

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PENGESAHAN	ii
HALAMAN PERNYATAAN	iii
HALAMAN PERSEMBAHAN	iv
HALAMAN MOTTO	v
PRAKATA	vi
DAFTAR ISI	ix
DAFTAR TABEL	xiii
DAFTAR GAMBAR	xiv
INTISARI	xvi
ABSTRACT	xvii
I PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Pembatasan Masalah	4
1.3. Tujuan dan Manfaat Penelitian	5
1.4. Tinjauan Pustaka	6
1.5. Metodologi Penelitian	9
1.6. Sistematika Penulisan	11
II LANDASAN TEORI	13
2.1. Aljabar Matriks	13
2.1.1. Ruang Vektor	13
2.1.2. Matriks	18
2.1.3. Determinan	23
2.1.4. Nilai Eigen dan Vektor Eigen	26
2.1.5. Matriks Simetris dan Positif Semidefinit	28
2.1.6. Norma dan Fungsi Jarak	30
2.1.7. Fungsi Kernel dan Kernel RBF	33
2.2. <i>Machine Learning</i>	36
2.2.1. <i>Supervised Learning</i> dan Klasifikasi	36
2.2.2. <i>Ensemble Learning</i> dan AdaBoost	37
2.2.3. Varian SAMME.R dan Pendekatan Probabilistik AdaBoost	39
2.2.4. Ketidakseimbangan Kelas	42
2.3. Optimasi	45

2.3.1.	Formulasi Masalah Optimasi	45
2.3.2.	Diferensiabilitas dan Keterbatasan Optimasi Berbasis Gradien	47
2.4.	Jaringan Saraf Tiruan	49
2.4.1.	Neuron dan Fungsi Aktivasi	50
2.4.2.	Layer dan <i>Forward Pass</i>	52
2.4.3.	Pelatihan Jaringan Saraf Tiruan dan <i>Backpropagation</i>	53
2.5.	<i>Explainable Artificial Intelligence</i>	55
2.5.1.	Taksonomi XAI	55
2.5.2.	<i>Counterfactual Explanation</i>	57
2.5.3.	Formulasi Optimasi CFE	60
2.5.4.	Properti Kualitas CFE	61
2.6.	Atrisi Karyawan	64
III	Integrasi CFProto berbasis Algoritma Genetika dan DPP	66
3.1.	Model Klasifikasi Atrisi Karyawan	66
3.1.1.	Prapemrosesan Data	66
3.1.2.	Kandidat Model dan Skema Validasi	68
3.1.3.	Ruang <i>Hyperparameter</i> dan Tuning	69
3.1.4.	Pencarian Ambang Batas Optimal	69
3.1.5.	Metrik Evaluasi Model Klasifikasi	69
3.2.	<i>Counterfactual Explanations Guided by Prototypes</i> (CFProto) . . .	71
3.2.1.	Fungsi <i>Loss</i> CFProto	72
3.2.2.	<i>Autoencoder</i> dan Plausibilitas	72
3.2.3.	Prototipe Laten	74
3.2.4.	Keterbatasan CFProto	75
3.3.	Algoritma Genetika	76
3.4.	<i>Determinantal Point Process</i>	78
3.4.1.	Definisi DPP dan <i>Kernel Matrix</i>	79
3.4.2.	Kernel <i>L-Ensemble</i>	79
3.4.3.	Seleksi Subset dengan <i>Greedy MAP Inference</i>	80
3.5.	Integrasi CFProto berbasis Algoritma Genetika dan DPP	82
3.5.1.	Arsitektur Pelatihan AE dan Prototipe Laten	83
3.5.2.	Pemilihan Instance untuk Pembangkitan CFE	84
3.5.3.	Fungsi <i>Fitness</i>	84
3.5.4.	Representasi dan Partisi Fitur	86
3.5.5.	Adaptasi Operator GA	87
3.5.6.	Konfigurasi Parameter	90

3.6. Metrik Evaluasi CFE	91
3.7. Alur Eksperimen	93
IV STUDI KASUS	96
4.1. Data Penelitian	96
4.2. Analisis Deskriptif dan Eksplorasi Data	99
4.3. <i>Prapemrosesan Data</i>	105
4.3.1. Pembagian Data Latih & Data Uji	105
4.3.2. Penanganan Pencilan (<i>Outlier</i>)	106
4.3.3. Encoding Variabel Kategorik	107
4.3.4. <i>Min-Max Scaling</i>	107
4.3.5. Penanganan Data Tidak Seimbang	107
4.4. Pemodelan Klasifikasi	108
4.4.1. Pemilihan Model Terbaik (<i>Model Selection</i>)	109
4.4.2. Optimasi Hyperparameter (<i>Hyperparameter Tuning</i>)	111
4.4.3. Penyesuaian Ambang Batas (<i>Threshold</i>)	112
4.4.4. Evaluasi Performa Model Terbaik	113
4.5. <i>Autoencoder</i> dan Prototipe Laten	115
4.5.1. Pelatihan <i>Autoencoder</i>	116
4.5.2. Evaluasi Kualitas Rekonstruksi	117
4.5.3. Konstruksi Prototipe Laten	118
4.6. Pembangkitan <i>Counterfactual Explanation</i> (CFE)	118
4.6.1. Partisi Fitur	119
4.6.2. Normalisasi Komponen Fungsi <i>Fitness</i>	120
4.7. Hasil <i>Counterfactual Explanation</i> Tanpa DPP	121
4.8. Hasil <i>Counterfactual Explanation</i> dengan DPP	125
4.9. Analisis Sensitivitas Bobot λ_2 , λ_3 , dan λ_4	129
4.10. Pembahasan	131
4.10.1. Pola Agregat Rekomendasi CFE	131
4.10.2. Contoh Pembahasan Penjelasan Kontrafaktual	132
V PENUTUP	138
5.1. Kesimpulan	138
5.2. Saran	139
DAFTAR PUSTAKA	141
A Hasil Eksperimen dan Tuning Parameter	145
B Kode Exploratory Data Analysis (EDA)	146
2.1. Instalasi dan Impor Library	146

2.2.	Pengambilan dan Prapemrosesan Data	147
2.3.	Analisis Deskriptif	148
2.4.	Visualisasi Data (EDA)	149
2.4.1.	Attrition berdasarkan Education Level	149
2.4.2.	Distribusi Monthly Income	151
2.4.3.	Attrition berdasarkan Job Role	151
2.4.4.	Environment Satisfaction per Job Role	153
2.4.5.	Dampak Business Travel terhadap Attrition	154
C	Kode Pemodelan Klasifikasi	157
3.1.	Instalasi dan Impor Library	157
3.2.	Pengambilan dan Prapemrosesan Data Lanjutan	158
3.3.	Pemilihan Model (Cross-Validation)	160
3.4.	Hyperparameter Tuning	163
3.5.	Optimasi Threshold (Threshold Moving)	164
3.6.	Evaluasi Model Final	165
3.7.	Penyimpanan Model	167
D	Kode Pembangkitan Penjelasan Kontrafaktual	169
4.1.	Persiapan dan Pelatihan Autoencoder	169
4.1.1.	Pengaturan Awal dan Prapemrosesan	169
4.1.2.	Membangun dan Melatih Autoencoder	172
4.1.3.	Penyimpanan Model Laten	173
4.2.	Fungsi Inti Algoritma Genetika dan DPP	173
4.2.1.	Model Wrapper dan Feature Helper	173
4.2.2.	Kalkulasi Jarak dan Fungsi Kebugaran (Fitness Function)	176
4.2.3.	Mutasi, Crossover, dan Seleksi Turnamen	178
4.2.4.	Implementasi Determinantal Point Process (DPP)	179
4.3.	Siklus Utama GA, Pengujian, dan Evaluasi Kualitatif	181
4.3.1.	Siklus Evolusi GA dan Seleksi Akhir DPP	181
4.3.2.	Eksekusi Pengujian Data	184
4.3.3.	Fungsi Evaluasi Metrik CFE	184
4.3.4.	Ekstraksi Analisis Kualitatif	186