

PROPOSAL

Prediksi Aktivitas Fisik Berdasarkan Data Sensor Menggunakan Metode SVM dan XGBoost



Dosen Pembimbing:

Hendri Kharisma S.Kom, M.T

Disusun Oleh:

Muhammad Labib (241552010009)

Program Studi: Teknik Informatika

Sekolah Tinggi Manajemen Informatika dan Komputer Tazkia

Jl. Raya Dramaga Blok Radar Baru No.8, RT.03/RW.03, Margajaya, Kecamatan

Bogor Baru, Kota Bogor, Jawa Barat 16116, Indonesia

2025

Kata Pengantar

Puji syukur penulis panjatkan ke hadirat Allah SWT atas segala limpahan rahmat, taufik, dan karunia-Nya sehingga penulis dapat menyelesaikan penyusunan Laporan Pengajuan Judul Penelitian ini dengan baik.

Laporan ini disusun sebagai salah satu persyaratan dalam tahap awal penyusunan jurnal penelitian pada Program Studi Teknik Informatika. Melalui laporan ini, penulis bermaksud menyampaikan gambaran umum mengenai topik penelitian yang diajukan, yang mencakup latar belakang masalah, tujuan penelitian, manfaat, batasan masalah, serta metodologi yang akan digunakan.

Penulis menyadari sepenuhnya bahwa laporan ini masih jauh dari sempurna. Oleh karena itu, kritik dan saran yang bersifat membangun sangat penulis harapkan demi penyempurnaan laporan ini di masa mendatang.

Akhir kata, penulis berharap semoga laporan ini dapat memberikan manfaat, baik bagi penulis sendiri maupun bagi pihak lain yang tertarik dengan bidang penelitian ini, serta menjadi langkah awal yang baik dalam pelaksanaan penelitian selanjutnya.

Bogor, 24 Oktober 2025

Penulis

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi sensor dan *wearable devices* memungkinkan aktivitas fisik manusia dapat direkam secara akurat dan real-time. Data tersebut berpotensi besar dimanfaatkan untuk mengenali pola gerak, memantau kebugaran, dan mendukung gaya hidup sehat berbasis data.

Daily and Sports Activities Dataset dari **UCI Machine Learning Repository** menyediakan data sensor dari berbagai aktivitas fisik, seperti berjalan, berlari, duduk, bersepeda, dan lainnya. Dengan menggunakan algoritma *machine learning* seperti **Support Vector Machine (SVM)** dan **Extreme Gradient Boosting (XGBoost)**, penelitian ini bertujuan untuk memprediksi jenis aktivitas yang dilakukan seseorang berdasarkan sinyal sensor.

Proyek ini penting karena dapat menjadi dasar pengembangan sistem *human activity recognition* (HAR), yang banyak digunakan dalam aplikasi kebugaran, fisioterapi digital, serta sistem pemantauan pasien lansia.

1.2 Rumusan Masalah

1. Bagaimana cara mengenali aktivitas fisik berdasarkan data sensor dari tubuh manusia?
2. Seberapa akurat algoritma SVM dan XGBoost dalam mengklasifikasikan aktivitas tersebut?
3. Fitur sensor mana yang paling berpengaruh terhadap hasil klasifikasi aktivitas?

1.3 Tujuan Penelitian

- Mengembangkan model *supervised learning* untuk mengenali aktivitas fisik dari data sensor.
- Membandingkan performa **SVM** dan **XGBoost** dalam klasifikasi aktivitas manusia.
- Mengidentifikasi fitur sensor yang paling berpengaruh terhadap prediksi aktivitas.

BAB II

DESKRIPSI DATASET

2.1 Dataset

- Nama Dataset: Daily and Sports Activities Dataset
- Sumber: UCI Machine Learning Repository
- URL: <https://archive.ics.uci.edu/dataset/256/daily+and+sports+activities>
- Jumlah Data: ±912.000 instance dari 8 subjek
- Jumlah Fitur: 45 atribut sensor (IMU di tangan, dada, dan kaki)
- Jumlah Kelas: 19 aktivitas berbeda
Contoh: Walking, Running, Sitting, Standing, Cycling, Jumping, Playing Ball, dll.
- Tipe Data: Time-series / Multivariate Sensor Data

Kolom	Deskripsi
Subject	ID unik untuk setiap partisipan penelitian.
Activity	Jenis aktivitas fisik yang dilakukan (misalnya: berjalan, duduk, berdiri, berlari, naik tangga, dll.).
Sensor_1 - Sensor_45	Data numerik hasil pembacaan sensor dari perangkat IMU (Inertial Measurement Unit) yang meliputi akselerasi, kecepatan sudut, dan orientasi tubuh pada berbagai bagian tubuh (torso, lengan, tungkai, dan kaki).
Class	Label aktivitas atau kategori hasil pengukuran yang digunakan sebagai variabel target untuk klasifikasi gaya hidup.

Catatan: Dataset Daily and Sports Activities dari UCI Machine Learning Repository berisi data aktivitas fisik yang dikumpulkan dari 8 subjek selama mereka melakukan 19 jenis aktivitas berbeda. Setiap aktivitas direkam menggunakan sensor IMU yang terpasang pada 5 bagian tubuh utama (torso, lengan kanan/kiri, dan kaki kanan/kiri). Dataset ini dapat dimanfaatkan untuk menganalisis pola aktivitas dan mengaitkannya dengan indikator kesehatan atau gaya hidup.

Variabel Target:

- Target (label) adalah jenis aktivitas yang dilakukan oleh subjek, dikategorikan ke dalam **19 kelas aktivitas fisik**.

BAB III

METODOLOGI PENELITIAN

3.1 Tahapan Penelitian

1. Pengumpulan dan Pemahaman Data

Mengunduh dataset dari UCI Repository dan melakukan eksplorasi awal terhadap setiap fitur dan struktur data sensor.

2. Pra-pemrosesan Data (Data Preprocessing)

- Menangani data kosong (*missing values*)
- Encoding label aktivitas menjadi numerik.
- Normalisasi fitur numerik menggunakan *StandardScaler*
- Membagi data menjadi train (80%) dan test (20%)

3. Pengembangan Model

- **Model 1:** Support Vector Machine (SVM)
- **Model 2:** XGBoost Classifier
- Proses tuning dilakukan menggunakan **GridSearchCV** untuk mendapatkan parameter terbaik.

4. Evaluasi Model

Evaluasi dilakukan menggunakan metrik: Accuracy, Precision, Recall, F1-score, ROC-AUC Score, Confusion Matrix

5. Analisis dan Interpretasi

- Membandingkan performa kedua algoritma.
- Menganalisis fitur sensor paling berpengaruh melalui *feature importance* (XGBoost).

3.2 Alur Penelitian

Alur kerja penelitian ini dapat digambarkan sebagai berikut:

1. Input Dataset
2. Preprocessing
3. Training Model
4. Evaluation
5. Visualization & Interpretation
6. Conclusion

3.3 Jadwal Kegiatan (2 Minggu)

Minggu	Kegiatan
1	Pengumpulan data, eksplorasi, dan pra-pemrosesan
2	Pelatihan model, evaluasi hasil, dan penulisan laporan.

3.4 Hasil yang Diharapkan

- Model klasifikasi aktivitas fisik berbasis data sensor dengan akurasi tinggi.
- Perbandingan kinerja objektif antara **SVM** dan **XGBoost**.
- Visualisasi hasil seperti *confusion matrix*, *ROC curve*, dan *feature importance*.

BAB IV

KESIMPULAN DAN RENCANA

4.1 Kesimpulan

Penelitian ini menunjukkan bagaimana **machine learning** dapat diterapkan untuk mengenali aktivitas fisik berdasarkan data sensor. Dengan algoritma **SVM** dan **XGBoost**, diharapkan dapat diperoleh model yang efisien, akurat, dan adaptif terhadap data multivariat seperti sinyal gerakan tubuh. **SVM** efektif pada dataset menengah dengan batas klasifikasi non-linear, sedangkan **XGBoost** unggul dalam menangani kompleksitas data dan meningkatkan akurasi melalui *boosting*.

4.2 Rencana

Rencana pengembangan meliputi:

- Optimasi model melalui *hyperparameter tuning* dan *cross-validation*.
- Eksperimen dengan algoritma tambahan seperti Random Forest dan Neural Network.
- Implementasi ke dalam aplikasi berbasis *Flask* atau *Streamlit* untuk prediksi interaktif.
- Analisis interpretabilitas menggunakan *SHAP values* atau *LIME*.
- Publikasi hasil ke platform riset atau repositori publik seperti GitHub.

Dengan fokus pada akurasi dan interpretabilitas, penelitian ini diharapkan mendukung pengembangan sistem **Human Activity Recognition** (HAR) berbasis *machine learning* dalam konteks gaya hidup sehat.

BAB V

TINJAUAN PUSTAKA

1. UCI Machine Learning Repository – Daily and Sports Activities Dataset.
<https://archive.ics.uci.edu/ml/datasets/daily+and+sports+activities>
2. Anguita, D., Ghio, A., Oneto, L., Parra, X., & Reyes-Ortiz, J.L. (2013).
A Public Domain Dataset for Human Activity Recognition Using Smartphones.
3. Zhang, Q., et al. (2022). *Human Activity Recognition Based on Sensor Data Using XGBoost and SVM Algorithms.*
4. Kuhn, M. & Johnson, K. (2019). *Applied Predictive Modeling.* Springer.