

DAFTAR ISI

halaman

HALAMAN JUDUL	
FORMULIR PERNYATAAN KEASLIAN DAN PERSETUJUAN UNGGAH	
TUGAS AKHIR	
PERSETUJUAN DOSEN PEMBIMBING SKRIPSI	
PERSETUJUAN TIM PENGUJI SKRIPSI	
ABSTRAK	vi
<i>ABSTRACT</i>	vii
KATA PENGANTAR.....	viii
DAFTAR ISI	x
DAFTAR GAMBAR	xiii
DAFTAR TABEL.....	xv
DAFTAR LAMPIRAN	xvi
 BAB I PENDAHULUAN	
1.1. Latar Belakang.....	1
1.2. Maksud dan Tujuan	4
1.3. Batasan Masalah	5
1.4. Metode Penelitian	5
1.5. Sistematika Penulisan	6
 BAB II LANDASAN TEORI	
2.1. Dasar Teori Tentang Kearsipan	8
2.1.1. Arsip Nasional Republik Indonesia (ANRI).....	8
2.1.2. Arsiparis.....	9
2.1.3. Jabatan Fungsional Arsiparis (JFA)	9
2.1.4. Sertifikasi Kompetensi Kearsipan	10
2.1.5. Sertifikasi Jabatan Fungsional Arsiparis (JFA).....	11
2.2. Dasar Teori Teknis Sains Data.....	12
2.2.1. <i>K-Means Clustering</i>	12
2.2.2. <i>Python</i>	15
2.2.3. <i>Jupyter Notebook</i>	15

2.2.4. <i>Anaconda</i>	16
2.2.5. <i>Machine Learning</i>	17
2.2.6. <i>Exploratory Data Analysis (EDA)</i>	18
2.2.7. <i>Dataset</i>	18
2.2.8. <i>Data Science</i>	19
2.2.9. <i>Scikit-learn</i>	20
2.2.10. <i>Predictive Analytics</i>	20
2.2.11. <i>Pandas</i>	21
2.2.12. <i>NumPy</i>	21
2.2.13. <i>Random Forest Regression</i>	22

BAB III METODE PENELITIAN

3.1. Deskripsi Sistem	23
3.1.1. <i>Datasets</i>	23
3.1.2. Perangkat	25
3.2. Alir Kerja.....	26
3.2.1. Akuisisi Data	26
3.2.2. Pembersihan Data	27
3.2.3. Penggabungan Data	27
3.2.4. <i>Exploratory Data Analysis</i>	28
3.2.5. Pengujian Data.....	28
3.2.6. Interpretasi Hasil.....	29
3.2.7. Evaluasi Hasil	29

BAB IV PELAKSANAAN PENELITIAN

4.1. Akuisisi Data	31
4.1.1. Pengumpulan Data.....	31
4.1.2. Rekontruksi Data	32
4.1.3. <i>Import Data</i>	33
4.2. <i>Exploratory Data Analysis</i>	34
4.2.1. Informasi <i>Dataset</i>	34
4.2.2. Pembersihan Data	36
4.2.3. Penggabungan Data	38
4.2.4. <i>Overview Dataprep Report</i>	40
4.2.5. <i>Preprocessing Data</i>	50

4.3.	<i>K-Means Clustering</i> dan Regresi <i>Random Forest</i>	63
4.3.1.	Pemodelan Algoritma <i>K-Means</i>	64
4.3.2.	<i>K-Means Clustering</i> dan Visualisasi	65
4.3.3.	Interpretasi Hasil dengan <i>Random Forest</i>	68
 BAB V ANALISIS DAN HASIL PENELITIAN		
5.1.	Analisis Penelitian	73
5.1.1.	Akuisisi Data	73
5.1.2.	<i>Exploratory Data Analysis</i>	74
5.1.3.	<i>K-Means Clustering</i>	75
5.1.4.	Regresi <i>Random Forest</i>	76
5.2.	Hasil Penelitian	78
5.2.1.	Metode <i>K-Means Clustering</i>	78
5.2.2.	Metode Regresi <i>Random Forest</i>	82
 BAB VI PENUTUP		
6.1.	Kesimpulan	83
6.2.	Saran	84
 DAFTAR PUSTAKA		
 DAFTAR SINGKATAN		
 LAMPIRAN		

DAFTAR GAMBAR

	halaman
Gambar 1.1 Grafik Data Kelulusan Sertifikasi JFA Tahun 2009-2023	2
Gambar 2.1 Logo <i>Jupyter</i>	16
Gambar 2.2 Logo <i>Anaconda Navigator</i>	17
Gambar 3.1 Diagram Alir Kerja	26
Gambar 4.1 <i>Database</i> Hasil Sertifikasi JFA	32
Gambar 4.2 <i>Path Dataset</i> di <i>Jupyter Notebook</i>	33
Gambar 4.3 <i>Import Dataset</i>	33
Gambar 4.4 Pembacaan <i>Dataset</i>	34
Gambar 4.5 Mencetak informasi dari setiap <i>dataset</i>	34
Gambar 4.6 Hasil cetak informasi dari <i>dataset</i> pada <i>data_1</i>	35
Gambar 4.7 Hasil cetak informasi dari <i>dataset</i> pada <i>data_5</i>	36
Gambar 4.8 Pembersihan <i>Dataset</i> Tahap 1	37
Gambar 4.9 Kode program penambahan nama kolom di <i>dataset</i>	38
Gambar 4.10 Kode program <i>merge dataset</i>	39
Gambar 4.11 Kode program menghilangkan <i>subset</i> "Nama"	39
Gambar 4.12 Hasil penggabungan <i>dataset</i>	40
Gambar 4.13 Kode program menjalankan <i>Overview Dataprep Report</i>	41
Gambar 4.14 <i>Overview Dataprep Report</i>	41
Gambar 4.15 <i>Bar chart</i> dari <i>missing values</i>	49
Gambar 4.16 Mencetak informasi data gabungan.....	50
Gambar 4.17 Kode program <i>replace</i> dan <i>convert</i> pada kolom Nilai	51
Gambar 4.18 Kode program menghapus kolom dan nilai kolom yang kosong....	51
Gambar 4.19 Kode program mencetak nilai unik pada kolom	52
Gambar 4.20 Kode program mencetak rangkuman nilai unik	55
Gambar 4.21 Hasil pembersihan <i>noise</i> nilai unik kolom "Jabatan"	56
Gambar 4.22 Kode program proses <i>encoding</i> variabel Jenis Sertifikasi	57
Gambar 4.23 Kode program proses <i>encoding</i> variabel Jabatan	58

Gambar 4.24 Kode program melihat hasil <i>encoding</i> variabel Jabatan.....	58
Gambar 4.25 Kode program untuk melihat hasil data gabungan.....	59
Gambar 4.26 Kode Program melihat nilai unik Pangkat, Angkatan, Jabatan	60
Gambar 4.27 Kode program proses <i>encoding</i> nilai Jabatan.....	61
Gambar 4.28 Kode program proses <i>encoding</i> nilai Pangkat.....	61
Gambar 4.29 Kode program proses <i>encoding</i> nilai Angkatan	61
Gambar 4.30 Kode program hasil data gabungan setelah <i>encoding</i>	62
Gambar 4.31 <i>Heatmap</i> data gabungan variabel Pangkat, Jabatan, dan Angkatan	63
Gambar 4.32 Kode program untuk mencetak isi dari <i>data_selected</i>	64
Gambar 4.33 Kode program untuk meng- <i>install library</i>	65
Gambar 4.34 Kode program menjalankan algoritma <i>k-means</i> dan <i>knee locator</i> ..	65
Gambar 4.35 Kode program untuk melakukan <i>clustering</i> dengan <i>k-means</i>	66
Gambar 4.36 Grafik metode <i>Elbow</i> untuk Optimal Kluster	67
Gambar 4.37 Visualisasi <i>clustering</i> menggunakan <i>k-means</i>	68
Gambar 4.38 Kode program mencari nilai akurasi <i>Random Forest</i>	70
Gambar 4.39 Kode program untuk mencari analisis <i>feature importance</i>	71
Gambar 4.40 Kode program mencetak <i>Confusion Matrix</i>	71
Gambar 4.41 Tampilan dari <i>Confusion Matrix</i>	72
Gambar 5.1 Grafik visualisasi <i>clustering</i> dengan <i>k-means</i>	78
Gambar 5.2 <i>Heatmap</i> korelasi antar kluster.....	79
Gambar 5.3 <i>Scatter Plot</i> (a) Pangkat-Nilai dan (b) Jabatan-Nilai	81

DAFTAR TABEL

	halaman
Tabel 2.1 Kelas Jabatan Fungsional Arsiparis.....	10
Tabel 3.1 Detail kolom <i>dataset</i> Sertifikasi JFA.....	24
Tabel 3.2 Perangkat yang digunakan dalam penelitian	25
Tabel 4.1 Hasil Rekontruksi <i>Dataset</i> Sertifikasi JFA.....	32
Tabel 4.2 Daftar <i>dataset</i> yang digabung	38
Tabel 4.3 <i>Overview dataprep report</i> dari setiap variabel	42
Tabel 4.4 Membersihkan <i>noise</i> nilai unik pada kolom	52
Tabel 4.5 Klasifikasi Peta Jabatan Fungsional Arsiparis	56

DAFTAR LAMPIRAN

halaman

Lampiran A: Environment Setup

<i>Installing Anaconda Distribution</i>	A-2
<i>Running Jupyter Notebook</i>	A-2
<i>Using Jupyter Notebook</i>	A-3

Lampiran B: Visualization Data

<i>Heatmap Data Gabungan</i>	B-2
<i>Heatmap Korelasi Variabel Kluster</i>	B-3
<i>Plot Persebaran Dataset 3-Dimensi</i>	B-4
<i>Plot Grafik Elbow Method</i>	B-5
<i>Scatter Plot pada Variabel Pangkat-Nilai dan Jabatan-Nilai</i>	B-5
<i>Confusion Matrix</i>	B-6

Lampiran C: Form & Similarity

<i>Form Final Project Guidance Monitoring</i>	C-2
<i>Form Similarity Check Clearance</i>	C-3
<i>Originality Report: BAB I</i>	C-4
<i>Originality Report: BAB II</i>	C-5
<i>Originality Report: BAB III</i>	C-6
<i>Originality Report: BAB IV</i>	C-7
<i>Originality Report: BAB V</i>	C-8
<i>Originality Report: BAB VI</i>	C-9
<i>Originality Report: FULL</i>	C-10

Lampiran D: Paper

Paper Final Project D-2

Lampiran E: Logbook Final Project

Logbook E-2

