

DAFTAR ISI

HALAMAN PENGESAHAN.....	ii
PERNYATAAN BEBAS PLAGIASI.....	iii
PRAKATA.....	iv
DAFTAR ISI.....	vi
DAFTAR GAMBAR.....	iv
DAFTAR TABEL.....	v
INTISARI.....	vi
ABSTRACT.....	vii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang Masalah.....	1
1.2 Rumusan Masalah.....	3
1.3 Batasan Masalah.....	3
1.4 Tujuan.....	4
1.5 Manfaat.....	4
1.6 Kebaruan Penelitian (Novelty).....	5
BAB II TINJAUAN PUSTAKA.....	6
BAB III LANDASAN TEORI.....	15
3.1 <i>Machine Learning</i> dan Peran Fitur dalam Klasifikasi.....	15
3.2 Seleksi Fitur.....	15
3.3 Kategori Metode Seleksi Fitur.....	16
3.3.1 <i>Filter Method</i>	16
3.3.2 <i>Wrapper Method</i>	16
3.3.3 <i>Embedded Method</i>	17
3.4 <i>Recursive Feature Elimination</i> (RFE).....	17
3.4.1 Mekanisme dan Tahapan Iterasi RFE.....	18
3.4.2 <i>Random Forest Classifier</i> sebagai Estimator RFE.....	19
3.5 Logistic Regression.....	22
3.6 <i>Support Vector Machine</i> (SVM).....	23

3.7	Metrik Evaluasi	24
3.8	Proses Rekrutmen Karyawan	25
BAB IV METODOLOGI PENELITIAN		27
4.1	Deskripsi Penelitian	27
4.2	Alur Kerja Penelitian.....	29
4.3	Alat dan Bahan.....	31
4.4	Arsitektur dan Perancangan Sistem Prediksi	31
4.4.1	Pengumpulan Dataset.....	31
4.4.2	Perancangan Pembagian Dataset (<i>Three-Way Split</i>).....	32
4.4.3	Perancangan <i>Fitur Engineering</i> menggunakan <i>Ordinal Encoding</i> dan <i>Multi-Hot Encoding</i>	33
4.4.4	Perancangan <i>K-Fold Cross Validation</i>	34
4.4.5	Perancangan Seleksi Fitur Menggunakan <i>Recursive Feature Elimination (RFE)</i>	35
4.4.6	Parameter dan Konfigurasi Eksperimen RFE	38
4.4.7	Protokol Pelatihan dan Evaluasi Model	39
4.5	Desain Skenario Eksperimen	41
4.5.1	Skenario 1 Baseline (Klasifikasi Tanpa Seleksi Fitur)	42
4.5.2	Skenario 2 Baseline dengan Validasi Silang (<i>K-Fold Cross Validation</i>) 43	
4.5.3	Skenario 3 RFE Global (Seleksi Fitur Berbasis <i>Gini Importance</i> Agregat) 44	
4.5.4	Skenario 4 RFE Per Posisi (Pendekatan Spesifik <i>One-vs-Rest</i>).....	45
BAB V IMPLEMENTASI.....		47
5.1	Persiapan Data.....	47
5.2	Implementasi Feature Encoding.....	49
5.2.1	<i>Ordinal Encoding</i> untuk Data Lama Bekerja dan Pendidikan.....	49
5.2.2	<i>Multi-Hot Encoding</i> untuk Data Keahlian	52
5.3	<i>K-Fold Cross Validation</i>	53
5.4	Seleksi Fitur <i>Recursive Feature Elimination (RFE)</i> Global	55

5.5	RFE Seleksi 10 Fitur Terbaik per Posisi (<i>One-vs-Rest</i>).....	57
5.5.1	Implementasi <i>One-vs-Rest (OVR)</i>	58
5.5.2	Penggabungan Fitur (<i>Union</i>).....	58
5.5.3	Transformasi Data Akhir.....	60
5.6	Pelatihan Model Final dan Evaluasi.....	60
5.7	Fungsi Prediksi Untuk Input Baru.....	62
BAB VI HASIL DAN PEMBAHASAN		68
6.1	Hasil Pre-processing dan Pembentukan Ruang Fitur.....	68
6.1.1	<i>Ordinal Encoding</i> untuk Data Lama Bekerja dan Pendidikan.....	68
6.1.2	<i>Multi-Hot Encoding</i> untuk Data Keahlian	68
6.2	Evaluasi dan Analisis Kinerja Model Tanpa K-Fold (<i>Validasi Hold-Out</i>)	69
6.2.1	Hasil Evaluasi Model <i>Logistic Regression</i>	69
6.2.2	Hasil Evaluasi Model <i>Support Vector Machine</i>	69
6.2.3	Hasil Evaluasi Model <i>Random Forest</i>	70
6.2.4	Analisis Kinerja Komparatif	71
6.2.5	Analisis Prediksi Kandidat Baru Validasi Hold Out.....	72
6.3	Evaluasi dan Analisis Kinerja Model dengan Validasi Silang K-Fold	74
(k=5)		
6.3.1	Hasil <i>K-Fold Cross Validation</i>	74
6.3.2	Hasil Evaluasi Model <i>Logistic Regression</i>	75
6.3.3	Hasil Evaluasi Model <i>Support Vector Machine</i>	76
6.2.6	Hasil Evaluasi Model <i>Random Forest</i>	77
6.3.4	Perbandingan Kinerja Akhir dan Generalisasi Model	77
6.3.5	Analisis Kelemahan Kelas <i>Full Stack Developer</i> (Fokus pada Data	
Ambiguitas)		78
6.3.6	Pembahasan dan Analisis Komparatif <i>K-Fold vs Hold-Out</i>	80
6.3.7	Analisis Hasil Prediksi <i>K-Fold vs Hold-Out</i>	81
6.4	Seleksi Fitur Menggunakan <i>Recursive Feature Elimination (RFE)</i>	82
6.4.1	Hasil Seleksi Fitur RFE (50 Fitur Terbaik).....	82

6.4.2	Tingkat Kepentingan Fitur Terpilih	84
6.4.3	Analisis Fitur-Fitur Paling Berpengaruh	84
6.4.4	Hasil Evaluasi Model Logistic Regression	86
6.4.5	Hasil Evaluasi Model <i>Support Vector Machine</i> (SVM)	88
6.4.6	Hasil Evaluasi Model <i>Random Forest</i>	89
6.4.7	Pembahasan dan Analisis Komparatif Kinerja Klasifikasi (Hold-Out/K-Fold vs. RFE 50 Fitur)	91
6.4.8	Analisis Prediksi Kandidat Baru (RFE Global 50 Fitur)	92
6.5	Seleksi Fitur RFE Per Posisi (10 Fitur Terbaik Per Kelas)	94
6.5.1	Analisis Seleksi Fitur OVR Untuk Memilih 10 Fitur Terbaik	94
6.5.2	Analisis Fitur Gabungan (<i>Union</i>) dan Efek <i>Sparsity</i>	95
6.5.3	Hasil Evaluasi Model Logistic Regression	95
6.5.4	Hasil Evaluasi Model <i>Support Vector Machine</i>	97
6.5.5	Hasil Evaluasi Model <i>Random Forest</i>	98
6.5.6	Analisis Perbandingan Kinerja Klasifikasi dan Dampak <i>Sparsity</i> ..	100
6.5.7	Ambiguitas Kelas <i>Full Stack Developer</i>	101
6.5.8	Analisis Prediksi Kandidat Baru (RFE 10 Fitur Per Posisi)	102
BAB VII KESIMPULAN DAN SARAN		104
7.1	Kesimpulan	104
7.2	Saran	105