
Higher Education Students Performance Evaluation

Muhammad Akmal Syarif

(Sistem Informasi,Stmik Tazkia Dramaga Bogor,schevencoxelxynon@gmail.com)

Rafly Ariel Hidayat

(Sistem Informasi,Stmik Tazkia Dramaga Bogor,raflyarielhidayat@gmail.com)

Abstrak

Penelitian ini bertujuan untuk memprediksi performa akademik mahasiswa dengan menggunakan algoritma Support Vector Machine (SVM). Data yang digunakan berasal dari Higher Education Students Performance Evaluation Dataset yang tersedia di UCI Machine Learning Repository (2019). Dataset ini memuat informasi akademik dan demografis mahasiswa seperti kehadiran, nilai tugas, ujian tengah semester, ujian akhir, serta latar belakang keluarga. Metode penelitian meliputi tahap pengumpulan data, pra-pemrosesan, pembagian data menjadi training dan testing set, pembangunan model SVM, serta evaluasi performa menggunakan metrik accuracy, precision, recall, dan F1-score.

Hasil yang diharapkan dari penelitian ini adalah model prediksi dengan akurasi tinggi yang dapat membantu pihak kampus dalam mengidentifikasi mahasiswa yang berisiko memiliki performa akademik rendah. Dengan demikian, hasil penelitian ini dapat menjadi dasar dalam pengambilan keputusan akademik berbasis data di lingkungan pendidikan tinggi.

1. Pendahuluan

Perkembangan teknologi informasi yang pesat telah memberikan dampak signifikan pada dunia pendidikan, khususnya dalam bidang analisis data akademik. Perguruan tinggi kini mampu mengumpulkan berbagai data terkait aktivitas belajar mahasiswa, seperti nilai ujian, kehadiran, partisipasi kelas, serta faktor demografis. Data tersebut dapat dimanfaatkan untuk memperoleh wawasan yang berguna dalam pengambilan keputusan akademik.

Salah satu tantangan utama dalam bidang pendidikan adalah bagaimana cara mendeteksi mahasiswa yang berpotensi memiliki performa rendah sejak dini agar dapat dilakukan intervensi yang tepat. Dalam konteks ini, pendekatan Machine Learning dapat dimanfaatkan untuk membangun model prediktif yang mampu memperkirakan performa akademik mahasiswa berdasarkan data yang tersedia.

Metode Support Vector Machine (SVM) dikenal memiliki kemampuan tinggi dalam memecahkan masalah klasifikasi dengan data berdimensi tinggi dan kompleks. Oleh karena itu, penelitian ini

bertujuan untuk menerapkan algoritma SVM dalam memprediksi performa akademik mahasiswa menggunakan dataset Higher Education Students Performance Evaluation dari UCI Machine Learning Repository yang diterbitkan pada tahun 2019.

2. Metodologi Penelitian

1. Sumber Data

Dataset yang digunakan berasal dari:

- Nama: Higher Education Students Performance Evaluation
- Sumber: UCI Machine Learning Repository
- Tahun: 2019
- Jumlah Data: 145 instance, 33 atribut
- Jenis Data: Data akademik dan demografis mahasiswa
- Tipe Analisis: Klasifikasi (Supervised Learning)

2. Variabel Penelitian

Fitur (input): Gender, Age, Attendance, Quiz, Project, Midterm, Final Exam, Study Hours, Participation, Parent Education Level, dll.

Label (output): Kategori performa akademik mahasiswa (Low, Medium, High).

3. Langkah-Langkah Penelitian

1. Data Collection: Mengunduh dataset dari UCI Machine Learning Repository.
2. Data Preprocessing: Menghapus missing values, encoding data kategorikal, dan normalisasi data.
3. Data Splitting: Membagi dataset menjadi train (80%) dan test (20%).
4. Model Training: Membangun model SVM menggunakan scikit-learn.
5. Model Evaluation: Menghitung Accuracy, Precision, Recall, F1-score, serta visualisasi Confusion Matrix.
6. Comparison: Membandingkan hasil SVM dengan hasil dari paper referensi.

4. Perangkat dan Tools

Bahasa Pemrograman: Python 3

Library: pandas, numpy, scikit-learn, matplotlib, seaborn

IDE: Google Colab / Jupyter Notebook

Sumber Dataset: UCI Machine Learning Repository

Kontrol Versi: GitHub

3. Hasil dan Pembahasan

1. Hasil Evaluasi Kinerja Support Vector Machine (SVM)

Sub-pokok ini menyajikan ringkasan hasil pengujian model **Support Vector Machine (SVM)** untuk memprediksi performa akademik mahasiswa.

- **Metode Klasifikasi:** Model dilatih menggunakan dataset *Higher Education Students Performance Evaluation* dari UCI Machine Learning Repository.
- **Proses Pengujian:** Data dibagi menjadi 80% untuk *training* dan 20% untuk *testing*.
- **Metrik Kinerja:** Kinerja model dievaluasi menggunakan metrik standar klasifikasi:
 - **Accuracy:** *[Nilai Akurasi]* %
 - **Precision:** *[Nilai Precision]*
 - **Recall:** *[Nilai Recall]*
 - **F1-score:** *[Nilai F1-score]*
- **Interpretasi Hasil:** Nilai akurasi sebesar *[Nilai Akurasi]* % menunjukkan bahwa model memiliki kemampuan yang baik dalam memprediksi performa akademik (Low, Medium, High). Hasil ini diharapkan dapat membantu institusi dalam mengidentifikasi mahasiswa yang berisiko memiliki performa rendah.

2. Perbandingan Kinerja SVM dengan Penelitian Terdahulu

Sub-pokok ini membandingkan kinerja model SVM dengan hasil penelitian sebelumnya yang menggunakan algoritma lain pada dataset yang sama.

Algoritma	Sumber Penelitian	Akurasi (%)
Decision Tree	Akçapınar et al. (2019)	<i>[Akurasi Decision Tree]</i>
Naive Bayes	Akçapınar et al. (2019)	<i>[Akurasi Naive Bayes]</i>
KNN	Akçapınar et al. (2019)	<i>[Akurasi KNN]</i>
Support Vector Machine (SVM)	Penelitian Ini	<i>[Nilai Akurasi dari 4.1]</i>

Kesimpulan

Proposal penelitian ini berfokus pada **prediksi performa akademik mahasiswa** menggunakan algoritma **Support Vector Machine (SVM)**.

- **Tujuan Utama:** Menerapkan metode Support Vector Machine (SVM) untuk membangun model *Supervised Learning* yang dapat memprediksi performa akademik mahasiswa.
- **Sumber Data:** Dataset yang digunakan adalah **Higher Education Students Performance Evaluation** dari UCI Machine Learning Repository (2019). Dataset ini mencakup informasi akademik (seperti kehadiran, nilai tugas, ujian tengah dan akhir) dan demografis mahasiswa.
- **Metodologi:** Penelitian ini melibatkan langkah-langkah mulai dari pengumpulan data, *pra-pemrosesan* (menghapus *missing values* dan *encoding* data kategorikal), pembagian data (80% *train*, 20% *test*), pelatihan model SVM, hingga evaluasi model.
- **Evaluasi Model:** Performa model akan diukur menggunakan metrik **Accuracy, Precision, Recall, dan F1-score**, serta akan dibandingkan dengan hasil penelitian terdahulu yang menggunakan algoritma Decision Tree, Naive Bayes, dan KNN.
- **Manfaat yang Diharapkan:** Hasil dari penelitian ini diharapkan dapat menghasilkan model prediksi dengan **akurasi tinggi**, yang dapat digunakan oleh institusi pendidikan untuk **mengidentifikasi mahasiswa yang berisiko** memiliki performa rendah sejak dulu, dan menjadi dasar bagi pengambilan keputusan akademik berbasis data.

Ucapan Terima Kasih

kami mengucapkan terima kasih yang sebesar-besarnya kepada berbagai pihak yang telah memberikan dukungan, bimbingan, dan fasilitas sehingga paper penelitian yang berjudul "Higher Education Students Performance Evaluation" ini dapat tersusun dengan baik.

Rasa terima kasih dan penghargaan ini secara khusus kami sampaikan kepada:

- Bapak Hendri Kharisma S.Kom, M.T , selaku Dosen Pengampu mata kuliah yang telah memberikan arahan, bimbingan, serta wawasan yang mendalam mengenai penerapan *Machine Learning* dalam analisis data akademik.
- Sekolah Tinggi Manajemen Informatika dan Komputer TAZKIA , atas fasilitas dan lingkungan akademik yang mendukung proses pembelajaran dan penelitian ini.
- Seluruh tim dan pihak terkait yang telah menyediakan serta mengelola Higher Education Students Performance Evaluation Dataset di UCI Machine Learning Repository.

Daftar Pustaka

1. Academic Performance Using Machine Learning Approaches. *Journal of Educational Technology & Society*, 22(1), 225–238.
2. Cortes, C., & Vapnik, V. (1995). Support-Vector Networks. *Machine Learning*, 20(3), 273–297.
3. UCI Machine Learning Repository. (2019). Higher Education Students Performance Evaluation Dataset. <https://archive.ics.uci.edu/dataset/856/higher+education+students+performance+evaluation>