

IMPLEMENTASI METODE KLASIFIKASI UNTUK IDENTIFIKASI JENIS KACANG KERING BERDASARKAN FITUR MORFOLOGI

Reyhan ¹⁾ Tri Sutrisno ²⁾

¹⁾ Teknik Informatika Universitas Tarumanagara
Jl. Letjen S.Parman No.1 Jakarta Barat 11440 Indonesia
email : reyhan.535210064@stu.untar.ac.id

²⁾ Sistem Informasi Universitas Tarumanagara
Jl. Letjen S.Parman No.1 Jakarta Barat 11440 Indonesia
email : tris@fti.untar.ac.id

ABSTRACT

Identification of dry bean species is a critical challenge in the agricultural industry, especially to ensure the quality and optimal utilization of the product. This study focuses on the application of a morphological feature-based classification method to automatically recognize dry bean species. Features such as size, shape, surface texture, and color are used as the basis for classification. Several classification algorithms, including K-Nearest Neighbors (KNN), Artificial Neural Network (ANN), Decision Tree, were tested to find the most effective algorithm in identifying the types of beans. The research dataset includes various types of dry beans with different morphological variations. The experimental results show that the morphological feature-based classification approach is able to achieve a high level of accuracy, with Artificial Neural Network (ANN) being the algorithm that shows the best performance. The implementation of this method is expected to provide practical solutions in managing the quality and processing of dry beans in the agricultural sector.

Key words

artificial neural networks, classification, dry bean, decision tree, k-nearest neighbors, morphological features

1. Pendahuluan

Identifikasi jenis kacang kering berdasarkan fitur morfologi telah menjadi tantangan penting dalam sektor pertanian, khususnya dalam menjaga kualitas produk dan memastikan efisiensi dalam proses distribusi. Penggunaan metode manual berbasis pengamatan visual tidak selalu efisien dan seringkali rawan kesalahan. Oleh karena itu, implementasi teknologi berbasis *machine learning* untuk klasifikasi kacang kering semakin menarik perhatian, mengingat kemampuannya untuk meningkatkan akurasi dan efisiensi. Alasan penggunaan

dataset kacang kering dilatar belakangi oleh keanekaragaman jenis kacang kering, pentingnya untuk mengetahui kualitas produk, efisiensi proses distribusi, mengurangi kesalahan manusia dalam klasifikasi secara manual, mengembangkan inovasi dalam industri pertanian dan menerapkannya dengan penggunaan teknologi, serta memberikan informasi yang berguna bagi petani dan produsen dalam memilih varietas kacang kering yang tepat untuk ditanam.

Penggunaan algoritma machine learning dalam klasifikasi kacang kering berdasarkan fitur morfologi terbukti efektif dalam berbagai penelitian. Klasifikasi berdasarkan fitur seperti panjang, lebar, bentuk, dan tekstur kacang dapat memberikan hasil yang lebih akurat dibandingkan metode manual. Algoritma K-Nearest Neighbors (KNN) menggunakan kedekatan sampel berdasarkan jarak dalam ruang fitur untuk memprediksi kelas kacang. Dalam beberapa penelitian, akurasi KNN berkisar di angka 87-90% untuk klasifikasi kacang kering. Kelebihan dari KNN adalah kesederhanaan implementasinya dan keandalannya untuk data yang memiliki fitur yang jelas [1] [2].

Algoritma Artificial Neural Networks (ANN), dengan kemampuan pembelajaran mendalam, menunjukkan performa yang lebih tinggi dalam mengenali pola morfologi kompleks pada kacang. ANN belajar melalui sejumlah lapisan tersembunyi yang menyaring fitur-fitur penting, sehingga cocok untuk masalah klasifikasi yang lebih rumit. Berdasarkan studi oleh Koklu dan Ozkan (2020), ANN mampu mencapai akurasi sekitar 91,73% dalam klasifikasi multiclass kacang kering [3] [4].

Selain itu, algoritma Decision Tree memiliki keunggulan dalam hal interpretasi yang mudah, dengan model yang membagi data berdasarkan atribut morfologi yang paling relevan. Decision Tree menunjukkan performa yang baik dalam penelitian dengan tingkat

akurasi mencapai 92,52%, menjadikannya algoritma yang efisien untuk aplikasi pertanian dan pangan[5] [4].

Algoritma lain seperti Support Vector Machine (SVM) juga digunakan dalam beberapa penelitian sebagai perbandingan dengan KNN, ANN, dan Decision Tree. SVM sering kali menghasilkan akurasi yang lebih tinggi, terutama untuk dataset yang memiliki margin pemisahan yang jelas antara kelas. Studi-studi terdahulu juga menunjukkan bahwa kombinasi algoritma ini dapat menghasilkan performa yang optimal dalam klasifikasi kacang kering[3] [2].

2. Metode Penelitian

2.1 Dataset

Dataset yang digunakan dalam penelitian ini merupakan data morfologi kacang kering yang berisi berbagai fitur geometris atau morfologis dari kacang. Setiap baris dalam dataset merepresentasikan satu sampel kacang, dengan kolom-kolom yang merepresentasikan fitur-fitur kacang tersebut, serta label kelas yang menunjukkan jenis kacang. Kacang kering merupakan salah satu produk pertanian yang penting dalam industri makanan, dan pengenalan jenis kacang kering menjadi penting untuk keperluan otomatisasi industri dan kualitas kontrol. Dataset yang digunakan mencakup berbagai fitur morfologi kacang kering, seperti area, perimeter, aspect ratio, convex area, dan fitur lainnya. Fitur-fitur ini memberikan informasi yang kaya mengenai ukuran dan bentuk kacang, yang menjadi dasar klasifikasi[6] [7]. Beberapa jenis kacang kering yang ada pada dataset yaitu diantaranya adalah Barbunya, Bombay, Cali, Dermason, Horoz, Seker, Sira.

Klasifikasi jenis kacang kering tidak hanya memiliki implikasi pada industri makanan, tetapi juga berperan penting dalam penelitian pertanian dan pemulihan varietas kacang yang lebih baik. Dalam konteks ini, dataset yang digunakan mencakup berbagai fitur morfologi kacang kering, seperti:

1. Area : Mengukur luas permukaan kacang. Area yang lebih besar dapat menunjukkan kacang yang lebih matang atau sehat[7].
2. Perimeter : Mengukur panjang tepi luar kacang, yang dapat memberikan informasi mengenai bentuk kacang[6].
3. Aspect Rasio : Rasio antara lebar dan tinggi kacang, yang membantu dalam membedakan jenis kacang berdasarkan bentuknya[7].
4. Convex Area : Luas dari bentuk cembung terkecil yang dapat membungkus kacang, memberikan informasi tambahan tentang morfologi kacang[7].

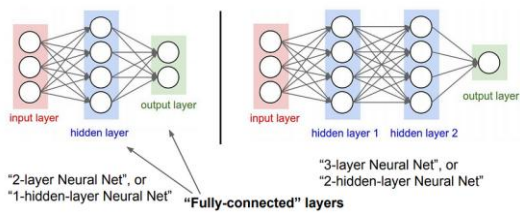
Fitur-fitur ini memberikan informasi yang kaya mengenai ukuran dan bentuk kacang, yang menjadi dasar klasifikasi. Pemilihan fitur morfologi yang tepat sangat penting, karena dapat meningkatkan akurasi model klasifikasi yang digunakan. Penggunaan algoritma klasifikasi seperti K-Nearest Neighbors (KNN), Artificial Neural Networks (ANN), dan Decision Tree dalam penelitian ini memungkinkan untuk menganalisis dataset dengan lebih mendalam. Algoritma-algoritma ini memiliki kemampuan untuk menangkap hubungan kompleks antara fitur-fitur morfologi dan label kelas, sehingga memungkinkan untuk mengidentifikasi jenis kacang dengan tingkat akurasi yang tinggi[6] [7].

Pentingnya dataset ini juga terletak pada aplikasinya dalam konteks industri. Dengan meningkatkan akurasi dalam klasifikasi jenis kacang kering, industri dapat lebih efisien dalam proses seleksi dan pemrosesan, serta memastikan kualitas produk akhir. Hal ini tidak hanya meningkatkan produktivitas, tetapi juga mengurangi limbah dan meningkatkan kepuasan konsumen[6]. Dengan demikian, penelitian ini tidak hanya berkontribusi pada pemahaman morfologi kacang kering, tetapi juga memberikan solusi praktis untuk tantangan dalam industri makanan dan pertanian[7] [8].

2.2 Artificial Neural Networks (ANN)

Artificial Neural Network (ANN) adalah salah satu jenis algoritma pembelajaran mesin yang terinspirasi oleh jaringan saraf biologis pada otak manusia. ANN digunakan untuk memodelkan hubungan yang kompleks dan non-linear antara input dan output dengan cara "belajar" dari data. ANN terdiri dari sejumlah besar unit sederhana yang dikenal sebagai neuron (atau node) yang saling terhubung. Beberapa keuntungan algoritma ini adalah kemampuan menangani data non-linear, adaptasi terhadap berbagai jenis data, dan kemampuan generalisasi. Kemudian untuk beberapa kekurangan algoritma ini adalah memerlukan data yang besar, waktu pelatihan yang lama, dan cenderung terkena masalah overfitting[9] [10] [11].

ANN bekerja melalui proses pembelajaran yang terbagi menjadi tiga lapisan utama: lapisan input, lapisan tersembunyi, dan lapisan output. Pada lapisan input, data mentah dimasukkan ke dalam model. Kemudian, data tersebut melewati lapisan tersembunyi, di mana sejumlah fungsi aktivasi diterapkan untuk memodelkan pola-pola yang kompleks dalam data. Fungsi aktivasi seperti sigmoid, ReLU (Rectified Linear Unit), dan tanh memungkinkan jaringan untuk belajar pola non-linear. Terakhir, lapisan output menghasilkan hasil prediksi dari model[9] [10] [11].

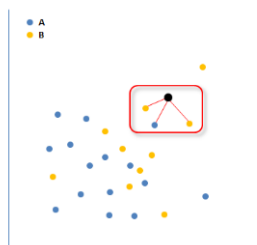


Gambar 1 Diagram Algoritma ANN (Sumber:

[https://yunusmuhammad007.medium.com/6-artificial-neural-network-ann-part-1-pengenalan-db487b8f8d85#:~:text=Gambar%201.%20Representasi%20Neural%20Network%20E2%80%94%20cs231n%20stanford%20\(Lecture%204%20E2%80%9392\)](https://yunusmuhammad007.medium.com/6-artificial-neural-network-ann-part-1-pengenalan-db487b8f8d85#:~:text=Gambar%201.%20Representasi%20Neural%20Network%20E2%80%94%20cs231n%20stanford%20(Lecture%204%20E2%80%9392))

2.3 K-Nearest Neighbors (KNN)

K-Nearest Neighbors adalah algoritma pembelajaran mesin yang digunakan untuk klasifikasi dan regresi. KNN adalah algoritma berbasis instance, di mana objek yang ingin diklasifikasikan akan diberikan label berdasarkan label mayoritas dari KKK tetangga terdekatnya dalam ruang fitur. Algoritma ini sederhana dan intuitif, serta tidak memerlukan pelatihan eksplisit sebelum digunakan untuk klasifikasi. Keuntungan algoritma ini adalah sederhana dan mudah dipahami, tidak memerlukan pelatihan, dan kinerja yang baik. Kemudian untuk beberapa kekurangan algoritma ini adalah memerlukan komputasi tinggi, sensitif terhadap noise data, dan hasil bergantung pada pemilihan K [12] [13] [14].



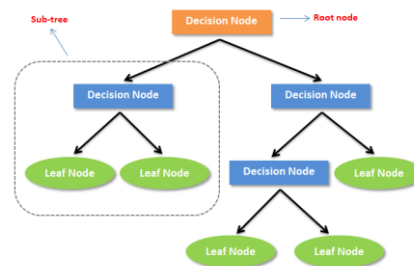
Gambar 2 Diagram Algoritma KNN (Sumber:

<https://ilmudatapy.com/algoritma-k-nearest-neighbor-knn-untuk-klasifikasi/#:~:text=Algoritma%20KNN%20atau%20K-Nearest%20Neighbor%20adalah%20salah%20satu%20algoritma%20yang>

2.4 Decision Tree

Decision Tree adalah algoritma pembelajaran mesin yang digunakan untuk klasifikasi dan regresi. Ini adalah model yang berbentuk pohon yang membagi dataset menjadi subset berdasarkan nilai dari fitur input. Setiap cabang dalam pohon merepresentasikan keputusan atau kondisi, sedangkan setiap daun merepresentasikan hasil atau kelas yang diprediksi. Keuntungan algoritma ini adalah mudah dipahami dan diinterpretasikan, tidak memerlukan Pra-pemrosesan yang rumit, dan fleksibel untuk digunakan. Kemudian beberapa kekurangan

algoritma ini adalah cenderung overfitting pada pelatihan data, sensitif terhadap data noise, dan kesulitan dalam menghitung hubungan non-linear [15] [16] [17].



Gambar 3 Diagram Algoritma Decision Tree (Sumber:

<https://ilmudatapy.com/penjelasan-algoritma-decision-tree/#:~:text=Kepopuleran%20algoritma%20decision%20tree%20dalam%20membangun%20model%20machine%20learning%20adalah>

2.5 Confusion Matrix

Confusion matrix adalah alat yang digunakan dalam analisis klasifikasi untuk mengevaluasi kinerja model prediktif. Ini memberikan gambaran tentang bagaimana model berperilaku dalam mengklasifikasikan data dengan membandingkan prediksi model terhadap label sebenarnya. Manfaat confusion matrix adalah visualisasi kinerja model, mengidentifikasi area untuk perbaikan, serta menghitung berbagai metrik kinerja yang membantu dalam memilih model terbaik. Beberapa komponen yang terdapat dalam confusion matrix yaitu True Positif (TP), True Negatif (TN), False Positif (FP), dan False Negatif (FN) [18] [19] [20].

2.6 Rumus Matematika

Berikut ini adalah beberapa rumus dari algoritma ANN, KNN, dan Decision Tree :

1. Artificial Neural Network (ANN) :

$$y = f\left(\sum_{i=1}^n w_i x_i + b\right) \tag{1}$$

Keterangan :

y = output dari neuron

f = fungsi aktivasi

w_i = bobot untuk input ke- i

x_i = nilai input ke- i

b = bias dari neuron

n = jumlah total fitur input masuk ke neuron

2. K-Nearest Neighbors (KNN) :

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \tag{2}$$

Keterangan :

d = jarak antara dua titik dalam ruang n -dimensi
 x dan y = titik dalam ruang n dimensi
 n = jumlah dimensi atau fitur dalam ruang x_i dan y_i = nilai dari komponen ke- i dan titik x dan y masing-masing

- 3. Decision Tree :
 - a. Gini Impurity :

$$Gini = 1 - \sum (p_i^2) \tag{3}$$

Keterangan :
 P_i = proporsi kelas i
 b. Entropy :

$$Entropy = - \sum (p_i \log_2(p_i)) \tag{4}$$

Keterangan :
 P_i = proporsi kelas i

- 4. Confusion Matrix :
 - a. Akurasi :

$$Akurasi = \frac{TP + TN}{TP + TN + FP + FN} * 100\% \tag{5}$$

- b. Presisi :

$$Presisi = \frac{TP}{TP + FP} * 100\% \tag{6}$$

- c. Recall :

$$Recall = \frac{TP}{TP + FN} * 100\% \tag{7}$$

3. Hasil Percobaan

Dalam pengujian data kacang kering ini, peneliti menggunakan 3 algoritma seperti yang sudah dijabarkan sebelumnya seperti : Artificial Neural Network (ANN), K-Nearest Neighbors (KNN), dan Decision Tree. Pada pengujian algoritma ini, data dibagi menjadi 2 yaitu data latih dan data uji. Komposisi data latih dan data uji menggunakan 70% : 30% dan 80% : 20%. Hasil dari pengujian ini menunjukkan bahwa algoritma ANN memiliki hasil terbaik dikarenakan pada komposisi 70% : 30% memiliki nilai akurasi sebesar 93% dan pada komposisi 80% : 20% memiliki nilai akurasi sebesar 93%, hal ini menunjukkan bahwa algoritma ANN mampu menangani pola morfologi kacang kering dengan baik, berkat kemampuannya dalam menangkap hubungan non-linear yang kompleks. Nilai akurasi precision dan recall pada algoritma ANN juga sangat baik yaitu bernilai 0.94 pada kedua rasio data, yang

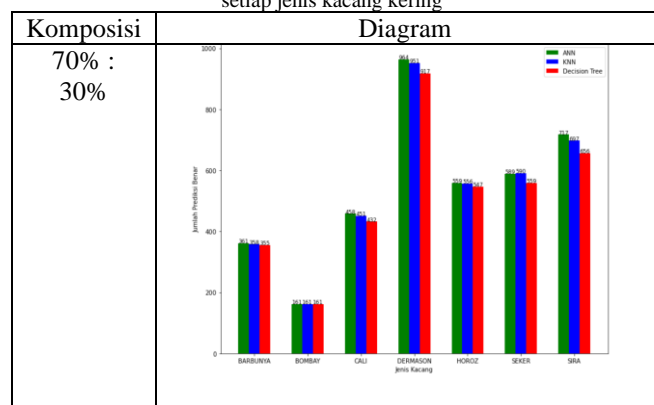
menunjukkan bahwa model ini sangat akurat dalam mengidentifikasi kelas yang benar serta jarang melakukan kesalahan dalam memprediksi kacang kering.

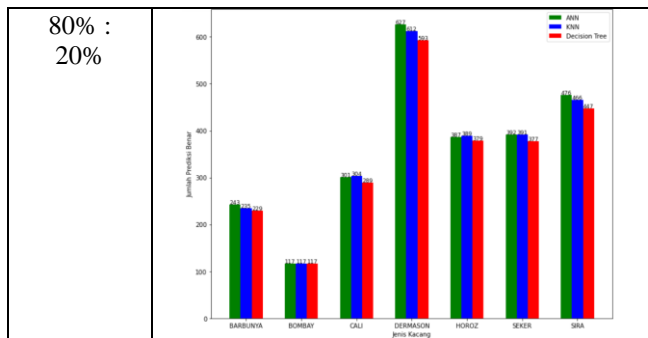
Kemudian hasil pengujian pada algoritma K-Nearest Neighbors pada komposisi 70% : 30% memiliki nilai akurasi sebesar 92% dan pada komposisi 80% : 20% memiliki nilai akurasi sebesar 92% juga. Hal ini menunjukkan bahwa algoritma KNN masih dianggap sangat kompetitif dibandingkan dengan algoritma yang lain dan andal dalam mengklasifikasikan kacang kering berdasarkan fitur morfologi. Kelebihan algoritma KNN adalah sederhana dan efektif akan tetapi sensitif terhadap pemilihan parameter K dan mungkin membutuhkan lebih banyak perhitungan untuk data yang besar. Nilai akurasi precision dan recall juga berada di angka 0.93 – 0.94 yang menunjukkan bahwa algoritma ini relatif baik dalam mengklasifikasikan sampel secara konsisten.

Setelah itu, hasil pengujian pada algoritma Decision Tree pada komposisi 70% : 30% dan 80% : 20% memiliki nilai akurasi sebesar 89%. Dibandingkan dengan kedua algoritma ANN dan KNN, nilai akurasi pada algoritma Decision Tree memiliki nilai yang lebih rendah. Kelebihan algoritma Decision Tree adalah interpretasinya yang mudah, akan tetapi sering kali cenderung overfitting pada data pelatihan, yang mungkin menjelaskan akurasi yang sedikit lebih rendah. Nilai akurasi precision dan recall pada algoritma Decision Tree menunjukkan hasil berada di angka 0.90 – 0.91 yang artinya algoritma ini masih mampu menangkap pola yang relevan dari data kacang kering, namun performanya tidak seoptimal algoritma ANN dan KNN.

Berikut ini adalah diagram batang perbandingan setiap algoritma dalam membaca tiap jenis kacang kering :

Tabel 1 Tabel Hasil Perbandingan Tiap Algoritma Dalam Membaca setiap jenis kacang kering





Pada komposisi 70% : 30%, algoritma ANN mampu membaca 361 data kacang kering jenis Barbunya, 161 data kacang kering jenis Bombay, 458 data kacang kering jenis Cali, 964 data kacang kering jenis Dermason, 559 data kacang kering jenis Horoz, 589 data kacang kering jenis Seker, dan 717 data kacang kering jenis Sira. Sedangkan algoritma KNN mampu membaca 358 data kacang kering jenis Barbunya, 161 data kacang kering jenis Bombay, 451 data kacang kering jenis Cali, 951 data kacang kering jenis Dermason, 556 data kacang kering jenis Horoz, 590 data kacang kering jenis Seker, dan 697 data kacang kering jenis Sira. Terakhir algoritma Decision Tree mampu membaca 355 data kacang kering jenis Barbunya, 161 data kacang kering jenis Bombay, 432 data kacang kering jenis Cali, 917 data kacang kering jenis Dermason, 547 data kacang kering jenis Horoz, 559 data kacang kering jenis Seker, dan 656 data kacang kering jenis Sira.

Pada komposisi 80% : 20%, algoritma ANN mampu membaca 243 data kacang kering jenis Barbunya, 117 data kacang kering jenis Bombay, 301 data kacang kering jenis Cali, 627 data kacang kering jenis Dermason, 387 data kacang kering jenis Horoz, 392 data kacang kering jenis Seker, dan 476 data kacang kering jenis Sira. Sedangkan algoritma KNN mampu membaca 235 data kacang kering jenis Barbunya, 117 data kacang kering jenis Bombay, 304 data kacang kering jenis Cali, 617 data kacang kering jenis Dermason, 389 data kacang kering jenis Horoz, 391 data kacang kering jenis Seker, 466 data kacang kering jenis Sira. Terakhir algoritma Decision Tree mampu membaca 229 data kacang kering jenis Barbunya, 117 data kacang kering jenis Bombay, 289 data kacang kering jenis Cali, 593 data kacang kering jenis Dermason, 379 data kacang kering jenis Horoz, 377 data kacang kering jenis Seker, dan 447 data kacang kering jenis Sira.

Berikut ini adalah tabel hasil perbandingan setiap algoritma dalam melakukan klasifikasi pada dataset kacang kering :

Tabel 1 Tabel Hasil Perbandingan Tiap Algoritma Dalam Melakukan Klasifikasi Kacang Kering

Data Latih : Data Uji	Algoritma	Precision	Recall	F1-Score	Akurasi
	ANN	0.94	0.94	0.94	0.93
	KNN	0.94	0.93	0.93	0.92
	Decision Tree	0.90	0.90	0.90	0.89

70% : 30%	Artificial Neural Network (ANN)	0.94	0.94	0.94	0.93
	K-Nearest Neighbors (KNN)	0.94	0.93	0.93	0.92
	Decision Tree	0.90	0.90	0.90	0.89
80% : 20%	Artificial Neural Network (ANN)	0.95	0.94	0.94	0.93
	K-Nearest Neighbors (KNN)	0.94	0.93	0.94	0.92
	Decision Tree	0.91	0.91	0.91	0.89

4. Kesimpulan

Penelitian ini membuktikan bahwa algoritma machine learning, khususnya algoritma Artificial Neural Network (ANN), K-Nearest Neighbors (KNN), dan Decision Tree mampu mengklasifikasikan jenis kacang kering dengan akurasi yang tinggi berdasarkan fitur morfologi. Algoritma ANN menunjukkan performa terbaik dengan akurasi 93%, diikuti KNN dengan akurasi 92%, sementara Decision Tree memiliki akurasi 89%. Hasil ini menunjukkan bahwa algoritma berbasis machine learning dapat menjadi solusi yang efektif dan efisien dalam klasifikasi kacang kering, membantu meningkatkan kualitas dan produktivitas di sektor pertanian dan industri pangan.

REFERENSI

- [1] G. Slowiński, "Dry Beans Classification Using Machine Learning," 2021. [Online]. Available: <http://ceur-ws.org>
- [2] M. Salauddin Khan *et al.*, "Comparison of multiclass classification techniques using dry bean dataset," *International Journal of Cognitive Computing in Engineering*, vol. 4, pp. 6–20, Jun. 2023, doi: 10.1016/j.ijcce.2023.01.002.
- [3] H. Vaidya, K. Swamy Karekal, K. V Prasad, and K. Kumar Swamy, "Multiclass Classification of Dry Beans using Artificial Neural Network," 2023. [Online]. Available: <https://www.researchgate.net/publication/373523354>
- [4] A. Mehta, P. Sengupta, D. Garg, H. Singh, and Y. S. Diamand, "Benchmarking the Effectiveness of Classification Algorithms and SVM Kernels for Dry Beans," Jul. 2023, [Online]. Available: <http://arxiv.org/abs/2307.07863>
- [5] E. Dönmez, S. Kılıçarslan, and A. Diker, "Classification of hazelnut varieties based on bigtransfer deep learning model," *European Food Research and Technology*, vol.

- 250, no. 5, pp. 1433–1442, May 2024, doi: 10.1007/s00217-024-04468-1.
- [6] M. Koklu and I. A. Ozkan, “Multiclass classification of dry beans using computer vision and machine learning techniques,” *Comput Electron Agric*, vol. 174, Jul. 2020, doi: 10.1016/j.compag.2020.105507.
- [7] S. Krishnan, S. K. Aruna, K. Kanagarathinam, and E. Venugopal, “Identification of Dry Bean Varieties Based on Multiple Attributes Using CatBoost Machine Learning Algorithm,” *Sci Program*, vol. 2023, 2023, doi: 10.1155/2023/2556066.
- [8] I. Wardhana, Musi Ariawijaya, Vandri Ahmad Isnaini, and Rahmi Putri Wirman, “Gradient Boosting Machine, Random Forest dan Light GBM untuk Klasifikasi Kacang Kering,” *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 6, no. 1, 2022, doi: 10.29207/resti.v6i1.3682.
- [9] H. Hananti and K. Sari, “Perbandingan Metode Support Vector Machine (SVM) dan Artificial Neural Network (ANN) pada Klasifikasi Gizi Balita,” *Seminar Nasional Official Statistics*, vol. 2021, no. 1, 2021, doi: 10.34123/semnasoffstat.v2021i1.1014.
- [10] E. Matel, F. Vahdatikhaki, S. Hosseinyalamdary, T. Evers, and H. Voordijk, “An artificial neural network approach for cost estimation of engineering services,” *International Journal of Construction Management*, vol. 22, no. 7, 2022, doi: 10.1080/15623599.2019.1692400.
- [11] K. Sinha, Z. Uddin, H. I. Kawsar, S. Islam, M. J. Deen, and M. M. R. Howlader, “Analyzing chronic disease biomarkers using electrochemical sensors and artificial neural networks,” 2023. doi: 10.1016/j.trac.2022.116861.
- [12] A. Almomany, W. R. Ayyad, and A. Jarrah, “Optimized implementation of an improved KNN classification algorithm using Intel FPGA platform: Covid-19 case study,” *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 6, 2022, doi: 10.1016/j.jksuci.2022.04.006.
- [13] M. Suyal and P. Goyal, “A Review on Analysis of K-Nearest Neighbor Classification Machine Learning Algorithms based on Supervised Learning,” 2022. doi: 10.14445/22315381/IJETT-V70I7P205.
- [14] S. R. Cholil, T. Handayani, R. Prathivi, and T. Ardianita, “Implementasi Algoritma Klasifikasi K-Nearest Neighbor (KNN) Untuk Klasifikasi Seleksi Penerima Beasiswa,” *IJCIT (Indonesian Journal on Computer and Information Technology)*, vol. 6, no. 2, 2021, doi: 10.31294/ijcit.v6i2.10438.
- [15] T. A. Q. Putri, A. Triayudi, and R. T. Aldisa, “Implementasi Algoritma Decision Tree dan Naïve Bayes Untuk Klasifikasi Sentimen Terhadap Kepuasan Pelanggan Starbucks,” *Journal of Information System Research (JOSH)*, vol. 4, no. 2, pp. 641–649, Jan. 2023, doi: 10.47065/josh.v4i2.2949.
- [16] M. B. Abdul Majid, Y. M. Cani, and U. Enri, “Penerapan Algoritma K-Means dan Decision Tree Dalam Analisis Prestasi Siswa Sekolah Menengah Kejuruan,” *Jurnal Sistem Komputer dan Informatika (JSON)*, vol. 4, no. 2, p. 355, Dec. 2022, doi: 10.30865/json.v4i2.5299.
- [17] F. Y. Pamuji and V. P. Ramadhan, “Komparasi Algoritma Random Forest dan Decision Tree untuk Memprediksi Keberhasilan Immunotherapy,” *Jurnal Teknologi dan Manajemen Informatika*, vol. 7, no. 1, 2021, doi: 10.26905/jtmi.v7i1.5982.
- [18] B. P. Pratiwi, A. S. Handayani, and S. Sarjana, “PENGUKURAN KINERJA SISTEM KUALITAS UDARA DENGAN TEKNOLOGI WSN MENGGUNAKAN CONFUSION MATRIX,” *Jurnal Informatika Upgris*, vol. 6, no. 2, 2021, doi: 10.26877/jiu.v6i2.6552.
- [19] M. Hasnain, M. F. Pasha, I. Ghani, M. Imran, M. Y. Alzahrani, and R. Budiarto, “Evaluating Trust Prediction and Confusion Matrix Measures for Web Services Ranking,” *IEEE Access*, vol. 8, 2020, doi: 10.1109/ACCESS.2020.2994222.
- [20] G. Zeng, “On the confusion matrix in credit scoring and its analytical properties,” *Commun Stat Theory Methods*, vol. 49, no. 9, 2020, doi: 10.1080/03610926.2019.1568485.