DAFTAR ISI

Hal

LEMBAR PENGESAHAN ..... **i**
HALAMAN PERNYATAAN ORISINALITAS ..... **ii**
HALAMAN PERSETUJUAN PUBLIKASI TUGAS AKHIR SKRIPSI UNTUK KEPENTINGAN AKADEMIS ..... **iii**
RIWAYAT HIDUP ..... **iv**
ABSTRAK ..... **v**
ABSTRACT ..... **vi**
KATA PENGANTAR ..... **vii**
DAFTAR ISI ..... **ix**
DAFTAR GAMBAR ..... **xii**
DAFTAR TABEL ..... **xiii**
DAFTAR LAMPIRAN ..... **xiv**
BAB I PENDAHULUAN **1**
1.1 Latar Belakang Masalah **1**
1.2 Identifikasi Masalah **9**
1.3 Ruang Lingkup Masalah **10**
1.4 Batasan Masalah **11**
1.5 Rumusan Masalah **11**
1.6 Tujuan Penelitian **12**
1.7 Manfaat Penelitian **12**
BAB II TINJAUAN PUSTAKA **14**
2.1 Diabetes **14**
2.2 Klasifikasi Biner **15**
2.3 Data Mining **16**
2.4 Decision Tree **17**
2.5 *Random Forest* **18**

2.6 *Confusion Matrix* 24
2.7 Penelitian yang Relevan 25
2.8 Metrik Evaluasi Kinerja Model 27
BAB III METODE PENELITIAN 28
3.1 Tempat dan Waktu Penelitian 28
3.2 Jenis Penelitian 28
3.3 Populasi dan Sampel 28
3.4 Teknik Pengumpulan Data 29
3.5 Prosedur Penelitian 30
BAB IV HASIL DAN PEMBAHASAN 34
4.1 Deskripsi Data 34
4.2 Pengumpulan *Dataset* 35
4.3 Penyeimbang Data: Teknik SMOTE 37
4.4 Praproses Data 40
4.4.1 Data Cleaning 40
4.4.2 Imputasi Data 41
4.4.3 Parsing Tekanan Darah 42
4.4.4 Encoding Data Kategorikal 43
4.4.5 Standardisasi Fitur (StandardScaler) 43
4.5 Pembangunan Model *Random Forest* 45
4.5.1 Rancangan Alur Proses Model 46
4.5.2 Arsitektur Algoritma *Random Forest* 50
4.5.3 Optimasi Hyperparameter 53
4.5.4 Pelatihan dan *Tuning* Model *Random Forest* 55
4.6 Analisis Hasil 57
4.6.1 Analisis Feature Importance 57
4.6.2 Evaluasi Model (5-*Fold Cross Validation)* 60
4.6.3 Metrik Evaluasi 61
4.6.4 Analisis Performa Model 63
4.7 Implementasi *Website* Prediksi Diabetes 64
4.8 Uji Black Box Testing 66
BAB V KESIMPULAN DAN SARAN 69

5.1 Kesimpulan 69
5.2 Saran 70
DAFTAR PUSTAKA 72
LAMPIRAN 78

UNIVERSITAS NEGERI MEDAN
UNIMED
THE Character Building UNIVERSITY

DAFTAR GAMBAR

Hal

Gambar 1.1 Jumlah Penderita Diabetes di Indonesia 4

Gambar 2.1 Struktur *Decision Tree* 17

Gambar 2.2 *Arsitektur Random Forest* 19

Gambar 3.1 Diagram Alir Proses Penelitian 30

Gambar 4.1 Jumlah Pasien Diabetes Berdasarkan Jenis Kelamin 35

Gambar 4.2 Tampilan Data Klinis Untuk *Dataset* 37

Gambar 4.3 Teknik SMOTE 38

Gambar 4.4 Penghapusan Kolom NO 41

Gambar 4.5 Konversi Bentuk Biner 43

Gambar 4.6 Perbandingan Distribusi Fitur Glukosa Darah Puasa Sebelum dan Sesudah Standardisasi 45

Gambar 4.7 Pembangunan Model *Random Forest* 46

Gambar 4.8 Save Model 49

Gambar 4.9 Arsitektur Algoritma *Random Forest* 50

Gambar 4.10 Hasil *RandomizedSearchCV* 54

Gambar 4.11 *Confusion Matrix* Model *Random Forest* 62

Gambar 4.12 Distribusi Nilai F1-*Score* Hasil RandomizedSearchCV 56

Gambar 4.13 Perbandingan Nilai F1-*Score* Rata-rata dan Terbaik 57

Gambar 4.14 Tampilan Awal *Website* 65

DAFTAR TABEL

Hal

Tabel 2.1 Confusion Matrix 24

Tabel 3.1 Dataset Klinis 29

Tabel 4.1 Atribut *Dataset* Penelitian 36

Tabel 4.2 Pemecahan Kolom 42

Tabel 4.3 *Encoding* Data Kategorikal 43

Tabel 4.4 Pembagian *Fold* 48

Tabel 4.5 Perhitungan 1 *Fold* 48

Tabel 4.6 *Output* Pohon 52

Tabel 4.7 Optimasi *Hyperparameter* Nilai yang diuji 54

Tabel 4.8 Hasil Terbaik 54

Tabel 4.9 Urutan 10 Fitur Terpenting Berdasarkan Model *Random Forest* 59

Tabel 4.10 Hasil Evaluasi Model *Random Forest* (5-*Fold Cross Validation*) 61

Tabel 4.11 Metrik Evaluasi 62

Tabel 4.12 Hasil Uji *Black Box Testing* Aplikasi Prediksi Diabetes 67

DAFTAR LAMPIRAN

Lampiran 1 Observasi Lapangan 78

Lampiran 2 Melaksanakan Penelitian 78

Lampiran 3 Surat Pernyataan Telah Melaksanakan Penelitian 79

Lampiran 4 Visual Struktur Decision pada Model Random Forest 80

Lampiran 5 Data Penelitian..... 81

Lampiran 6 Source Code Google Colab 84