

*Laporan 5*

**PRAKTIKUM PEMBELAJARAN MESIN**

**“Support Vector Machine (SVM)”**



Asisten :

- 1. Muhammad Syawal Idil Fitrrah Baharuddin, S.Kom**
- 2. Nur Annisa Safitri Yusuf, S.Kom**

Oleh

Nama : Ahmad Ikbal Djaya

Nim : 60200120073

Kelas : A

**LABORATORIUM KOMPUTER TERPADU**

**JURUSAN TEKNIK INFORMATIKA**

**FAKULTAS SAINS DAN TEKNOLOGI**

**UNIVERSITAS ISLAM NEGERI ALAUDDIN MAKASSAR**

**2023**

## **LEMBAR PENGESAHAN/ACC PRAKTIKUM**

Laporan praktikum “PEMBELAJARAN MESIN” dengan judul “**Support Vector Machine (SVM)**” yang disusun oleh :

Nim : 60200120073

Nama : Ahmad Ikbal Djaya

Kelas : A

Telah diperiksa dan dikonsultasikan kepada Asisten I dan Asisten II maka dinyatakan diterima.

Gowa, Juni 2023

**Asisten I**

**Asisten II**



Muhammad Syawal Idil Fitrah Baharuddin, S.Kom

Nur Annisa Safitri Yusuf, S.Kom

Mengetahui,  
**Laboran Terkait**

**Abdurrahman Ramli, S.Kom**

### LEMBAR PENILAIAN LAPORAN PRAKTIKUM

Nama : Ahmad Ikbal Djaya

NIM : 60200120073

Praktikum : Pembelajaran Mesin

Judul praktikum : **Support Vector Machine (SVM)**

No	Komponen penilaian	Nilai
I.	Nilai Pelaksanaan Asistensi ke 1	(100)
	A. Kedisiplinan	(20)
	B. Keterampilan menggunakan Komputer/Software	(20)
	C. Keaktifan	(20)
	D. Nilai laporan	(40)
II.	Nilai Pelaksanaan Asistensi ke 2	(100)
	A. Kedisiplinan	(20)
	B. Keterampilan menggunakan Komputer/Software	(20)
	C. Keaktifan	(20)
	D. Nilai laporan	(40)
III.	Total Nilai laporan praktikum (Asistensi 1 + Asistensi 2)/2	(100) 80

Tanggal praktikum : Juni 2023

Dikumpulkan tanggal : Juni 2023

Dikoreksi tanggal : Juni 2023

Asisten Praktikum,



**Muhammad Syawal Idil Fitrrah Baharuddin, S.Kom**

## **BAB V**

### ***SUPPORT VECTOR MACHINE (SVM)***

#### ***A. Tujuan Praktikum***

Mahasiswa mampu mengetahui *Support Vector Machine (SVM)* dan mampu mengimplementasikannya dalam klasifikasi.

#### ***B. Teori Dasar***

Machine Learning dianggap sebagai subfield dari Kecerdasan Buatan dan berkaitan dengan pengembangan teknik dan metode yang memungkinkan komputer untuk belajar. Secara sederhana, pengembangan algoritma yang memungkinkan mesin untuk belajar dan melakukan tugas dan aktivitas. Machine learning memiliki banyak persamaan dengan statistika. Seiring berjalannya waktu, banyak teknik dan metodologi yang dikembangkan untuk tugas machine learning (Jakkula, 2011).

Support Vector Machine (SVM) pertama kali diperkenalkan pada tahun 1992 oleh Boser, Guyon, dan Vapnik di COLT-92. Support vector machines (SVM) adalah sekumpulan metode pembelajaran terawasi yang digunakan untuk klasifikasi dan regresi. SVM termasuk dalam keluarga klasifikasi linier generalisasi. Dalam kata lain, Support Vector Machine (SVM) adalah alat klasifikasi dan prediksi regresi yang menggunakan teori machine learning untuk memaksimalkan akurasi prediksi sambil secara otomatis menghindari over-fit terhadap data. Mesin vektor pendukung dapat didefinisikan sebagai sistem yang menggunakan ruang hipotesis dari fungsi linier dalam ruang fitur berdimensi tinggi, dilatih dengan algoritma pembelajaran dari teori optimisasi yang menerapkan bias pembelajaran yang berasal dari teori pembelajaran statistik. SVM menjadi terkenal ketika, dengan menggunakan peta piksel sebagai input, memberikan akurasi yang sebanding dengan jaringan saraf yang rumit dengan fitur yang lebih rumit dalam tugas pengenalan tulisan tangan. SVM juga digunakan untuk banyak aplikasi, seperti analisis tulisan tangan, analisis wajah, dan sebagainya, terutama untuk klasifikasi pola dan aplikasi berbasis regresi. Dasar-dasar Support Vector Machines

(SVM) telah dikembangkan oleh Vapnik dan menjadi populer karena banyak fitur menjanjikan seperti kinerja empiris yang lebih baik. Formulasi ini menggunakan prinsip Structural Risk Minimization (SRM), yang terbukti lebih unggul, dibandingkan dengan prinsip Empirical Risk Minimization (ERM) yang digunakan oleh jaringan saraf konvensional. SRM meminimalkan batas atas risiko yang diharapkan, sedangkan ERM meminimalkan kesalahan pada data pelatihan. Perbedaan ini memberikan SVM kemampuan yang lebih besar dalam generalisasi, yang merupakan tujuan dalam pembelajaran statistik. SVM awalnya dikembangkan untuk memecahkan masalah klasifikasi, tetapi belakangan ini telah diperluas untuk memecahkan masalah regresi (Jakkula, 2011).

Machine Learning dianggap sebagai subfield dari Kecerdasan Buatan dan berkaitan dengan pengembangan teknik dan metode yang memungkinkan komputer untuk belajar. Secara sederhana, pengembangan algoritma yang memungkinkan mesin untuk belajar dan melakukan tugas dan aktivitas. Machine learning memiliki banyak persamaan dengan statistika. Seiring berjalannya waktu, banyak teknik dan metodologi yang dikembangkan untuk tugas machine learning .

Klasifikasi merupakan operasi untuk memisahkan beragam entitas kedalam beberapa kelas. Pengklasifikasian merupakan pelatihan atau pembelajaran terhadap fungsi  $f$  (target/label) yang memetakan setiap atribut  $x$  ke satu dari jumlah label kelas  $y$  yang tersedia. Apabila target kelas sudah diketahui maka proses klasifikasi termasuk dalam supervised dan apabila dataset belum memiliki target kelas maka termasuk unsupervised contohnya adalah proses kluster (Sihombing, 2020).

Support Vector Machines (SVM) adalah algoritma supervised learning yang sangat kuat yang digunakan untuk klasifikasi atau regresi. SVM dapat digunakan untuk prediksi numerik serta klasifikasi. SVM adalah algoritma yang berfungsi untuk pemetaan nonlinier dan mengubah data pelatihan asli ke dimensi yang lebih tinggi. Di dalam dimensi baru ini, SVM mencari linear optimal dengan memisahkan hyperplane. Dengan pemetaan nonlinier yang sesuai ke dalam dimensi yang cukup tinggi, data dari dua kelas selalu dapat dipisahkan oleh hyperplane. SVM

menemukan hyperplane tersebut dengan menggunakan vektor dukungan dan margin. Dua sifat khusus dari SVM yaitu (1) mencapai generalisasi yang tinggi dengan memaksimalkan margin, dan (2) mendukung pembelajaran yang efisien dari fungsi nonlinier pada trik kernel sehingga membuat kinerja generalisasinya baik dalam menyelesaikan masalah pengenalan pola. Pada permasalahan klasifikasi SVM mencoba untuk mencari garis pemisah yang optimal yang diekspresikan sebagai kombinasi linier dari subset data training dengan menyelesaikan masalah keterbatasan linier pemrograman kuadrat (QP) dengan margin maksimum antara dua kelas (Sihombing, 2020).

Pengklasifikasi SVM dilakukan dengan prosedur dua langkah: Pertama, vektor data sampel dipetakan ("diproyeksikan") ke ruang dimensi. Dimensi ruang ini secara signifikan lebih besar dari dimensi ruang data asli. Kemudian, algoritma digunakan untuk menemukan hyperplane di ruang ini dengan margin terbesar yang memisahkan kelas data. Hal itu menunjukkan bahwa keakuratan klasifikasi biasanya hanya akan turun tipis pada proyeksi tertentu. Ide dasar SVM adalah memaksimalkan batas hyperplan (maximal margin hyperplane). Konsep klasifikasi dengan SVM dapat dijelaskan secara sederhana sebagai usaha untuk mencari hyperplane terbaik yang berfungsi sebagai pemisah dua buah kelas data pada ruang input (Sihombing, 2020).

SVM merupakan salah satu metode untuk mengklasifikasikan data dimana memilah data menggunakan hyperplanes [6]. SVM dapat digunakan untuk menghasilkan beberapa hyperplanes sehingga data dibagi menjadi beberapa bagian dan setiap bagian hanya berisi satu jenis data. Metode SVM umumnya berguna sekali untuk data yang memiliki non-regularity atau data yang distribusinya tidak diketahui [7]. Konsep SVM menitikberatkan pada risk minimization, yaitu untuk mengestimasi suatu fungsi dengan cara meminimalkan batas atas dari generalization error, sehingga SVM mampu mengatasi overfitting (Karim, 2020).

### C. Alat dan Bahan

1. Laptop/PC
2. Google Drive
3. Google Colaboratory
4. Mendeley

### D. Hasil dan Pembahasan

#### Soal:

Implementasikan *support vector machine* pada *python* dan terapkan semua kernel yang ada pada *support vector machine*.

#### Pembahasan:

Langkah – langkah melakukan klasifikasi menggunakan *SVM*

1. Buka *google Colaboratory*
2. Langkah pertama Memuat dataset kanker payudara menggunakan *scikit-learn's `datasets.load\_breast\_cancer()`*.

```
from sklearn import datasets  
cancer = datasets.load_breast_cancer()
```

**Gambar 1.** Import dataset dari *library*

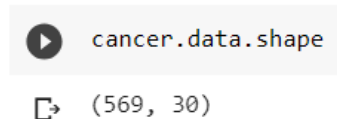
3. Selanjutnya mencetak nama fitur dan label untuk mengetahui variabel *feature/independen* dan nama target.

```
print("Features: ", cancer.feature_names)  
print("Labels: ", cancer.target_names)
```

Features: ['mean radius' 'mean texture' 'mean perimeter' 'mean area'  
'mean smoothness' 'mean compactness' 'mean concavity'  
'mean concave points' 'mean symmetry' 'mean fractal dimension'  
'radius error' 'texture error' 'perimeter error' 'area error'  
'smoothness error' 'compactness error' 'concavity error'  
'concave points error' 'symmetry error' 'fractal dimension error'  
'worst radius' 'worst texture' 'worst perimeter' 'worst area'  
'worst smoothness' 'worst compactness' 'worst concavity'  
'worst concave points' 'worst symmetry' 'worst fractal dimension']  
Labels: ['malignant' 'benign']

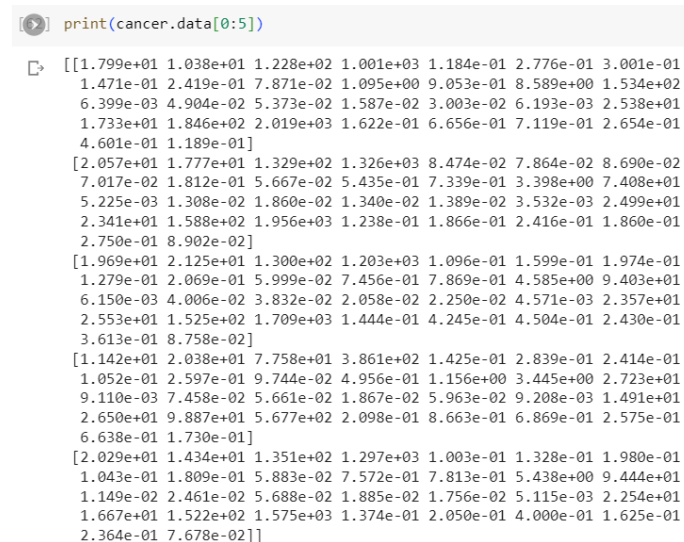
**Gambar 2.** Mencetak nama *feature* dan target

4. Menampilkan bentuk data untuk memeriksa dimensi datanya (Baris dan Kolom).



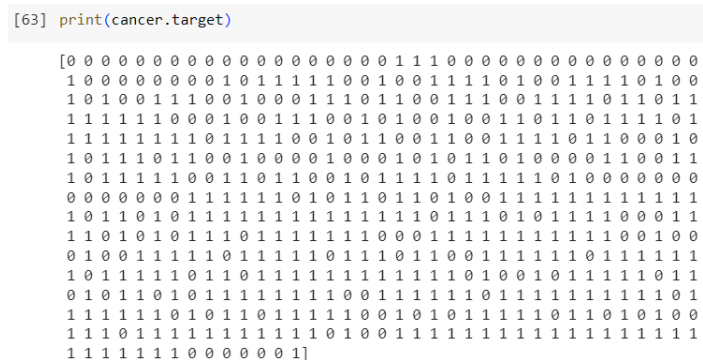
**Gambar 3.** Menampilkan dimensi data

5. Mencetak lima sampel pertama dari data yang memiliki 569 baris dan 30 kolom akan dilihat 5 data pertama dari variabel “*features*”



**Gambar 4.** Mencetak 5 data pertama

6. Mencetak label target untuk mengetahui dan memberikan kelas untuk mengidentifikasi kategori dalam melatih model (0 = *Malignant*, 1 = *Benign*).



**Gambar 5.** Mencetak label target



7. Memisahkan dataset menjadi data latih dan data uji menggunakan `train_test_split()` dari `sklearn.model_selection`. Disini data dibagi sebanyak 75% : 25% dimana 75% untuk data latih dan 25% untuk data uji.

```
from sklearn.model_selection import train_test_split
x_train, x_test, y_train, y_test = train_test_split(cancer.data, cancer.target, test_size=0.25, random_state=123)
```

**Gambar 6.** Split data

8. Membuat objek `svm.SVC` dengan kernel *linear*. Lalu Melatih model SVM menggunakan data latih menggunakan `clf.fit(x_train, y_train)`. Kemudian Memprediksi label untuk data uji menggunakan `y_pred = clf.predict(x_test)`.

```
from sklearn import svm
clf = svm.SVC(kernel='linear')
clf.fit(x_train, y_train)
y_pred = clf.predict(x_test)
```

**Gambar 7.** Membuat objek klasifikasi

9. Menghitung matriks kebingungan (*confusion matrix*) agar memudahkan dalam mengetahui apakah terdapat kesalahan dalam pengklasifikasiannya dan dari *output* yang dihasilkan dapat dilihat bahwa terdapat 2 kesalahan dalam pengklasifikasian menggunakan algoritma SVM, yaitu 2 pasien diklasifikasikan dalam pasien pengidap kanker jinak, tetapi dalam keadaan sebenarnya, pasien mengalami kanker ganas.

```
from sklearn.metrics import confusion_matrix
confusion_matrix(y_test, y_pred)

array([[52,  2],
       [ 0, 89]])
```

**Gambar 8.** Menghitung *confusion matrix*

10. Langkah terakhir yaitu Mencetak laporan klasifikasi (*classification report*) untuk melihat akurasi dari hasil pengklasifikasian menggunakan algoritma SVM. Akurasi disini yaitu proporsi jumlah prediksi benar.

```

▶ from sklearn.metrics import classification_report
  print(classification_report(y_test, y_pred))

```

	precision	recall	f1-score	support
0	1.00	0.96	0.98	54
1	0.98	1.00	0.99	89
accuracy			0.99	143
macro avg	0.99	0.98	0.99	143
weighted avg	0.99	0.99	0.99	143

**Gambar 9.** Mencetak *classification report*

Kernel linear adalah salah satu jenis kernel yang digunakan dalam algoritma *Support Vector Machines* untuk pemrosesan data. Kernel *linear* cocok untuk dataset yang memiliki pemisahan yang linear antar kelas. Jika dataset memiliki pola pemisahan yang dapat dijelaskan dengan garis lurus atau *hipersuperplanar*, kernel *linear* dapat memberikan hasil yang baik dalam klasifikasi.

11. Kemudian lakukan hal yang sama seperti diatas dengan mencoba mengganti kernel yang digunakan, disini kernelnya akan diganti menjadi *Radial Basis Function*.

```

[87] from sklearn import svm
      clf = svm.SVC(kernel='rbf')
      clf.fit(x_train, y_train)
      y_pred = clf.predict(x_test)

```

```

[88] from sklearn.metrics import confusion_matrix
      confusion_matrix(y_test, y_pred)

```

```

array([[43, 11],
       [ 0, 89]])

```

```

[89] from sklearn.metrics import classification_report
      print(classification_report(y_test, y_pred))

```

	precision	recall	f1-score	support
0	1.00	0.80	0.89	54
1	0.89	1.00	0.94	89
accuracy			0.92	143
macro avg	0.95	0.90	0.91	143
weighted avg	0.93	0.92	0.92	143

**Gambar 10.** Mengganti kernel menjadi *RBF*

kernel *RBF* (*Radial Basis Function*) adalah jenis kernel yang digunakan dalam metode *Support Vector Machine* (*SVM*) untuk mengubah ruang data asli menjadi ruang fitur yang memiliki dimensi yang lebih tinggi. Pada penerapan kernel *RBF* pada kasus kanker payudara, kernel ini sering memberikan hasil yang baik karena kemampuannya dalam menangani pemisahan *non-linear* yang kompleks. Meskipun demikian, dalam beberapa percobaan, akurasi yang dihasilkan oleh kernel *RBF* mungkin lebih rendah dibandingkan dengan kernel *linear*. Kernel *RBF* memiliki fleksibilitas yang baik dan mampu menyesuaikan pemetaan sesuai dengan kebutuhan data, sehingga dapat memberikan hasil yang lebih baik dalam kasus – kasus di mana pemisahan data lebih kompleks daripada pemisahan linear yang sederhana.

12. Kemudian kernelnya akan diganti lagi dengan kernel lainnya yaitu kernel polinomial.

```
[108] from sklearn import svm
      clf = svm.SVC(kernel='poly', degree=3)
      clf.fit(x_train, y_train)
      y_pred = clf.predict(x_test)
```

```
[109] from sklearn.metrics import confusion_matrix
      confusion_matrix(y_test, y_pred)
```

```
array([[43, 11],
       [ 0, 89]])
```

```
from sklearn.metrics import classification_report
print(classification_report(y_test, y_pred))
```

	precision	recall	f1-score	support
0	1.00	0.80	0.89	54
1	0.89	1.00	0.94	89
accuracy			0.92	143
macro avg	0.95	0.90	0.91	143
weighted avg	0.93	0.92	0.92	143

**Gambar 11.** Mengganti kernel menjadi polinomial

Kernel polinomial adalah kernel yang memetakan data ke dalam ruang fitur yang memiliki dimensi yang lebih tinggi menggunakan fungsi basis polinomial. Jika dataset tidak dapat dipisahkan dengan garis lurus atau

*hipersuperplanar*, kernel polinomial dapat membantu SVM dalam memodelkan pola pemisahan yang kompleks. Hasil penerapan kernel ini menghasilkan akurasi yang sama dengan Kernel *RBF* pada kedua parameternya. Kernel ini cocok untuk pemisahan non-linear dengan bentuk pola yang kompleks.

13. Kemudian lakukan pergantian kernel dengan kernel yang terakhir yaitu kernel *sigmoid*.

```
[129] from sklearn import svm
      clf = svm.SVC(kernel='sigmoid')
      clf.fit(x_train, y_train)
      y_pred = clf.predict(x_test)
```

```
[130] from sklearn.metrics import confusion_matrix
      confusion_matrix(y_test, y_pred)
```

```
array([[ 8, 46],
       [38, 51]])
```

```
[131] from sklearn.metrics import classification_report
      print(classification_report(y_test, y_pred))
```

	precision	recall	f1-score	support
0	0.17	0.15	0.16	54
1	0.53	0.57	0.55	89
accuracy			0.41	143
macro avg	0.35	0.36	0.35	143
weighted avg	0.39	0.41	0.40	143

**Gambar 12.** Mengganti kernel menjadi *sigmoid*

Kernel *sigmoid* menggunakan fungsi sigmoid untuk memetakan data ke dalam ruang fitur yang memiliki dimensi yang lebih tinggi. Kernel ini dapat membantu SVM dalam memodelkan pemisahan yang kompleks dengan pola non-linear yang curam. Kernel sigmoid kurang umum digunakan dalam kasus kanker payudara dibandingkan dengan kernel lainnya. Akurasi yang dihasilkan pada penerapan ini menghasilkan nilai yang rendah karena kernel ini kurang cocok dengan dataset yang digunakan dengan nilai 0.35 dan 0.39 pada precision.

### ***E. Kesimpulan***

Support Vector Machine (SVM) adalah metode pembelajaran yang digunakan untuk klasifikasi dan regresi. SVM dapat mengatasi data kompleks dan tidak linier dengan menggunakan fungsi kernel, serta menerapkan prinsip SRM untuk menghindari overfitting. SVM telah sukses digunakan dalam berbagai bidang, seperti pengenalan pola dan analisis wajah.

#### ***F. Daftar Pustaka***

- Jakkula, V. (2011). Tutorial on Support Vector Machine (SVM). *School of EECS, Washington State University*, 1–13.  
<http://www.ccs.neu.edu/course/cs5100f11/resources/jakkula.pdf>
- Karim, A. (2020). Perbandingan Prediksi Kemiskinan di Indonesia Menggunakan Support Vector Machine (SVM) dengan Regresi Linear. *Jurnal Sains Matematika Dan Statistika*, 6(1), 107.  
<https://doi.org/10.24014/jsms.v6i1.9259>
- Sihombing, P. R. (2020). Perbandingan Metode Artificial Neural Network (ANN) dan Support Vector Machine (SVM) untuk Klasifikasi Kinerja Perusahaan Daerah Air Minum (PDAM) di Indonesia. *Jurnal Ilmu Komputer*, 13(1), 9. <https://doi.org/10.24843/jik.2020.v13.i01.p02>