

PEMANFAATKAN BAYESIAN SEARCH PADA SUPPORT VECTOR MACHINE (SVM) UNTUK KLASIFIKASI TANAMAN HERBAL AFRODISIAK

Puguh Jayadi^{1)*}, Weka Sidha Bhagawan²⁾, Jofanza Denis Aldida¹⁾

¹⁾Teknik Informatika, Fakultas Teknik, Universitas PGRI Madiun

²⁾Farmasi, Fakultas Ilmu Kesehatan dan Sains, Universitas PGRI Madiun

email: puguh.jayadi@unipma.ac.id^{1)*}

Abstrak

Penelitian di bidang klasifikasi tanaman herbal telah berkembang pesat seiring meningkatnya kebutuhan akan identifikasi sumber bahan alami dengan potensi farmakologis. Salah satu kategori penting adalah tanaman herbal afrodisiak, yang memiliki kandungan senyawa aktif untuk meningkatkan vitalitas dan fungsi reproduksi. Namun, proses klasifikasi tanaman ini sering menghadapi tantangan karena keterbatasan data, kompleksitas fitur, dan ketidaktepatan model klasifikasi konvensional. Model Support Vector Machine (SVM) telah banyak digunakan karena kemampuannya dalam menangani data non-linear dan berdimensi tinggi, tetapi performanya bergantung pada pemilihan hiperparameter yang optimal. Selama ini, penentuan hiperparameter sering dilakukan secara manual atau menggunakan pencarian grid yang memakan waktu dan kurang efisien. Penelitian ini mengusulkan pendekatan Bayesian Search Optimization untuk mengoptimalkan parameter SVM dalam tugas klasifikasi tanaman herbal afrodisiak. Metode ini secara adaptif mengeksplorasi ruang parameter dan mempelajari hubungan probabilistik antar variabel untuk menemukan kombinasi optimal. Hasil eksperimen menunjukkan peningkatan signifikan terhadap model baseline, dengan kenaikan precision sebesar 122.34%, recall sebesar 36.87%, dan f1-score sebesar 79.40%. Peningkatan ini membuktikan bahwa integrasi Bayesian Search dapat secara substansial memperbaiki kemampuan generalisasi model SVM. Penelitian ini berimplikasi pada pengembangan sistem klasifikasi berbasis kecerdasan buatan yang lebih akurat dan efisien dalam identifikasi tanaman herbal potensial, serta membuka peluang penerapan pada bidang fitofarmaka dan eksplorasi bahan alam masa depan.

Kata Kunci: Klasifikasi, Afrodisiak, Support Vector Machine, Bayesian Search Optimization



This is an open access article under the [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/)

PENDAHULUAN

Tanaman obat telah menjadi objek penting dalam penelitian multidisiplin, terutama karena kontribusinya terhadap pengembangan obat berbasis alam yang berkelanjutan (Bhagawan et al., 2022). Dalam dua dekade terakhir, pendekatan berbasis teknologi informasi telah diintegrasikan ke dalam eksplorasi sumber daya hayati, termasuk klasifikasi spesies dan identifikasi senyawa bioaktif (Asafo-Agyei et al., 2023). Klasifikasi otomatis berbasis pembelajaran mesin semakin populer, karena mampu mengatasi keterbatasan pengamatan manual yang memerlukan keahlian khusus dan waktu yang tidak efisien (Azadnia et al., 2024). Terutama dalam bidang etnofarmakologi, identifikasi yang akurat terhadap tanaman herbal sangat penting untuk menjamin kemurnian spesies dan efektivitas farmakologisnya.

Support Vector Machine (SVM) merupakan salah satu algoritma klasifikasi yang banyak digunakan dalam bioinformatika dan farmasi komputasional. Algoritma ini efektif untuk dataset berdimensi tinggi dan dikenal memiliki kemampuan generalisasi yang kuat. SVM juga terbukti unggul dalam berbagai aplikasi pengenalan pola seperti klasifikasi daun tanaman, ekstraksi morfologi, hingga diagnosis molekuler (Nhut et al., 2024). Dalam konteks tanaman obat, SVM sering diandalkan untuk mengelompokkan spesies berdasarkan fitur morfologis atau spektral, terutama ketika data jumlahnya terbatas namun kompleksitas tinggi. Namun, keberhasilan model sangat dipengaruhi oleh konfigurasi hiperparameternya, seperti parameter regulasi dan fungsi kernel.

Meskipun SVM telah banyak digunakan, tantangan utama terletak pada proses tuning hiperparameter yang optimal. Metode pencarian hiperparameter tradisional seperti grid search dan random search membutuhkan banyak iterasi dan tidak adaptif terhadap pola performa model yang muncul (Feurer & Hutter, 2019). Hal ini menjadi hambatan besar saat menangani ruang parameter yang luas dan bersifat non-linier. Bayesian Optimization hadir sebagai solusi alternatif yang lebih efisien dan cerdas dalam menjelajah ruang parameter, dengan memanfaatkan model probabilistik sebagai fungsi surrogate untuk memprediksi performa konfigurasi berikutnya (Bergstra & Bengio, 2012). Namun, meskipun telah banyak diaplikasikan pada deep learning dan bioinformatika, pemanfaatan pendekatan ini dalam klasifikasi tanaman herbal, khususnya afrodisiak, masih jarang dilaporkan.

Studi ini mengusulkan integrasi Bayesian Search Optimization ke dalam pipeline Support Vector Machine untuk meningkatkan akurasi klasifikasi tanaman obat afrodisiak. Ruang pencarian dibangun untuk mengeksplorasi parameter regulasi (C) dan skema pembobotan kelas, yang dikombinasikan dengan validasi silang untuk menjamin generalisasi. Proses optimisasi dilakukan menggunakan pendekatan Gaussian Process dengan fungsi akuisisi Expected Improvement. Pemilihan metode ini didasarkan pada efisiensi evaluasi serta kemampuannya dalam menemukan konfigurasi optimal dalam jumlah iterasi yang relatif kecil, sesuai dengan kondisi data terbatas namun kaya fitur yang umum dijumpai pada domain fitofarmaka.

Dengan memanfaatkan integrasi antara metode margin-based seperti SVM dan pendekatan eksplorasi parameter berbasis Bayesian, pendekatan ini diharapkan dapat memberikan kontribusi terhadap sistem klasifikasi tanaman obat yang lebih akurat dan hemat sumber daya. Selain relevan untuk kepentingan akademik, pendekatan ini juga potensial untuk diadaptasi dalam sistem praktis seperti platform mobile untuk identifikasi tanaman secara real-time maupun sistem pendukung pengambilan keputusan di bidang industri herbal dan kesehatan masyarakat berbasis produk alam.

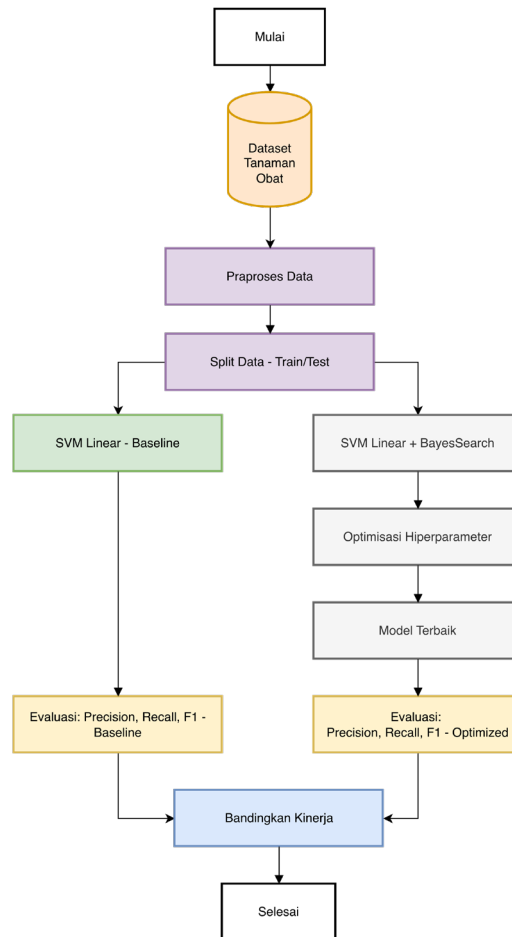
METODE PENELITIAN

Metodologi penelitian pada Gambar 1 dimulai dengan tahap pengumpulan dataset yang berisi fitur-fitur tanaman obat afrodisiak. Dataset seperti pada Tabel 1 dikumpulkan berdasarkan beberapa penelitian terkait (Bhagawan et al., 2022, 2023, 2024). Dataset kemudian melalui proses pra-proses yang meliputi pembersihan data, normalisasi fitur numerik menggunakan standardisasi, serta encoding label target (Tran et al., 2024). Setelah data siap, dilakukan pembagian data ke dalam dua subset, yaitu data pelatihan dan data pengujian, dengan teknik stratified split untuk menjaga proporsi kelas yang seimbang.

Tabel 1. Dataset Afrodisiak

| Family | Species | Nama Lokal | Provinsi | Kabupaten | Claims |
|----------------|-------------------------------|------------|----------------|-------------------|--------------------|
| Acanthaceae | <i>Acorus calamus l.</i> | Jerangau | Aceh | Aceh Singkit | Gangguan Vitalitas |
| Acanthaceae | <i>Coriandrum sativum l.</i> | Ketumbar | Aceh | Aceh Singkit | Organ Pria |
| Acanthaceae | <i>Areca catechu l.</i> | Pinang | Aceh | Aceh Selatan | Gangguan Vitalitas |
| Alliaceae | <i>Ageratum conyzoides l.</i> | Babandotan | Aceh | Kota Subulussalam | Obat Kuat |
| ... | ... | ... | ... | ... | ... |
| Zygophyllaceae | <i>Piper sp.</i> | Sirih | Sumatera Utara | Deli Serdang | Gangguan Vitalitas |

Model awal (baseline) dibangun menggunakan Support Vector Machine (SVM) dengan kernel linear dan parameter default dari pustaka scikit-learn (Shembo et al., 2024). Model baseline ini dilatih menggunakan data pelatihan, kemudian dievaluasi menggunakan data uji dengan tiga metrik utama: precision, recall, dan F1-score.



Gambar 1. Metodologi Penelitian

Sebagai upaya untuk meningkatkan kinerja model, dilakukan optimisasi hiperparameter menggunakan pendekatan Bayesian Search. Proses ini dilakukan pada model SVM yang sama, dengan ruang pencarian untuk parameter regulasi (C) dalam skala log dan opsi pembobotan kelas ($class_weight$) (Kiflie et al., 2024). Proses optimisasi dilakukan menggunakan BayesSearchCV yang mengintegrasikan Gaussian Process sebagai surrogate model dan strategi Expected Improvement untuk fungsi akuisisi.

Model hasil optimisasi kemudian dievaluasi ulang menggunakan data uji yang sama, dan hasilnya dibandingkan dengan model baseline. Perbandingan dilakukan secara kuantitatif berdasarkan precision, recall, dan F1-score. Tahapan ini bertujuan untuk mengetahui sejauh mana optimisasi berbasis Bayesian Search mampu meningkatkan kinerja model klasifikasi pada domain tanaman obat afrodisiak.

HASIL DAN PEMBAHASAN

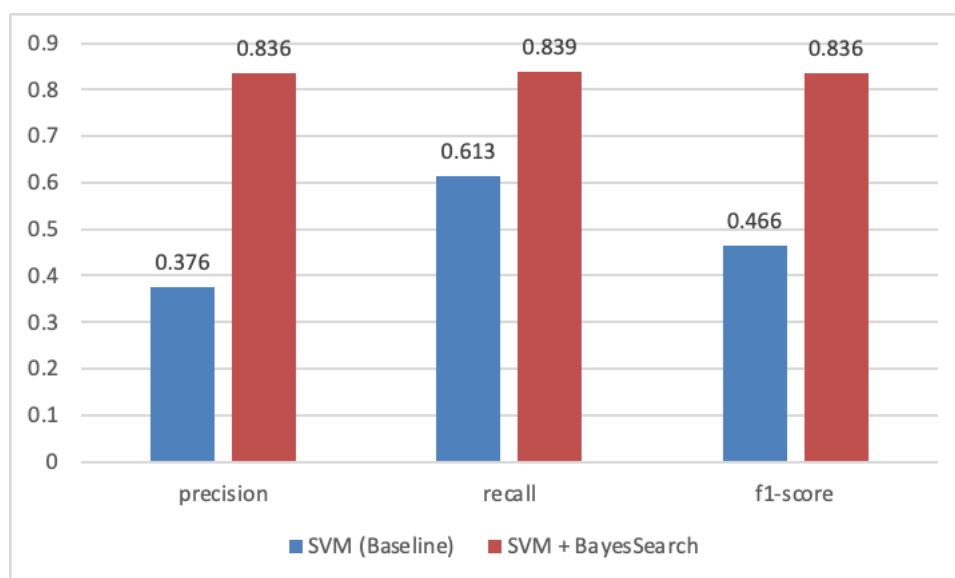
Evaluasi performa model dilakukan menggunakan tiga metrik utama, yaitu precision, recall, dan f1-score, untuk mengukur kualitas klasifikasi dari model SVM linear sebelum dan sesudah dioptimasi menggunakan Bayesian Search. Hasil evaluasi tersebut disajikan pada Tabel 2, yang menunjukkan bahwa seluruh metrik mengalami peningkatan yang sangat signifikan setelah proses optimisasi. Precision model baseline hanya berada pada nilai 0.376, namun meningkat drastis menjadi 0.836 setelah BayesSearch diterapkan, menunjukkan bahwa kemampuan model dalam menghasilkan prediksi positif yang akurat menjadi jauh lebih tinggi. Recall model juga mengalami peningkatan dari 0.613 menjadi 0.839, yang berarti model menjadi lebih sensitif dalam mendeteksi kelas positif secara benar. F1-score sebagai metrik gabungan dari precision dan recall turut melonjak dari 0.466 menjadi 0.836, mencerminkan peningkatan kinerja klasifikasi secara menyeluruh dan seimbang. Nilai persentase

peningkatan masing-masing metrik, yakni precision sebesar 122.3%, recall sebesar 36.9%, dan f1-score sebesar 79.4%, menggarisbawahi dampak besar dari proses optimisasi terhadap efektivitas model.

Tabel 2. Hasil evaluasi

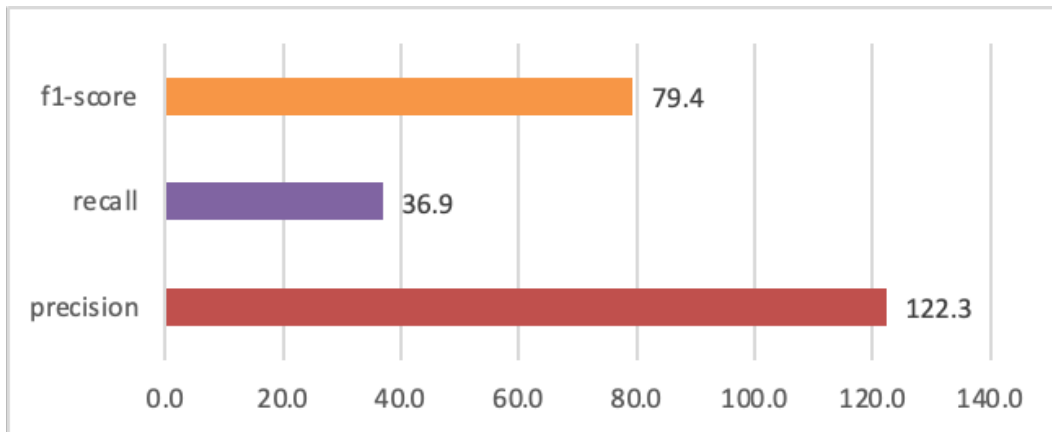
| | SVM (Baseline) | SVM + BayesSearch | Peningkatan % |
|------------------|----------------|-------------------|---------------|
| precision | 0.376 | 0.836 | 122.3 |
| recall | 0.613 | 0.839 | 36.9 |
| f1-score | 0.466 | 0.836 | 79.4 |

Gambar 2 menyajikan visualisasi perbandingan kinerja model antara SVM baseline dan SVM yang telah dioptimasi. Tiga batang berwarna biru mewakili performa awal model, sedangkan tiga batang berwarna merah menggambarkan performa setelah optimisasi. Perbedaan tinggi batang yang signifikan secara visual langsung memperlihatkan keberhasilan metode BayesSearch dalam meningkatkan nilai ketiga metrik. Precision dan f1-score memperlihatkan peningkatan yang paling menonjol, dengan batang merah hampir dua kali lebih tinggi dibandingkan batang biru pada metrik precision.



Gambar 2. Perbandingan Evaluasi

Gambar 3 melengkapi penyajian hasil dengan menampilkan grafik batang horizontal yang menggambarkan peningkatan relatif masing-masing metrik dalam bentuk persentase. Peningkatan precision yang mencapai 122.3% divisualkan dengan batang merah horizontal yang sangat panjang, diikuti oleh peningkatan f1-score sebesar 79.4% yang ditampilkan dalam warna oranye, dan peningkatan recall sebesar 36.9% yang divisualisasikan dengan batang berwarna ungu. Grafik ini menegaskan bahwa precision adalah metrik yang paling terdampak secara positif oleh proses optimisasi, sementara recall juga tetap mengalami perbaikan yang substansial. Secara keseluruhan, kombinasi dari hasil numerik dan visualisasi grafik tersebut memberikan gambaran yang kuat bahwa penerapan Bayesian Search Optimization pada SVM linear secara signifikan meningkatkan kualitas klasifikasi data tanaman herbal afrodisiak dalam seluruh dimensi evaluasi performa.



Gambar 3. Persentase peningkatan

Peningkatan performa yang signifikan pada precision, recall, dan f1-score menunjukkan bahwa optimisasi hiperparameter dengan Bayesian Search mampu menemukan konfigurasi terbaik secara efisien dibandingkan pendekatan default. Hal ini mendukung temuan sebelumnya bahwa metode Bayesian mampu mengeksplorasi ruang pencarian secara adaptif dan hemat evaluasi. Dalam konteks klasifikasi tanaman obat afrodisiak, peningkatan precision sangat penting untuk meminimalkan prediksi salah terhadap spesies non-target, sedangkan peningkatan recall memastikan keberhasilan identifikasi spesies yang relevan. Keseimbangan yang dicapai melalui f1-score yang tinggi mencerminkan stabilitas model terhadap distribusi data yang mungkin tidak seimbang, menjadikan pendekatan ini ideal untuk aplikasi bioinformatika berbasis data terbatas dan fitur kompleks.

SIMPULAN DAN SARAN

Penelitian ini membahas penerapan Bayesian Search Optimization pada model Support Vector Machine (SVM) linear untuk meningkatkan performa klasifikasi tanaman herbal afrodisiak. Dengan pendekatan optimisasi berbasis probabilistik, model yang dihasilkan menunjukkan peningkatan signifikan pada ketiga metrik evaluasi utama, yaitu precision, recall, dan f1-score. Peningkatan precision sebesar 122.3%, recall sebesar 36.9%, dan f1-score sebesar 79.4% menandakan keberhasilan BayesSearch dalam menemukan konfigurasi hiperparameter optimal. Temuan ini memperkuat relevansi SVM yang dioptimasi dalam konteks klasifikasi bioinformatika, khususnya pada domain dengan data yang terbatas dan variabel input yang kompleks.

Untuk pengembangan selanjutnya, penelitian ini dapat diperluas dengan melibatkan dataset yang lebih besar dan beragam, termasuk berbagai bagian tanaman serta kondisi lingkungan yang berbeda. Selain itu, evaluasi model dapat ditingkatkan melalui pengujian dengan metrik lain seperti akurasi keseluruhan, AUC-ROC, dan penggunaan confusion matrix. Kombinasi pendekatan optimisasi lain seperti Optuna atau Hyperband juga dapat diuji untuk membandingkan efisiensi dan efektivitas proses tuning. Penerapan metode ini pada data lapangan secara real-time dan integrasi dengan sistem berbasis Internet of Things (IoT) atau mobile image recognition juga berpotensi menjadi arah studi lanjutan yang aplikatif.

DAFTAR PUSTAKA

Asafo-Agyei, T., Appau, Y., Barimah, K. B., & Asase, A. (2023). Medicinal plants used for management of diabetes and hypertension in Ghana. *Heliyon*, 9(12), e22977. <https://doi.org/https://doi.org/10.1016/j.heliyon.2023.e22977>

- Azadnia, R., Noei-Khodabadi, F., Moloudzadeh, A., Jahanbakhshi, A., & Omid, M. (2024). Medicinal and poisonous plants classification from visual characteristics of leaves using computer vision and deep neural networks. *Ecological Informatics*, 82, 102683. <https://doi.org/https://doi.org/10.1016/j.ecoinf.2024.102683>
- Bergstra, J., & Bengio, Y. (2012). Random Search for Hyper-Parameter Optimization. *Journal of Machine Learning Research*, 13(10), 281–305. <http://jmlr.org/papers/v13/bergstra12a.html>
- Bhagawan, W. S., Ekasari, W., & Agil, M. (2023). Ethnopharmacology of medicinal plants used by the Tenggerese community in Bromo Tengger Semeru National Park, Indonesia. *Biodiversitas Journal of Biological Diversity*, 24(10). <https://doi.org/10.13057/biodiv/d241028>
- Bhagawan, W. S., Ekasari, W., & Agil, M. (2024). Ethnobotanical survey of herbal steam baths among the Tenggerese community in Bromo Tengger Semeru National Park, Indonesia. *IOP Conference Series: Earth and Environmental Science*, 1352(1), 012103. <https://doi.org/10.1088/1755-1315/1352/1/012103>
- Bhagawan, W. S., Suproborini, A., Putri, D. L. P., Nurfatma, A., & Putra, R. T. (2022). Ethnomedicinal study, phytochemical characterization, and pharmacological confirmation of selected medicinal plant on the northern slope of Mount Wilis, East Java, Indonesia. *Biodiversitas Journal of Biological Diversity*, 23(8). <https://doi.org/10.13057/biodiv/d230855>
- Feurer, M., & Hutter, F. (2019). *Hyperparameter Optimization* (pp. 3–33). https://doi.org/10.1007/978-3-030-05318-5_1
- Kiflie, M. A., Sharma, D. P., & Haile, M. A. (2024). Deep learning for Ethiopian indigenous medicinal plant species identification and classification. *Journal of Ayurveda and Integrative Medicine*, 15(6), 100987. <https://doi.org/https://doi.org/10.1016/j.jaim.2024.100987>
- Nhut, D. T. N., Tan, T. D., Quoc, T. N., & Hoang, V. T. (2024). Medicinal plant recognition based on Vision Transformer and BEiT. *Procedia Computer Science*, 234, 188–195. <https://doi.org/https://doi.org/10.1016/j.procs.2024.02.165>
- Shembo, A. K., Ayichew, S. S., Stiers, I., Geremew, A., & Carson, L. (2024). Classification and ordination analysis of wild medicinal plants in Ada'a district, Ethiopia: Implication for sustainable conservation and utilization. *Ecological Frontiers*, 44(4), 809–819. <https://doi.org/https://doi.org/10.1016/j.ecofro.2024.04.002>
- Tran, T. P., Ud Din, F., Brankovic, L., Sanin, C., Hester, S. M., & Duc Hoang Le, M. (2024). Incremental and Zero-Shot Machine Learning for Vietnamese Medicinal Plant Image Classification. *Procedia Computer Science*, 246, 606–615. <https://doi.org/https://doi.org/10.1016/j.procs.2024.09.469>