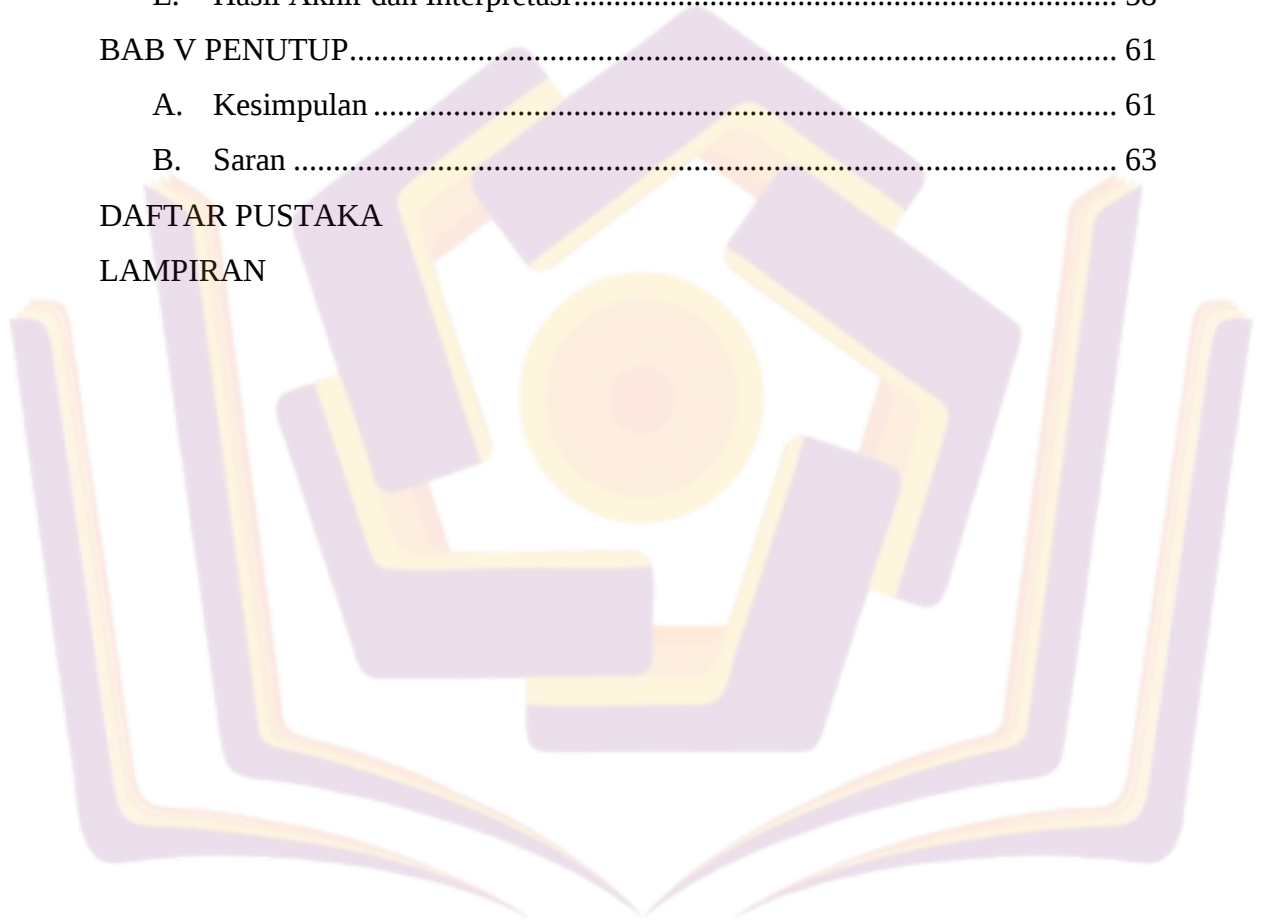


DAFTAR ISI

HALAMAN SAMPUL	i
HALAMAN JUDUL.....	ii
HALAMAN PERSETUJUAN.....	iii
HALAMAN PENGESAHAN.....	iv
HALAMAN PERNYATAAN KEASLIAN	v
HALAMAN PERSEMBAHAN	vi
HALAMAN MOTTO	vii
KATA PENGANTAR	viii
DAFTAR ISI.....	x
DAFTAR TABEL	xii
DAFTAR GAMBAR	xiii
DAFTAR ISTILAH	xiv
DAFTAR LAMPIRAN.....	xv
INTISARI.....	xvii
<i>ABSTRACT</i>	xviii
BAB I PENDAHULUAN	1
A. Latar Belakang Masalah	1
B. Rumusan Masalah.....	8
C. Batasan Masalah	8
D. Tujuan Penelitian	9
E. Manfaat Penelitian	9
BAB II TINJAUAN PUSTAKA.....	11
A. Landasan Teori.....	11
B. Penelitian Sebelumnya.....	17
BAB III METODE PENELITIAN.....	24
A. Tempat dan Waktu Penelitian.....	24
B. Metode Pengumpulan Data.....	24
C. Alat dan Bahan Penelitian.....	25

D. Konsep Penelitian	25
BAB IV HASIL DAN PEMBAHASAN	35
A. Pengumpulan Data	35
B. Pelabelan RoBERTa model HuggingFace.....	36
C. Pra-pemrosesan dan Persiapan Data	37
D. <i>Cross Validation</i> dan Evaluasi Model	48
E. Hasil Akhir dan Interpretasi.....	58
BAB V PENUTUP.....	61
A. Kesimpulan	61
B. Saran	63
DAFTAR PUSTAKA	
LAMPIRAN	



DAFTAR TABEL

Tabel 2. 1 Penelitian Sebelumnya.....	21
Tabel 3. 1 Komentar Masyarakat tentang Danantara di media sosial X.....	27
Tabel 4. 1 Komentar Masyarakat tentang Danantara di media sosial X.....	35
Tabel 4. 2 Contoh proses <i>Cleaning</i> dan <i>Case Folding</i>	39
Tabel 4. 3 Contoh proses <i>Tokenizing</i>	40
Tabel 4. 4 Contoh proses <i>Stopword Removal</i>	42
Tabel 4. 5 Contoh proses <i>Stemming</i>	43
Tabel 4. 6 Hasil Evaluasi per <i>Fold</i> Model SVM.....	50
Tabel 4. 7 Laporan Klasifikasi <i>fold</i> terbaik SVM berdasarkan akurasi	52
Tabel 4. 8 Hasil evaluasi dari model <i>Random Forest</i>	53
Tabel 4. 9 Laporan Klasifikasi Fold Terbaik Model <i>Random Forest</i>	55
Tabel 4. 10 <i>Fold</i> Terbaik dari SVM dan <i>Random Forest</i>	57

DAFTAR GAMBAR

Gambar 3. 1 Kerangka Berpikir	26
Gambar 3. 2 Tahapan <i>Preprocessing</i>	29
Gambar 3. 3 Diagram Batang Distribusi Label.....	31
Gambar 4. 1 Distribusi jumlah data berdasarkan hasil pelabelan	37
Gambar 4. 2 Hasil Proses TF-IDF.....	44
Gambar 4. 3 Hasil sebelum penerapan <i>SMOTE</i>	46
Gambar 4. 4 Hasil setelah penerapan <i>SMOTE</i>	47
Gambar 4. 5 <i>Confusion Matrix SVM Fold Terbaik</i>	51
Gambar 4. 6 <i>Confusion Matrix Random Forest Fold Terbaik</i>	54
Gambar 4. 7 Perbandingan <i>Fold Terbaik Model SVM dan Random Forest</i>	56

DAFTAR ISTILAH

1. *SVM (Support Vector Machine)*
Algoritma machine learning untuk klasifikasi yang bekerja dengan mencari hyperplane optimal yang memisahkan data dua kelas secara maksimal.
2. *Random Forest*
Algoritma ensemble learning berbasis kumpulan pohon keputusan (decision tree) yang menghasilkan klasifikasi berdasarkan voting mayoritas.
3. *Klasifikasi Sentimen*
Proses pengelompokan opini atau komentar menjadi dua kategori utama, yaitu sentimen positif dan negatif, berdasarkan isi teks.
4. *Media Sosial X*
Platform media sosial berbasis microblogging yang sebelumnya dikenal sebagai Twitter, digunakan sebagai sumber data opini publik.
5. *TF-IDF (Term Frequency-Inverse Document Frequency)*
Metode pembobotan kata berdasarkan frekuensi kata dalam dokumen dan jaraknya dalam keseluruhan dokumen, digunakan untuk merepresentasikan fitur teks.
6. *SMOTE (Synthetic Minority Oversampling Technique)*
Teknik oversampling yang digunakan untuk mengatasi ketidakseimbangan kelas dengan membuat data sintesis pada kelas minoritas.
7. *Preprocessing*
Tahapan awal pengolahan data teks meliputi *cleaning*, *tokenizing*, *stopword removal*, dan *stemming* untuk menyiapkan data sebelum klasifikasi.
8. *Stemming*
Proses mengubah kata turunan menjadi bentuk dasarnya untuk menyamakan kata-kata yang memiliki makna serupa.
9. *Stopword Removal*
Proses penghapusan kata-kata umum seperti "yang", "dan", "di" yang tidak memiliki kontribusi penting dalam klasifikasi teks.

10. *Tokenizing*

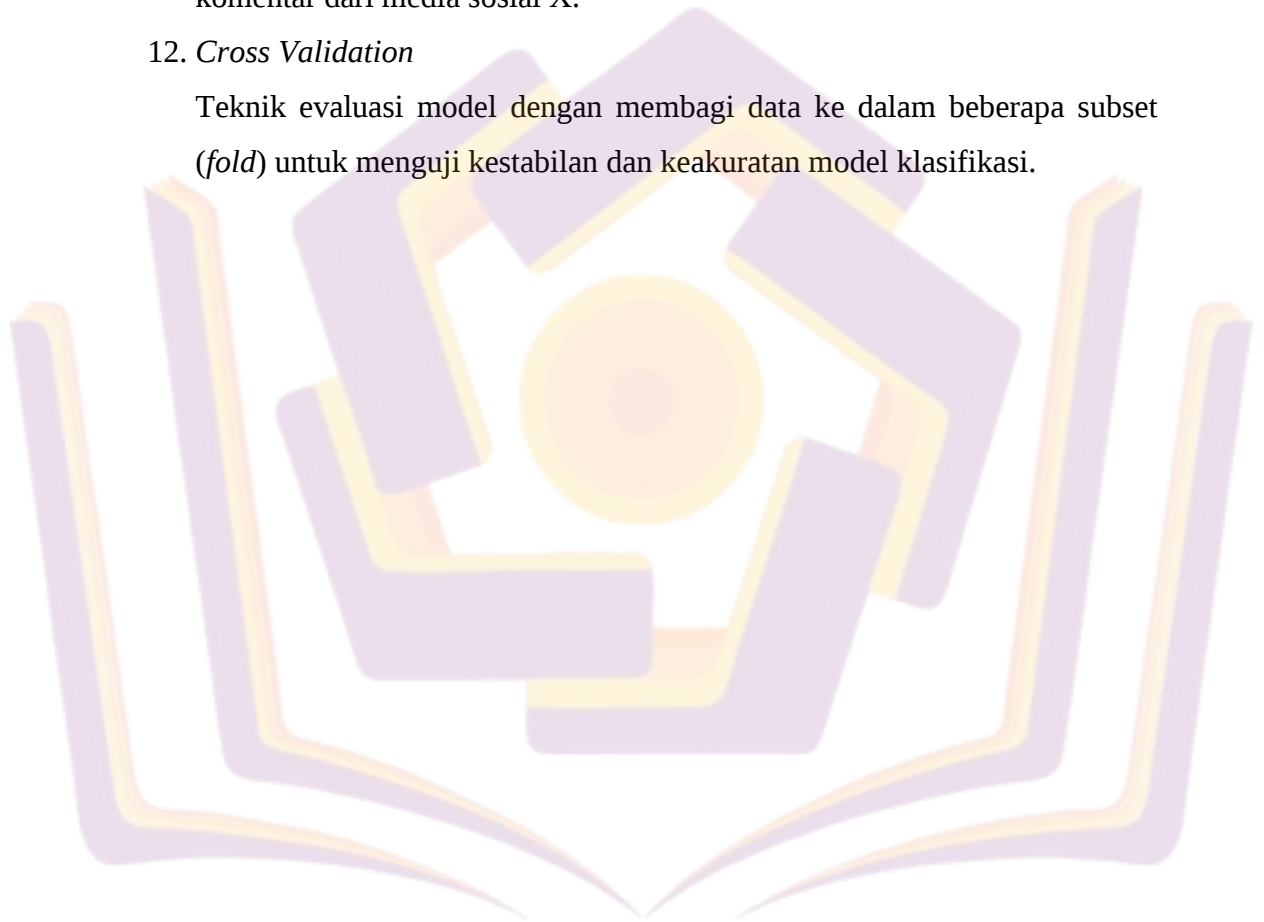
Proses memecah teks menjadi unit-unit kecil atau *token*, biasanya kata, sebagai langkah awal dalam pemrosesan data teks.

11. *Web Scraping*

Teknik pengambilan data secara otomatis dari situs web menggunakan program tertentu, dalam penelitian ini digunakan untuk mengumpulkan data komentar dari media sosial X.

12. *Cross Validation*

Teknik evaluasi model dengan membagi data ke dalam beberapa subset (*fold*) untuk menguji kestabilan dan keakuratan model klasifikasi.



DAFTAR LAMPIRAN

Lampiran 1. Kartu Bimbingan Dosen Pembimbing 1

Lampiran 2. Kartu Bimbingan Dosen Pembimbing 2

Lampiran 3. *Upload dan Load Dataset*

Lampiran 4. Preprocessing Data

Lampiran 5. Pelabelan dataset

Lampiran 6. TF-IDF Vectorization

Lampiran 7. SMOTE Oversampling

Lampiran 8. Cross Validation dan Evaluasi

