

Evaluasi Kinerja Berbagai Algoritma Machine Learning dalam Prediksi Data Klasifikasi

Abstrak

Penelitian ini bertujuan untuk mengevaluasi kinerja beberapa algoritma machine learning pada kasus klasifikasi, seperti Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Random Forest, dan Naive Bayes. Model diuji menggunakan satu atau beberapa dataset klasifikasi standar dan dibandingkan berdasarkan metrik akurasi, precision, recall, dan F1-score. Hasil menunjukkan bahwa setiap algoritma memiliki keunggulan berbeda tergantung karakteristik data, seperti jumlah fitur, sebaran kelas, dan tingkat noise. Temuan ini diharapkan dapat menjadi panduan awal dalam pemilihan algoritma klasifikasi yang sesuai untuk kebutuhan pengembang sistem berbasis machine learning.

Kata kunci: machine learning, klasifikasi, evaluasi kinerja, SVM, Random Forest

Pendahuluan

Perkembangan teknologi informasi mendorong pemanfaatan machine learning dalam berbagai bidang seperti kesehatan, keuangan, pendidikan, hingga e-commerce. Salah satu tugas paling umum dalam machine learning adalah klasifikasi, yaitu memetakan data ke dalam kelas-kelas tertentu berdasarkan pola yang dipelajari dari data latih. Namun, terdapat banyak algoritma klasifikasi yang dapat digunakan, dan masing-masing memiliki karakteristik, kompleksitas, serta performa yang berbeda.

Tanpa evaluasi yang sistematis, pemilihan algoritma sering kali hanya berdasarkan kebiasaan atau contoh di internet, bukan pada kecocokan dengan karakteristik data yang digunakan. Oleh karena itu, penelitian ini berfokus pada evaluasi kinerja beberapa algoritma klasifikasi populer untuk memberikan gambaran perbandingan performa dan membantu peneliti maupun praktisi dalam menentukan algoritma yang lebih sesuai untuk jenis data tertentu.(Alfaruk et al., 2023; Mohadib et al., 2025; Muslim et al., 2025; Rasyidin Alfaruk, 2024; Saddam RA et al., 2023)

Metode Penelitian

Penelitian ini menggunakan pendekatan eksperimen kuantitatif dengan tahapan sebagai berikut:

1. Pengumpulan data

- Menggunakan satu atau lebih dataset klasifikasi publik (misalnya data penyakit, data pelanggan, atau data transaksi).
- Dataset dibersihkan dari data duplikat dan data yang tidak lengkap.

2. **Pra-pemrosesan data**

- Menangani missing value (menghapus atau imputasi).
- Normalisasi atau standarisasi fitur numerik.
- Encoding untuk fitur kategorikal (misalnya one-hot encoding).

3. **Pembagian data**

- Data dibagi menjadi data latih dan data uji, misalnya dengan perbandingan 80:20 atau menggunakan k-fold cross validation (misalnya k = 5 atau 10).

4. **Pembangunan model**

- Membangun model klasifikasi dengan beberapa algoritma:
 - Support Vector Machine (SVM)
 - K-Nearest Neighbor (KNN)
 - Random Forest
 - Naive Bayes
- Setiap model dilatih menggunakan data latih dengan parameter awal (default) dan/atau hasil tuning sederhana (grid search atau random search).

5. **Evaluasi model**

- Mengukur performa model pada data uji menggunakan metrik:
 - Akurasi
 - Precision
 - Recall
 - F1-score
 - Confusion matrix
- Membandingkan hasil semua algoritma secara tabular dan grafik.

6. **Analisis dan interpretasi**

- Menjelaskan perbedaan performa antar algoritma berdasarkan karakteristik dataset dan parameter model.

Hasil

Secara umum, hasil eksperimen (yang nanti kamu isi angka detail-nya) bisa digambarkan seperti ini:

- Random Forest cenderung memberikan akurasi dan F1-score tertinggi pada dataset dengan jumlah fitur yang cukup banyak dan pola non-linear.
- SVM menunjukkan performa yang sangat baik pada data yang relatif bersih, terpisah dengan margin yang jelas, terutama setelah proses normalisasi fitur.
- KNN memberikan hasil yang cukup baik pada dataset berukuran kecil hingga menengah, namun performanya menurun pada data berukuran sangat besar karena beban komputasi saat prediksi.
- Naive Bayes bekerja cukup baik pada data teks atau data dengan asumsi independensi fitur yang relatif terpenuhi, dengan waktu komputasi pelatihan yang sangat cepat.

Hasil ini biasanya ditampilkan dalam bentuk tabel berisi nilai akurasi, precision, recall, dan F1-score untuk masing-masing algoritma, lalu diperkuat dengan grafik seperti bar chart atau line chart.

Pembahasan

Berdasarkan hasil evaluasi, terlihat bahwa tidak ada satu algoritma yang selalu unggul untuk semua jenis dataset. Random Forest unggul pada data yang kompleks dan bervariasi karena kemampuannya membangun banyak pohon keputusan dan melakukan voting, sehingga lebih tahan terhadap overfitting dibandingkan pohon tunggal. Sementara itu, SVM sangat bergantung pada pemilihan kernel dan parameter regulasi; jika parameter disetel dengan baik, SVM dapat memberikan hasil sangat baik pada data dengan margin pemisah yang jelas. KNN relatif mudah diimplementasikan dan dipahami, namun sangat sensitif terhadap skala fitur dan nilai k yang dipilih. Selain itu, biaya komputasi saat prediksi cukup tinggi karena perlu menghitung jarak terhadap semua data latih. Naive Bayes, meskipun sering kali lebih sederhana, dapat menjadi pilihan yang efektif untuk masalah tertentu seperti klasifikasi teks berlabel (spam, sentimen, dll.) karena struktur datanya cocok dengan asumsi algoritma. Dari sudut pandang praktis, pemilihan algoritma tidak hanya melihat akurasi tertinggi, tetapi juga mempertimbangkan waktu pelatihan, waktu inferensi, kebutuhan sumber daya komputasi, serta kemudahan interpretasi model.

Kesimpulan

Penelitian ini menyimpulkan bahwa kinerja algoritma machine learning untuk tugas klasifikasi sangat dipengaruhi oleh karakteristik dataset dan parameter model yang digunakan. Random Forest dan SVM cenderung memberikan performa yang tinggi pada banyak kasus, namun KNN dan Naive Bayes tetap relevan terutama untuk data dengan ukuran tertentu dan kebutuhan

komputasi yang berbeda. Peneliti atau praktisi disarankan untuk tidak hanya mengandalkan satu algoritma, tetapi melakukan eksperimen awal dengan beberapa model, disertai evaluasi menggunakan metrik yang beragam. Penelitian lanjutan dapat mengembangkan studi ini dengan menambahkan lebih banyak algoritma, melakukan tuning hiperparameter yang lebih mendalam, atau menguji pada domain data spesifik seperti kesehatan, keuangan, dan pendidikan.

Daftar Pustaka

Alfaruk, S. R., Qoyum, A. A., Sugiono, S., & Haryono, W. (2023). PERANCANGAN APLIKASI BUKU INDUK SISWA (BISA) BERBASIS WEB MENGGUNAKAN METODE EXTREME PROGRAMMING DI MTS MATHLA'UL ANWAR BAROS. *Journal Information & Computer*, 1(1), 18–24. <https://doi.org/10.32493/jicomisc.v1i1.26724>

Mohadib, M., Widjdi, A. F., Fathoni, A., & Faruq, S. R. Al. (2025). Strengthening Cyber Resilience through IoT Best Practices: Future Applications. *Formosa Journal of Science and Technology*, 4(1), 349–368. <https://doi.org/10.55927/fjst.v4i1.13260>

Muslim, A. C., Hutahaean, H. A., & Alfaruk, S. R. (2025). Prediksi Kebutuhan Energi Listrik di Banten: Proyeksi dan Skenario Transisi Energi Berkelanjutan. *Blend Sains Jurnal Teknik*, 4(2), 280–287. <https://doi.org/10.56211/blendsains.v4i2.1231>

Rasyidin Alfaruk, S. (2024). Pelatihan Web CBT untuk Ujian Siswa pada Guru MTs Mathla'ul Anwar Baros. *Jurnal Pengabdian Dharma Laksana*, 6(2), 534–538. <https://doi.org/10.32493/j.pdl.v6i2.52969>

Saddam RA, Angga Pranata, Sugiono, Rizki Zulanggara, Nur Halimah, Sri Nur H, Rosdiana SM, Nurhalim, & Aprina Handayani. (2023). Sniffing and Spoofing in Computer Security. *International Journal of Integrative Sciences*, 2(6), 881–886. <https://doi.org/10.55927/ijis.v2i6.4528>