# DAFTAR ISI

HALAMAN SAMPUL i

HALAMAN JUDUL ii

HALAMAN PERNYATAAN ORISINALITAS iii

ABSTRAK iv

KATA PENGANTAR v

DAFTAR ISI vii

DAFTAR TABEL xi

DAFTAR GAMBAR xiv

DAFTAR LAMPIRAN xvi

BAB 1 PENDAHULUAN 1

1.1 Latar Belakang 1

1.2 Rumusan Masalah 3

1.3 Ruang Lingkup 4

1.4 Tujuan dan Manfaat 4

1.4.1 Tujuan 4

1.4.2 Manfaat 5

1.5 Metode Penelitian 5

1.5.1 Metode Pengumpulan Data 5

1.5.2 Metode Perancangan Sistem 6

1.5.3 Metode Evaluasi 7

1.6 Sistematika Penulisan 7

BAB 2 TINJAUAN REFERENSI 9

2.1 Landasan Teori 9

2.1.1 Algoritma 9

2.1.2 *Artificial Intelligence* 10

2.1.3 *Natural Language Processing* (NLP) 10

2.1.4 *Machine Learning* 12

2.1.4.1 Pengertian *Machine Learning* 12

2.1.4.2 Algoritma *Machine Learning* 13

2.1.4.2.1 *Support Vector Machine* 13

2.1.4.2.2 Naïve Bayes 16

2.1.4.2.3 *Logistic Regression* 17

2.1.4.2.4 *Gradient Boosting* 18

2.1.4.2.5 *Linear Discriminant Analysis* 19

2.1.4.3 *Library Machine Learning* 19

2.1.4.3.1 Scikit-learn 19

2.1.4.3.2 NumPy 20

2.1.4.3.3 Pandas 20

2.1.4.3.4 Imbalanced-Learn 20

2.1.4.3.5 NLTK (*Natural Language Toolkit*) 20

2.1.5 *Deep Learning* 21

2.1.5.1 Pengertian *Deep Learning* 21

2.1.5.2 Arsitektur *Deep Learning* 22

2.1.5.2.1 MLP (*Multilayer Perceptron*) 22

2.1.5.2.2 CNN (*Convolutional Neural Network*) 23

2.1.5.2.3 LSTM (*Long Short-Term Memory*) 25

2.1.5.2.4 GRU (*Gated Recurrent Unit*) 26

2.1.5.3 *Library Deep Learning* 26

2.1.5.3.1 Keras 26

2.1.5.3.2 Theano 27

2.1.6 My*Personality* 27

2.1.7 *Resampling* 27

2.1.8 Python 28

2.1.9 Eclipse IDE 29

2.1.10 Notepad++ 30

2.1.11 Microsoft Excel 2010 30

2.1.12 XAMPP 30

2.1.13 Apache HTTP Server 30

2.1.14 *Web Service* 30

2.1.15 *The Big Five Traits* 33

2.1.16 LIWC 36

2.1.17 SPLICE 38

2.2 Penelitian Terkait 39

2.2.1 *Personality and Patterns of Facebook Usage* 39

2.2.2 *Recognising Personality Traits Using Facebook Status Updates* 43

2.2.3 *Personality, Gender, and Age in the Language of Social Media:*

*The Open-Vocabulary Approach* 47

2.2.4 Sistem Prediksi Kepribadian “*The Big Five Traits*” Dari Data

Twitter 53

2.2.5 *Personality Prediction Based on Twitter Information in Bahasa* 56

2.2.6 *Personality Traits Recognition on Social Network - Facebook* 61

2.2.7 *Deep Learning-Based Document Modeling for Personality Detection*

*from Text* 63

BAB 3 METODOLOGI 67

3.1 Analisis Masalah 67

3.2 Usulan Pemecahan Masalah 68

3.3 Kerangka Berpikir 69

3.3.1 *Dataset Collecting* 69

3.3.2 *Data Preprocessing* 75

3.3.3 *Feature Selection* 81

3.3.4 *Model Learning Process* 94

3.3.5 *Prediction Model Testing* 96

3.3.6 *Final Result of Prediction Based on Big Five Model* 97

3.4 Rancangan Layar 97

3.4.1 Rancangan Layar *Homepage* 98

3.4.2 Rancangan Layar *Privacy Policy* 98

3.4.3 Rancangan Layar *Result* 99

BAB 4 HASIL PENELITIAN 101

4.1 *Testing Environment* 101

4.1.1 Spesifikasi Sistem 101

4.1.1.1 *Device* 101

4.1.1.2 *Server* 101

4.1.2 Perangkat Lunak 101

4.1.3 *Library* 101

4.2 Hasil 102

4.2.1 Distribusi Data 102

4.2.2 Faktor Skenario Percobaan 103

4.2.2.1 *Feature Selection* 103

4.2.2.2 *Resampling* 103

4.2.3 Implementasi *Machine Learning* 104

4.2.3.1 Skenario 104

4.2.3.2 Dataset my*Personality* 106

4.2.3.2.1 Skenario 1 (Percobaan tanpa menggunakan *Features*

*Selection* dan tanpa *Resampling*) 106

4.2.3.2.2 Skenario 2 (Percobaan dengan menggunakan *Features*

*Selection*) 108

4.2.3.2.3 Skenario 3 (Percobaan dengan menggunakan

*Undersampling*) 110

4.2.3.2.4 Skenario 4 (Percobaan dengan menggunakan

*Oversampling*) 112

4.2.3.2.5 Skenario 5 (Percobaan dengan menggunakan *Features*

*Selection* dan dengan *Undersampling*) 114

4.2.3.2.6 Skenario 6 (Percobaan dengan menggunakan *Features*

*Selection* dan dengan *Oversampling*) 116

4.2.3.2.7 Kesimpulan Hasil Implementasi *Machine Learning*

pada *Dataset* my*Personality* 118

4.2.3.3 Dataset Manual Gathering 119

4.2.3.3.1 Skenario 7 (Percobaan tanpa menggunakan *Features*

*Selection* dan tanpa *Resampling*) 119

4.2.3.3.2 Skenario 8 (Percobaan dengan menggunakan *Features*

*Selection*) 121

4.2.3.3.3 Skenario 9 (Percobaan dengan menggunakan

*Undersampling*) 122

4.2.3.3.4 Skenario 10 (Percobaan dengan menggunakan

*Oversampling*) 123

4.2.3.3.5 Skenario 11 (Percobaan dengan menggunakan *Features*

*Selection* dan dengan *Undersampling*) 125

4.2.3.3.6 Skenario 12 (Percobaan dengan menggunakan *Features*

*Selection* dan dengan *Oversampling*) 126

4.2.3.3.7 Kesimpulan Hasil Implementasi *Machine Learning*

pada *Dataset Manual Gathering* 128

4.2.3.4 Dataset Gabungan 129

4.2.3.4.1 Skenario 13 (Percobaan tanpa menggunakan *Features*

*Selection* dan tanpa *Resampling*) 129

4.2.3.4.2 Skenario 14 (Percobaan dengan menggunakan *Features*

*Selection*) 131

4.2.3.4.3 Skenario 15 (Percobaan dengan menggunakan

*Undersampling*) 132

4.2.3.4.4 Skenario 16 (Percobaan dengan menggunakan

*Oversampling*) 134

4.2.3.4.5 Skenario 17 (Percobaan dengan menggunakan *Features*

*Selection* dan dengan *Undersampling*) 136

4.2.3.4.6 Skenario 18 (Percobaan dengan menggunakan *Features*

*Selection* dan dengan *Oversampling*) 137

4.2.3.4.7 Kesimpulan Hasil Implementasi *Machine Learning*

pada *Dataset* Gabungan 139

4.2.3.5 Kesimpulan Implementasi *Machine Learning* 140

4.2.4 Implementasi *Deep Learning* 142

4.2.4.1 Skenario 142

4.2.4.2 *Dataset* my*Personality* 143

4.2.4.2.1 Skenario 1 (Percobaan tanpa menggunakan

*Resampling*) 143

4.2.4.2.2 Skenario 2 (Percobaan dengan menggunakan

*Undersampling*) 144

4.2.4.2.3 Skenario 3 (Percobaan dengan menggunakan

*Oversampling*) 145

4.2.4.2.4 Kesimpulan Hasil Implementasi *Deep Learning*

pada *Dataset* my*Personality* 146

4.2.4.3 Dataset Manual Gathering 147

4.2.4.3.1 Skenario 4 (Percobaan tanpa menggunakan

*Resampling*) 147

4.2.4.3.2 Skenario 5 (Percobaan dengan menggunakan

*Undersampling*) 148

4.2.4.3.3 Skenario 6 (Percobaan dengan menggunakan

*Oversampling*) 149

4.2.4.3.4 Kesimpulan Hasil Implementasi *Deep Learning*

pada *Dataset Manual Gathering* 150

4.2.4.4 Dataset Gabungan 150

4.2.4.4.1 Skenario 7 (Percobaan tanpa menggunakan

*Resampling*) 151

4.2.4.4.2 Skenario 8 (Percobaan dengan menggunakan

*Undersampling*) 151

4.2.4.4.3 Skenario 9 (Percobaan dengan menggunakan

*Oversampling*) 153

4.2.4.4.4 Kesimpulan Hasil Implementasi *Deep Learning*

pada *Dataset* Gabungan 154

4.2.4.5 Kesimpulan Implementasi *Deep Learning* 155

4.2.5 Tampilan Layar Aplikasi 156

4.3 Evaluasi 162

4.3.1 Evaluasi Subjektif 162

4.3.2 Evaluasi Objektif 164

4.3.2.1 Sistem Prediksi Kepribadian “*The Big Five Personality*”

Dari Data Twitter 164

4.3.2.2 *Personality Traits Recognition on Social Network -*

*Facebook* 166

4.3.2.3 *Deep Learning-Based Document Modeling for Personality*

*Detection from Text* 169

BAB 5 SIMPULAN DAN SARAN 173

5.1 Simpulan 173

5.2 Saran 174

REFERENSI 175

LAMPIRAN L1

RIWAYAT HIDUP

SURAT SURVEI