

# project\_SVM

December 1, 2025

## 1 PROJECT MACHINE LEARNING – Support Vector Machine (SVM)

### 1.1 Studi Kasus: Prediksi Kelulusan Mahasiswa

Durasi pengerjaan: **100 menit** Mode pengerjaan: **Python-based (Google Colab / Jupyter Notebook)** Pengumpulan: **Upload ke GitHub** masing-masing mahasiswa

---

## 2 1. Judul Proyek

**Prediksi Kelulusan Mahasiswa Menggunakan Support Vector Machine (SVM)** (*Supervised Classification Project*)

---

## 3 2. Deskripsi Singkat Proyek

Mahasiswa akan melakukan analisis data kelulusan mahasiswa menggunakan algoritma **Support Vector Machine (SVM)**. Dataset yang digunakan berisi berbagai atribut mahasiswa (misalnya IPK, jumlah SKS, lama studi, status keaktifan, dll) dan label “**Lulus / Belum Lulus**”.

Mahasiswa diminta untuk:

1. Melakukan eksplorasi dataset
2. Melakukan preprocessing data
3. Membangun model klasifikasi SVM
4. Melakukan evaluasi model
5. Menyimpan notebook dalam GitHub masing-masing
6. Menuliskan README sebagai dokumentasi

Durasi pengerjaan adalah **100 menit**, sehingga langkah-langkah tugas dibuat jelas dan terarah.

---

## 4 3. Dataset

File: [datakelulusanmahasiswa.csv](#)

(Link menuju dataset)

Mahasiswa diminta:

- Membaca dataset menggunakan **pandas**
- Mengidentifikasi fitur-fitur berikut (contoh asumsi, disesuaikan dataset asli):
  - IPK
  - Total SKS
  - Umur
  - Status kehadiran
  - Lama studi
  - Keaktifan organisasi
  - Nilai mata kuliah tertentu
  - Label: **Lulus** / **Belum Lulus**

Jika dataset memiliki kolom yang berbeda, mahasiswa menyesuaikan secara eksploratif.

---

## 5 4. Instruksi Proyek (Langkah Per Langkah – 100 Menit)

---

### 5.1 Bagian A – Setup & Loading Dataset (10 menit)

Mahasiswa wajib:

- Import pandas, numpy, sklearn
  - Load file dataset
  - Tampilkan 5 baris awal
  - Deskripsikan kolom fitur dan label
  - Laporkan missing values
- 

### 5.2 Bagian B – Exploratory Data Analysis (15 menit)

Mahasiswa harus melakukan minimal:

1. Statistik deskriptif (mean, std, min, max)
  2. Visualisasi sederhana:
    - Histogram IPK
    - Countplot status lulus vs tidak lulus
  3. Menjawab pertanyaan:
    - “Apakah distribusi IPK berbeda antara mahasiswa lulus dan tidak lulus?”
    - “Apakah ada fitur-fitur yang tampak dominan menentukan kelulusan?”
-

### 5.3 Bagian C – Preprocessing Data (15 menit)

Mahasiswa wajib:

1. Menangani missing values
2. Encoding data kategorikal (OneHot / Label Encoding)
3. Feature scaling menggunakan StandardScaler
4. Train-test split (60:40, 75:25, 80:20 dan 90:10)

**Catatan:** SVM sangat sensitif terhadap skala fitur → scaling wajib.

---

### 5.4 Bagian D – Training Model SVM (25 menit)

Mahasiswa harus:

1. Melatih dua model:
    - **SVM Linear**
    - **SVM RBF Kernel**
  2. Melakukan hyperparameter tuning dasar:
    - $C = \{0.1, 1, 10\}$
    - $\text{gamma} = \{\text{'scale'}, 0.1, 1\}$
    - $\text{kernel} = \{\text{'linear'}, \text{'rbf'}\}$
  3. Menampilkan:
    - Confusion matrix
    - Classification report
    - Nilai akurasi
    - Precision, recall, F1-score
  4. Menjawab pertanyaan analitis:
    - “Model mana yang lebih baik? Linear atau RBF?”
    - “Bagaimana pengaruh parameter C terhadap decision boundary?”
    - “Apakah dataset ini tampak linearly separable?”
- 

### 5.5 Bagian E – Model Interpretation (10 menit)

Mahasiswa wajib memberikan:

- Penjelasan fitur paling berpengaruh
  - Interpretasi apakah mahasiswa ber-IPK rendah memiliki kecenderungan tidak lulus
  - Kesimpulan umum berdasarkan prediksi model
- 

### 5.6 Bagian F – Deployment (Sederhana) (5 menit)

Mahasiswa harus membuat sebuah fungsi Python:

```
predict_status(ipk, sks, umur, lamastudi, ... )
```

yang akan menampilkan prediksi **Lulus / Tidak Lulus** menggunakan model yang sudah dilatih.

Tidak perlu web deployment. Cukup fungsi sederhana di notebook.

---

## 5.7 Bagian G – Dokumentasi & GitHub Upload (10 menit)

Mahasiswa harus mengupload:

1. File notebook (SVM\_kelulusan.ipynb)
  2. Folder /data berisi dataset
  3. File README.md berisi:
    - Judul project
    - Deskripsi dataset
    - Langkah pengerjaan
    - Hasil evaluasi model
    - Cara menjalankan notebook
    - Identitas mahasiswa
- 

## 6 5. Rubrik Penilaian (Total 100%)

---

Komponen	Bobot
Exploratory Analysis	15%
Preprocessing & Feature Scaling	20%
Pembangunan Model SVM	25%
Evaluasi Model & Interpretasi	20%
Dokumentasi & GitHub	20%

---

## 7 6. Deliverables (wajib dikumpulkan)

Mahasiswa harus mengupload ke GitHub:

1. **Notebook lengkap:** SVM\_kelulusan.ipynb
  2. **Dataset:** datakelulusanmahasiswa.csv
  3. **README.md**
  4. **Screenshot hasil prediksi**
  5. Link GitHub diserahkan melalui Edlink
-

## 8 8. Note

“Pastikan repository kalian terlihat profesional, rapi, mudah dipahami. Perhatikan penamaan file, struktur folder, dan README. Project ini melatih kalian berpikir analitis, mengimplementasikan algoritma SVM sesuai RPS Machine Learning UNDA, serta mendokumentasikan pekerjaan secara bertanggung jawab sesuai CPL Prodi Sistem Informasi.”