DAFTAR ISI

ABSTRAK i
ABSTRACT ii
LEMBAR PENGESAHAN iii
LEMBAR PERNYATAAN ORISINALITAS iv
KATA PENGANTAR v
DAFTAR ISI vi
DAFTAR GAMBAR x
DAFTAR TABEL xii
DAFTAR ISTILAH xiv
BAB I PENDAHULUAN 1
I.1 Latar Belakang 1
I.2 Perumusan Masalah 4
I.3 Tujuan Penelitian 4
I.4 Batasan Penelitian 4
I.5 Manfaat Penelitian 5
BAB II TINJAUAN PUSTAKA 6
II.1 *Learning Management System* 6
II.2 *Data Mining* 6
II.3 *Text Mining* 7
II.4 *Cross Industry Standard Process for Data Mining* 7
II.5 Klasifikasi 9
II.6 Klasifikasi Multi-label 9
II.7 *Random Forest* 9
II.8 *K-Nearest Neighbors* 11
II.9 *Text Preprocessing* 13
II.9.1 *Case Folding* 13
II.9.2 *Tokenizing* 13
II.9.3 *Stopword Removal* 14
II.9.4 *Stemming* 14
II.10 TF-IDF 15
II.10.1 *Term Frequency* 15

II.10.2 *Inverse Document Frequency* 15
II.10.3 *Term Frequency-Inverse Document Frequency* 16
II.11 *Multilabel Problem Transformation* 16
II.11.1 *Binary Relevance* 16
II.11.2 *Classifier Chain* 17
II.11.3 *Label Powerset* 17
II.12 *Confusion Matrix* 18
II.13 *K-Fold Cross Validation* 20
II.14 *Wordcloud* 20
II.15 *State of the Art* 21
BAB III METODOLOGI PENELITIAN 25
III.1 Pengembangan Model Konseptual 25
III.2 Sistematika Penyelesaian Masalah 25
III.2.1 *Business Understanding* 27
III.2.2 *Data Understanding* 27
III.2.3 *Data Preparation* 28
III.2.4 *Modeling* 29
III.2.5 *Evaluation* 29
III.3 *Deployment* 30
III.4 Alasan Pemilihan Metode 30
BAB IV ANALISIS DAN PERANCANGAN 31
IV.1 *Data Understanding* 31
IV.1.1 Pengumpulan Data 31
IV.1.2 Menentukan Label Topik 32
IV.1.3 Pemilihan Atribut 32
IV.1.4 Penyaringan Data 32
IV.2 *Data Preparation* 33
IV.2.1 Penghapusan Tag HTML 34
IV.2.2 *Case Folding* 35
IV.2.3 *Tokenizing* 35
IV.2.4 *Stopword Removal* 36
IV.2.5 *Stemming* 37

IV.2.6 Pembobotan TF-IDF 37
IV.2.7 *Splitting Data* 41
IV.3 *Data Modeling* 41
IV.3.1 *Problem Transformation Method* 42
IV.3.2 *Random Forest* 44
IV.3.3 *K-Nearest Neighbors* 47
IV.3.4 *Hyperparameter Tuning* 47
IV.4 *Evaluation* 49
IV.4.1 *Confusion Matrix* 50
IV.4.2 *Classification Report* 52
IV.5 *Deployment* 54
IV.6 Skenario Pengujian 54
BAB V IMPLEMENTASI DAN PENGUJIAN 55
V.1 Evaluasi Klasifikasi Multi-Label 55
V.1.1 Pengujian *Random Forest* 55
V.1.2 Pengujian K-NN 57
V.1.3 Evaluasi Perbandingan Algoritma 58
V.1.4 Evaluasi *K-Fold Cross Validation* 59
V.2 *Dashboard* Model Klasifikasi 60
V.3 Analisis *Wordcloud* 62
V.3.1 Topik Puisi 63
V.3.2 Topik Ulasan 63
V.3.3 Topik Kalimat 64
V.3.4 Topik Kalimat_efektif 64
V.3.5 Topik Perbaikan_kalimat 65
V.3.6 Topik Ulasan_kalimat 65
V.3.7 Topik Paragraf 66
V.3.8 Topik Majas 66
V.3.9 Topik Larik 67
V.3.10 Topik Teks_diskusi 67
V.3.11 Topik Cerpen 68
BAB VI KESIMPULAN DAN SARAN 69

VI.1 Kesimpulan 69
VI.2 Saran 70
DAFTAR PUSTAKA 71
LAMPIRAN 78