

# IMPLEMENTASI ALGORITMA MACHINE LEARNING UNTUK DETEKSI PERFORMA AKADEMIK MAHASISWA

*By* Mitra Novitri Waruwu

**IMPLEMENTASI ALGORITMA *MACHINE LEARNING*  
UNTUK DETEKSI PERFORMA AKADEMIK  
MAHASISWA**

**SKRIPSI**



Oleh  
**MITRA NOVITRI WARUWU**  
NIM 202117036

**5**  
**PROGRAM STUDI PENDIDIKAN MATEMATIKA**  
**FAKULTAS KEGURUAN DAN ILMU PENDIDIKAN**  
**UNIVERSITAS NIAS**  
**2024**

**IMPLEMENTASI ALGORITMA *MACHINE LEARNING*  
UNTUK DETEKSI PERFORMA AKADEMIK  
MAHASISWA**

**SKRIPSI**

Diajukan Kepada  
Universitas Nias  
Untuk Memenuhi Salah Satu Persyaratan Menyelesaikan  
Program Sarjana Pendidikan

Oleh  
**MITRA NOVITRI WARUWU**  
NIM 202117036

**PROGRAM STUDI PENDIDIKAN MATEMATIKA  
FAKULTAS KEGURUAN DAN ILMU PENDIDIKAN  
UNIVERSITAS NIAS  
2024**

# **BAB 1**

## **PENDAHULUAN**

### **1.1. Latar Belakang**

Pendidikan adalah suatu cara mentransformasikan pengetahuan agar manusia mampu mengembangkan potensi dirinya. Pendidikan mendorong setiap orang untuk berkembang dan beradaptasi dengan perubahan zaman, seperti kemajuan dalam bidang teknologi. Dalam UU No. 20 tahun 2023 tentang Sistem Pendidikan Nasional dijelaskan Pendidikan adalah usaha sadar dan terencana dari manusia untuk mengembangkan potensi – potensi yang ada pada dirinya, yaitu rohani dan jasmani. Pendidikan tersebut ditempuh pada suatu Lembaga satuan Pendidikan seperti perguruan tinggi. Dalam dunia pendidikan tinggi, performa keberhasilan akademik mahasiswa merupakan aspek yang sangat penting dan menjadi perhatian utama bagi institusi pendidikan. Menurut Fatimah et al., (2021), mendeteksi performa akademik mahasiswa secara akurat merupakan tugas penting yang dihadapi oleh institusi pendidikan tinggi saat ini. Kemampuan untuk mengidentifikasi faktor-faktor yang berkontribusi terhadap keberhasilan atau kegagalan akademik mahasiswa dapat membantu institusi dalam memberikan dukungan dan intervensi yang tepat waktu.

Secara tradisional, penilaian performa akademik mahasiswa dilakukan dengan menganalisis faktor-faktor seperti nilai ujian, kehadiran kuliah, dan partisipasi kelas. Namun, pendekatan ini seringkali terbatas dan hanya memberikan gambaran sebagian dari keseluruhan faktor yang mempengaruhi keberhasilan akademik. Seperti yang disoroti oleh Munjirin & Iswinarti (2023), mengatakan bahwa pemahaman yang mendalam tentang faktor-faktor yang mempengaruhi keberhasilan akademik mahasiswa sangat penting untuk merancang strategi intervensi yang efektif.

Performa akademik mahasiswa adalah indikator kunci keberhasilan pengelolaan program studi. Deteksi performa akademik dapat membantu pengelola program studi memonitor dan mengambil tindakan proaktif terhadap mahasiswa yang berpotensi mengalami kesulitan. Menurut Nurhidayat et al., (2021), menyatakan bahwa performa akademik mahasiswa umumnya dilihat dari index prestasi mahasiswa (IPK) dan kelulusan tepat waktu. Sejalan dengan itu, menurut

Li et al., (2023), performa akademik dipengaruhi oleh beberapa faktor yaitu : lingkungan belajar, motivasi, latar belakang dan dukungan keluarga, regulasi diri, kondisi Kesehatan fisik dan mental, kecerdasan dan bakat alami, manajemen waktu dan metode belajar yang efektif. Studi sebelumnya telah mencoba mendeteksi performa akademik mahasiswa dengan berbagai metode. Contohnya, menggunakan demografi mahasiswa, kehadiran kelas, akses internet, kepemilikan komputer, dan jumlah mata kuliah yang diambil menggunakan analisa tradisional. Penggunaan data akademik yang besar dan belum terolah dapat menurunkan kualitas dalam mendeteksi performa akademik mahasiswa. Hal ini terjadi karena tidak mendapatkan keakuratan hasil yang pasti dan tidak mendapatkan informasi terkait performa akademik mahasiswa secara detail dan lengkap.

Dengan kemajuan teknologi dan perkembangan di bidang kecerdasan buatan, terbuka peluang untuk menggunakan pendekatan baru dalam mendeteksi performa akademik mahasiswa. Studi sebelumnya telah mendeteksi performa akademik dengan menggunakan metode *Visual-prediction* dan pendekatan pemahaman domain dan analisis visual data. *Visual-prediction* memiliki kelemahan yaitu model kurang mampu memvisualisasikan pola yang relevan dan overfitting kurang efektif dalam mendeteksi performa akademik. Kelemahan yang dialami oleh model *Visual-prediction*, membuat hasil prediksi kurang akurat. Untuk itu, semakin berkembangnya teknologi, *machine learning* menjadi solusi yang dapat digunakan dalam memprediksi performa akademik mahasiswa. *Machine learning* merupakan cabang dari kecerdasan buatan yang memungkinkan komputer untuk belajar dari data dan mengidentifikasi pola-pola tersembunyi yang sulit dideteksi oleh manusia. Dalam konteks ini, *machine learning* dapat digunakan untuk mengekstrak informasi dari kumpulan data yang besar untuk mendeteksi pola dan tren yang mungkin tidak terlihat secara langsung (Banerjee et al., 2021).

*Machine learning* dapat menjadi solusi untuk mengatasi tantangan ini dengan membantu dalam klasifikasi dan deteksi kemampuan akademik mahasiswa. Teknik *machine learning* telah terbukti sangat efektif dalam menganalisis data yang kompleks dan mengungkap pola-pola tersembunyi yang sulit dideteksi oleh manusia (Fatimah et al., 2021). Dengan memanfaatkan kekuatan *machine learning*, institusi pendidikan dapat memanfaatkan data historis mahasiswa untuk

mengembangkan model prediktif yang akurat. Seperti yang dinyatakan oleh Khasanah et al., (2022), prediksi performa akademik mahasiswa menggunakan *machine learning* dapat memberikan wawasan berharga bagi para pengambil keputusan dalam merancang intervensi yang tepat waktu dan meningkatkan kualitas pembelajaran. Hal ini sejalan dengan pentingnya memberikan dukungan yang sesuai kepada mahasiswa yang mungkin menghadapi risiko kegagalan akademik. Selain itu, implementasi *machine learning* dalam memprediksi performa akademik mahasiswa juga dapat memberikan kontribusi terhadap pengembangan sistem pendukung keputusan yang lebih efektif bagi para pembuat kebijakan di institusi pendidikan. Dengan memahami faktor-faktor yang mempengaruhi keberhasilan akademik, perguruan tinggi dapat mengalokasikan sumber daya dengan lebih baik dan merancang program yang sesuai dengan kebutuhan mahasiswa.

Oleh karena itu, penelitian ini bertujuan untuk mengeksplorasi implementasi *machine learning* untuk deteksi performa akademik mahasiswa dengan menggunakan dataset yang kaya dan mencakup berbagai faktor yang relevan. Dengan menerapkan teknik *machine learning* yang sesuai, penelitian ini diharapkan dapat memberikan solusi praktis untuk meningkatkan keberhasilan akademik mahasiswa dan berkontribusi dalam pengembangan metode pembelajaran yang lebih efektif dan terpersonalisasi. Selain itu, Dengan menggabungkan berbagai sumber data dan menggunakan pendekatan *machine learning* yang tepat, diharapkan perguruan tinggi dapat memperoleh pemahaman yang lebih mendalam tentang dinamika performa akademik mahasiswa dan mengambil tindakan yang diperlukan untuk mendukung kesuksesan mereka dalam mencapai tujuan akademik dan profesional.

Berdasarkan uraian diatas, penulis tertarik untuk memprediksi performa akademik mahasiswa dengan menggunakan model *machine learning*. Untuk itulah calon peneliti mengambil judul untuk penelitian yaitu “**Implementasi Algoritma *Machine Learning* Untuk Deteksi Performa Akademik Mahasiswa**”

## 1.2. Identifikasi Masalah

Berdasarkan latar belakang masalah yang telah diuraikan diatas, dapat diidentifikasi masalah sebagai berikut:

- a. Kesulitan dalam mengidentifikasi dan menganalisis secara akurat faktor-faktor yang mempengaruhi performa akademik mahasiswa dengan metode tradisional.
- b. Keterbatasan dalam memprediksi performa akademik mahasiswa secara dini dan akurat.
- c. Kesulitan dalam mengidentifikasi mahasiswa yang berisiko mengalami kegagalan akademik.

## 1.3. Batasan Masalah

Agar materi dan pembahasan penelitian seperti yang diharapkan dan tidak meluas, maka penelitian ini diberikan batasan. Beberapa batasan masalah pada penelitian sebagai berikut:

- a. Penelitian ini akan berfokus pada implementasi *machine learning* untuk memprediksi performa akademik mahasiswa, dengan menggunakan dataset yang kaya dan mencakup berbagai faktor yang relevan.
- b. Penelitian ini akan menggunakan dataset yang dikumpulkan dari data primer baru.
- c. Penelitian ini akan berfokus pada tingkat pendidikan tinggi Universitas Nias.

## 1.4. Rumusan Masalah

Berdasarkan latar belakang dan identifikasi masalah yang telah dimuat diatas maka rumusan masalah dari penelitian ini adalah: “Bagaimana mengimplementasikan algoritma *machine learning* untuk memprediksi performa akademik mahasiswa dengan akurat?”

## 1.5. Tujuan Penelitian

Berdasarkan paparan dari rumusan masalah di atas, maka tujuan penelitian ini adalah untuk mengetahui sejauh mana implementasi algoritma *machine learning* untuk deteksi performa akademik mahasiswa dengan akurat.

## **1.6. Manfaat Penelitian**

### **1.6.1 Manfaat Teoritis**

Melalui penelitian ini dapat memberikan kontribusi baru pada pengetahuan tentang penerapan *machine learning* dan dapat digunakan sebagai referensi dalam mengetahui performa akademik mahasiswa prodi Pendidikan matematika melalui algoritma *machine learning*.

### **1.6.2 Manfaat Praktis**

- a. Untuk peneliti, manfaatnya yakni ilmu yang diperoleh selama kuliah digunakan untuk mendukung persiapan transisi ke dunia kerja dan juga untuk menambah pandangan yang lebih luas tentang algoritma *machine learning*.
- b. Bagi Program Studi Pendidikan matematika, dapat digunakan sebagai bahan penulisan ilmiah bagi pembaca dan sebagai referensi bagi mahasiswa terkait deteksi performa akademik mahasiswa serta dapat digunakan sebagai bahan referensi untuk meningkatkan jumlah dan kualitas mahasiswa pada program studi Pendidikan matematika.
- c. Bagi instansi, dapat dijadikan sebagai bahan penelitian kampus Universitas Nias untuk lebih memperhatikan performa akademik mahasiswa sehingga mampu mengidentifikasi mahasiswa yang berisiko mengalami kesulitan belajar dan membutuhkan intervensi diri.

## **BAB II**

### **KAJIAN TEORI**

#### **2.1 Kajian Teori**

##### **2.1.1 Performa Akademik**

###### **a. Pengertian Performa Akademik**

Performa akademik mahasiswa menjadi salah satu faktor penting untuk kelulusan mahasiswa. Menurut Fatimah et al., (2021), mengungkapkan bahwa performa akademik atau prestasi akademik merupakan suatu hal yang menunjukkan pencapaian suatu tujuan karena usaha belajar yang optimal. Prestasi akademik berfokus pada nilai yang dicapai dalam proses pembelajaran. Sejalan dengan pendapat tersebut, Firman Akbar & Rahmaddeni (2022), mengungkapkan bahwa performa akademik adalah suatu keberhasilan yang telah diraih dan dapat mengakibatkan peningkatan akademis. Selain itu, Christian (2022), mendefinisikan bahwa performa adalah keberhasilan atau pencapaian seseorang dalam melakukan sesuatu. Performa dapat didefinisikan sebagai tingkat pencapaian hasil atau tingkat pencapaian suatu tujuan di dalam dunia pekerjaan atau pendidikan. Performa akademik mengacu pada hasil akhir yang dicapai oleh seseorang sebagai keberhasilan selama pendidikan. Sejalan dengan itu, Yustira et al., (2021), mendefinisikan bahwa performa akademik adalah evaluasi hasil belajar dari suatu proses, seperti nilai pelajaran, nilai ujian, dan lain-lain. Performa akademik diukur dengan nilai akhir belajar.

Berdasarkan beberapa definisi yang telah dikemukakan, maka dapat disimpulkan bahwa performa akademik yaitu suatu hal yang menunjukkan keberhasilan atau tingkat pencapaian akademik seseorang yang dinyatakan dalam bentuk nilai yang dicapai dalam proses pembelajaran sebagai bukti telah melaksanakan tugas akademik selama menempuh pendidikan di perguruan tinggi.

Performa akademik menurut Briones et al., (2022), umumnya dinilai berdasarkan pada Indeks Prestasi Kumulatif (IPK), yang merupakan rata-rata nilai mata kuliah yang ditempuh selama studi. Namun, performa akademik tidak hanya

sebatas IPK. Dimensi performa akademik mahasiswa yang lebih luas meliputi:

- Kemampuan akademik: Meliputi daya nalar, daya kritis, kemampuan memecahkan masalah, dan penguasaan materi perkuliahan.
- Keterampilan belajar: Meliputi strategi belajar, manajemen waktu, dan kemampuan riset.
- Motivasi belajar: Meliputi semangat, tekad, dan minat belajar.
- Sikap belajar: Meliputi disiplin, tanggung jawab, dan etos belajar.
- Keterlibatan dalam pembelajaran: Meliputi partisipasi aktif di kelas, mengerjakan tugas, dan mengikuti kegiatan ekstrakurikuler yang relevan dengan bidang studi.
- Kemampuan interpersonal: Meliputi komunikasi efektif, kerjasama tim, dan kepemimpinan.

## **b. Faktor-faktor yang Mempengaruhi Performa Akademik**

Faktor-faktor yang Mempengaruhi Performa Akademik Menurut Briones et al., (2022), faktor penentu kinerja mahasiswa dapat dijelaskan menggunakan pendekatan teori atribusi yang menyatakan bahwa dua kategori dasar atribusi yang melekat pada diri seorang mahasiswa yang akan menentukan performanya, yaitu atribusi yang bersifat internal dan yang bersifat eksternal.

### **1. Faktor Internal**

Faktor internal meliputi:

- a. Kecerdasan atau intelegensi, yaitu kemampuan menyesuaikan diri dengan keadaan baru dengan menggunakan alat-alat berpikir. Setiap orang memiliki tingkat IQ yang berbeda-beda. Seseorang yang memiliki IQ 110 - 140 dapat digolongkan cerdas, dan yang memiliki IQ 140 ke atas tergolong jenius. Golongan ini mempunyai potensi untuk dapat menyelesaikan pendidikan di Perguruan Tinggi. Seseorang yang memiliki IQ kurang dari 90 tergolong lemah mental, mereka inilah yang banyak mengalami kesulitan belajar.
- b. Bakat, yaitu potensi atau kecakapan dasar yang dibawa sejak lahir. Setiap individu mempunyai bakat yang berbeda-beda. Seseorang akan lebih mudah mempelajari sesuatu yang sesuai dengan

bakatnya. Apabila seseorang harus mempelajari sesuatu yang tidak sesuai dengan bakatnya, ia akan cepat bosan, mudah putus asa dan tidak senang. Hal-hal tersebut akan tampak pada anak suka mengganggu kelas, berbuat gaduh, tidak mau pelajaran sehingga nilainya rendah.

- c. Minat, yaitu rasa lebih suka maupun rasa keterikatan pada suatu hal atau aktivitas, tanpa ada yang menyuruh. Tidak adanya minat seorang anak terhadap suatu pelajaran akan timbul kesulitan belajar. Belajar yang tidak ada minatnya mungkin tidak sesuai dengan bakatnya, tidak sesuai dengan kebutuhannya, tidak sesuai dengan kecakapan dan akan menimbulkan problem pada diri anak. Ada tidaknya minat terhadap suatu pelajaran dapat dilihat dari cara anak mengikuti pelajaran, lengkap tidaknya catatan dan aktif tidaknya dalam proses pembelajaran.
- d. Motivasi, yaitu rangsangan, dorongan ataupun pembangkit tenaga yang dimiliki seseorang atau sekelompok masyarakat yang mau berbuat dan bekerjasama secara optimal dalam melaksanakan sesuatu yang telah direncanakan untuk mencapai tujuan yang telah ditetapkan. Motivasi sebagai faktor dalam (batin) berfungsi menimbulkan, mendasari dan mengarahkan perbuatan belajar. Motivasi dapat menentukan baik tidaknya dalam mencapai tujuan, sehingga semakin besar motivasinya akan semakin besar kesuksesan belajarnya. Seorang yang besar motivasinya akan giat berusaha, tampak gigih, tidak mau menyerah dan giat membaca buku-buku untuk meningkatkan prestasinya. Sebaliknya mereka yang motivasinya lemah, tampak acuh tak acuh, mudah putus asa perhatiannya tidak tertuju pada pelajaran, suka mengganggu kelas dan sering meninggalkan pelajaran. Akibatnya mereka banyak mengalami kesulitan belajar.
- e. Kelelahan, kelelahan jasmani dan rohani dapat mempengaruhi prestasi belajar dan agar siswa belajar dengan baik haruslah menghindari jangan sampai terjadi kelelahan dalam belajarnya

seperti lemah lunglainya tubuh. Sehingga perlu diusahakan kondisi yang bebas dari kelelahan rohani seperti memikirkan masalah yang berarti tanpa istirahat, mengerjakan sesuatu karena terpaksa tidak sesuai dengan minat dan perhatian. Ini semua besar sekali pengaruhnya terhadap pencapaian kinerja akademik siswa.

## 2. Faktor Eksternal

Faktor eksternal meliputi:

- a. Faktor keluarga, faktor ini sangat berperan aktif bagi para pelajar. contohnya antara lain: cara orang tua mendidik, relasi antara anggota keluarga, keadaan keluarga, pengertian orang tua, keadaan ekonomi keluarga, latar belakang kebudayaan dan suasana rumah.
- b. Faktor lingkungan kampus, faktor ini dapat berupa cara dosen mengajar, alat-alat pelajaran, kurikulum, waktu kuliah, interaksi dosen dan mahasiswa, disiplin kuliah, dan media pendidikan.
- c. Faktor masyarakat, faktor yang mempengaruhi kinerja mahasiswa antara lain teman bergaul, kegiatan lain di luar sekolah dan cara hidup di lingkungan masyarakatnya.

### 2.1.2 Machine Learning

#### a. Pengertian Machine Learning

*Machine Learning* (ML) merupakan salah satu varian dari *sistem Artificial Intelligence* (AI) atau kecerdasan buatan yang memungkinkan komputer dapat belajar tanpa diprogram secara eksplisit. Secara umum, pekerjaan *machine Learning* (ML) yang seringkali digunakan adalah untuk mengklasifikasikan satu permasalahan menjadi beberapa kelompok. Dalam kehidupan sehari-hari, objek dapat diidentifikasi dengan mudah oleh manusia, namun belum tentu dapat dijelaskan secara spesifik. Maka diperlukan *machine learning* dalam mengenali, mengidentifikasi, ataupun memprediksi data tertentu dengan mempelajari histori (Buslim & Iswara, 2019). Sejalan dengan itu, Nurhidayat et al., (2021) mengemukakan bahwa *machine learning* adalah teknik untuk melakukan inferensi

terhadap data dengan pendekatan matematis. Disebut model matematis karena keilmuan *machine learning* merupakan turunan dari rumusan matematika dan statistika. Selain itu, Firdaus et al., (2023), menyatakan bahwa algoritma *machine learning* memberikan pendekatan inovatif dalam analisis data berdasarkan pola yang terdeteksi yang umumnya digunakan dalam mengklasifikasikan.

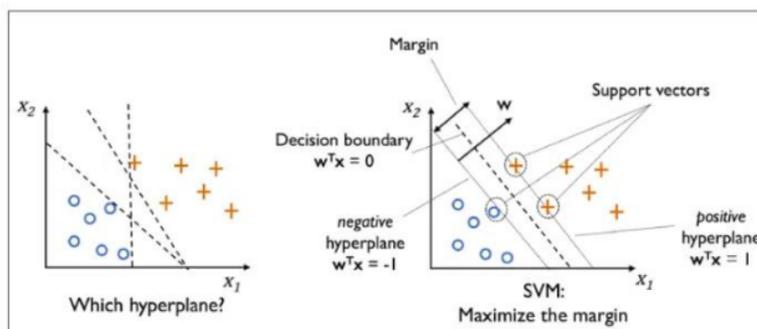
Dari uraian diatas, dapat disimpulkan bahwa *Machine Learning* adalah bagian dari AI yang dapat mengolah data tanpa di program secara implisit dan memberikan pendekatan inovatif dalam menganalisis data serta memberikan solusi untuk menyelesaikan permasalahan yang kompleks.

Menurut Al Husaini et al., (2022), machine learning adalah bidang yang berkaitan dengan sibernetika dan ilmu komputer yang terus mengalami perkembangan. Selanjutnya dikemukakan bahwa pada tahun 1940-an, saat pertama kali mesin komputer dibuat, para ahli hanya berpikir tentang cara mempercepat komputasi. Seiring berkembangnya waktu, ada banyak pendekatan baru untuk komputasi. Namun, tidak ada rencana untuk membuat program komputer yang memungkinkan komputer untuk belajar seperti manusia. Sejalannya waktu yang kemungkinan berada pada tahun 1957, 1959, dan 1960, seorang psikolog yang bernama Frank Rosenblatt dari *Cornell University*, menemukan ide mengembangkan suatu mesin program komputer berdasarkan kerja sistem saraf manusia, sehingga menciptakan sebuah kelompok yang membuat mesin untuk mengenali huruf – huruf alfabet. Mesin yang disebut “*Perceptron*” oleh penciptanya, menggunakan sinyal analog dan diskrit serta menyertakan elemen ambang batas yang mengubah sinyal analog menjadi sinyal diskrit. Pada awal tahun 1960-an, beberapa kelompok terlibat dalam pendeteksian dan pengujian sistem pengenalan pembelajaran yang semakin diteliti dengan berbagai teori. Dari panjangnya proses yang dilakukan, hingga di temukanlah beberapa vektor – vektor yang menjadi cikal bakal AI. Pada tahun 1964, algoritma yang terus diuji dan di kembangkan, tidak mengkhianati hasil sehingga pembelajaran mesin di temukan titik terangnya dengan ditemukannya algoritma SVM (*Support vector machine*). Dari tahun 1964–sekarang, perkembangan ML semakin berkembang dan dapat digunakan dengan sangat baik.

## b. Support Vector Machine (SVM)

Dalam machine learning, salah satu teknik klasifikasi yang lebih unggul adalah Support Vector Machine (SVM) dibandingkan metode klasifikasi lainnya. SVM termasuk dalam kategori pembelajaran terawasi (supervised learning), di mana algoritma ini mengembangkan suatu fungsi yang memetakan input ke output yang diinginkan, dan umumnya digunakan untuk menyelesaikan masalah klasifikasi dan regresi, baik linier maupun non-linier. Disebut pembelajaran terawasi karena kualitas hasil yang diperoleh sangat bergantung pada kesesuaian antara input dan output yang diberikan, sehingga diperlukan validasi atau pengawasan sebelumnya oleh pengguna.

Konsep Support Vector Machine (SVM) secara sederhana dapat dijelaskan sebagai upaya menemukan hyperplane terbaik yang berfungsi sebagai pemisah antara dua kelas dalam input space. Dalam input space ini, terdapat pola yang merupakan anggota dari dua kelas, yaitu  $+1$  dan  $-1$ , serta terdapat berbagai alternatif garis pemisah (discrimination boundaries). Margin adalah jarak antara hyperplane tersebut dengan pola terdekat dari masing-masing kelas, di mana pola terdekat ini dikenal sebagai support vector. Proses pencarian lokasi hyperplane ini menjadi inti dari pembelajaran dalam SVM (Kang et al., 2020). Selaras dengan itu, Nurhidayat et al. (2021) menjelaskan bahwa teknik SVM digunakan untuk menemukan fungsi pemisah yang optimal, di mana hyperplane tersebut dapat berbentuk garis pada dua dimensi atau berupa bidang datar pada dimensi yang lebih tinggi.



Gambar 2.1 *Hyperplane* yang memisahkan dua kelas positif dan negatif

3  
Pada Gambar 2.1, hyperplane yang ditemukan oleh SVM diilustrasikan seperti pada Gambar 1, di mana posisinya berada di tengah antara dua kelas. Ini menunjukkan bahwa jarak antara hyperplane dan objek-objek data yang berbeda dengan kelas yang berdekatan (terluar) yang diberi tanda bulat kosong dan positif adalah seimbang. Dalam SVM, objek data terluar yang paling dekat dengan hyperplane disebut sebagai support vector. Objek yang menjadi support vector ini paling sulit untuk diklasifikasikan karena posisinya hampir tumpang tindih (overlap) dengan kelas lainnya. Karena sifatnya yang krusial, hanya support vector ini yang diperhitungkan oleh SVM untuk menentukan hyperplane yang paling optimal.

### 2.1.3 Python

Python adalah bahasa pemrograman tingkat tinggi yang dikenal dengan keserbagunaan dan kemudahan penggunaannya di berbagai bidang seperti biologi komputasi, ilmu komputer, dan matematika (Harris et al., 2020). Bahasa pemrograman ini banyak digunakan untuk komputasi numerik yang efisien melalui struktur seperti larik NumPy, yang memungkinkan implementasi komputasi numerik yang kompleks dengan cara yang mudah (Walt et al., 2011). Fleksibilitas Python meluas ke aplikasi dalam pembelajaran mesin, di mana alat seperti Pythran dan Numba digunakan untuk mengoptimalkan kinerja dalam tugas-tugas komputasi (Harris et al., 2020). Selain itu, kemampuan Python telah dimanfaatkan di berbagai bidang seperti pemrosesan data neuroimaging, analisis data besar, dan bahkan dalam pengembangan kontrak pintar untuk aplikasi blockchain (Cao, 2023).

Popularitas bahasa ini dalam penelitian ilmiah dan teknik terbukti dalam penggunaannya untuk tugas-tugas seperti memvisualisasikan data urutan, menganalisis sifat permukaan bahan biologis, dan melakukan penelitian psikologis dengan data besar (Rickert et al., 2021). Peran Python dalam memfasilitasi komputasi statistik melalui pustaka seperti panda semakin menggarisbawahi signifikansinya dalam analisis dan pemodelan data (Bühler et al., 2022).

Penggunaan Python dalam berbagai konteks telah menunjukkan fleksibilitas dan keandalan bahasa pemrograman ini. Dalam bidang Data Science, Python digunakan untuk pengolahan data, mengimplementasikan algoritma

analisis, dan memvisualisasikan hasil (Priyatno, 2023). Selain itu, Python juga digunakan dalam pengembangan aplikasi seperti *Research Data Annotation (RDA)* untuk melakukan anotasi khusus menggunakan *library Python* (Supriyanto, 2024). Di sisi lain, dalam konteks pengembangan aplikasi mobile seperti pada TensorFlow, Python digunakan sebagai *API front-end* untuk membangun aplikasi berkinerja tinggi (Negara, 2024).

Python juga diterapkan dalam berbagai proyek penelitian dan teknis. Misalnya, dalam pengembangan algoritma A\* Path Planning pada base station robot sepak bola beroda, Python digunakan tanpa mengubah fitur-fitur lain yang telah ada pada base station sebelumnya (Safitri, 2023). Selanjutnya, dalam menganalisis sentimen masyarakat terhadap penggunaan e-commerce, Python digunakan dengan *library pandas* untuk membaca dan menampilkan dataset dalam bentuk tabel (Kusuma, 2023). Penggunaan python juga menjadi pilihan utama dalam analisis data, seperti dalam menerapkan model regresi linier untuk memprediksi harga mobil bekas di India. Selain itu, python juga digunakan dalam berbagai aplikasi praktis, seperti dalam mengembangkan sistem penghitungan kapasitas ruang parkir menggunakan pustaka OpenCV untuk menghitung jumlah mobil dan ketersediaan ruang parkir (Sardjono dkk., 2023). Selain itu, python bahkan digunakan dalam proyek-proyek inovatif seperti mendeteksi penggunaan masker wajah dengan ESP32Cam menggunakan OpenCV dan TensorFlow (Dharma et al., 2022). Dengan kemudahan penggunaan, aplikasi yang luas, dan dukungan pustaka yang kaya, python telah menjadi bahasa pemrograman yang sangat populer dan efisien di berbagai bidang.

Python telah mendapatkan popularitas yang signifikan di bidang pembelajaran mesin karena keserbagunaannya, kemudahan penggunaannya, dan berbagai macam pustaka dan kerangka kerja yang tersedia untuk mendukung berbagai tugas pembelajaran mesin. Meningkatnya jumlah peneliti dan pengembang yang menggunakan python untuk komputasi ilmiah, analitik data, dan aplikasi pembelajaran mesin adalah bukti dari adopsi python yang meluas (Awar et al., 2021). Python bersama dengan pustaka seperti NumPy, Pandas, dan Scikit-learn, telah menyederhanakan penerapan teknik pembelajaran mesin yang sudah

ada. Evolusi dan peningkatan berkelanjutan dari pustaka python disebabkan oleh komunitas pengembang dan peneliti yang aktif (Fassbender, 2019).

Dominasi python dalam pembelajaran mesin didukung oleh dukungannya yang luas untuk pustaka komputasi larik seperti NumPy, yang berfungsi sebagai lapisan interoperabilitas antara berbagai pustaka komputasi larik, menyediakan kerangka kerja yang fleksibel untuk analisis ilmiah dan industri (Harris et al., 2020). Selain itu, kompatibilitas python dengan antarmuka data umum seperti larik NumPy meningkatkan interoperabilitasnya dalam ekosistem pembelajaran mesin (Albertsson et al., 2020). Peran python dalam pembelajaran mesin meluas ke pengembangan alat dan kerangka kerja yang disesuaikan dengan tugas-tugas pembelajaran mesin tertentu. Misalnya, ada kotak alat python yang dirancang khusus untuk mengimplementasikan algoritma utama untuk pembelajaran inkremental kelas, yang dapat meringankan beban kerja para peneliti di komunitas pembelajaran mesin (Zhou et al., 2021).

```
from sklearn import datasets
from sklearn.model_selection import train_test_split
from sklearn.svm import SVC
from sklearn.metrics import accuracy_score

# Load dataset
iris = datasets.load_iris()
X = iris.data
y = iris.target

# Bagi data menjadi training dan testing
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3, random_state=42)

# Buat dan latih model
model = SVC(kernel='linear')
model.fit(X_train, y_train)

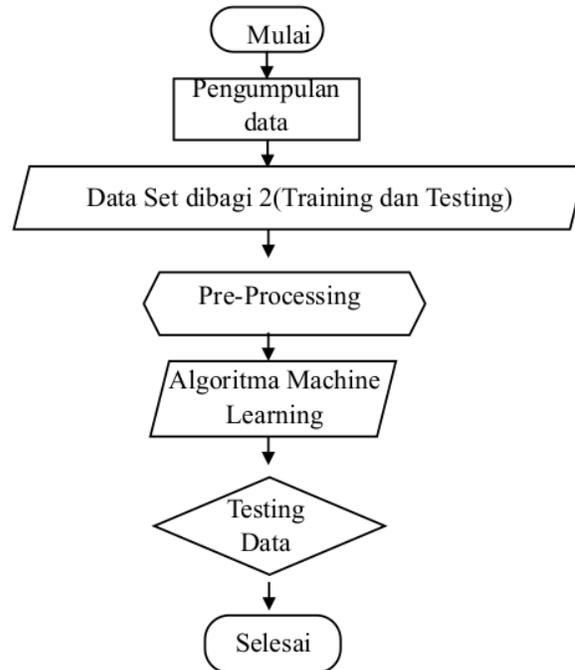
# Prediksi dan evaluasi
y_pred = model.predict(X_test)
accuracy = accuracy_score(y_test, y_pred)
print(f"Akurasi: {accuracy}")
```

Gambar 2.2 Model penerapan Python dalam Mendeteksi Akurasi

## 2.2 Kerangka Berpikir

Adapun kerangka berpikir yang digunakan dalam penelitian ini yaitu:

Kerangka berpikir pada penelitian ini yaitu :



## **BAB III**

### **METODOLOGI PENELITIAN**

#### **3.1 Jenis Penelitian**

Penelitian ini adalah jenis penelitian terapan di bidang pemrograman. Penelitian terapan merupakan penelitian dengan temuan yang dapat diterapkan untuk pemecahan masalah yang dapat menjamin objektivitas dalam pengumpulan fakta. Data yang digunakan dalam penelitian ini adalah data primer karena data diperoleh dari hasil responden mahasiswa aktif selama menempuh Pendidikan matematika yang datanya akan diolah menjadi data set. Data yang telah terolah menjadi data set kemudian dibagi dalam 2 bagian data yaitu data latih 70% dan data uji 30% (Septarian & Nugroho, 2023). Hasil data set akan dilakukan *pre-processing* dengan pemanfaatan *machine learning*.

Metode penelitian ini adalah proses kegiatan mengumpulkan data primer mahasiswa program studi Pendidikan matematika Universitas Nias, yang selanjutnya akan dilakukan klasifikasi berdasarkan pembagian data set dan mengimplementasikan *machine learning* sehingga mampu menginterpretasikan hasil sesuai dengan tujuan penelitian.

### 3.2 Variabel Penelitian

Dalam penelitian ini, adapun variabel yang digunakan yaitu sebagai berikut:

**Tabel 3.1 Tabel Variabel Dependen**

Variabel <i>Dependen</i>	Kategori	Skala
Nilai IPK (Academic Performance)	<2,76	1
	2,76 – 3,00	2
	3,01 – 3,50	3
	>3,50	4

**Tabel 3.2 Tabel Variabel Independen**

Variabel Independen	Kategori	Skala
Jenis Kelamin (X1)	Laki – laki	1
	Perempuan	2
Jenis Sekolah Asal	Negeri	1
	Swasta	2
Jurusan di Sekolah Asal	IPA	1
	IPS	2
	Lainnya	3
Kabupaten/Kota asal	Kabupaten Nias	1
	Kabupaten Nias barat	2
	Kabupaten Nias Selatan	3
	Kabupaten Nias Utara	4
	Kota Gunungsitoli	5
	Lainnya	6
Beasiswa	Pernah/Sedang menerima	1
	Tidak Pernah Menerima	2
Pilihan Jurusan Di Perkuliahan	Jurusan Pilihan Pertama dan Diminati	1
	Bukan Jurusan Pertama	2
Durasi Belajar Per hari	< 1 Jam	1

	1-2 Jam	2
	2 – 3 Jam	3
	3 – 4 Jam	4
	> 4 Jam	5
Presensi	Selalu Hadir 16 Kali Pertemuan	1
	1 – 2 kali Tidak hadir	2
	3 – 4 Kali Tidak Hadir	3
	> 4 Kali Tidak Hadir	4
Jumlah mata kuliah/Semester	< 10 SKS	1
	10 – 15 SKS	2
	16 – 20 SKS	3
	20 – 23 SKS	4
	24 SKS	5
Penyerahan Tugas Mata kuliah	Selalu diserahkan tepat waktu	1
	Selalu diserahkan walau terlambat	2
	Kadang – kadang menyerahkan	3
	Tidak pernah menyerahkan	4
Motivasi belajar	Motivasi belajar sangat tinggi	1
	Tidak ada motivasi belajar	2
Tekanan Akademik	Tidak ada	
	Stress dalam beberapa mata kuliah	
	Stress berlebihan	
Bimbingan Belajar	Pernah/Sedang Mengikuti	1
	Tidak pernah	2
Jenis Tempat Tinggal	Kos	1
	Bersama Orang Tua	2
	Bersama Wali	3
	Asrama	4
Status Pekerjaan Mahasiswa	Kuliah dan bekerja	1
	Kuliah dan tidak bekerja	2
Organisasi	Aktif Organisasi	1

	Tidak aktif organisasi	2
Tingkat Pendidikan Orang Tua	Tidak Sekolah	1
	SD/Sederajat	2
	SMP/Sederajat	3
	SMA/Sederajat	4
	Pendidikan Tinggi (Diploma, Sarjana, Magister, Doktor)	5
Pendapatan orang tua	< Rp 1.000.000	1
	Rp 1.000.000 s/d Rp 1.499.999	2
	Rp 1.500.000 s/d Rp 1.999.999	3
	Rp 2.000.000 s/d Rp 2.499.999	4
	≥ Rp 2.500.000	5
Memiliki Laptop	Ya	1
	Tidak	2
Memiliki Kendaraan	Ya	1
	Tidak	2

### 3.3 Subjek Penelitian

Yang menjadi subjek dalam penelitian ini yaitu seluruh mahasiswa aktif Program Studi Pendidikan Matematika Universitas Nias. Mahasiswa Program Studi Pendidikan Matematika berjumlah 201 orang terdiri dari 40 orang semester II, 47 orang semester IV, 67 orang semester VI dan 47 orang semester VIII.

### 3.4 Instrumen Penelitian

Instrumen penelitian dalam penelitian ini adalah kuesioner yang dibagikan dan diisi oleh subjek. Menurut Lawuna (2023), Kuesioner merupakan teknik pengumpulan data melalui responden untuk mendapatkan informasi melalui angket. Pernyataan-pernyataan yang diajukan oleh peneliti berupa pertanyaan terbuka dan tertutup. Pada kuesioner tertutup, setiap pertanyaan sudah tersedia alternatif jawaban sehingga responden tinggal memilih diantara jawaban yang tersedia. Sedangkan kuesioner terbuka, peneliti memberi kesempatan kepada untuk

memberikan jawaban secara objektif. Adapun kisi-kisi dari instrument penelitian ini yaitu sebagai berikut:

**Tabel 3.2 Kisi – kisi Instrumen untuk Mahasiswa**

<b>Variabel</b>	<b>Indikator</b>	<b>Nomor Pertanyaan</b>
Data Pribadi Responden	Nama	
	NIM	
	Jenis Kelamin	
	Usia	
	Alamat	
	Kontak	
Tingkat ekonomi	Penerima beasiswa	
	Pendidikan Orang tua	
	Pekerjaan orang tua	
	Pendapatan Orang tua	
	Pekerjaan diluar perkuliahan	
Kesehatan Mental	Motivasi Kuliah	
	Beban selama Kuliah	
	Kecemasan selama kuliah	
Riwayat Sekolah Asal	Nama sekolah	
	Jurusan	
	Tahun Lulus	
Riwayat Perkuliahan	Tahun Masuk / tahun Angkatan	
	Semester berjalan	
	Jumlah Mata kuliah yang di kontrak	
	Jumlah Tugas Mata Kuliah	
	Jumlah Pertemuan/Semester	
	Jumlah Presensi	
	Lama waktu Belajar/ matkul	
	Jumlah matkul yang di kontrak	
	Peralatan Elektronik	
	Jarak tempat Tinggal ke kampus	
	kendaraan	
	Rata rata Nilai Tugas	
	IPS	

### 3.5 Teknik pengumpulan data

Dalam penelitian ini, Teknik pengumpulan data yang akan dipakai dalam penelitian ini adalah berupa:

a. Dokumentasi

Dokumentasi merupakan Teknik pengumpulan data dengan cara memperoleh informasi dari macam – macam sumber tertulis atau dokumen yang ada pada responden. Menurut Sugiyono (2015:329), dokumentasi merupakan Teknik pengumpulan data untuk memperoleh data dan informasi dalam bentuk buku, arsip, dokumen, tulisan angka dan gambar yang berupa laporan serta keterangan yang dapat mendukung penelitian.

b. Kuesioner

2  
Angket atau kuesioner merupakan teknik pengumpulan data yang dilakukan dengan cara memberi seperangkat pertanyaan atau pernyataan tertulis kepada responden untuk dijawab. Tipe pertanyaan dalam angket dibagi menjadi dua, yaitu: terbuka dan tertutup. Pertanyaan terbuka adalah pertanyaan yang mengharapkan responden untuk menuliskan jawabannya berbentuk uraian tentang sesuatu hal. Sebaliknya pertanyaan tertutup adalah pertanyaan yang mengharapkan jawaban singkat atau mengharapkan responden untuk memilih salah satu alternatif jawaban dari setiap pertanyaan yang telah tersedia. Setiap pertanyaan angket yang mengharapkan jawaban berbentuk data nominal, ordinal, interval, dan rasio, adalah bentuk pertanyaan tertutup (Sugiyono 2017:143).

### 3.6 Teknik Analisis Data

Pada penelitian ini, data dianalisis menggunakan analisis kuantitatif karena sebagian besar datanya bersifat numerik dan metode yang digunakan adalah metode matematika statistika dan berperan di balik layar menggunakan Bahasa pemrograman python dengan model *Support Vector Machine*. Tahapan yang dilakukan dalam analisis datanya adalah:

- a. Pengumpulan data, Yang dimana data akan dikumpulkan dengan cara memberikan angket kepada mahasiswa yang merupakan sampel

penelitian. Data yang akan terkumpul adalah data akademik dan data demografis mahasiswa.

- b. Pre-pemrosesan data, dalam pre-pemrosesan data yang akan dilakukan adalah pembersihan data yang tidak konsisten, melakukan encoding dengan cara data diubah kedalam format yang sesuai dengan model SVM. Setelah itu, data dibagi menjadi 2 bagian yaitu data latih dan data uji dengan skala 70:30.
- c. Implementasi SVM, Menggunakan Library scikit-learn di python dan memilih kernel SVM yang sesuai yaitu SVM linear. Setelah itu, model SVM dilatih menggunakan data latih yang telah diproses. Proses ini melibatkan penggunaan algoritma optimasi untuk meminimalkan error dan meningkatkan akurasi.
- d. Evaluasi Model, dimana model yang telah dilatih di evaluasi menggunakan data uji. Metrik evaluasi yang dimungkinkan untuk digunakan adalah metrik akurasi, presisi dan recall, F1-Score dan Kurva ROC.
- e. Deployment model, yang dimana Setelah model diuji dan menunjukkan hasil yang memuaskan, model dapat di-deploy untuk digunakan dalam sistem yang sebenarnya.

### 3.7 Tempat dan Jadwal Penelitian

#### 1. Tempat Penelitian

Penelitian ini dilaksanakan di Universitas Nias yang beralamat di Jln. Yos<sup>5</sup> Sudarso No. 118/E-S. Ombolata Ulu, Kota Gunungsitoli, Sumatera Utara, Nias KodePos 22812.

#### 2. Jadwal Penelitian

Adapun jadwal penelitian ini diperkirakan pada semester genap tahun akademik 2023/2024. Pemaparan jadwal penelitian bisa dilihat pada tabel berikut:

No.	Kegiatan	Waktu Kegiatan			
		November 2023	April - Mei 2024	Juni 2024	Juni – Juli 2024
1	Pengajuan Judul	√			
2	Pengumpulan Literatur		√		
3	Seminar Proposal Penelitian			√	
4	Penelitian				√

## BAB IV PEMBAHASAN

### 4.1 Hasil Penelitian

#### 4.1.1 Data

Menurut Kamus Besar Bahasa Indonesia, data ialah keterangan yang benar dan nyata, keterangan atau bahan nyata yang dapat di jadikan dasar kajian atau informasi dalam bentuk yang dapat diproses oleh komputer. Dalam penelitian ini, sebelum proses perhitungan data dilakukan, terlebih dahulu dilakukan pengumpulan data. Pengumpulan data dilakukan dengan menyebarkan kuesioner dengan bantuan google form kepada mahasiswa sebagai responden melalui group Whatsapp maupun dengan menghubungi langsung kontak tesponden. Penyebaran kuesioner dilakukan dari tanggal 29 Juli – 03 Agustus 2024. Data yang diperoleh terdapat pada lampiran 1 dengan jumlah responden terdiri atas 193 responden mahasiswa Pendidikan mahasiswa Universitas Nias.

### 4.2. Implementasi Algoritma Machine Learning

#### 4.2.1 Pengumpulan Data

Data yang telah di kumpulkan, berisikan data mahasiswa sebanyak 193 mahasiswa dengan mengisi kuoesioner denga jumlah pertanyaan sebanyak 55 pertanyaan terkait kehidupan perkuliahan mahasiswa. Berikut gambaran singkat data mahasiswa yang telah dikumpulkan:

	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	SEMESTER	IPK	Saya mera	Nilai ujian	Saya menc	Saya jaran	Saya hadir	Saya sering	Saya mera	Saya terlib	Saya berin	Saya terlib	Saya memi	Kegiatan e
2	II	3,32	2	3	2	5	4	4	4	5	2	1	2	3
3	II	3,06	3	1	3	4	5	4	4	4	3	3	3	3
4	II	2,64	1	3	1	1	2	3	3	3	1	3	1	3
5	II	2,89	3	3	3	3	4	3	4	4	3	3	3	3
6	II	3,25	5	5	4	4	5	4	5	5	4	4	3	4
7	II	3,01	1	2	1	4	4	3	4	4	3	2	2	4
8	II	3,55	5	4	4	4	4	4	4	4	3	3	3	4
9	II	3,58	4	4	3	5	5	3	3	4	3	2	2	3
10	II	2,86	4	4	3	3	4	3	4	3	2	3	3	3
11	II	3,4	3	3	3	4	5	4	4	4	3	4	3	3
12	II	2,5	3	3	3	3	4	4	4	4	3	3	2	4
13	II	3,06	2	4	2	2	2	4	4	3	3	3	2	5
14	II	3,18	4	3	3	3	5	3	4	4	3	2	1	3
15	II	3,08	4	4	4	4	5	4	4	4	3	4	3	5
16	II	3,18	4	4	4	4	5	3	4	4	3	4	3	4
17	II	3,14	2	3	3	5	4	5	5	5	4	2	4	5
18	II	3,59	4	4	4	4	4	4	4	4	4	4	4	4

Gambar 4.1 Hasil Angket Mahasiswa

Dari data yang telah di kumpulkan data telah di susun sedemikian rupa dengan memisahkan data mahasiswa setiap semesternya. Tujuannya adalah untuk melihat perbandingan kenaikan atau penurunan nilai IPK di setiap semesternya.

#### 4.2.2 Data Pre-processing

Data yang telah di kumpulkan di bersihkan dari data yang bersifat acak menjadi data diskrit. Pada proses pembersihan data yang telah di lakukan, telah terjadi penghapusan data untuk kolom nomor baris data dan nama mahasiswa, dan telah terjadi penyeimbangan data terhadap kolom semester yang dimana data sebelumnya berisi angka romawi di ubah menjadi data yang bersifat numerik. Berikut data yang telah di bersihkan:

1	SEMESTER	IPK	Per1	Per2	Per3	Per4	Per5	Per6	Per7	Per8	Per9	Per10	Per11	Per12	Per13	Per14	Per
2	2	3,32	2	3	2	5	4	4	4	5	2	1	2	3	3	2	
3	2	3,06	3	1	3	4	5	4	4	3	3	3	3	3	3	3	
4	2	2,64	1	3	1	1	2	3	3	3	1	3	1	3	1	3	
5	2	2,89	3	3	3	3	4	3	4	4	3	3	3	3	4	3	
6	2	3,25	5	5	4	4	5	4	5	5	4	4	3	4	4	4	
7	2	3,01	1	2	1	4	4	3	4	4	3	2	2	4	4	4	
8	2	3,55	5	4	4	4	4	4	4	4	3	3	3	4	4	3	
9	2	3,58	4	4	3	5	5	3	3	4	3	2	2	3	3	4	
10	2	2,86	4	4	3	3	4	3	4	3	2	3	3	3	3	3	
11	2	3,4	3	3	3	4	5	4	4	4	3	4	3	3	4	4	
12	2	2,5	3	3	3	3	4	4	4	4	3	3	2	4	4	4	
13	2	3,06	2	4	2	2	2	4	4	3	3	3	2	5	3	3	
14	2	3,18	4	3	3	3	5	3	4	4	3	2	1	3	3	3	
15	2	3,08	4	4	4	4	5	4	4	4	3	4	3	5	5	3	
16	2	3,18	4	4	4	4	5	3	4	4	3	4	3	4	3	3	
17	2	3,14	2	3	3	5	4	5	5	5	4	2	4	5	4	4	
18	2	3,59	4	4	4	4	4	4	4	4	4	4	4	4	4	4	
19	2	3,25	4	4	4	4	5	4	4	4	4	4	2	2	4	4	
20	2	3,39	3	5	5	5	5	5	4	5	3	5	3	3	5	3	
21	2	2,91	3	3	4	4	4	4	5	4	3	3	4	3	4	5	4

Gambar 4.2 Data yang telah di bersihkan

Setelah data di bersihkan, data di bagi menjadi data training dan data testing dengan skala 70:30.

#### 4.2.3 Pemodelan Support Vector Machine

Data yang telah mengalami pembersihan, maka Langkah selanjutnya yang dilakukan adalah memodelkan data menggunakan model SVM. Dalam hal ini, tahapan dalam pemodelan data, dilakukan menggunakan aplikasi Visual Studio Code dengan Bahasa pemrograman Python. Dataset yang digunakan pada pemodelan ini sebanyak 193 data. Bahasa pemrograman python digunakan untuk mengetahui tingkat akurasi algoritma SVM. Langkah – Langkah dalam pemodelan adalah sebagai berikut:

##### 4.2.3.1 Import dataset

Sebelum dataset di import, maka hal yang seharusnya dilakukan adalah mengumpulkan library yang akan digunakan dalam pemodelan. Library yang dikumpulkan adalah library yang akan digunakan dalam pemograman. Selain mengimport library, hal yang harus diimport adalah mengimport modul yang sinkron dengan library yang akan digunakan

```

no1.py > ...
1 import pandas as pd
2 import seaborn as sns
3 import matplotlib.pyplot as plt
4 import numpy as np
5 from sklearn.model_selection import train_test_split
6 from sklearn.preprocessing import StandardScaler
7 from sklearn.preprocessing import KBinsDiscretizer
8 from sklearn.preprocessing import LabelEncoder
9 from sklearn import svm
10 from sklearn.metrics import accuracy_score, classification_report, confusion_matrix
11
12 #load data
13 df = pd.read_excel("Data Mahasiswa FIXS.xlsx")
14 print(df)
15 print(df.head())
16 print(df.info())
17 print(df.columns)
18
--

```

Gambar 4.3 Import dataset

Hasil data yang telah di import adalah sebagai berikut:

```

PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS
PS C:\Users\USER\data pandas numpy> & "C:/Program Files/Python312/python.exe" "c:/Users/USER/data pandas numpy/no1.py"
SEMESTER  IPK  Per1  Per2  ...  Per52  Per53  Per54  Per55
0         2  3.32  2    3  ...    4     5     4     3
1         2  3.06  3    1  ...    3     3     1     3
2         2  2.64  1    3  ...    3     5     1     5
3         2  2.89  3    3  ...    3     3     3     3
4         2  3.25  5    5  ...    4     4     4     4
..
188      8  3.13  2    3  ...    3     5     5     4
189      8  3.19  3    4  ...    5     5     5     4
190      8  2.84  4    4  ...    4     4     5     5
191      8  2.89  1    3  ...    3     5     1     3
192      8  3.05  4    4  ...    4     4     5     5

[193 rows x 57 columns]
SEMESTER  IPK  Per1  Per2  Per3  ...  Per51  Per52  Per53  Per54  Per55
0         2  3.32  2    3    2  ...    3     4     5     4     3
1         2  3.06  3    1    3  ...    3     3     3     1     3
2         2  2.64  1    3    1  ...    3     3     5     1     5
3         2  2.89  3    3    3  ...    3     3     3     3     3
4         2  3.25  5    5    4  ...    4     4     4     4     4

[5 rows x 57 columns]

```

Gambar 4.4 Hasil Load data

#### 4.2.3.2 Membuat DataFrame

Setelah mengimport data, Langkah selanjutnya adalah membuat DataFrame dengan kodingan sebagai berikut:

```

20 #Scale data
21 scaler = StandardScaler()
22 df_scaled = scaler.fit_transform(df[['SEMESTER', 'IPK', 'Per1', 'Per2', 'Per3', 'Per4', 'Per5', 'Per6',
23 'Per7', 'Per8', 'Per9', 'Per10', 'Per11', 'Per12', 'Per13', 'Per14',
24 'Per15', 'Per16', 'Per17', 'Per18', 'Per19', 'Per20', 'Per21', 'Per22',
25 'Per23', 'Per24', 'Per25', 'Per26', 'Per27', 'Per28', 'Per29', 'Per30',
26 'Per31', 'Per32', 'Per33', 'Per34', 'Per35', 'Per36', 'Per37', 'Per38',
27 'Per39', 'Per40', 'Per41', 'Per42', 'Per43', 'Per44', 'Per45', 'Per46',
28 'Per47', 'Per48', 'Per49', 'Per50', 'Per51', 'Per52', 'Per53', 'Per54',
29 'Per55']])
30
31 #split data
32 x = df_scaled[:, :-1]
33 y = df_scaled[:, -1]
34 x_train, x_test, y_train, y_test = train_test_split(x, y, test_size=0.3, random_state=0)
35
36 label_encoder = LabelEncoder()
37 label_encoder.fit(y_train)
38 y_encoded = label_encoder.transform(y_train)
39 y_test_encoded = label_encoder.transform(y_test)
40
41 bins = [0,2,0,2.7,3.2,4,0]
42 labels = ['kurang', 'cukup', 'baik', 'sangat baik']
43 df['IPK_Binned'] = pd.cut(df['IPK'], bins=bins, labels=labels, include_lowest = True)
44 y_train_binned = np.digitize(y_train, bins)-1
45 y_test_binned = np.digitize(y_test, bins)-1
46

```

Gambar 4.5 DataFrame

Setelah membuat kode tersebut maka hasil yang di dapatkan adalah sebagai berikut:

```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 193 entries, 0 to 192
Data columns (total 57 columns):
#   Column      Non-Null Count  Dtype
---  -
0   SEMESTER    193 non-null    int64
1   IPK          193 non-null    float64
2   Per1        193 non-null    int64
3   Per2        193 non-null    int64
4   Per3        193 non-null    int64
5   Per4        193 non-null    int64
6   Per5        193 non-null    int64
7   Per6        193 non-null    int64
8   Per7        193 non-null    int64
9   Per8        193 non-null    int64
10  Per9        193 non-null    int64
11  Per10       193 non-null    int64
12  Per11       193 non-null    int64
13  Per12       193 non-null    int64
14  Per13       193 non-null    int64
15  Per14       193 non-null    int64
16  Per15       193 non-null    int64
17  Per16       193 non-null    int64
18  Per17       193 non-null    int64
19  Per18       193 non-null    int64
20  Per19       193 non-null    int64
21  Per20       193 non-null    int64
22  Per21       193 non-null    int64
23  Per22       193 non-null    int64
24  Per23       193 non-null    int64
25  Per24       193 non-null    int64
26  Per25       193 non-null    int64
27  Per26       193 non-null    int64
28  Per27       193 non-null    int64
29  Per28       193 non-null    int64
30  Per29       193 non-null    int64
31  Per30       193 non-null    int64
32  Per31       193 non-null    int64
33  Per32       193 non-null    int64
34  Per33       193 non-null    int64
35  Per34       193 non-null    int64
36  Per35       193 non-null    int64
37  Per36       193 non-null    int64

```

Gambar 4.6 hasil Run DataFrame

Dari hasil ini, dapat diketahui bahwa Range index datanya sebanyak 193 entries, kolom data sebanyak 57 kolom dan tipe data yaitu dtypes: float64(1), int64(56).

#### 4.2.3.3 Evaluasi Model

Dalam evaluasi model, digunakan kode sebagai berikut:

```
46
47 #Create and train SVM model
48 svm_model = svm.SVC(kernel='linear',C=1.0, gamma=0.1)
49 svm_model.fit(x_train,y_train_binned)
50
51 #Model Evaluation
52 y_pred = svm_model.predict(x_test)
53 accuracy = accuracy_score(y_test_binned, y_pred)
54 print("accuracy:",accuracy)
55 print(f"Akurasinya: {accuracy:.2f}")
56 print("Classification Report:")
57 print(classification_report(y_test_binned, y_pred))
58 print("Confusion Matrix:")
59 print(confusion_matrix(y_test_binned, y_pred))
60
```

Gambar 4.7 codingan evaluasi model

### 4.3 Hasil

Hasil pemodelan yang menggunakan algoritma machine learning dengan model SVM, terlihat dari hasil akurasi sebagai berikut:

```
dtype='object')
accuracy: 0.7758620689655172
Akurasinya: 0.78
Classification Report:
      precision    recall  f1-score   support

   -1         0.67     0.70     0.68         20
    0         0.84     0.82     0.83         38

 accuracy          0.78         0.78         0.78         58
 macro avg         0.75         0.76         0.75         58
 weighted avg         0.78         0.78         0.78         58

Confusion Matrix:
[[14  6]
 [ 7 31]]
DE C:\Users\HSEB\data reader\svm\svm
```

Gambar 4.8 Hasil Evaluasi Model

Dari hasil evaluasi model diatas dapat terlihat bahwa, kinerja model Support Vector Machine (SVM) dalam mengklasifikasi data IPK dievaluasi menggunakan berbagai metrik. Akurasi model ditemukan sebesar 0,7759, yang menunjukkan bahwa sekitar 77,59% dari instance diklasifikasi dengan benar.

Analisis yang lebih rinci tentang kinerja klasifikasi disajikan dalam laporan klasifikasi. Presisi, recall, dan F1-score untuk setiap kelas disajikan, serta dukungan (yaitu jumlah instance) untuk setiap kelas. Hasil menunjukkan bahwa model mencapai presisi 0,67 dan recall 0,70 untuk kelas yang diberi label -1, serta presisi 0,84 dan recall 0,82 untuk kelas yang diberi label 0. F1-score, yang merupakan rata-rata harmonik dari presisi dan recall, adalah 0,68 untuk kelas -1 dan 0,83 untuk kelas 0. Rata-rata makro dan rata-rata terbobot dari presisi, recall, dan F1-score juga disajikan. Rata-rata makro memberikan bobot yang sama untuk setiap kelas, sedangkan rata-rata terbobot mengambil akun dukungan untuk setiap kelas. Hasil menunjukkan bahwa rata-rata makro F1-score adalah 0,75, sedangkan rata-rata terbobot F1-score adalah 0,78.

Confusion matrix menyajikan pandangan yang lebih rinci tentang kinerja klasifikasi. Matriks menunjukkan bahwa dari 20 instance kelas -1, 14 diklasifikasi dengan benar, sedangkan 6 salah diklasifikasi sebagai kelas 0. Demikian pula, dari 38 instance kelas 0, 31 diklasifikasi dengan benar, sedangkan 7 salah diklasifikasi sebagai kelas -1. Secara keseluruhan, hasil menunjukkan bahwa model SVM dapat mencapai tingkat akurasi yang wajar dalam mengklasifikasi data IPK, dengan sedikit bias ke kelas 0. Namun, kinerja model dapat ditingkatkan dengan mengoptimalkan hyperparameter atau menjelajahi algoritma pembelajaran mesin lainnya.

## **BAB V**

### **SARAN DAN KESIMPULAN**

#### **5.1 Kesimpulan**

Berdasarkan hasil analisis, implementasi machine learning menggunakan Support Vector Machine (SVM) untuk deteksi performa akademik mahasiswa menunjukkan akurasi sebesar 0,7759. Hasil ini menunjukkan bahwa model SVM dapat memprediksi performa akademik mahasiswa dengan cukup akurat. Namun, hasil klasifikasi report menunjukkan bahwa model SVM memiliki bias ke kelas 0, yang berarti bahwa model lebih baik dalam memprediksi mahasiswa dengan performa akademik yang baik daripada mahasiswa dengan performa akademik yang kurang baik. Matriks kebingungan juga menunjukkan bahwa model SVM memiliki tingkat kesalahan yang cukup tinggi dalam memprediksi mahasiswa dengan performa akademik yang kurang baik. Oleh karena itu, perlu dilakukan pengoptimalan model dan penggunaan teknik lain untuk meningkatkan akurasi prediksi.

#### **5.2 Saran**

Berdasarkan hasil analisis, beberapa saran dapat diberikan untuk meningkatkan akurasi deteksi performa akademik mahasiswa menggunakan machine learning:

- a) Penggunaan teknik feature selection untuk memilih fitur yang paling relevan dengan performa akademik mahasiswa.
- b) Penggunaan teknik oversampling atau undersampling untuk mengatasi masalah imbalance data.
- c) Penggunaan model machine learning lain, seperti Random Forest atau Gradient Boosting, untuk membandingkan kinerja dengan model SVM.
- d) Penggunaan teknik ensemble untuk menggabungkan hasil prediksi dari beberapa model machine learning.
- e) Penggunaan data tambahan, seperti data demografi atau data psikologis, untuk meningkatkan akurasi prediksi.

Dengan demikian, implementasi machine learning untuk deteksi performa akademik mahasiswa dapat menjadi lebih akurat dan efektif dalam membantu institusi pendidikan untuk mengidentifikasi mahasiswa yang memerlukan bantuan akademik.



# IMPLEMENTASI ALGORITMA MACHINE LEARNING UNTUK DETEKSI PERFORMA AKADEMIK MAHASISWA

---

ORIGINALITY REPORT

---

# 14%

SIMILARITY INDEX

---

## PRIMARY SOURCES

---

1	<a href="http://epub.imandiri.id">epub.imandiri.id</a> Internet	193 words — 3%
2	<a href="http://digilibadmin.unismuh.ac.id">digilibadmin.unismuh.ac.id</a> Internet	175 words — 3%
3	<a href="http://jurnal.unprimdn.ac.id">jurnal.unprimdn.ac.id</a> Internet	81 words — 1%
4	<a href="http://docplayer.info">docplayer.info</a> Internet	68 words — 1%
5	<a href="http://www.j-cup.org">www.j-cup.org</a> Internet	64 words — 1%
6	<a href="http://repository.unja.ac.id">repository.unja.ac.id</a> Internet	51 words — 1%
7	<a href="http://ejurnal.pps.ung.ac.id">ejurnal.pps.ung.ac.id</a> Internet	49 words — 1%
8	<a href="http://www.gunadarma.ac.id">www.gunadarma.ac.id</a> Internet	38 words — 1%
9	<a href="http://eprints.undip.ac.id">eprints.undip.ac.id</a> Internet	30 words — 1%

---

EXCLUDE QUOTES ON  
EXCLUDE BIBLIOGRAPHY ON

EXCLUDE SOURCES < 1%  
EXCLUDE MATCHES OFF