

## BAB 4

### IMPLEMENTASI DAN PENGUJIAN

#### 4.1. Implementasi

Tahap implementasi merupakan tahap setelah analisis dan perancangan sistem, hasil analisis akan di implementasikan ke dalam program agar selanjutnya bisa dilakukan pengujian. Bahasa pemrograman yang digunakan adalah *PHP*.

##### 4.1.1. Implementasi Hardware

Spesifikasi *hardware* yang digunakan pada implementasi dan pengujian dalam penelitian ini dapat dilihat pada tabel 4.1 *Hardware yang Digunakan*.

**Tabel 4.1 Hardware yang Digunakan**

<i>Hardware</i>	Spesifikasi
<i>Processor</i>	<i>AMD Ryzen 3 – 1200</i>
<i>RAM</i>	8 GB
<i>Harddisk</i>	1 TB
<i>Monitor</i>	19" (1366 x 768)
<i>Keyboard &amp; Mouse</i>	<i>Standard</i>

##### 4.1.2. Implementasi Software

Spesifikasi *software* yang digunakan pada implementasi dan pengujian dalam penelitian ini dapat dilihat pada tabel 4.2 *Software yang Digunakan*.

**Tabel 4.2 Software yang Digunakan**

Software	Spesifikasi
<i>Operating System</i>	<i>Microsoft Windows 10</i>
<i>Development Kit</i>	XAMPP

### 4.1.3. Implementasi Database

Basis data yang dibangun pada sistem disesuaikan dengan perancangan basis data yang telah dirancang sebelumnya, basis data yang dibangun menggunakan MySQL. Pada sistem ini digunakan tabel data training, tabel kelas, tabel data uji, tabel akurasi, tabel akurasi\_ig, tabel ekstraksi fitur dan tabel information gain yang digunakan untuk mengolah data.

#### 1. Tabel Data Training

Nama Tabel : data\_training

Jumlah Field : 6

Primary Key : id\_dokumen\_training

Tabel data\_training digunakan untuk menyimpan kumpulan dokumen yang akan dilakukan training. Kemudian tabel data\_training dibangun dengan atribut berupa id\_dokumen\_training, nama\_dokumen, text, type, tahun dan status seperti pada Tabel 4.3.

**Tabel 4.3 Tabel Data Training**

Nama Tabel	Query Tabel
data_training	<pre>create table "data_training"( `id_dokumen_training` int(12) primary key, `nama_dokumen` text, `text_dokumen` text, `type` enum('sampul', 'abstrak'), `tahun` varchar(5), `status` enum('0', '1'));</pre>

#### 2. Tabel Kelas

Nama Tabel : kelas

Jumlah Field : 2

Primary Key : id\_kelas

Tabel kelas digunakan untuk menyimpan data kelas kategori untuk ekstraksi yang akan digunakan pada proses pengklasifikasian data uji dan data training. Tabel kelas dibangun dengan atribut berupa id\_kelas, dan nama\_kelas seperti pada Tabel 4.4.

**Tabel 4.4 Tabel Kelas**

Nama Tabel	Query Tabel
kelas	<pre>create table "kelas"( `id_kelas` int(10) primary key, `nama_kelas` varchar(255));</pre>

### 3. Tabel Data Uji

Nama Tabel : data\_uji

Jumlah Field : 6

Primary Key : id\_dokumen\_training

Tabel data\_uji digunakan untuk menyimpan kumpulan dokumen yang akan dilakukan pengujian. Kemudian tdata\_uji dibangun dengan atribut berupa id\_dokumen\_uji, nama\_dokumen, text, type, tahun dan status seperti pada Tabel 4.5.

**Tabel 4.5 Tabel Data Uji**

Nama Tabel	Query Tabel
data_uji	<pre>create table "data_uji"( `id_dokumen_uji` int(12) primary key, `nama_dokumen` text, `text_dokumen` text, `type` enum('sampul', 'abstrak'),</pre>

	<pre> `tahun` varchar(5),  `status` enum('0', '1'); </pre>
--	--

#### 4. Tabel Hasil Uji

Nama Tabel : tb\_hasil\_uji

Jumlah Field : 6

Primary Key : id

Tabel tb\_hasil\_uji digunakan untuk menyimpan data hasil pengujian. Kemudian tb\_hasil\_uji dibangun dengan atribut berupa id\_hasil\_uji, data, kelas\_hasil\_uji dan kelas\_sebenarnya seperti pada Tabel 4.6.

**Tabel 4.6 Tabel Hasil Uji**

Nama Tabel	Query Tabel
tb_data_uji	<pre> create table "tb_hasil_uji"( `id_hasil_uji` int12) primary key, `data` text, `kelas_hasil_uji` integer(10), `kelas_sebenarnya` integer(10)); </pre>

#### 4.1.4. Implementasi Sequential Query Language (SQL)

Sequential Query Language (SQL) yang digunakan pada pembangunan sistem disesuaikan dengan kebutuhan sistem yang telah dirancang sebelumnya. Fungsi-fungsi SQL yang digunakan menggunakan MySQL. Nama fungsi dan sql yang digunakan pada pembangunan sistem pengklasifikasi kategori sampel dan abstrak dapat dilihat pada Tabel 4.7.

**Tabel 4.7 Implementasi SQL**

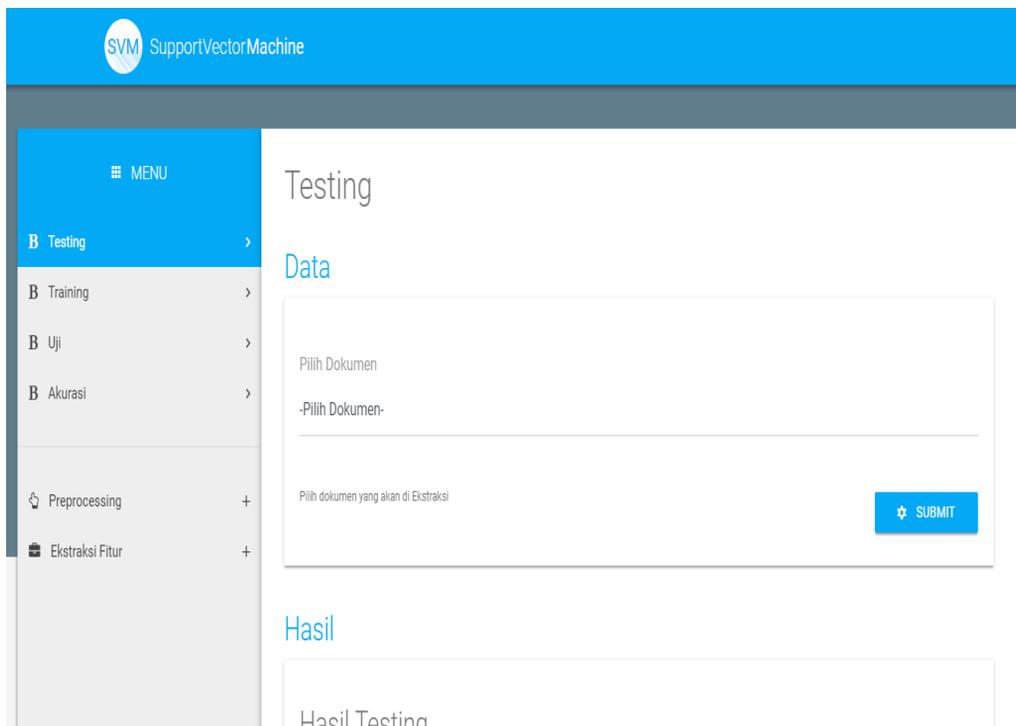
Nama Fungsi	Query
Menambah Data Dokumen	<pre>\$sql = "INSERT INTO tb_dokumen (id_dokumen, nama_dokumen, text_dokumen, tipe, tahun, status) VALUES ( '\$.sid_dokumen.', '\$.nama_dokumen.', \$.stext_dokumen.', '\$.stipe.', '\$.stahun.', '\$.status.' )";</pre>
Menampilkan Semua Dokumen	<pre>sql = "SELECT * FROM tb_dokumen";</pre>
Menampilkan Dokumen Yang Dipilih	<pre>sql = "SELECT * FROM tb_dokumen WHERE `id_dokumen` = '\$.sid_dokumen.' ";</pre>
Menambah Data Kelas	<pre>\$sql = "INSERT INTO tb_kelas (id_kelas, nama_kelas) VALUES ( '\$.sid_kelas.', '\$.nama_kelas.' )";</pre>
Menampilkan Semua Kelas	<pre>sql = "SELECT * FROM tb_kelas";</pre>
Menampilkan Kelas Yang Dipilih	<pre>sql = "SELECT * FROM tb_kelas WHERE `id_kelas` = '\$.sid_kelas.' ";</pre>
Mengupdate Dokumen Yang Sudah Dittraining	<pre>sql = "UPDATE tb_dokumen SET status = '1' WHERE id_dokumen = '\$.sid_dokumen.' ";</pre>

#### 4.1.5. Implementasi Antarmuka

Implementasi antarmuka sistem dibuat berdasarkan perancangan antarmuka pada bab 3. Pada perangkat lunak system Ekstraksi Informasi Karya Tulis Ilmiah Dengan SVM terdapat tujuh implementasi antarmuka, yaitu:

##### 4.1.5.1. Implementasi Antarmuka Halaman Testing

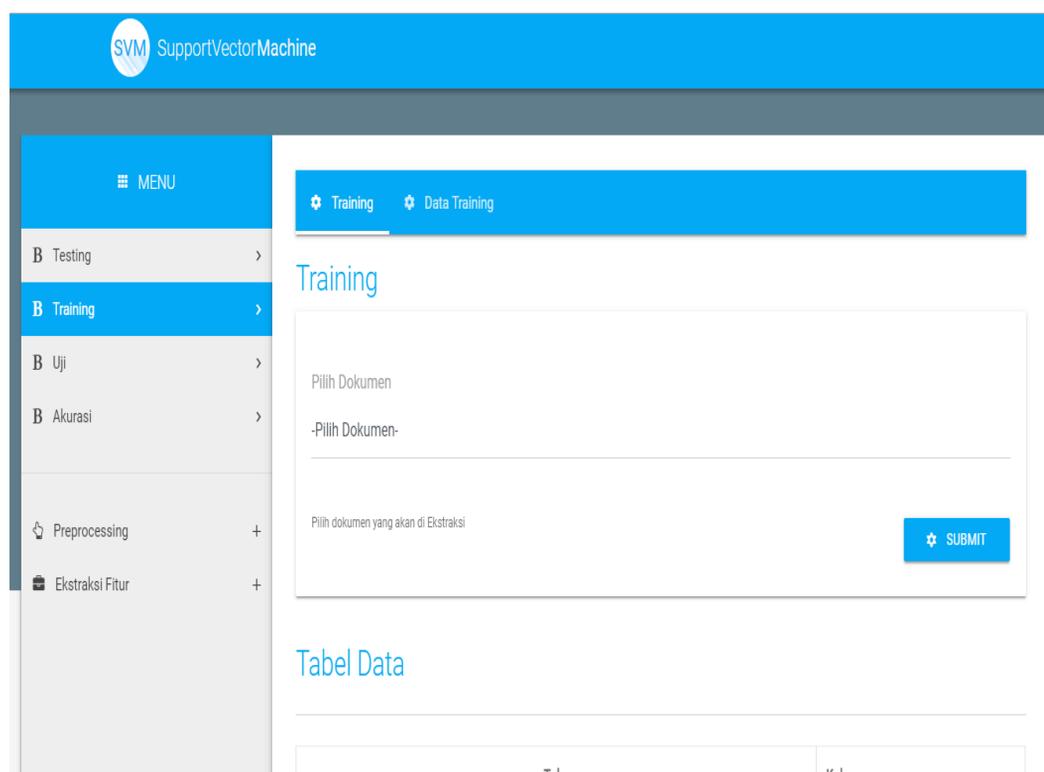
Antarmuka *testing* adalah tampilan awal pada saat membuka perangkat lunak. Pada implementasi antarmuka ini terdapat halaman nama sistem dan menu sistem berbentuk *navigator bar*. Terdapat tampilan untuk upload dokumen yang akan ditesing dan hasil dari testing dengan SVM. Implementasi antarmuka *Testing* dapat dilihat pada Gambar 4.1.



**Gambar 4.1** Implementasi Antarmuka Halaman Testing

#### 4.1.5.2. Implementasi Antarmuka Halaman Training

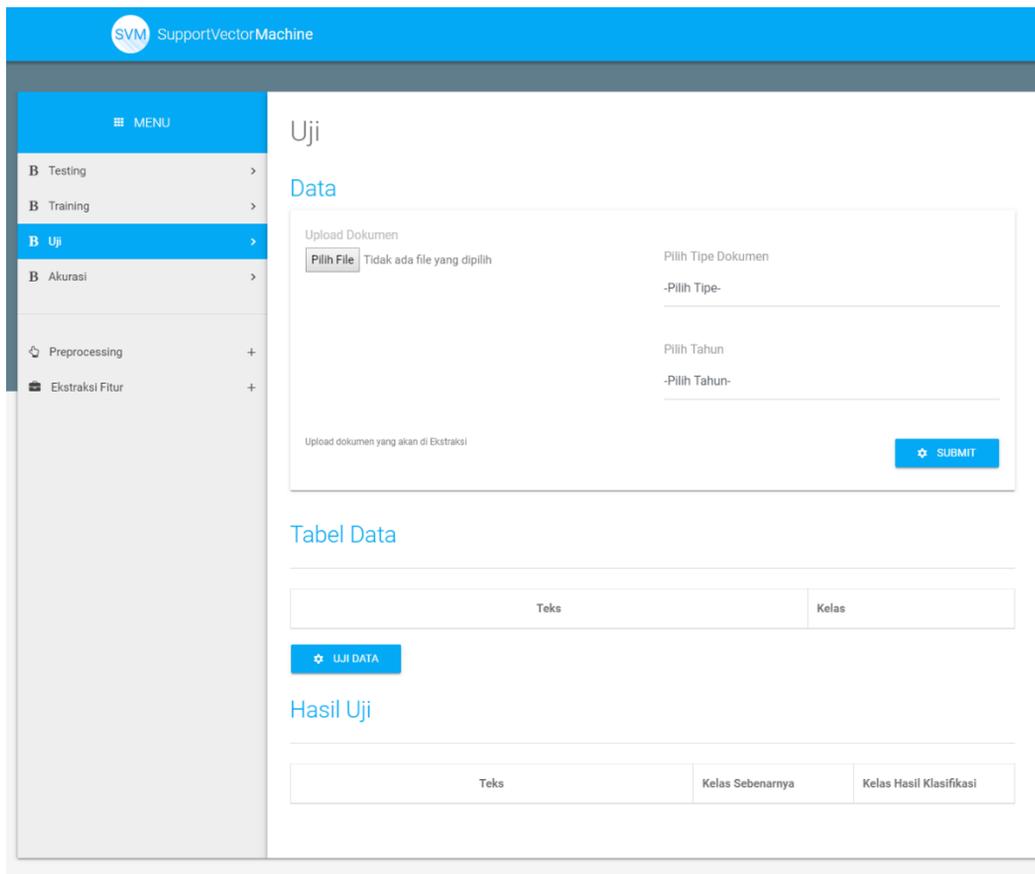
Antarmuka *training* adalah tampilan untuk melakukan proses training dokumen yang diupload. Pada implementasi antarmuka ini terdapat halaman nama sistem dan menu sistem berbentuk *navigator bar*. Terdapat tampilan upload dokumen pada database yang akan dilakukan proses training dan pelabelan setiap token hasil *preprocessing*. Implementasi antarmuka untuk halaman *training* dapat dilihat pada Gambar 4.2.



**Gambar 4.2 Implementasi Antarmuka Halaman Training**

#### 4.1.5.3. Implementasi Antarmuka Halaman Uji

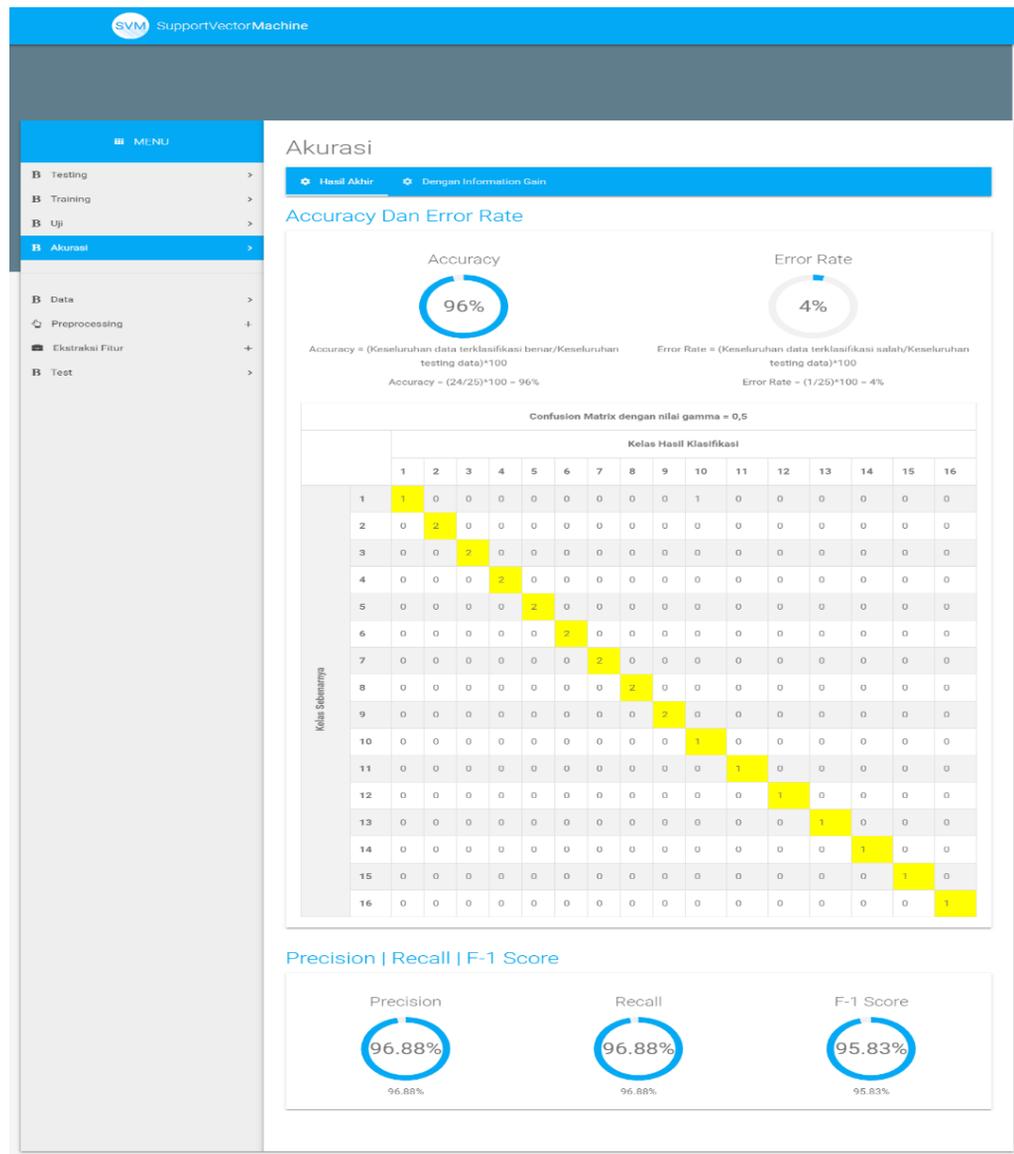
Antarmuka *uji* adalah tampilan untuk melakukan proses pengujian dokumen yang diupload. Pada implementasi antarmuka ini terdapat halaman nama sistem dan menu sistem berbentuk *navigator bar*. Terdapat tampilan upload dokumen pada database yang akan dilakukan proses pengujian dan pelabelan setiap token hasil *preprocessing*. Implementasi antarmuka untuk halaman *uji* dapat dilihat pada Gambar 4.3.



**Gambar 4.3 Implementasi Antarmuka Halaman Uji**

#### **4.1.5.4. Implementasi Antarmuka Halaman Akurasi**

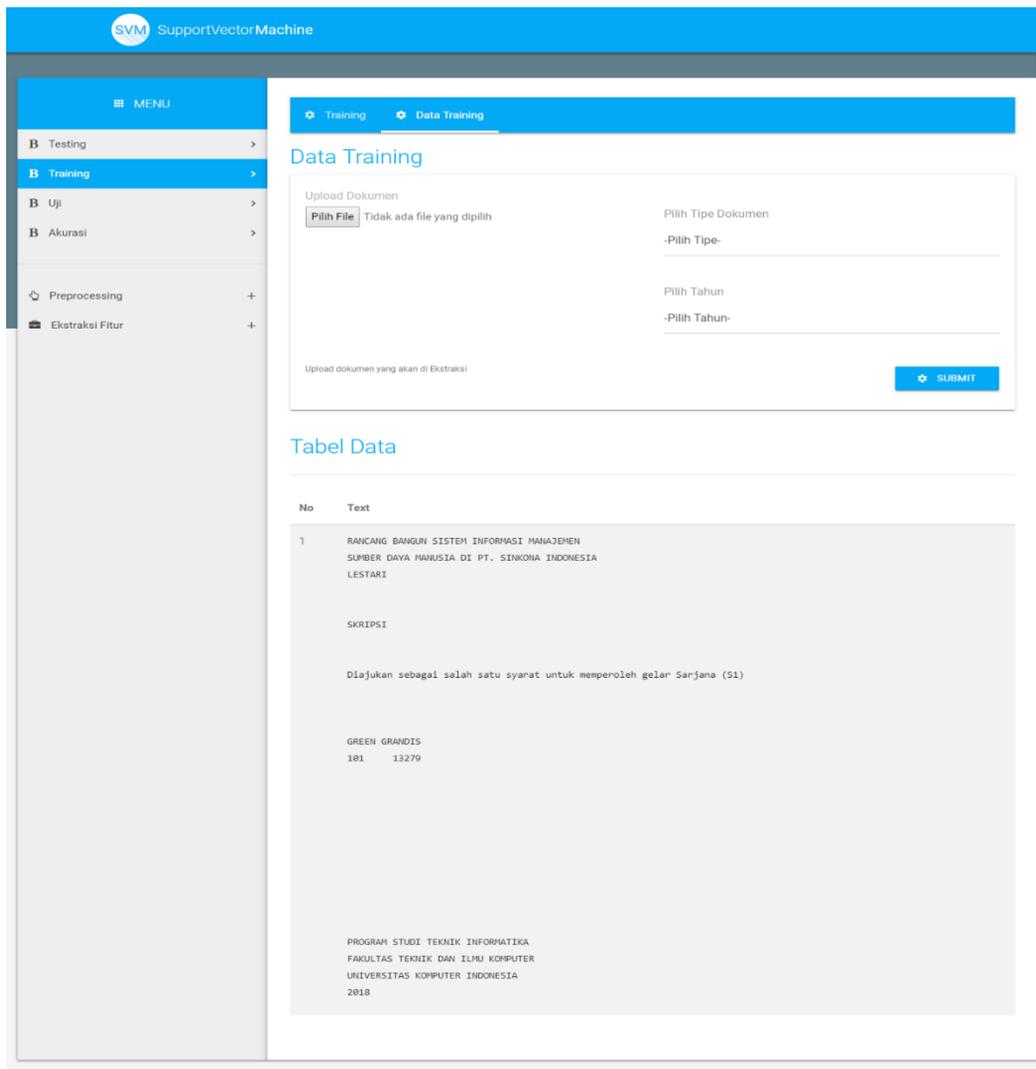
Antarmuka *akurasi* adalah tampilan untuk menampilkan hasil akurasi dari metode klasifikasi dengan svm dan information gain. Pada implementasi antarmuka ini terdapat halaman nama sistem dan menu sistem berbentuk *navigator bar*. Terdapat tampilan 2 panel untuk akurasi tanpa information gain dan dengan information gain. Implementasi antarmuka untuk halaman *akurasi* dapat dilihat pada Gambar 4.4.



**Gambar 4.4 Implementasi Antarmuka Halaman Akurasi**

#### 4.1.5.5. Implementasi Antarmuka Halaman Data Training

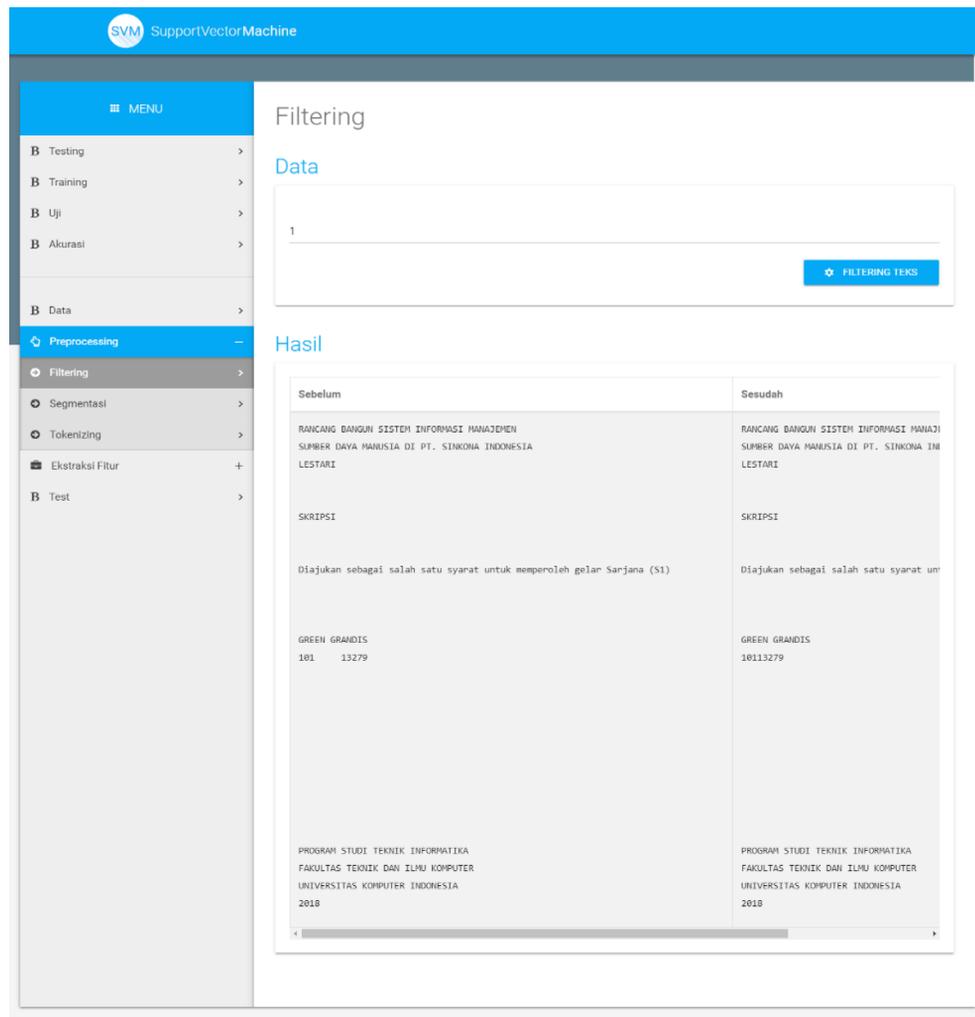
Antarmuka *data training* adalah tampilan untuk menampilkan semua data pada database data training. Pada implementasi antarmuka ini terdapat halaman nama sistem dan menu sistem berbentuk *navigator bar*. Terdapat tampilan upload dokumen pada database dan data yang ada pada database. Implementasi antarmuka untuk halaman *data training* dapat dilihat pada Gambar 4.5.



**Gambar 4.5 Implementasi Antarmuka Halaman Data Training**

#### 4.1.4.6. Implementasi Antarmuka Halaman Preprocessing

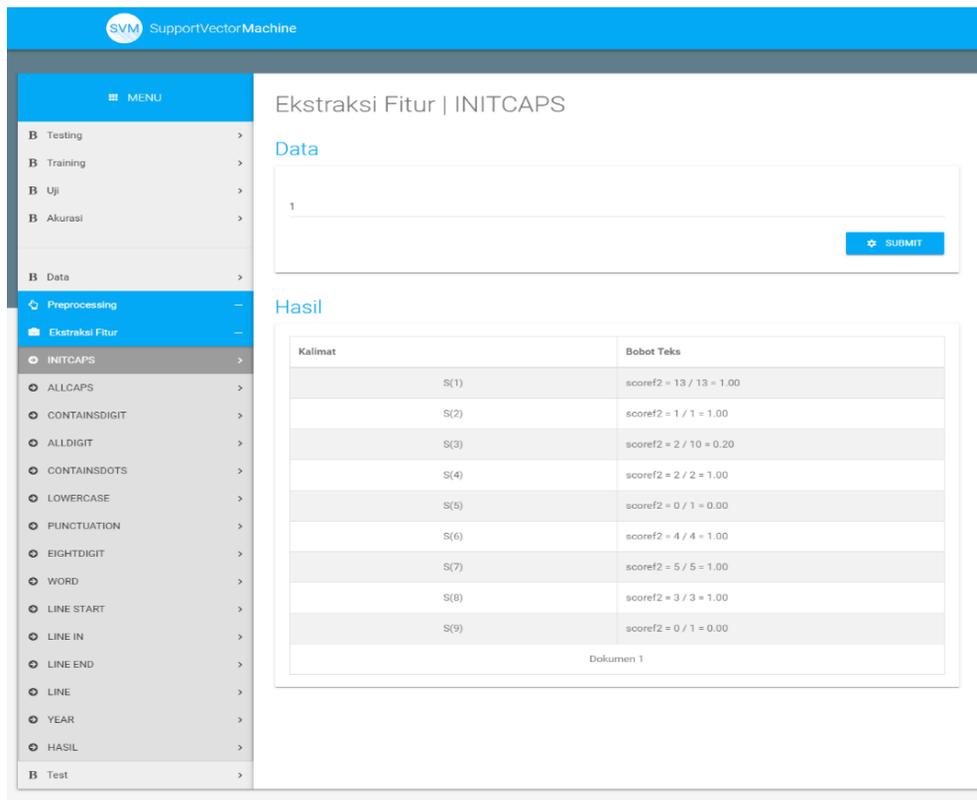
Antarmuka *preprocessing* adalah tampilan untuk melakukan proses *preprocessing*. Pada implementasi antarmuka ini terdapat halaman nama sistem dan menu sistem berbentuk *navigator bar*. Terdapat tampilan pilih jumlah dokumen pada database yang akan dilakukan proses *preprocessing* berdasarkan proses *preprocessing* yang dipilih. Implementasi antarmuka untuk halaman *preprocessing* dapat dilihat pada Gambar 4.6.



**Gambar 4.6 Implementasi Antarmuka Halaman Preprocessing**

#### 4.1.5.6. Implementasi Antarmuka Halaman Ekstraksi Fitur

Antarmuka *ekstraksi fitur* adalah tampilan untuk melakukan proses *ekstraksi fitur*. Pada implementasi antarmuka ini terdapat halaman nama sistem dan menu sistem berbentuk *navigator bar*. Terdapat tampilan pilih jumlah dokumen pada database yang akan dilakukan proses *ekstraksi fitur* berdasarkan proses *ekstraksi fitur* yang dipilih. Implementasi antarmuka untuk halaman *ekstraksi fitur* dapat dilihat pada Gambar 4.7.



**Gambar 4.7 Implementasi Antarmuka Halaman Ekstraksi Fitur**

#### 4.2. Pengujian Akurasi

Setelah dilakukan proses training dengan SVM, tahap lainnya adalah dilakukan pengujian terhadap hasil klasifikasi metode SVM. Pengujian dilakukan terhadap 40 dokumen uji yang terdiri dari 20 dokumen sampel dan 20 dokumen abstrak, setelah dilakukannya proses preprocessing didapatkan total keseluruhan sebanyak 322 dataset yang kemudian dilakukan pelabelan seperti pada Tabel 3.24. Setelah dilakukan proses pelabelan, hasil klasifikasi akan dihitung dengan menggunakan rumus *confusion matrix* untuk mencari akurasi keseluruhan dengan rumus 2.32 dan *error rate* dengan rumus 2.33. Berikut pada Tabel 3.26 adalah kelas hasil klasifikasi dan kelas sebenarnya sebagai pembandingan untuk melakukan pengujian.

Kemudian akan dihitung tingkat akurasi dengan hasil klasifikasi yang sesuai telah didapatkan sebanyak 292 data. Tingkat akurasi keseluruhan dihitung dengan rumus 2.32 sebagai berikut.



8	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0
12	0	0	6	0	0	0	0	0	0	0	0	15	0	0	0	0
13	0	0	0	0	0	0	0	0	0	0	0	0	19	0	0	0
14	0	0	0	0	0	0	0	0	0	0	0	0	1	20	0	0
15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20	0
16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20

Kemudian akan dihitung tingkat *precision*, *recall* dan *f-1 score* dengan rumus 2.34 sampai 2.36 berdasarkan hasil klasifikasi 16 kelas.

$$Precision = \frac{\frac{2}{20} + \frac{20}{20} + \dots + \frac{20}{20}}{16} * 100\% = \frac{15.006}{16} * 100\% = 93.79\%$$

$$Recall = \frac{\frac{2}{2} + \frac{20}{24} + \dots + \frac{20}{20}}{16} * 100\% = \frac{14.518}{5} * 100\% = 0.36 * 100\% \\ = 90.74 \%$$

$$f1 - Score = \frac{\frac{2 * 0.1 * 1}{0.1 + 1} + \frac{2 * 1 * 0.83}{1 + 0.83} + \dots + \frac{2 * 1 * 1}{1 + 1}}{16} * 100\% \\ = \frac{14.27}{16} * 100\% = 89.21 \%$$

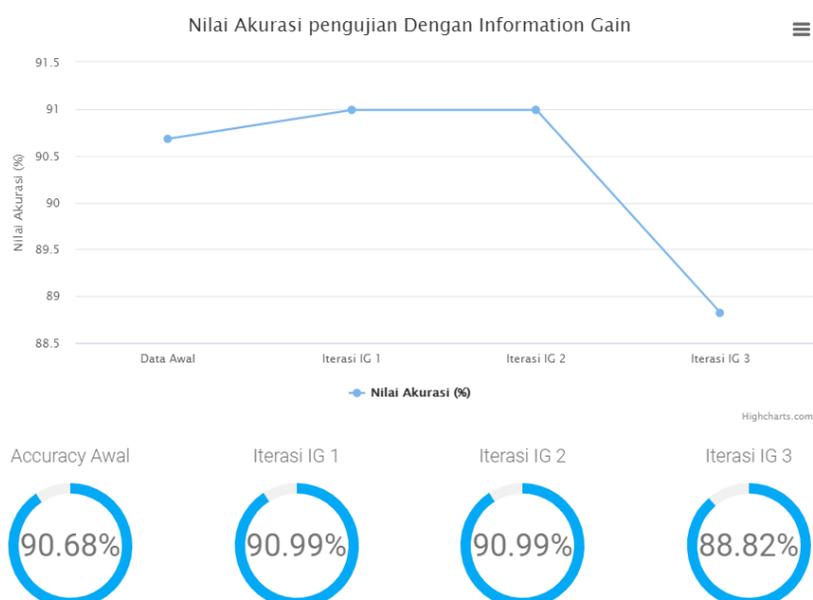
Didapatkan hasil dari pengujian dengan nilai  $\gamma=0,5$  adalah tingkat akurasi sebesar 90.68% sementara *error rate* sebesar 9.32% dan nilai *precision* sebesar 93.79% lalu nilai *recall* sebesar 90.74 % sedangkan nilai *f-1 score* sebesar 89.21%.

### 4.3. Pengujian Akurasi Dengan Information Gain

Setelah dilakukan proses klasifikasi tanpa seleksi fitur, kemudian dilakukan perhitungan akurasi dengan seleksi fitur information gain sampai akurasi lebih rendah dari akurasi sebelumnya. Untuk hasil akurasi dengan penyeleksian fitur dapat dilihat pada Tabel 4.9.

**Tabel 4.9 Akurasi Dengan Information Gain**

Seleksi Fitur	Akurasi Awal	Iterasi Ke – 1	Iterasi Ke – 2	Iterasi Ke – 3
Akurasi (%)	90,68 %	90,99 %	90,99 %	88,82 %



**Gambar 4.8 Hasil Akurasi Dengan Information Gain**

Dari hasil perhitungan akurasi dengan information gain diatas, didapatkan peningkatan akurasi sebesar 0,31 % pada iterasi IG ke-2 dan mengalami penurunan pada iterasi IG ke-3. Untuk perbandingan akurasi tanpa information gain dan akurasi dengan information gain dapat dilihat pada Tabel 4.10.

**Tabel 4.10 Perbandingan Tingkat Akurasi**

<b>Algoritma</b>	<b>Tanpa Information Gain</b>	<b>Dengan Information Gain</b>
<b>SVM</b>	90,68 %	90,99 %

#### **4.4. Kesimpulan Hasil Pengujian**

Hasil Pengujian yang diperoleh memiliki beberapa penyebab. Analisis yang dilakukan terhadap hasil pengujian dengan bentuk pengujian *Confusion Matrix* menyimpulkan bahwa yang pertama adalah karena data masukan yang digunakan oleh sistem adalah file teks, maka terdapat keterbatasan pada fitur yang digunakan. Kedua, berdasarkan pengamatan yang dilakukan, dampak yang mempengaruhi pada kelas hasil klasifikasi tidak sesuai disebabkan oleh fitur kelas Judul Penelitian (Sampul) dan Judul Halaman Abstrak yang memiliki beberapa data dengan pembobotan fitur yang serupa. Keempat, tidak adanya penggunaan ekstraksi fitur yang dapat membedakan kedua dokumen secara spesifik,. Sedangkan, berdasarkan pengamatan, nilai akurasi yang diperoleh paling besar adalah 90.99%.