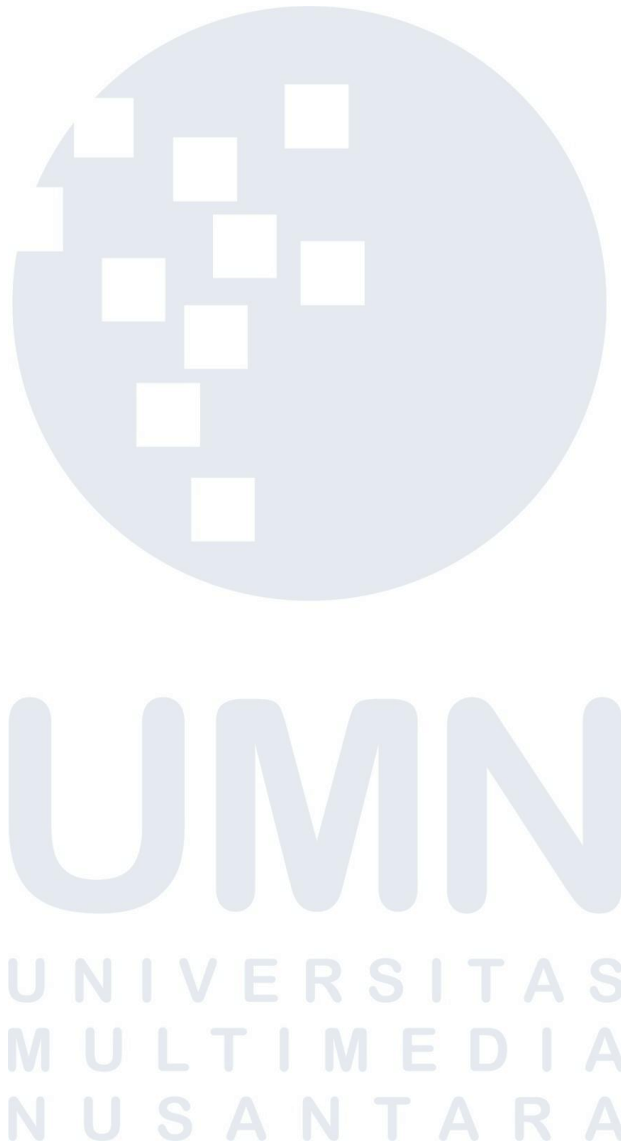
Tabel 2. 1 Penelitian Terdahulu	10
Tabel 2. 2 Matriks Kriteria Penetapan Peringkat LDR [35]	22
Tabel 2. 3 Matriks Kriteria Penetapan Peringkat NPL [35]	22
Tabel 2. 4 Matriks Kriteria Penetapan Peringkat GCG [35]	23
Tabel 2. 5 Matriks Kriteria Penetapan Peringkat BOPO [36]	23
Tabel 2. 6 Matriks Kriteria Penetapan Peringkat ROA [36]	24
Tabel 2. 7 Matriks Kriteria Penetapan Peringkat CAR [35]	24
Tabel 3. 7 Perbandingan Framework	35
Tabel 4. 1 Kegunaan Library Python	46
Tabel 4. 2 Deskripsi Kolom pada Dataset	47
Tabel 4. 3 Perbandingan Tipe Data sebelum dan sesudah Konversi	52
Tabel 4. 4 Hasil Evaluasi K-Means	64
Tabel 4. 5 Hasil Evaluasi K-Medoids	66
Tabel 4. 6 Rata-Rata Nilai Indikator RGEC per Cluster	70
Tabel 4. 7 Pemetaan Cluster terhadap Kategori Kesehatan BPR	70
Tabel 4. 8 Classification Report	75
Tabel 4. 9 Perbandingan Akurasi Data pada Model Random Forest	80



DAFTAR GAMBAR

Gambar 1. 1 Jumlah Bank Perekonomian Rakyat per Tahun	2
Gambar 1. 2 Perkembangan Total Aset BPR.....	3
Gambar 3. 1 Alur Penelitian.....	36
Gambar 4. 1 Inisialisasi Library	45
Gambar 4. 2 Read Dataset.....	46
Gambar 4. 3 Deskripsi Dataset.....	47
Gambar 4. 4 Visualisasi Missing Values Dataset	49
Gambar 4. 5 Pemeriksaan Data Duplikat	50
Gambar 4. 6 Penghapusan Kolom No.....	51
Gambar 4. 7 Code Proses Konversi Tipe Data	51
Gambar 4. 8 Deskripsi Statistik Data	53
Gambar 4. 9 Visualisasi Distribusi Rasio Keuangan BPR.....	54
Gambar 4. 10 Visualisasi Distribusi Rasio Keuangan BPR.....	55
Gambar 4. 11 Visualisasi Korelasi antar Rasio Keuangan BPR.....	56
Gambar 4. 12 Code Program Transformasi Rasio RGEC.....	58
Gambar 4. 13 Code Program Konversi ke Dataset	58
Gambar 4. 14 Dataset Hasil Transformasi Rasio ke Peringkat Kesehatan (PK 1–5)	59
Gambar 4. 15 Visualisasi Distribusi Nilai Setelah Transformasi	59
Gambar 4. 16 Proses Standardisasi Data.....	61
Gambar 4. 17 Visualisasi Distribusi Data Setelah Standardisasi	61
Gambar 4. 18 Metode Elbow untuk Menentukan Jumlah Cluster Optimal	63
Gambar 4. 19 Code Program Implementasi K-Means dan Perhitungan Evaluasi	64
Gambar 4. 20 Visualisasi Clustering K-Means (PCA 2D)	65
Gambar 4. 21 Kode Program Implementasi K-Medoids dan Perhitungan Evaluasi	66
Gambar 4. 22 Visualisasi Clustering K-Medoids (PCA 2D).....	67
Gambar 4. 23 Distribusi Nilai Indikator RGEC per Cluster	68
Gambar 4. 24 Code Program Export Dataset Hasil Clusterisasi.....	72
Gambar 4. 25 Import Dataset Klasifikasi.....	73
Gambar 4. 26 Proses pemisahan antara Fitur dan Target.....	73
Gambar 4. 27 Proses Label Encoding Kategori Kesehatan	74
Gambar 4. 28 Pembagian Data Train dan Test	74
Gambar 4. 29 Proses Standarisasi Fitur	74
Gambar 4. 30 Accuracy Random Forest Classifier.....	75
Gambar 4. 31 Confusion Matrix Random Forest.....	76
Gambar 4. 32 Kurva ROC Multi-Class.....	76
Gambar 4. 33 Features Importance	77
Gambar 4. 34 Proses Penyimpanan Model dan Objek Pendukung.....	78
Gambar 4. 35 Data Baru Testing.....	79
Gambar 4. 36 Akurasi Testing Data Baru.....	80

Gambar 4. 37 Tampilan Awal Aplikasi Sistem Prediksi Tingkat Kesehatan BPR82	
Gambar 4. 38 Tampilan Pratinjau Data Hasil Prediksi	82
Gambar 4. 39 Tampilan Visualisasi Distribusi Prediksi Kesehatan	83
Gambar 4. 40 Tampilan Fitur Filter, Pratinjau Tabel Terfilter, dan Opsi Unduh Hasil	84



DAFTAR RUMUS

Rumus 2. 1 LDR (<i>Loan to Deposit Rasio</i>).....	22
Rumus 2. 2 NPL (<i>Non-Performing Loan</i>).....	22
Rumus 2. 3 BOPO (<i>Operational Expense to Operating Income Rasio</i>)	23
Rumus 2. 4 ROA (<i>Return on Assets</i>)	24
Rumus 2. 5 CAR (<i>Capital Adequacy Rasio</i>).....	24
Rumus 2. 6 StandardScaler	27
Rumus 2. 7 Rata-rata (Mean).....	27
Rumus 2. 8 Standar Deviasi (Standard Deviation)	27
Rumus 2. 9 Akurasi (<i>Accuracy</i>).....	30
Rumus 2. 10 Presisi (<i>Precision</i>).....	31
Rumus 2. 11 Recall (Sensitivitas atau True Positive Rate).....	31
Rumus 2. 12 F1-Score.....	31



DAFTAR LAMPIRAN

Lampiran A Turnitin Similarity Report	99
Lampiran B Form Konsultasi Bimbingan.....	111
Lampiran C Formulir Penggunaan Perangkat Kecerdasan Artifisial (AI).....	112

