

DAFTAR ISI

HALAMAN JUDUL	
PERNYATAAN DAN PERSETUJUAN UNGGAH TUGAS AKHIR.....	ii
PERSETUJUAN DOSEN PEMBIMBING TUGAS AKHIR	iii
PERSETUJUAN TIM PENGUJI TUGAS AKHIR.....	iv
ABSTRAK	v
ABSTRACT	vi
KATA PENGANTAR.....	vii
DAFTAR ISI.....	ix
DAFTAR GAMBAR.....	xi
DAFTAR TABEL	xiii
BAB I PENDAHULUAN	1
1.1 LATAR BELAKANG.....	1
1.2 RUMUSAN MASALAH	3
1.3 TUJUAN PENELITIAN	3
1.4 BATASAN MASALAH	4
1.5 MANFAAT PENELITIAN	4
1.6 SISTEMATIKA PENULISAN	4
BAB II LANDASAN TEORI	6
2.1 KANKER PARU-PARU.....	6
2.1.1 JENIS-JENIS KANKER PARU-PARU	7
2.1.2 DIAGNOSIS KANKER PARU-PARU	8
2.2 MACHINE LEARNING	9
2.2.1 JENIS-JENIS MACHINE LEARNING	11
2.2.2 DATA PREPROCESSING	12
2.3 METODE KLASIFIKASI.....	14
2.4 K-NEAREST NEIGHBOR (KNN).....	16
2.5 RANDOM FOREST	20
2.5 PENELITIAN TERDAHULU	23
BAB III METODOLOGI PENELITIAN	26
3.1 KERANGKA PIKIR	26

3.2 METODE PENGUMPULAN DATA	27
3.3 TAHAPAN PENELITIAN.....	27
3.4 METODE PENYELESAIAN	29
3.4.1 METODE PENYELESAIAN K-NEAREST NEIGHBOR.....	30
3.4.2 METODE PENYELESAIAN RANDOM FOREST	31
3.5 METODE PENGUKURAN AKURASI	32
BAB IV HASIL DAN PEMBAHASAN	35
4.1 HASIL PENGUMPULAN DATA	35
4.2 HASIL PENELITIAN	38
4.2.1 DATA PREPROCESSING	39
4.2.2 IMPLEMENTASI PROSES KLASIFIKASI DI ORANGE	66
4.3 PEMBAHASAN	72
BAB V KESIMPULAN DAN SARAN	76
5.1 KESIMPULAN	76
5.2 SARAN	76
DAFTAR PUSTAKA	77
LAMPIRAN.....	A-1
LAMPIRAN A: DATASET PENYAKIT KANKER PARU-PARU.....	A-1

DAFTAR GAMBAR

Gambar 2. 1 Algoritma K-Nearest Neighbor (KNN).....	19
Gambar 2. 2 Algoritma Random Forest.....	22
Gambar 3. 1 Kerangka Pikir	26
Gambar 3. 2 Tahapan Penelitian	28
Gambar 4. 1 Mapping Proses Penghapusan Data Duplikat	39
Gambar 4. 2 Atribut Index pada fitur Edit Domain	41
Gambar 4. 3 Atribut Age pada fitur Edit Domain.....	42
Gambar 4. 4 Atribut Gender pada fitur Edit Domain	43
Gambar 4. 5 Atribut Air Pollution pada fitur Edit Domain	44
Gambar 4. 6 Atribut Alcohol use pada fitur Edit Domain	45
Gambar 4. 7 Atribut Dust Allergy pada fitur Edit Domain.....	46
Gambar 4. 8 Atribut Occupational Hazards pada fitur Edit Domain.....	47
Gambar 4. 9 Atribut Genetic Risk pada fitur Edit Domain.....	48
Gambar 4. 10 Atribut Chronic Lung Disease pada fitur Edit Domain.....	49
Gambar 4. 11 Atribut Balanced Diet pada fitur Edit Domain.....	50
Gambar 4. 12 Atribut Obesity pada fitur Edit Domain	51
Gambar 4. 13 Atribut Smoking pada fitur Edit Domain.....	52
Gambar 4. 14 Atribut Passive Smoker pada fitur Edit Domain.....	53
Gambar 4. 15 Atribut Chest Pain pada fitur Edit Domain	54
Gambar 4. 16 Atribut Coughing of Blood pada fitur Edit Domain	55
Gambar 4. 17 Atribut Fatigue pada fitur Edit Domain	56
Gambar 4. 18 Atribut Weight Loss pada fitur Edit Domain.....	57
Gambar 4. 19 Atribut Shortness of Breath pada fitur Edit Domain.....	58
Gambar 4. 20 Atribut Wheezing pada fitur Edit Domain	59
Gambar 4. 21 Atribut Swallowing Difficulty pada fitur Edit Domain	60
Gambar 4. 22 Atribut Clubbing of Finger Nails pada fitur Edit Domain	61
Gambar 4. 23 Atribut Frequent Cold pada fitur Edit Domain	62
Gambar 4. 24 Atribut Dry Cough pada fitur Edit Domain	63
Gambar 4. 25 Atribut Snoring pada fitur Edit Domain.....	64
Gambar 4. 26 Atribut Level pada fitur Edit Domain	65
Gambar 4. 27 Atribut Patient Id pada fitur Edit Domain.....	66
Gambar 4. 28 Mapping proses klasifikasi dengan Orange	67
Gambar 4. 29 Proses Select Columns	68
Gambar 4. 30 Proses Data Sampler.....	69
Gambar 4. 31 Tuning algoritma K-Nearest Neighbor	70
Gambar 4. 32 Tuning algoritma Random Forest	70
Gambar 4. 33 Hasil akurasi dari prediksi (tahap training).....	71
Gambar 4. 34 Hasil akurasi dari prediksi (tahap testing).....	72
Gambar 4. 35 Hasil Confusion Matrix algoritma K-Nearest Neighbor.....	73
Gambar 4. 36 Hasil Confusion Matrix algoritma Random Forest.....	73

Gambar 4. 37 Hasil nilai Area under the ROC Curve (AUC), accuracy, precision, recall, dan F-1 Score (tahap testing) 74



DAFTAR TABEL

Tabel 2. 1 Penelitian Terdahulu	23
Tabel 3. 1 Tabel Confusion Matrix	33
Tabel 4. 1 Tabel Dataset Awal	35
Tabel 4. 2 Keterangan Atribut Dataset	37

