

DAFTAR ISI

HALAMAN PENGESAHAN PEMBIMBING	i
HALAMAN PENGESAHAN PENGUJI.....	ii
SURAT PERNYATAAN KARYA ASLI TUGAS AKHIR	iii
PERNYATAAN BEBAS PLAGIASI	iv
ABSTRAK	v
ABSTRACT	vi
KATA PENGANTAR	vii
DAFTAR TABEL	xi
DAFTAR GAMBAR.....	xiii
DAFTAR RUMUS	xiv
BAB I PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah	3
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian.....	3
1.6 Tahapan Penelitian.....	4
1.7 Sistematika Penulisan.....	5
BAB II TINJAUAN LITERATUR.....	6
2.1 Konsep Mobilitas Penumpang Transportasi Publik	6
2.2 Kriteria dan Perencanaan Rute Angkutan Umum	6
2.3 Data Tap In dan Tap Out TransJakarta	8
2.4 Data Mining.....	9
2.5 Clustering	10
2.5.1 K-Means	10
2.5.2 ST-DBSCAN	10
2.6 Gaussian Mixture Model	11
2.7 Estimasi Parameter Menggunakan Expectation-Maximization	11
2.8 Seleksi Model Menggunakan Bayesian Information Criterion	12
2.9 Struktur Matriks Kovarians dalam Gaussian Mixture Model	13
2.10 Association Rule Mining	14
2.11 Metrik Support, Confidence, dan Lift	14
2.12 Algoritma FP-Growth.....	16
2.13 Kombinasi Gaussian Mixture Model dan Association Rule Mining.....	16
2.14 State of the Art.....	17
BAB III METODOLOGI PENELITIAN	20
3.1 Kerangka Penelitian.....	20
3.2 Sumber dan Karakteristik Data	21
3.3 Data Preparation	22
3.3.1 Data Quality Assessment	24
3.3.2 Transformasi Variabel Waktu.....	24
3.3.3 Filter Durasi Perjalanan	25
3.3.4 Filter Jam Operasional.....	25
3.3.5 Feature Engineering.....	25
3.3.6 Imputasi Nilai Hilang	26
3.3.7 Pembersihan Data dan Outlier	26
3.3.8 Normalisasi Variabel.....	27

3.4 Pemodelan Gaussian Mixture Model (GMM).....	27
3.4.1 Feature Matrix (Data Numerik)	29
3.4.2 Spesifikasi Model GMM	29
3.4.3 Estimasi Parameter Expectation-Maximization (EM).....	30
3.4.4 Seleksi Model (BIC)	31
3.4.5 Penugasan Cluster Probabilistik	32
3.4.6 Evaluasi Kualitas Clustering.....	33
3.4.7 Interpretasi Cluster.....	33
3.5 Association Rule Mining (ARM)	34
3.5.1 Penentuan Struktur Transaksi	35
3.5.2 Transformasi ke Format Basket	35
3.5.3 Pembentukan Transactions Object.....	36
3.5.4 Analisis Frekuensi Item	36
3.5.5 Penerapan Algoritma FP-Growth	37
3.5.6 Evaluasi Association Rules.....	38
3.5.7 Analisis Pola Asosiasi per Cluster	39
3.6 Analisis Integratif	39
3.7 Perangkat dan Lingkungan Implementasi	42
3.8 Pengujian Grey Box	43
BAB IV HASIL DAN PEMBAHASAN	44
4.1 Hasil Data Preparation.....	44
4.1.1 Data Quality Assessment	44
4.1.2 Transformasi Variabel Waktu.....	47
4.1.3 Filter Durasi Perjalanan	47
4.1.4 Filter Jam Operasional.....	48
4.1.5 Feature Engineering.....	48
4.1.6 Imputasi Nilai Hilang	50
4.1.7 Pembersihan Data dan Outlier	50
4.1.8 Normalisasi Variabel.....	51
4.2 Hasil Pemodelan Gaussian Mixture Model (GMM)	52
4.2.1 Representasi Data untuk Pemodelan	53
4.2.2 Estimasi Parameter Expectation-Maximization (EM).....	53
4.2.3 Seleksi Model dan Penentuan Jumlah Cluster Optimal.....	53
4.2.4 Penugasan Cluster Probabilistik	55
4.2.5 Evaluasi Kualitas Clustering.....	55
4.2.6 Output Pemodelan dan Interpretasi Cluster.....	56
4.3 Perbandingan K-Means dan Gaussian Mixture Model	62
4.4 Hasil Association Rule Mining (ARM).....	64
4.4.1 Representasi Data Transaksi	65
4.4.2 Pembentukan Transactions Object.....	65
4.4.3 Analisis Frekuensi Item	66
4.4.4 Penerapan Algoritma FP-Growth	67
4.4.5 Evaluasi Association Rules.....	67
4.4.6 Analisis Pola Asosiasi Per Cluster	68
4.5 Hasil Pengujian Grey Box	69
4.6 Pembahasan	70
4.6.1 Profil Temporal Kluster	70
4.6.2 Perbandingan Dominasi Rute Antar Kluster.....	71
4.6.3 Visualisasi Spasial Keterhubungan Koridor	72

4.6.4 Implikasi dan Keterbatasan Temuan.....	74
BAB V PENUTUP	76
5.1 Kesimpulan.....	76
5.2 Saran.....	76
DAFTAR PUSTAKA	78



DAFTAR TABEL

Tabel 2.1 Parameterisasi Matriks Kovarians pada mclust1	30
Tabel 2.2 State of the Art	34
Tabel 3.1 Struktur Atribut Dataset	39
Tabel 3.2 Contoh data dari data raw	39
Tabel 3.3 Ilustrasi Kondisi Dataset Awal	41
Tabel 3.4 Contoh Transformasi Variabel Waktu Sebelum Transformasi	41
Tabel 3.5 Contoh Transformasi Variabel Waktu Sesudah Transformasi	41
Tabel 3.6 Contoh Filter Durasi Perjalanan Sebelum Filter	42
Tabel 3.7 Contoh Filter Durasi Perjalanan Sesudah Filter	42
Tabel 3.8 Contoh Filter Jam Operasional Sebelum Filter	42
Tabel 3.9 Contoh Filter Jam Operasional Sesudah Filter	42
Tabel 3.10 Contoh Variabel Hasil Rekayasa Fitur	42
Tabel 3.11 Contoh Proses Sebelum Imputasi	43
Tabel 3.12 Contoh Proses Sesudah Imputasi	43
Tabel 3.13 Ilustrasi data sebelum penanganan outlier	43
Tabel 3.14 Ilustrasi data sesudah penanganan outlier	43
Tabel 3.15 Contoh Sebelum Normalisasi Variabel	44
Tabel 3.16 Contoh Sesudah Normalisasi Variabel	44
Tabel 3.17 Contoh Hasil Data Preparation	45
Tabel 3.18 Variabel Pemodelan GMM	46
Tabel 3.19 Spesifikasi Eksperimen GMM	46
Tabel 3.20 Ilustrasi Evaluasi Model Berbasis Bayesian Information Criterion	48
Tabel 3.21 Perubahan Nilai BIC (Delta BIC)	48
Tabel 3.22 Ilustrasi Probabilitas Keanggotaan Cluster	49
Tabel 3.23 Ilustrasi Evaluasi Kualitas Clustering	50
Tabel 3.24 Ilustrasi Ringkasan Karakteristik Cluster	50
Tabel 3.25 Ilustrasi Struktur Data Transaksi	52
Tabel 3.26 Ilustrasi Struktur Data Basket	52
Tabel 3.27 Ilustrasi Representasi Transactions Object	53
Tabel 3.28 Frekuensi Kemunculan Data pada Dataset	54
Tabel 3.29 Ilustrasi Data Transaksi	54
Tabel 3.30 Ilustrasi Frequent Itemsets	55
Tabel 3.31 Ilustrasi Evaluasi Association Rules	55
Tabel 3.32 Ilustrasi Pola Dominan Per Cluster	56
Tabel 3.33 Rancangan Pengujian Grey Box	60
Tabel 4.1 Sepuluh Atribut Utama	62
Tabel 4.2 Perubahan Jumlah Data pada Setiap Tahap Data Preparation	62
Tabel 4.3 Variabel untuk Pemodelan Gaussian Mixture Model	70
Tabel 4.4 Variabel untuk Pemodelan Gaussian Mixture Model	71

Tabel 4.5 Ringkasan Representasi Data Transaksi untuk ARM	83
Tabel 4.6 Ringkasan Evaluasi Association Rule Mining.....	85
Tabel 4.7 Hasil Pengujian Grey Box	88



DAFTAR GAMBAR

Gambar 2.1 Titik Pemberhentian Layanan 1R	24
Gambar 2.2 Proses tap in di dalam bus	25
Gambar 2.3 Proses tap in di halte BRT	25
Gambar 2.4 Model Kovarians Gaussian Mixture Model	30
Gambar 3.1 Alur Penelitian	37
Gambar 3.2 Diagram Alur Data Preparation	40
Gambar 3.3 Diagram Alur Pemodelan Gaussian Mixture Model (GMM).....	45
Gambar 3.4 Proses Iterasi Algoritma Expectation-Maximization.....	47
Gambar 3.5 Diagram Alur Association Rule Mining	51
Gambar 3.6 Wireframe Halaman Peta dan Koridor	57
Gambar 3.7 Wireframe Halaman Association Rules	58
Gambar 3.8 Wireframe Halaman GMM Clustering	58
Gambar 4.1 Top 10 missing value pada dataset awal.....	62
Gambar 4.2 Top 10 koridor dengan frekuensi perjalanan tertinggi.....	63
Gambar 4.3 Top 10 halte asal dengan frekuensi tap-in tertinggi.....	64
Gambar 4.4 Top 10 halte tujuan dengan frekuensi tap-out tertinggi.....	65
Gambar 4.5 Distribusi jam tap-in pada hari kerja dan akhir pekan.....	65
Gambar 4.6 Hasil filter jam operasional.....	66
Gambar 4.7 Distribusi intensitas perjalanan per pengguna per hari.....	67
Gambar 4.8 Hasil feature engineering pada dimensi hari	67
Gambar 4.9 Hasil imputasi nilai hilang pada atribut koridor	68
Gambar 4.10 Funnel data cleaning	69
Gambar 4.11 Normalisasi z-score pada variabel kontinu.....	69
Gambar 4.12 Heatmap korelasi antar fitur pada dataset final	70
Gambar 4.13 Kurva BIC terhadap jumlah cluster	71
Gambar 4.14 Perubahan BIC antar jumlah cluster	71
Gambar 4.15 Contoh probabilitas posterior hasil penugasan cluster	73
Gambar 4.16 Silhouette score per jumlah cluster.....	73
Gambar 4.17 Profil detail cluster hasil Gaussian Mixture Model	74
Gambar 4.18 Distribusi ukuran cluster hasil GMM	74
Gambar 4.19 Heatmap karakteristik fitur per cluster	76
Gambar 4.20 Pola jam tap-in per cluster	77
Gambar 4.21 Profil weekend dan commuter per cluster	78
Gambar 4.22 Proporsi Pengaruh Relatif Variabel pada Hasil GMM	78
Gambar 4.23 Perbandingan nilai BIC eksperimen GMM dua variabel	80
Gambar 4.24 Elbow plot K-Means.....	81
Gambar 4.25 Scatter plot K-Means	81
Gambar 4.26 Scatter plot Gaussian Mixture Model.....	82
Gambar 4.27 Frekuensi item dominan per cluster.....	84

Gambar 4.28 Jumlah aturan terpilih pada setiap cluster.....	85
Gambar 4.29 Top origin-destination patterns per cluster	87
Gambar 4.30 Profil temporal klaster berdasarkan kategori jam.....	89
Gambar 4.31 Heatmap dominansi rute pada setiap cluster	90
Gambar 4.32 Peta keterhubungan koridor pada aturan asosiasi.....	90
Gambar 4.33 Tampilan Ringkasan Association Rules per Cluster	91
Gambar 4.34 Tampilan Detail Rule Association Rules	91
Gambar 4.35 Tampilan Distribusi Cluster GMM	91



DAFTAR RUMUS

Rumus 2.1 Bentuk aturan asosiasi	27
Rumus 2.2 Support	28
Rumus 2.3 Confidence	28
Rumus 2.4 Lift	28
Rumus 3.2 Log-likelihood model Gaussian Mixture Model	43
Rumus 3.3 Bayesian Information Criterion	44
Rumus 3.4 Probabilitas keanggotaan cluster	44

