

DAFTAR ISI

HALAMAN PENGESAHAN.....	iii
PERNYATAAN BEBAS PLAGIASI.....	iv
HALAMAN MOTO DAN PERSEMBAHAN	v
KATA PENGANTAR.....	vi
DAFTAR ISI	ix
DAFTAR GAMBAR	xii
DAFTAR TABEL.....	xiii
DAFTAR LAMPIRAN	xv
INTISARI.....	xvi
ABSTRACT.....	xvii
BAB I PENDAHULUAN	1
1.1. Latar Belakang dan Permasalahan	1
1.2. Pembatasan Masalah	4
1.3. Tujuan Penelitian.....	4
1.4. Tinjauan Pustaka	5
1.5. Metode Penelitian.....	8
1.6. Sistematika Penulisan	8
BAB II LANDASAN TEORI	10
2.1. <i>Data Mining</i>	10
2.2. <i>Machine Learning</i>	10
2.3. Vektor	11
2.4. Matriks	13
2.4.1. Pengertian Matriks	13
2.4.2. Operasi Matriks	14
2.4.3. Jenis-Jenis Matriks	16
2.5. Variabel Random.....	18
2.5.1. Pengertian Variabel Random.....	18
2.5.2. Ekspektasi dan Variansi.....	21
2.6. Deteksi <i>Outlier</i>	23
2.6.1. <i>Z-score</i>	23
2.6.2. <i>Isolation Forest</i>	24
2.7. Jarak <i>Euclidean</i>	24

2.8.	Analisis Statistika Multivariat.....	25
2.8.1.	Tujuan Analisis Statistika Multivariat.....	25
2.8.2.	Matriks Data Multivariat.....	26
2.9.	Analisis Kluster	27
2.9.1.	Pengertian Analisis Kluster	27
2.9.2.	Metode Dasar Analisis Kluster.....	28
2.9.3.	Standarisasi Data.....	31
2.10.	Asumsi Analisis Kluster	32
2.10.1.	Sampel Data yang Representatif.....	32
2.10.2.	Tidak ada Multikolinearitas Antarvariabel.....	32
BAB III PENERAPAN METODE ISOLATION FOREST PADA DETEKSI DAN PENANGANAN OUTLIER DALAM ANALISIS KLAS TER MENGGUNAKAN ALGORITMA K-MEANS, K-MEANS++, DAN K-MEDOIDS.....		34
3.1.	Deteksi <i>Outlier</i>	34
3.1.1.	<i>Z-score</i>	35
3.1.2.	<i>Isolation Forest</i>	36
3.2.	Analisis Kluster <i>K-Means</i>	45
3.3.	Analisis Kluster <i>K-Means++</i>	51
3.4.	Analisis Kluster <i>K-Medoids</i>	53
3.5.	Penentuan Jumlah Kluster Optimal.....	61
3.5.1.	Metode Elbow	61
3.5.2.	Metode <i>Silhouette Coefficient</i>	62
3.5.3.	Metode <i>Calinski-Harabasz Index</i>	64
3.5.4.	Metode <i>Davies Bouldin Index</i>	65
3.6.	Metrik Evaluasi Kluster <i>Dunn Index</i>	67
BAB IV STUDI KASUS		69
4.1.	Deskripsi Data.....	69
4.2.	Pengujian Asumsi Analisis Kluster	70
4.2.1.	Sampel Data yang Representatif.....	70
4.2.2.	Tidak Ada Multikolinearitas Antarvariabel.....	70
4.3.	Standarisasi Data.....	71
4.4.	Deteksi <i>Outlier</i> Awal Menggunakan <i>Z-score</i>	71
4.5.	Analisis Kluster Tanpa <i>Isolation Forest</i>	72
4.5.1.	Penentuan Jumlah Kluster Optimal Algoritma <i>K-Means</i>	72

4.5.2.	Pengelompokan Data Algoritma <i>K-Means</i>	74
4.5.3.	Penentuan Jumlah Kluster Optimal Algoritma <i>K-Means++</i>	75
4.5.4.	Pengelompokan Data Algoritma <i>K-Means++</i>	77
4.5.5.	Penentuan Jumlah Kluster Optimal Algoritma <i>K-Medoids</i>	78
4.5.6.	Pengelompokan Data Algoritma <i>K-Medoids</i>	80
4.6.	Deteksi <i>Outlier</i> Menggunakan <i>Isolation Forest</i>	81
4.7.	Analisis Kluster Setelah Melalui <i>Isolation Forest</i>	84
4.7.1.	Penentuan Jumlah Kluster Optimal Algoritma <i>K-Means</i>	84
4.7.2.	Pengelompokan Data Algoritma <i>K-Means</i>	86
4.7.3.	Penentuan Jumlah Kluster Optimal Algoritma <i>K-Means++</i>	88
4.7.4.	Pengelompokan Data Algoritma <i>K-Means++</i>	90
4.7.5.	Penentuan Jumlah Kluster Optimal Algoritma <i>K-Medoids</i>	92
4.7.6.	Pengelompokan Data Algoritma <i>K-Medoids</i>	94
4.8.	Perbandingan Hasil Kluster.....	95
4.9.	Profilisasi Kluster Final.....	98
4.10.	Meninjau Sebaran Tenaga Kesehatan di Indonesia.....	103
BAB V	PENUTUP.....	105
5.1.	Kesimpulan	105
5.2.	Saran.....	106
DAFTAR	PUSTAKA	107
LAMPIRAN		110