

## BAB V

### PENGUJIAN DAN ANALISA

#### 5.1. Implementasi Pengujian

Pada pengujian implementasi algoritma SVM pada penentuan potensi bencana tsunami akibat gempa bumi ini menggunakan dataset dari *National Geophysical and Data Center* yang berjumlah 140 data. Terdapat beberapa macam pengujian yang dilakukan terhadap tingkat akurasi, yaitu:

1. Pengujian data latih terhadap tingkat akurasi di mana berfungsi untuk mengetahui seberapa besar jumlah data latih yang digunakan untuk memberikan hasil yang optimal. Pengujian ini dilakukan dengan menggunakan data latih yang berbeda.
2. Pengujian pengaruh nilai variabel atau parameter pada *Sequential Training* SVM. Pengujian ini dilakukan untuk menemukan kombinasi nilai variabel atau parameter sehingga mendapatkan akurasi terbaik.
3. Pengujian pengaruh penggunaan *kernel*. Pengujian ini dilakukan untuk menemukan *kernel* yang cocok digunakan untuk penentuan potensi tsunami sehingga mendapatkan akurasi terbaik.

##### 5.1.1. Pengujian Jumlah Data Latih terhadap Tingkat Akurasi

Pada pengujian kali ini menggunakan data latih yang berbeda. Jumlah data latih yang digunakan antara lain 50, 60, 70, 80, 90, dan 100. Sedangkan jumlah data uji yang digunakan 40 data dan isi data yang digunakan tetap. Nilai parameter *Sequential Training* yang digunakan tetap yaitu nilai  $\lambda = 0.1$ ,  $\gamma = 0.01$ ,  $C = 1$ , dan  $\epsilon = 0.0001$ , dan menggunakan kernel *Polynomial Degree*.

### 5.1.1.1. Hasil Pengujian Jumlah Data Latih terhadap Tingkat Akurasi

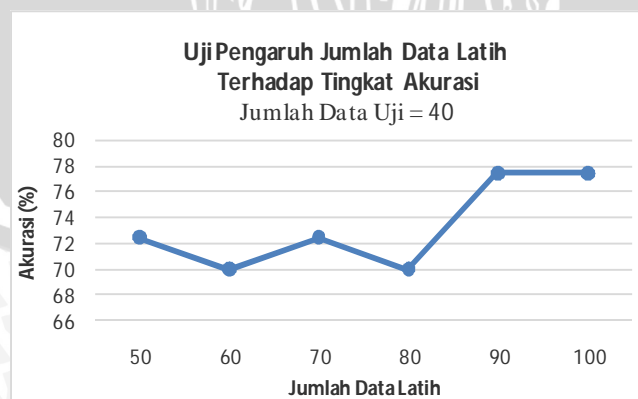
Berdasarkan Tabel 5.1 didapatkan akurasi yang berbeda di mana akurasi tertinggi didapatkan pada skenario 5 dan 6 yaitu dengan jumlah data latih 90 dan 100 dengan akurasi sebesar 77.50%.

Tabel 5.1 Hasil akurasi pengujian jumlah data latih terhadap tingkat akurasi

Skenario	Data Latih	Data Uji	Akurasi
1	50	40	72.5 %
2	60	40	70.0 %
3	70	40	72.5 %
4	80	40	70.0 %
5	90	40	77.5 %
6	100	40	77.5 %

### 5.1.1.2. Analisa Jumlah Data Latih terhadap Tingkat Akurasi

Pada analisa hasil pengaruh jumlah data latih terhadap akurasi ini digunakan data latih dengan jumlah yang berbeda-beda dan data uji yang sama. Berdasarkan skenario uji coba yang dilakukan pada semua data latih yang berbeda-beda tersebut didapatkan hasil akurasi tertinggi sebesar 77.5% di mana akurasi tersebut didapatkan pada skenario 5 dan 6 dengan menggunakan data latih masing-masing sebanyak 90 dan 100. Hal ini karena semakin banyak data latih yang digunakan maka dapat memberikan proses *learning* pada algoritma menjadi semakin baik. Hasil pengujian jumlah data latih terhadap jumlah akurasi ditampilkan dalam sebuah grafik yang dapat dilihat pada Gambar 5.1.



Gambar 5.1 Grafik Hasil Akurasi Pengujian Data Latih

### 5.1.2. Pengujian Pengaruh Nilai Variabel atau Parameter pada *Sequential Training SVM* terhadap Tingkat Akurasi

Pengujian ini dilakukan untuk mengetahui pengaruh nilai variabel pada perhitungan *Sequential Training SVM* terhadap tingkat akurasi. Data latih dan data uji yang digunakan masing-masing sebesar 100 dan 40 data. *Kernel* yang digunakan pada pengujian ini adalah *Polynomial Degree*.

#### 5.1.2.1. Hasil Pengaruh Nilai $\lambda$ (lambda) terhadap Tingkat Akurasi

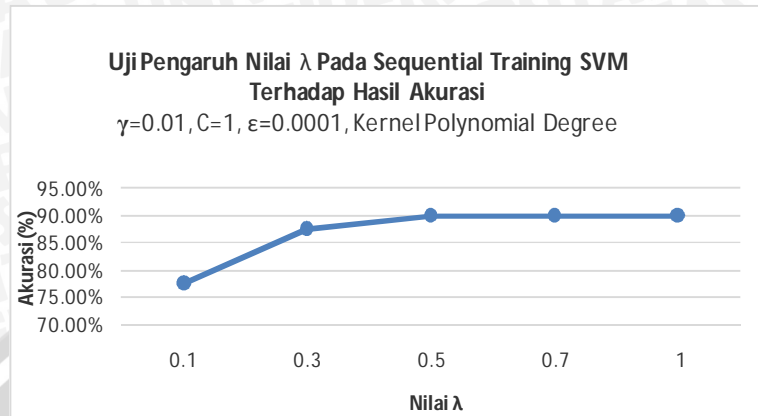
Pengujian nilai variabel  $\lambda$  (lambda) pada *Sequential Training SVM* dengan menggunakan nilai variabel tetap yaitu  $\gamma$  (gamma) = 0.01,  $C = 1$ ,  $\epsilon$  (epsilon) = 0.0001. Dari hasil pengujian yang ditunjukkan pada Tabel 5.2 didapatkan akurasi yang berbeda-beda di mana akurasi tertinggi didapatkan pada skenario 3, 4, dan 5 dengan nilai  $\lambda$  (lambda) yaitu 0.5, 0.7, dan 1 yang memiliki akurasi sebesar 90%.

Tabel 5.2 Hasil akurasi dari pengaruh nilai variabel  $\lambda$  (lambda).

Skenario	$\lambda$	$\gamma$	$\epsilon$	C	Akurasi
1	0.1	0.01	0.0001	1	77.5%
2	0.3	0.01	0.0001	1	87.5%
3	0.5	0.01	0.0001	1	90%
4	0.7	0.01	0.0001	1	90%
5	1	0.01	0.0001	1	90%

#### 5.1.2.2. Analisa Pengaruh Nilai $\lambda$ (lambda) terhadap Tingkat Akurasi

Pada analisa kali ini yaitu analisa nilai variabel atau parameter dari  $\lambda$  (lambda). Pada pengujian nilai parameter  $\lambda$  (lambda) yang telah dilakukan menghasilkan semakin besar nilai  $\lambda$  (lambda) maka akurasi yang dihasilkan cenderung meningkat. Hal ini dapat disebabkan karena nilai  $\lambda$  (lambda) digunakan untuk menambah jarak atau *space* antar data sehingga semakin besar nilai dapat memberikan kemudahan *hyperplane* dalam membedakan kelas positif dan negatif. Gambar 5.4 menunjukkan grafik pengaruh nilai hasil akurasi  $\lambda$  (lambda) terhadap tingkat akurasi.



Gambar 5.2 Grafik Hasil Akurasi Pengujian Nilai  $\lambda$

Sumber: [Pengujian]

### 5.1.2.3. Hasil Pengaruh Nilai Konstanta $\gamma$ (gamma) terhadap Tingkat Akurasi

Selanjutnya dilakukan pengujian terhadap nilai konstanta  $\gamma$  (gamma). Tabel 5.3 menunjukkan hasil akurasi dari pengujian nilai konstanta  $\gamma$  (gamma) pada *Sequential Training SVM* dengan menggunakan nilai variabel tetap yaitu  $\lambda$  (lambda) = 1,  $C = 1$ ,  $\epsilon$  (epsilon) = 0.0001

Tabel 5.3 Hasil akurasi dari pengaruh nilai konstanta  $\gamma$  (gamma).

Skenario	$\lambda$	$\gamma$	$\epsilon$	C	Akurasi
1	1	0.01	0.0001	1	90%
2	1	0.005	0.0001	1	90%
3	1	0.001	0.0001	1	90%
4	1	0.0005	0.0001	1	90%
5	1	0.0001	0.0001	1	92.5%

Berdasarkan Tabel 5.3 ditunjukkan bahwa hasil akurasi terbaik sebesar 92.5% pada skenario 5. Dengan demikian tingkat akurasi terbaik didapatkan dengan nilai konstanta  $\gamma$  (gamma) yaitu 0.0001.

#### 5.1.2.4. Analisa Pengaruh Nilai Konstanta $\gamma$ (gamma) terhadap Tingkat Akurasi

Pada pengujian parameter konstanta  $\gamma$  (gamma) didapatkan akurasi yang berbeda di mana akurasi yang menunjukkan penurunan didapatkan skenario 5. Pada metode *sequential training* SVM ini, konstanta  $\gamma$  (gamma) digunakan untuk mengontrol kecepatan proses *learning* di mana kecepatan tersebut ditunjukkan oleh jumlah iterasi yang dibutuhkan untuk mencapai konvergensi. Hasil pengujian menunjukkan bahwa semakin kecil nilai konstanta  $\gamma$  (gamma) maka iterasi yang diperlukan semakin banyak dan dapat memberikan akurasi yang lebih baik. Hal ini disebabkan karena nilai konstanta  $\gamma$  (gamma) mempengaruhi perhitungan  $\delta\alpha$ . Nilai  $\delta\alpha$  tersebut akan berkurang yang selisihnya tidak jauh dari nilai konstanta  $\gamma$  (gamma). Sehingga memerlukan proses iterasi lebih banyak untuk mencapai nilai  $\delta\alpha$  yang kurang dari  $\epsilon$  (epsilon). Tetapi dengan iterasi yang lebih banyak tersebut memberikan algoritma untuk dapat melakukan proses *learning* menjadi lebih baik. Sehingga dapat memberikan akurasi yang lebih baik.



Gambar 5.3 Grafik Hasil Akurasi Pengujian Konstanta  $\gamma$  (gamma)

#### 5.1.2.5. Hasil Pengaruh Nilai C terhadap Tingkat Akurasi

Selanjutnya dilakukan pengujian terhadap nilai C. Tabel 5.3 menunjukkan hasil akurasi dari pengujian nilai C pada *Sequential Training* SVM dengan menggunakan nilai variabel tetap yaitu  $\lambda$  (lambda) = 1,  $\gamma$  = 0.0001,  $\epsilon$  (epsilon) = 0.0001.

Tabel 5.4 Hasil akurasi dari pengaruh nilai C.

Skenario	$\lambda$	$\gamma$	$\epsilon$	C	Akurasi
1	1	0.0001	0.0001	10	92.5%
2	1	0.0001	0.0001	1	92.5%
3	1	0.0001	0.0001	0.1	92.5%
4	1	0.0001	0.0001	0.01	92.5%

Berdasarkan Tabel 5.4 ditunjukkan bahwa pada semua skenario menghasilkan akurasi yang sama. Tingkat akurasi yang didapatkan sebesar 92.5%.

#### 5.1.2.6. Analisa Pengaruh Nilai C terhadap Tingkat Akurasi

Pada pengujian parameter C didapatkan akurasi yang sama tetapi iterasi yang berhenti berbeda. Pada metode *sequential training SVM* ini, nilai C digunakan untuk mengontrol *tradeoff* antara *margin* dan *error* klasifikasi. Tetapi pada pengujian yang telah dilakukan, nilai C tidak berpengaruh terhadap akurasi. Hal ini disebabkan karena data yang diuji dengan nilai C pada pengujian tersebut memiliki *error* klasifikasi yang tidak berpengaruh pada *margin*. Sehingga akurasi yang didapatkan selalu sama.

#### 5.1.2.7. Hasil Pengaruh Nilai $\epsilon$ (epsilon) terhadap Tingkat Akurasi

Selanjutnya dilakukan pengujian terhadap nilai variabel  $\epsilon$  (epsilon). Tabel 5.5 menunjukkan hasil akurasi dari pengujian nilai variabel  $\epsilon$  (epsilon) pada *Sequential Training SVM* dengan menggunakan nilai variabel tetap yaitu  $\lambda$  (lambda) = 1,  $\gamma$  (gamma) = 0.0001, C = 0.01.

Tabel 5.5 Hasil akurasi dari pengaruh nilai variabel  $\epsilon$  (epsilon) 1.

Skenario	$\lambda$	$\gamma$	$\epsilon$	C	Akurasi
1	1	0.0001	0.0001	0.01	92.5%
2	1	0.0001	0.00001	0.01	92.5%
3	1	0.0001	0.000001	0.01	92.5%

Pada Tabel 5.5 akurasi yang dihasilkan sama. Sehingga dilakukan pengujian lagi dengan data latih dan data uji yang diambil secara acak dari *dataset*. Skenario pengujian tersebut ditunjukkan pada Tabel 5.6.

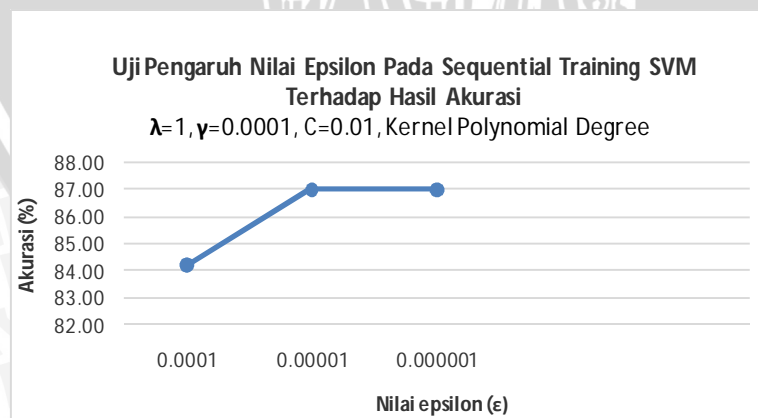
Tabel 5.6 Hasil akurasi dari pengaruh nilai variabel  $\epsilon$  (epsilon) 2.

Skenario	$\epsilon = 0.0001$	$\epsilon = 0.00001$	$\epsilon = 0.000001$
1	87.5%	80%	80%
2	80%	90%	90%
3	85%	92.5%	92.5%
4	90%	87.5%	87.5%
5	82.5%	82.5%	82.5%
6	80%	90%	90%
Rata-rata	84.17	87%	87%

Tabel 5.6 menunjukkan hasil akurasi terbaik sebesar 87% dengan nilai  $\epsilon = 0.00001$  dan  $\epsilon = 0.000001$ .

### 5.1.2.8. Analisa Pengaruh Nilai $\epsilon$ (epsilon) terhadap Tingkat Akurasi

Berdasarkan uji coba yang telah dilakukan ditunjukkan bahwa nilai variabel  $\epsilon$  (epsilon) yang semakin kecil menghasilkan akurasi yang semakin meningkat. Hal ini disebabkan karena nilai variabel  $\epsilon$  (epsilon) yang semakin kecil dapat menghasilkan nilai  $\alpha$  yang optimal sehingga dapat menentukan *support vector* yang tepat. Support vector yang tepat tersebut dapat menghasilkan *hyperplane* (garis pemisah) untuk memisahkan kedua kelas tersebut semakin optimal.



Gambar 5.4 Grafik Hasil Akurasi Pengujian Nilai  $\epsilon$  (epsilon)

### 5.1.3. Pengujian Pengaruh Jenis Kernel terhadap Tingkat Akurasi

Pada pengujian kali ini menggunakan 2 jenis kernel yaitu *Polynomial Degree* dan *Gaussian RBF*. Nilai parameter *Sequential Training* yang digunakan tetap yaitu dengan nilai  $\lambda = 1$ ,  $\gamma = 0.01$ ,  $C = 1$ , dan  $\epsilon = 0.00001$ . Jumlah data latih yang digunakan 100 data dan jumlah data uji yang digunakan 40 data tetapi isi data latih dan data uji tersebut didapatkan dari proses random dari *dataset* atau data secara keseluruhan yang berjumlah 140 data.

#### 5.1.3.1. Hasil Pengujian Pengaruh Jenis Kernel terhadap Tingkat Akurasi

Berdasarkan Tabel 5.7 ditunjukkan bahwa rata-rata akurasi yang dihasilkan oleh kernel *Polynomial Degree* sebesar 85.50%. Sedangkan rata-rata yang dihasilkan kernel *Gaussian RBF* sebesar 88.75%.

Tabel 5.7 Hasil akurasi dari pengaruh jenis *kernel*.

Uji Coba	<i>Polynomial Degree</i>	<i>Gaussian RBF</i>
1	77.50%	82.50%
2	85%	85%
3	92.50%	95%
4	77.50%	85%
5	80%	87.50%
6	90%	90%
7	92.50%	92.50%
8	77.50%	85%
9	95%	97.50%
10	87.50%	87.50%
<b>Rata-rata</b>	<b>85.50%</b>	<b>88.75%</b>

#### 5.1.3.2. Analisa Penujian Pengaruh Jenis Kernel terhadap Tingkat Akurasi

Pada tabel 5.7 menunjukkan bahwa *kernel Gaussian RBF* menghasilkan akurasi konstan diatas 80%. Dilihat dari rata-rata akurasinya, *kernel Gaussian RBF* lebih tinggi dari pada *kernel Polynomial Degree*. karena pada tabel 5.7 menunjukkan bahwa akurasi *kernel Gaussian RBF* selalu berada di atas 80% dan memiliki rata-rata akurasi yang lebih tinggi yaitu sebesar 88.75%. Hal ini disebabkan karena pada saat proses pemetaan data dalam dimensi yang lebih tinggi, *kernel Gaussian RBF* merepresentasikan sebaran data yang lebih baik dibandingkan *kernel Polynomial Degree*.