

DAFTAR ISI

| | Halaman |
|--|-------------------------------------|
| SKRIPSI..... | i |
| LEMBAR PENGESAHAN NASKAH SKRIPSI..... | iii |
| SURAT PERNYATAAN TENTANG ORISINALITAS | Error! Bookmark not defined. |
| PEDOMAN PENGGUNAAN SKRIPSI | vi |
| KATA PENGANTAR | vii |
| UCAPAN TERIMAKASIH..... | viii |
| ABSTRAK | x |
| ABSTRACT..... | xi |
| DAFTAR ISI..... | xii |
| DAFTAR GAMBAR | xv |
| DAFTAR TABEL..... | xvii |
| BAB I PENDAHULUAN..... | 1 |
| 1.1 Latar Belakang..... | 1 |
| 1.2 Rumusan Masalah | 4 |
| 1.3 Tujuan..... | 4 |
| 1.4 Manfaat..... | 4 |
| 1.5 Batasan Masalah..... | 5 |
| BAB II TINJAUAN PUSTAKA..... | 6 |
| 2.1 <i>Microblogging</i> | 6 |
| 2.2 Pendidikan | 6 |
| 2.3 <i>Machine Learning</i> | 7 |
| 2.4 Analisis Sentimen..... | 9 |

| | | |
|-----------------------------------|--|----|
| 2.5 | <i>K-Nearest Neighbor (KNN)</i> | 10 |
| 2.6 | TF-IDF..... | 14 |
| 2.7 | <i>Crowdsourcing Labelling</i> | 15 |
| 2.8 | <i>K-Fold Cross Validation</i> | 16 |
| 2.9 | <i>Confusion Matrix</i> | 17 |
| 2.10 | <i>Application Programming Interface (API)</i> | 20 |
| 2.11 | Sistem Informasi Geografis (SIG)..... | 20 |
| 2.12 | <i>Unified Modelling Language (UML)</i> | 20 |
| BAB III METODE PENELITIAN..... | | 23 |
| 3.1 | Pengumpulan Data <i>Tweet</i> | 24 |
| 3.2 | <i>Labelling Data</i> | 25 |
| 3.3 | <i>Preprocessing Data</i> | 26 |
| 3.4 | Klasifikasi dengan Model KNN | 28 |
| 3.5 | Evaluasi Model KNN | 29 |
| 3.6 | Perancangan Sistem..... | 30 |
| 3.7 | Implementasi Sistem | 30 |
| BAB IV HASIL DAN PEMBAHASAN | | 31 |
| 4.1 | Pengumpulan Data <i>Tweet</i> | 31 |
| 4.2 | <i>Labelling Data</i> | 35 |
| 4.3 | <i>Preprocessing Data</i> | 38 |
| 4.4 | Klasifikasi dengan Model KNN | 45 |
| 4.5 | Evaluasi Model KNN | 50 |
| 4.6 | Perancangan Sistem..... | 54 |
| 4.7 | Implementasi Sistem | 60 |
| BAB V KESIMPULAN DAN SARAN..... | | 70 |

| | | |
|----------------------|-----------------|----|
| 5.1 | Kesimpulan..... | 70 |
| 5.2 | Saran..... | 71 |
| DAFTAR PUSTAKA | | 72 |
| LAMPIRAN..... | | 77 |

DAFTAR GAMBAR

Gambar 2.1 Perbedaan *traditional programming* dengan *machine learning* 8

Gambar 2.2 Ilustrasi *clustering* 9

Gambar 2.3 Implementasi KNN 11

Gambar 2.4 *5-Fold Cross Validation*.....**Error! Bookmark not defined.**

Gambar 3.1 Tahapan Penelitian 23

Gambar 3.2 Alur pengumpulan data 24

Gambar 3.3 Tahap *cleaning* data 25

Gambar 3.4 Tahap *labelling* data 26

Gambar 4.1 Contoh hasil pembuatan aplikasi baru 31

Gambar 4.2 Grafik jumlah sentiment masyarakat di setiap provinsi 37

Gambar 4.3 Matriks hasil perhitungan Term Frequency (TF)..... 43

Gambar 4.4 Hasil perhitungan Inverse Document Frequency (IDF)..... 44

Gambar 4. 5 Hasil perhitungan TF-IDF..... 44

Gambar 4.6 Grafik skor K terbaik..... 50

Gambar 4.7 Confusion Matrix 51

Gambar 4.8 Hasil penghitungan parameter pengukuran KNN 54

Gambar 4.9 Use case diagram sistem deteksi persebaran sentimen masyarakat .. 55

Gambar 4.10 Activity diagram login 56

Gambar 4.11 Activity diagram ubah label klasifikasi..... 57

Gambar 4.12 Activity diagram menambahkan data training 57

Gambar 4.13 Activity diagram training ulang model klasifikasi..... 58

Gambar 4.14 Activity diagram scrap data..... 59

Gambar 4. 15 Activity diagram lihat geovisualisasi sebaran data 59

Gambar 4.16 Tampilan utama sistem..... 61

Gambar 4.17 Menu pilih provinsi 62

Gambar 4.18 Tampilan setelah memilih salah satu provinsi 62

Gambar 4.19 Tampilan tidak ada data pada provinsi yang dipilih 63

Gambar 4.20 Total data sentimen di seluruh Indonesia..... 63

Gambar 4.21 Total masing-masing sentimen pada data seluruh Indonesia..... 64

| | |
|---|----|
| Gambar 4.22 Tren data sentimen zonasi sekolah sejak tahun 2016-2020 | 64 |
| Gambar 4.23 Jumlah data provinsi Jawa Barat | 65 |
| Gambar 4.24 Jumlah data provinsi Jawa Timur..... | 65 |
| Gambar 4. 25 Jumlah data provinsi Jawa Tengah..... | 66 |
| Gambar 4. 26 Tampilan login | 66 |
| Gambar 4.27 Tampilan dasbor pakar | 67 |
| Gambar 4.28 Grafik pie, batang, dan garis | 67 |
| Gambar 4.29 Tampilan data training | 68 |
| Gambar 4.30 Tampilan hasil klasifikasi..... | 68 |
| Gambar 4.31 Tampilan ubah label | 69 |

DAFTAR TABEL

| | |
|--|----|
| Tabel 2.1 Contoh confusion matrix..... | 17 |
| Tabel 2.2 Confusion matrix dengan data positif dan negatif | 18 |
| Tabel 2.3 Komponen use case diagram..... | 21 |
| Tabel 2.4 Komponen activity diagram..... | 22 |
| Tabel 3.1 Contoh case folding | 27 |
| Tabel 3.2 Contoh tokenization | 27 |
| Tabel 3.3 Contoh penghapusan stopwords..... | 28 |
| Tabel 3.4 Contoh stemming | 28 |
| Tabel 4.1 Penjelasan atribut data | 33 |
| Tabel 4.2 Contoh data hasil scraping dan cleaning..... | 34 |
| Tabel 4.3 Contoh proses labelling..... | 35 |
| Tabel 4.4 Jumlah sentimen masyarakat di provinsi yang memiliki data sentimen | 36 |
| Tabel 4.5 Contoh hasil case folding | 38 |
| Tabel 4.6 Contoh hasil tokenization..... | 39 |
| Tabel 4.7 Contoh hasil penghapusan stopwords | 41 |
| Tabel 4.8 Contoh hasil stemming..... | 42 |
| Tabel 4.9 Perbandingan 10-Fold | 48 |
| Tabel 4.10 Perbandingan akurasi nilai K | 49 |