

## BAB 4

### HASIL PENELITIAN

#### 4.1 RINGKASAN HASIL PENELITIAN

Ringkasan hasil penelitian ini yaitu pengambilan data *tweet* dari Twitter menggunakan Jupyter Notebook dengan *keyword* perpindahan ibu kota baru menggunakan *library snsrape* selama 1 Juli 2020 hingga 2 Agustus 2024 yang berjumlah jumlah sebanyak 3.000 data. Menggunakan data mentah sebanyak 1000 data dari 3000 data mentah yang diperoleh. Setelah data mentah diperoleh dilanjutkan dengan tahap Preprocessing untuk membersihkan data dan dilabeli secara manual. Data yang digunakan pada proses training sebanyak 788 data *tweet*. Sedangkan untuk data testing yang digunakan sebanyak 198 *tweet* dengan rincian 119 data positif, dan 79 data negatif. Dari data training ini dilakukan pengujian, dan menghasilkan nilai akurasi sebesar 77,77%, sedangkan akurasi tahap testing sebesar 77,27%. Tahapan dalam penelitian ini adalah pengambilan data, preprocessing, training data, TF-IDF, testing data, dan klasifikasi.

#### 1.2 AKURASI ALGORITMA

Pada penelitian ini menggunakan data training sebanyak 788 data, dengan rincian 469 data berlabel positif dan 319 data berlabel negatif. Hasil pada perhitungan *confusion matrix* pada data training dapat dilihat pada tabel 4.1

**Tabel 4.1** Confusion Matrix Data Training

| Kelas Pediksi | Kelas Aktual |         |
|---------------|--------------|---------|
|               | Positif      | Negatif |
| Positif       | 389          | 157     |
| Negatif       | 80           | 162     |

Pada tabel 4.1 diatas menampilkan nilai TP sebesar 389, nilai FN sebesar 157, nilai FP sebesar 80, dan nilai TN sebesar 162. Hasil *accuracy* dan *f1-score* pada data *training* dapat dilihat pada tabel 4.2

**Tabel 4.2** Akurasi Data Training

| Akurasi Data <i>Training</i> | Hasil  |
|------------------------------|--------|
| <i>Accuracy</i>              | 77,27% |
| <i>F1-Score</i>              | 77,12% |

Pada tabel 4.2 diatas menampilkan tabel hasil perhitungan *confusion matrix* data *training* yang menghasilkan nilai *accuracy* sebesar 77,27% dan nilai *f-measure* bernilai 77,12%.

Tahap selanjutnya yaitu tahap testing, dengan menggunakan data *testing* sebanyak 198 data *tweet* dengan rincian 119 data berlabel positif, dan 79 data berlabel negatif. Berikut merupakan hasil dari *confusion matrix* data testing dapat dilihat pada tabel 4.3

**Tabel 4.3** *Confusion Matrix Data Testing*

| Kelas Pediksi | Kelas Aktual |         |
|---------------|--------------|---------|
|               | Positif      | Negatif |
| Positif       | 92           | 17      |
| Negatif       | 27           | 62      |

Pada tabel 4.3 diatas menghasilkan nilai TP sebesar 92, nilai FN sebesar 17, nilai FP sebesar 27, nilai TN sebesar 62. Nilai akurasi dari data testing dapat dilihat pada tabel 4.4

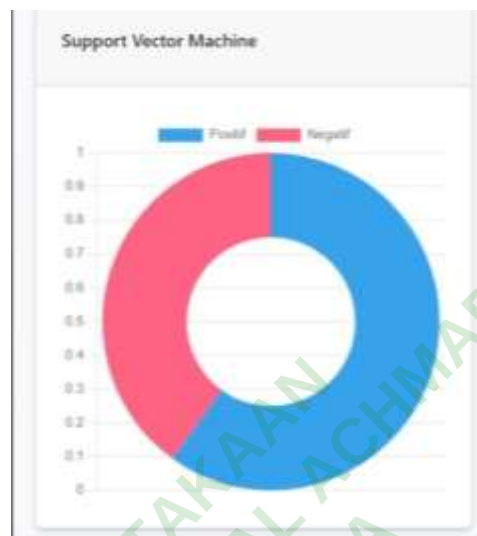
**Tabel 4.4** Akurasi Data Testing

| Akurasi Data Testing | Hasil  |
|----------------------|--------|
| <i>Accuracy</i>      | 77,77% |
| <i>F1-Score</i>      | 77,60% |

Pada tabel 4.4 diatas menampilkan tabel akurasi perhitungan pada data testing, menghasilkan nilai *accuracy* sebesar 77,77% dan nilai *f-measure* sebesar 77,60%.

### 4.3 EVALUASI DAN KLASIFIKASI

Hasil dari klasifikasi ini dibagi menjadi dua sentimen, yaitu sentiment positif dan sentimen negatif. Grafik hasil dari keseluruhan data dapat dilihat pada grafik lingkaran pada gambar 4.1



**Gambar 4.1** Grafik klasifikasi sentimen

Pada gambar 4.1 diatas menampilkan grafik hasil klasifikasi sentimen dengan hasil data *tweet* positif siap berjumlah 588, dan jumlah data sentimen negatif adalah 398.

### 4.4 FITUR-FITUR SISTEM

#### 4.4.1 Fitur Dashboard

Fitur Dashboard terdapat halaman Dasboard yang berisi Support Vector Machine, Confusion Matrix Data Training dan Confusion Matrix Data Testing. Tampilan halaman ini dapat dilihat pada Gambar 4.2

**Gambar 4.3** Tampilan halaman Data

Pada Gambar 4.3 diatas terdapat Pengambilan 1000 data dari Twitter namun pada gambar di atas hanya dapat menampilkan 10 data saja.

### 4.4.2 Fitur Data

Fitur Data terdapat halaman yang berisi sebuah data – data yang telah diambil dari twitter yang dapat dilihat pada Gambar 4.3

**Gambar 4.3** Tampilan halaman Data

Pada Gambar 4.3 diatas terdapat Pengambilan 1000 data dari Twitter namun pada gambar di atas hanya dapat menampilkan 10 data saja.

#### 4.4.3 Fitur Cleansing

Fitur Cleansing sebuah halaman yang melakukan pembersihan data yang telah diunggah lebih dari satu kali. Tampilan Cleansing dapat dilihat pada Gambar 4.4



**Gambar 4.4** Tampilan halaman Cleansing

Pada Gambar 4.4 menampilkan halaman Cleansing yang berisikan data yang akan digunakan untuk tahap Tokenizing.

#### 4.4.4 Fitur Tokenizing

Fitur Tokenizing Tahap tokenizing berfungsi untuk menampilkan data berupa kata yang telah pecahkan atau melakukan *root* yang dapat dilihat pada Gambar 4.5

**Gambar 4.6** Tampilan halaman Stopword Remove

Pada Gambar 4.6 menampilkan halaman Stopword Remove yang berisi data yang tidak memiliki arti akan dihapus secara otomatis.

Fitur Stopword Remove fitur ini berfungsi untuk penghilang kata henti menghapus kata-kata yang tidak memiliki arti. Tampilan halaman ini dapat dilihat pada Gambar 4.6

pada Gambar 4.6

**Gambar 4.6** Tampilan halaman Stopword Remove

Pada Gambar 4.6 menampilkan halaman Stopword Remove yang berisi data yang tidak memiliki arti akan dihapus secara otomatis.



Pada Gambar 4.8 diatas terdapat sebuah halaman yang menampilkan halaman untuk melakukan pelabelan data yang akan digunakan proses data training dan testing.

#### 4.4.8 Fitur Training Data

Fitur Training data adalah halaman yang berisikan hasil data TF-IDF yang telah diambil. Tampilan Training Data dapat di lihat pada Gambar 4.9

**Gambar 4.9** Tampilan halaman Training Data

Pada Gambar 4.9 diatas terdapat sebuah tampilan hasil dari training data yang meliputi hasil data trainingnya 788.

#### 4.4.9 Fitur Testing Data

Fitur Testing data yaitu sebuah halaman yang menampilkan halaman hasil TF-IDF data testing. Tampilan testing data dapat dilihat pada Gambar 4.10

**Gambar 4.10** Tampilan Hasil Testing Data

Pada Gambar 4.10 berisi data testing berjumlah 198 data label, dengan 119 sentimen positif dan 79 sentimen negatif.

#### 4.4.10 Pengujian SVM Model

Fitur hasil menampilkan diagram lingkaran dari hasil perhitungan menggunakan SVM. Tampilan fitur hasil dapat dilihat pada gambar 4.11



**Gambar 4.11** Tampilan Hasil Pengujian Support Verctor Machine

Pada gambar 4.11 di atas menampilkan hasil perhitungan algoritma SVM beserta akurasi data training serta data testing.

PERPUSTAKAAN  
UNIVERSITAS JENDERAL ACHMAD YANI  
YOGYAKARTA