

DAFTAR ISI

| | |
|---|------|
| HALAMAN JUDUL..... | ii |
| HALAMAN PENGESAHAN | iii |
| HALAMAN PERNYATAAN..... | iv |
| PRAKATA | vi |
| DAFTAR ISI | vii |
| DAFTAR TABEL..... | x |
| DAFTAR GAMBAR | xi |
| ABSTRAK | xiii |
| ABSTRACT | xiv |
| BAB I PENDAHULUAN | 1 |
| 1.1 Latar Belakang Masalah | 1 |
| 1.2. Rumusan Masalah..... | 4 |
| 1.3. Batasan Masalah..... | 4 |
| 1.4. Tujuan Penelitian..... | 4 |
| 1.5. Manfaat Penelitian | 4 |
| 1.6. Metodologi Penelitian..... | 4 |
| 1.7. Sistematika Penulisan | 5 |
| BAB II TINJAUAN PUSTAKA | 6 |
| BAB III LANDASAN TEORI | 11 |
| 3.1. <i>Twitter Application Program Interface (Twitter API)</i> | 11 |
| 3.2. <i>Preprocessing</i> | 11 |
| 3.2.1 Menghapus URL..... | 11 |
| 3.2.2 <i>Squeeze Whitespace</i> | 12 |
| 3.2.3 <i>Case Folding</i> | 12 |
| 3.2.4 Menghapus <i>Hashtags, Mentions & RT (Retweet)</i> | 13 |

| | |
|--|-----------|
| 3.2.5 <i>Split Punctuation</i> | 13 |
| 3.2.6 Tokenisasi..... | 14 |
| 3.3 POS Tagging | 14 |
| 3.4 Pengenalan Entitas Bernama (NER)..... | 18 |
| 3.4.1 <i>BIO Tagging</i> /Segmentasi Entitas..... | 19 |
| 3.4.2 Multinomial Naïve Bayes (MNB) | 21 |
| 3.4.3 Conditinonal Random Field (CRF) | 23 |
| 3.5 Definisi dan Jenis Bencana | 24 |
| 3.6 <i>Performance Measure</i> | 27 |
| BAB IV RANCANGAN PENELITIAN | 28 |
| 4.1 Deskripsi Umum..... | 28 |
| 4.2 Pengumpulan Data..... | 29 |
| 4.3 <i>Preprocessing</i> | 30 |
| 4.3.1 Menghapus URL..... | 30 |
| 4.3.2 Menghapus Retweet(s), Hashtag(s), & Mention(s) | 31 |
| 4.3.3 <i>Squeeze Whitespace</i> | 31 |
| 4.3.4 Menghapus <i>Duplicate</i> | 31 |
| 4.3.5 Melakukan <i>Lowercase</i> | 31 |
| 4.3.6 Melakukan Tokenisasi | 31 |
| 4.4 Klasifikasi Manual..... | 32 |
| 4.5 <i>Postagging</i> | 32 |
| 4.6 <i>BIO Tagging Semi-Manual</i> | 32 |
| 4.7 <i>Feature Extraction</i> | 33 |
| 4.8 Pengenalan Entitas Bernama (NER)..... | 33 |
| 4.9 Evaluasi Hasil Prediksi Model | 34 |
| BAB V IMPLEMENTASI | 36 |
| 5.1 Spesifikasi <i>Hardware dan Software</i> | 36 |
| 5.2 Pengumpulan Data..... | 36 |
| 5.3 <i>Preprocessing</i> | 37 |
| 5.4 Klasifikasi Manual..... | 39 |

| | |
|---|-----------|
| 5.5 <i>Postagging</i> | 39 |
| 5.6 <i>BIO Tagging</i> Semi-Manual | 40 |
| 5.7 <i>Feature Extraction</i> | 43 |
| 5.8 Pengenalan Entitas Bernama (NER)..... | 44 |
| 5.9 Evaluasi..... | 47 |
| BAB VI HASIL DAN PEMBAHASAN | 50 |
| 6.1 Hasil Pengumpulan Data..... | 50 |
| 6.2 Hasil <i>Preprocessing & Cleansing</i> | 51 |
| 6.3 Hasil Klasifikasi Manual..... | 52 |
| 6.4 Hasil <i>Postagging</i> | 53 |
| 6.5 Hasil <i>BIO Tagging</i> Semi-Manual | 54 |
| 6.6 <i>Feature Extraction</i> | 55 |
| 6.7 Hasil Pengujian NER..... | 57 |
| 6.7.1 Hasil <i>Multinomial Naïve Bayes (MNB)</i> | 57 |
| 6.7.2 Hasil <i>Conditional Random Field (CRF)</i> | 58 |
| 6.8 Pembahasan..... | 63 |
| BAB VII KESIMPULAN DAN SARAN | 65 |
| 7.1 Kesimpulan | 65 |
| 7.2 Saran..... | 65 |