UNIVERSITAS
GADJAH MADA
PENGARUH REDUKSI FITUR DENGAN PRINCIPAL COMPONENTS ANALYSIS (PCA) TERHADAP PERFORMA AUTOMATED ESSAY SCORING (AES) BAHASA INDONESIA
ANDIKA RAHIM D, Drs. Sri Mulyana, M.Kom.
Universitas Gadjah Mada, 2020 | Diunduh dari http://etd.repository.ugm.ac.id/

DAFTAR ISI

HALAMAN JUDUL i

HALAMAN PENGESAHAN ii

HALAMAN PERNYATAAN iii

HALAMAN MOTTO iv

PRAKATA v

DAFTAR ISI vii

DAFTAR TABEL x

DAFTAR GAMBAR xi

INTISARI xii

ABSTRACT xiii

I PENDAHULUAN 1

1.1 Latar Belakang Masalah 1
1.2 Rumusan Masalah 2
1.3 Batasan Masalah 3
1.4 Tujuan Penelitian 3
1.5 Manfaat Penelitian 3
1.6 Sistematika Penulisan 4

II KAJIAN PUSTAKA 5

III LANDASAN TEORI 8

3.1 *Automated Essay Scoring* 8
3.2 *Text Preprocessing* 9
3.2.1 *Folding Cases* 9
3.2.2 *Tokenizing* 9
3.2.3 *Stemming* 9

UNIVERSITAS
GADJAH MADA
PENGARUH REDUKSI FITUR DENGAN PRINCIPAL COMPONENTS ANALYSIS (PCA) TERHADAP PERFORMA AUTOMATED ESSAY SCORING (AES) BAHASA INDONESIA
ANDIKA RAHIM D, Drs. Sri Mulyana, M.Kom.
Universitas Gadjah Mada, 2020 | Diunduh dari http://etd.repository.ugm.ac.id/

3.2.4 *Stop Word Removal* 9
3.3 Ekstraksi Fitur 10
3.3.1 *N-gram* 10
3.3.2 *P*embobotan Term Frequency-Inverse Documents Frequency (TF-IDF) 10
3.3.3 *Vector Space Model* (VSM) 10
3.3.4 *Singular Value Decomposition* (SVD) 11
3.3.5 *Principal Components Analysis* (PCA) 13
3.4 *Classifier* 13
3.4.1 *Decision Tree* 13
3.4.2 *Random Forest* (RF) 14
3.4.3 *Adaptive Boosting* (AdaBoost) 15
3.4.4 *K-Nearest Neighbors*(KNN) 15
3.4.5 *Support Vector Machine* (SVM) 16
3.5 Evaluasi 17

IV METODOLOGI PENELITIAN 19
4.1 Deskripsi Penelitian 19
4.2 Dataset 20
4.3 Tahapan Penelitian 21
4.3.1 Pengumpulan Data 21
4.3.2 *Preprocessing* 21
4.3.3 TF-IDF 22
4.3.4 PCA 23
4.3.5 *Classifier* 25
4.3.6 Pelatihan dan Pengujian 30
4.3.7 Analisis 30

V IMPLEMENTASI 31
5.1 Implementasi Akuisisi Data 31
5.2 Implementasi Folding Cases 32
5.3 Implementasi Tokenizing 33
5.4 Implementasi Stop Word Removal 33
5.5 Implementasi Steming 34
5.6 Implementasi TF-IDF 34
5.7 Implementasi PCA 35

UNIVERSITAS
GADJAH MADA
PENGARUH REDUKSI FITUR DENGAN PRINCIPAL COMPONENTS ANALYSIS (PCA) TERHADAP PERFORMA AUTOMATED ESSAY SCORING (AES) BAHASA INDONESIA
ANDIKA RAHIM D, Drs. Sri Mulyana, M.Kom.
Universitas Gadjah Mada, 2020 | Diunduh dari http://etd.repository.ugm.ac.id/

5.8 Implementasi *Classifier* 36
5.8.1 Implementasi 10-Fold Cross Validation 36
5.8.2 Implementasi *Classifier* 36
5.9 Implementasi Pelatihan dan Pengujian 37
5.10 Impelementasi Simulasi 38

VI HASIL DAN PEMBAHASAN **40**
6.1 Hasil Akusisi data 40
6.2 Hasil *Folding Cases* 40
6.3 Hasil *Tokenizing* 41
6.4 Hasil *Stop Word Removal* 41
6.5 Hasil Steming 41
6.6 Hasil TF-IDF 42
6.7 Hasil Simulasi Pelatihan dan Pengujian 42
6.7.1 *Decision Tree* 43
6.7.2 *Random Forest* 44
6.7.3 *Adaptive Boosting* 46
6.7.4 *K-Nearest Neighbors* 47
6.7.5 *Support Vector Machine* 49
6.8 Perbandingan antar *Classifier* 50

VIIPENUTUP **54**
7.1 Kesimpulan 54
7.2 Saran 54

DAFTAR PUSTAKA **56**