

DAFTAR ISI

| | |
|---|------|
| HALAMAN PENGESAHAN..... | iii |
| PERNYATAAN..... | iv |
| HALAMAN MOTTO DAN PERSEMBAHAN..... | v |
| PRAKATA..... | vii |
| DAFTAR ISI..... | ix |
| DAFTAR TABEL..... | xii |
| DAFTAR GAMBAR..... | xiv |
| INTISARI..... | xvi |
| ABSTRACT..... | xvii |
| BAB I PENDAHULUAN..... | 1 |
| 1.1. Latar Belakang..... | 1 |
| 1.2. Rumusan Masalah..... | 2 |
| 1.3. Batasan Masalah..... | 2 |
| 1.4. Tujuan Penelitian..... | 2 |
| 1.5. Manfaat Penelitian..... | 3 |
| BAB II TINJAUAN PUSTAKA..... | 4 |
| BAB III LANDASAN TEORI..... | 7 |
| 3.1. <i>Text mining</i> | 7 |
| 3.2. Twitter..... | 8 |
| 3.3. <i>Text Document Classification</i> | 8 |
| 3.4. <i>Clickbait</i> | 10 |
| 3.5. <i>Text Preprocessing</i> | 11 |
| 3.5.1. Tokenisasi..... | 12 |
| 3.5.2. <i>Stemming</i> | 12 |
| 3.6. <i>Term Frequency-Inverse Document Frequency</i> | 13 |
| 3.7. Support Vector Machine..... | 15 |
| 3.8. Multinomial Naïve Bayes..... | 19 |
| 3.9. Evaluasi Performa..... | 20 |
| 3.9.1. Akurasi..... | 20 |
| 3.9.2. Presisi..... | 20 |
| 3.9.3. Recall..... | 21 |

| | | |
|---------------------------------------|---|----|
| 3.9.4. | <i>F1-Score</i> | 21 |
| BAB IV ANALISIS DAN PERANCANGAN | | 22 |
| 4.1. | Analisis Kebutuhan | 22 |
| 4.2. | Tahapan Penelitian | 22 |
| 4.3. | Rancangan Pengambilan data | 24 |
| 4.4. | Rancangan Algoritme | 26 |
| 4.4.1. | Pelabelan Data | 28 |
| 4.4.2. | <i>Preprocessing</i> | 28 |
| 4.4.3. | <i>Ekstraksi Fitur</i> | 35 |
| 4.4.4. | <i>Pembagian dataset</i> | 35 |
| 4.4.5. | <i>Training</i> | 36 |
| 4.4.6. | <i>Testing</i> | 37 |
| 4.4.7. | Perhitungan Performa | 38 |
| 4.5. | Rancangan Pengujian | 38 |
| 4.5.1. | Pengujian Kemampuan untuk Menerima Tweet Baru | 38 |
| 4.5.2. | Pengujian Klasifikasi <i>Tweet</i> Berita | 38 |
| BAB V IMPLEMENTASI | | 39 |
| 5.1. | Implementasi Akuisisi <i>Dataset</i> | 39 |
| 5.2. | Implementasi <i>Preprocessing</i> | 41 |
| 5.2.1. | <i>Case Folding</i> | 41 |
| 5.2.2. | <i>Retweet Removal</i> | 42 |
| 5.2.3. | <i>Mention Removal</i> | 42 |
| 5.2.4. | <i>Hashtag Removal</i> | 42 |
| 5.2.5. | <i>URLs Removal</i> | 43 |
| 5.2.6. | <i>Non-alpha-numeric Removal</i> | 43 |
| 5.2.7. | <i>Stemming</i> | 44 |
| 5.3. | Implementasi Ekstraksi fitur | 45 |
| 5.4. | Implementasi Pembagian <i>Dataset</i> | 46 |
| 5.5. | Implementasi <i>Training</i> | 48 |
| 5.5.1. | <i>Training SVM</i> | 48 |
| 5.5.2. | <i>Training MNB</i> | 52 |
| 5.6. | Implementasi <i>Testing</i> | 52 |
| 5.6.1. | <i>Testing SVM</i> | 52 |
| 5.6.2. | <i>Testing MNB</i> | 53 |

| | |
|--|-----------|
| 5.7. Implementasi Perhitungan Performa..... | 58 |
| BAB VI HASIL DAN PEMBAHASAN | 60 |
| 6.1. Hasil Akuisisi <i>Dataset</i> dan Pelabelan Manual..... | 60 |
| 6.2. Hasil <i>Preprocessing</i> | 61 |
| 6.3. Hasil Ekstraksi Fitur..... | 62 |
| 6.4. Hasil Pembagian <i>Dataset</i> | 63 |
| 6.5. Hasil <i>Training</i> | 64 |
| 6.5.1. <i>Training SVM</i> | 64 |
| 6.5.2. <i>Training MNB</i> | 66 |
| 6.6. Hasil <i>Testing</i> | 66 |
| 6.6.1. <i>Testing SVM</i> | 66 |
| 6.6.2. <i>Testing MNB</i> | 67 |
| 6.7. Hasil Perhitungan Performa | 71 |
| BAB VII KESIMPULAN DAN SARAN | 73 |
| 7.1. Kesimpulan | 73 |
| 7.2. Saran..... | 73 |
| DAFTAR PUSTAKA | 74 |