

Coffee Aroma Classification Based on an Electronic Nose: Comparative Evaluation of DNN, SVM, and ANN with IQR and Wavelet Preprocessing

Muhammad Yulvi Aditya Pradana¹, Ifnu Wisma Dwi Prasetya², Ita Aristia Sa'ida³

^{1,2,3} Teknik Informatika, Universitas Nahdlatul Ulama Sunan Giri

¹zulviadityapradana@gmail.com, ²ifnudwi867@gmail.com, ³itaaristia@unugiri.ac.id

Article Info

Article history:

Received 2026-01-08

Revised 2026-03-10

Accepted 2026-04-10

Keyword:

Electronic Nose, coffee aroma classification, Deep Neural Network, Support Vector Machine, Artificial Neural Network, Discrete Wavelet Transform

ABSTRACT

Coffee aroma assessment is commonly performed through cupping tests, which are subjective and highly dependent on human perception. This study proposes an objective approach for multi-class coffee aroma classification using a low-cost Electronic Nose (E-Nose) and a comparative evaluation of Deep Neural Network (DNN), Support Vector Machine (SVM), and Artificial Neural Network (ANN). Aroma data were acquired from an ESP32-based sensor array consisting of MQ-9, MQ-135, and DHT22, resulting in 1,503 samples from 10 Arabica and Robusta coffee classes. Two preprocessing scenarios were evaluated: (1) raw data + Interquartile Range (IQR) + Discrete Wavelet Transform (DWT) + Min–Max scaling, and (2) raw data + Min–Max scaling. The models were evaluated using an 80:20 train–test split and 5-fold cross-validation on the training set. The results show that DNN achieved the best performance in both scenarios, reaching 91.45% accuracy in the first scenario and 80.07% in the second. SVM also improved substantially from 61.46% to 81.78% after applying IQR and DWT. In contrast, ANN performed better in the second scenario, achieving 73.72% accuracy compared to 61.43% in the first. These findings indicate that preprocessing effectiveness depends on the classification model, while DNN remains the most robust and consistent model for multi-class coffee aroma classification based on E-Nose tabular data.



This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

I. PENDAHULUAN

Kopi merupakan komoditas penting di Indonesia yang berperan dalam penguatan ekonomi wilayah dan pengembangan kawasan berbasis pertanian. Pengembangan kopi sebagai komoditas unggulan, khususnya dalam konteks agropolitan, menekankan pentingnya peningkatan mutu produk dan daya saing melalui pengelolaan komoditas secara berkelanjutan [1]. Dalam kerangka tersebut, kualitas kopi, terutama atribut aroma, menjadi faktor kunci yang menentukan penerimaan pasar dan nilai tambah produk [2].

Penilaian mutu kopi di lapangan umumnya masih mengandalkan evaluasi sensori, seperti *cupping test*, yang bertumpu pada persepsi panelis manusia. Meskipun metode ini digunakan secara luas sebagai standar industri, beberapa kajian menunjukkan bahwa evaluasi sensori memiliki keterbatasan karena dipengaruhi subjektivitas penilai, kondisi fisiologis, serta variasi pengalaman antar panelis. Akibatnya,

konsistensi hasil penilaian menjadi lebih sulit dijaga ketika jumlah sampel meningkat [2].

Sejalan dengan kebutuhan akan metode penilaian yang lebih objektif, *Electronic Nose* (E-Nose) berkembang sebagai pendekatan instrumentasi untuk mengkarakterisasi aroma melalui respons *array* sensor gas yang diproses menggunakan teknik pengenalan pola [3]. Pada konteks kopi, E-Nose telah digunakan untuk menilai profil dan intensitas aroma biji kopi secara sistematis melalui pemodelan berbasis *machine learning*, termasuk pendekatan *fingerprint* aroma dan jaringan saraf [3], [4]. Selain algoritma pemodelan, aspek perancangan perangkat E-Nose seperti tipe ruang (*chamber*) dan posisi sensor juga dilaporkan memengaruhi sensitivitas dan konsistensi akuisisi sinyal aroma [5].

Namun demikian, penerapan E-Nose pada analisis aroma menghadapi tantangan utama pada kualitas sinyal sensor gas. Sinyal E-Nose cenderung terpengaruh oleh fluktuasi lingkungan dan dinamika respons sensor, sehingga sering

mengandung *noise*, *drift*, serta nilai ekstrem (*outlier*) yang dapat mengganggu proses klasifikasi. Pengembangan metode deteksi *outlier* berbasis penyaringan adaptif pada sinyal E-Nose dilaporkan mampu meningkatkan kualitas data dengan capaian *balanced accuracy* sebesar 0,904, yang menunjukkan pentingnya penanganan *outlier* sebelum tahap pemodelan [6]. Temuan ini sejalan dengan kajian kualitas data deret waktu pada dataset gas dan cuaca, yang menegaskan bahwa deteksi *outlier* dan penanganan data hilang merupakan langkah krusial untuk meningkatkan reliabilitas data sebelum analisis lanjutan [7].

Selain penanganan *outlier*, peningkatan kualitas sinyal E-Nose juga dapat dilakukan melalui transformasi fitur. *Discrete Wavelet Transform* (DWT) banyak digunakan untuk memisahkan komponen informasi dan gangguan pada domain waktu-frekuensi, sehingga sinyal yang dihasilkan menjadi lebih representatif untuk proses klasifikasi. Sejumlah studi menunjukkan bahwa DWT efektif dalam meningkatkan kualitas representasi sinyal dan mendukung peningkatan performa klasifikasi pada berbagai aplikasi [8], [9].

Sebagai pembeda, penelitian ini tidak hanya menitikberatkan pada eliminasi *outlier*, melainkan juga mengintegrasikan *Discrete Wavelet Transform* (DWT) sebagai tahap transformasi waktu-frekuensi untuk memperkaya representasi fitur sinyal. Perbedaan ini tampak bila dibandingkan dengan penelitian Ifnu Wisma Dwi Prastyana dan Sarno yang menerapkan *signal processing* melalui dua tahap, yakni penghilangan *noise* menggunakan *interquartile range* (IQR) dan pemulusan sinyal menggunakan *moving average*, kemudian dilanjutkan dengan ekstraksi fitur statistik, seleksi fitur, serta pengujian beberapa pengklasifikasi berbasis *machine learning* [10]. Dengan memasukkan DWT, penelitian ini berupaya menekan pengaruh *noise* dan dinamika sinyal yang belum sepenuhnya tertangkap oleh pemulusan sederhana, sehingga tumpang tindih pola volatil antarkelas diharapkan dapat dipisahkan lebih tegas pada ruang fitur.

Tahap pra-pemrosesan berikutnya yang berpengaruh terhadap stabilitas pembelajaran adalah normalisasi data. *Min-Max scaling* digunakan untuk memetakan nilai fitur ke dalam rentang yang seragam sehingga perbedaan skala antar fitur tidak mendominasi proses pelatihan model. Pada studi E-Nose untuk klasifikasi madu, penerapan normalisasi *Min-Max* dalam rangkaian pemrosesan dilaporkan menghasilkan akurasi klasifikasi sebesar 93,6%, yang menunjukkan bahwa normalisasi berperan penting dalam meningkatkan kinerja model [11].

Di samping pra-pemrosesan, strategi evaluasi model turut menentukan reliabilitas hasil klasifikasi. Evaluasi berbasis pemisahan data (*hold-out split*) digunakan untuk menguji kemampuan generalisasi model terhadap data yang tidak terlihat saat pelatihan, sementara *cross-validation* digunakan untuk menilai kestabilan performa pada variasi pembagian data. Pendekatan evaluasi tersebut banyak digunakan pada penelitian klasifikasi karena mampu memberikan estimasi performa model yang lebih representatif dan mengurangi

ketergantungan hasil terhadap satu skema pembagian data tertentu [12].

Lebih lanjut, kompleksitas pola sinyal multisensor mendorong penggunaan model dengan kapasitas representasi yang lebih tinggi. *Deep Neural Network* (DNN), termasuk arsitektur bertingkat (*stacked*), mampu memodelkan hubungan nonlinier yang kompleks pada data E-Nose. Pada studi deteksi penyakit pernapasan berbasis E-Nose dari keringat aksila, penerapan *stacked* DNN menghasilkan akurasi pengujian sebesar 0,934, yang menunjukkan keunggulan pendekatan *deep learning* dalam klasifikasi berbasis sinyal aroma [13].

Selain pendekatan *deep learning*, berbagai algoritma pembelajaran mesin juga telah digunakan dalam klasifikasi aroma berbasis *Electronic Nose*. Metode *Support Vector Machine* (SVM) dilaporkan mampu memberikan performa klasifikasi yang baik pada sistem pengenalan aroma kopi. Penelitian oleh Erwanto dkk. menunjukkan bahwa penggunaan SVM pada sistem *Electronic Nose* mampu mencapai akurasi klasifikasi sebesar 82,2% dalam membedakan beberapa varietas kopi berdasarkan profil aromanya [14].

Di sisi lain, metode *Artificial Neural Network* (ANN) juga banyak diterapkan dalam klasifikasi aroma kopi berbasis *Electronic Nose* karena kemampuannya dalam memodelkan hubungan nonlinier antara respons sensor dan karakteristik aroma. Penelitian oleh Astuti dkk. menunjukkan bahwa kombinasi *Electronic Nose* dan ANN mampu mencapai akurasi hingga 98,8% pada klasifikasi profil tingkat sangrai kopi [15]. Namun demikian, penelitian tersebut hanya membedakan beberapa kategori tingkat sangrai sehingga kompleksitas klasifikasinya relatif lebih rendah dibandingkan klasifikasi multikelas berdasarkan jenis dan asal kopi seperti pada penelitian ini.

Berdasarkan karakteristik data pada penelitian ini, fitur masukan berbentuk data tabular dengan tiga atribut numerik utama, yaitu MQ-9, MQ-135, dan suhu (Temp C), tanpa struktur spasial seperti citra dan tanpa ketergantungan sekuensial panjang yang dimodelkan secara eksplisit. Oleh karena itu, pendekatan berbasis multilayer perceptron seperti *Deep Neural Network* (DNN) dinilai lebih sesuai karena berfokus pada pembelajaran relasi nonlinier antar fitur tabular [16].

Meskipun berbagai studi telah menunjukkan potensi *Electronic Nose* untuk klasifikasi aroma kopi, masih terdapat beberapa celah penelitian yang belum banyak dibahas secara terpadu. Pertama, sebagian penelitian terdahulu berfokus pada jumlah kelas yang relatif terbatas, seperti intensitas aroma atau tingkat sangrai, sehingga tingkat kompleksitas klasifikasinya masih lebih rendah dibandingkan klasifikasi multikelas berdasarkan jenis dan asal kopi. Kedua, studi terdahulu umumnya menekankan satu aspek tertentu, misalnya pemodelan klasifikasi atau pra-pemrosesan sinyal, tetapi belum banyak yang mengevaluasi secara komparatif pengaruh pipeline pra-pemrosesan berurutan terhadap beberapa model klasifikasi pada data E-Nose kopi berbentuk

tabular. Ketiga, perbandingan langsung antara DNN, SVM, dan ANN pada dua skenario pra-pemrosesan, yaitu dengan dan tanpa kombinasi IQR dan DWT, masih terbatas dilaporkan pada konteks klasifikasi aroma kopi multikelas. Oleh karena itu, penelitian ini menempatkan kontribusinya pada evaluasi komparatif model dan skenario pra-pemrosesan secara sistematis untuk menilai sejauh mana IQR dan DWT berkontribusi terhadap peningkatan kualitas representasi fitur dan performa klasifikasi.

Berdasarkan uraian tersebut, penelitian ini bertujuan mengevaluasi performa tiga model klasifikasi, yaitu Deep Neural Network (DNN), Support Vector Machine (SVM), dan Artificial Neural Network (ANN), pada klasifikasi aroma kopi multikelas berbasis Electronic Nose.

Evaluasi dilakukan pada dua skenario, yaitu skenario dengan pra-pemrosesan berurutan berupa Interquartile Range (IQR), Discrete Wavelet Transform (DWT), dan normalisasi Min–Max, serta skenario perbandingan yang hanya menerapkan normalisasi Min–Max. Untuk memperoleh estimasi performa yang lebih stabil dan tetap menjaga pemisahan data uji secara independen, penelitian ini menggunakan train–test split 80:20 dan K-Fold Cross Validation dengan $k = 5$ pada data latih.

Dengan rancangan tersebut, penelitian ini diharapkan dapat menunjukkan pengaruh pipeline pra-pemrosesan terhadap kualitas representasi fitur sekaligus mengidentifikasi model yang paling efektif untuk klasifikasi aroma kopi multikelas.

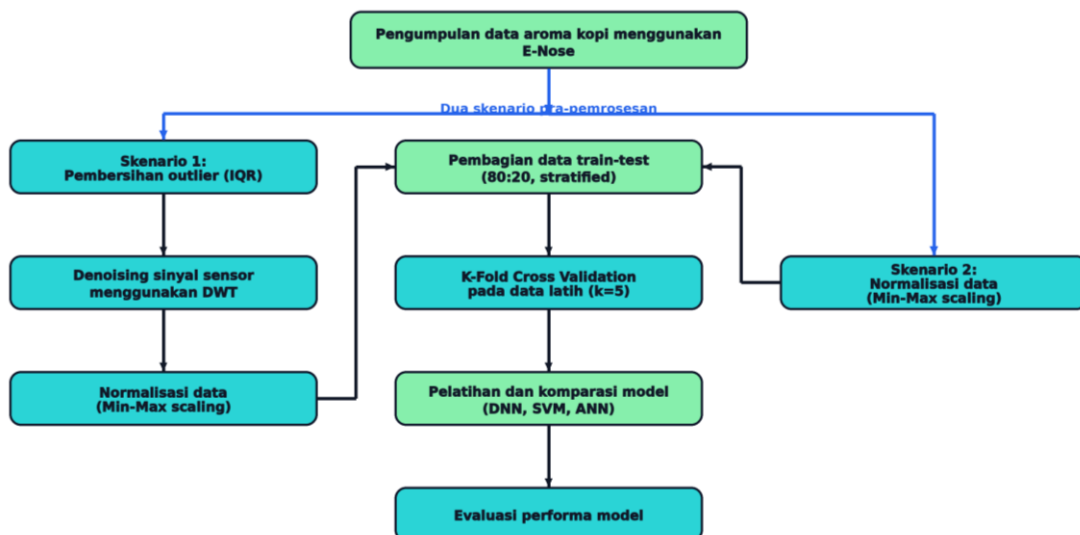
Bagian ini menjelaskan tahapan penelitian untuk membangun sistem klasifikasi aroma kopi berbasis *Electronic Nose* (E-Nose) serta mengevaluasi performa beberapa model klasifikasi, yaitu *Deep Neural Network* (DNN), *Support Vector Machine* (SVM), dan *Artificial Neural Network* (ANN). Alur penelitian dimulai dari akuisisi data menggunakan perangkat E-Nose, dilanjutkan dengan pra-pemrosesan data yang meliputi pembersihan *outlier*, denoising sinyal, dan normalisasi, kemudian pemodelan dan evaluasi performa menggunakan skema validasi serta metrik klasifikasi. Penyusunan tahapan secara bertingkat diperlukan karena data sensor gas umumnya bersifat fluktuatif dan dapat mengandung nilai ekstrem, sehingga perlu distabilkan terlebih dahulu sebelum dipelajari oleh model pembelajaran mesin agar tidak menghasilkan keputusan yang bias atau *overfitting* [17]. Ringkasan alur penelitian ditunjukkan pada Gambar 1.

Selain skenario utama yang menerapkan pembersihan *outlier* menggunakan IQR dan denoising sinyal menggunakan *Discrete Wavelet Transform* (DWT), penelitian ini juga menguji skenario perbandingan tanpa penerapan IQR dan DWT. Pada kedua skenario tersebut, data selanjutnya dinormalisasi menggunakan *Min–Max scaling* sebelum diproses oleh tiga model klasifikasi, yaitu DNN, SVM, dan ANN. Rancangan ini bertujuan untuk menganalisis pengaruh tahapan pra-pemrosesan terhadap kualitas representasi fitur sekaligus membandingkan performa antar model klasifikasi pada data aroma kopi multikelas. Dengan demikian, dapat dievaluasi sejauh mana kontribusi IQR dan DWT dalam meningkatkan kestabilan sinyal, kualitas fitur, dan kemampuan generalisasi masing-masing model.

II. METODE

Diagram Alur Eksperimen Klasifikasi Aroma Kopi

Dua skenario pra-pemrosesan: (1) IQR + DWT + Min-Max, dan (2) Min-Max saja



Gambar 1. Alur Penelitian

Seluruh model pada kedua skenario dievaluasi menggunakan pembagian data, metrik evaluasi, dan prosedur

validasi yang sama agar perbandingan performa dilakukan secara adil pada kondisi eksperimen yang setara.

A. Pengumpulan Data Aroma Kopi Menggunakan E-Nose

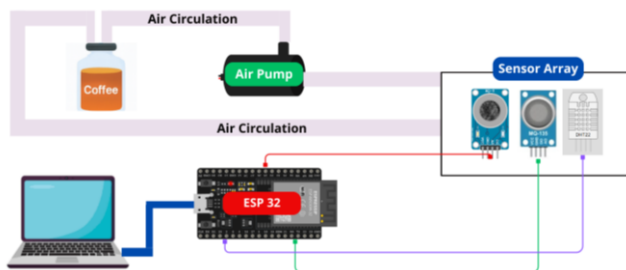
1) Desain Perangkat dan Konfigurasi Sensor

Akuisisi aroma kopi dilakukan menggunakan perangkat *E-Nose* berbasis *ESP32* yang membaca keluaran sensor gas *MQ-9* dan *MQ-135*, serta sensor lingkungan *DHT22* (suhu-kelembapan). Prinsip *E-Nose* memanfaatkan pola gabungan respons *multi-sensor* (*sensor array*) sebagai *fingerprint* aroma [18]. Pola gabungan ini umumnya lebih informatif dibanding satu sensor tunggal karena tiap sensor memiliki sensitivitas berbeda sehingga saling melengkapi dalam membentuk ciri aroma [14].

Pemilihan sensor *MQ-9*, *MQ-135*, dan *DHT22* pada penelitian ini didasarkan pada pertimbangan bahwa aroma kopi merupakan campuran senyawa volatil kompleks yang tidak direpresentasikan oleh satu komponen gas tunggal. *MQ-135* digunakan karena sensitif terhadap berbagai senyawa gas dan *volatile organic compounds* (*VOC*) yang umum muncul pada aroma kopi, sedangkan *MQ-9* melengkapi respons terhadap komponen gas reduktif dan hidrokarbon aromatik. *DHT22* tidak digunakan untuk mengidentifikasi senyawa aroma secara langsung, tetapi untuk memantau kondisi lingkungan pengukuran, khususnya suhu, yang dapat memengaruhi stabilitas respons sensor *MOS*. Dengan kombinasi ini, sistem tidak dimaksudkan untuk mengukur komposisi kimia kopi secara spesifik, melainkan menangkap pola gabungan respons sensor sebagai *fingerprint* aroma yang dapat dibedakan oleh model klasifikasi.

Selain itu, beberapa studi klasifikasi kopi berbasis *E-Nose* menekankan bahwa aroma dari sampel yang berbeda dapat menghasilkan respons yang saling mendekati, sehingga konsistensi rancangan sistem (*chamber*, aliran udara, dan posisi sensor) harus dijaga agar perbedaan sinyal benar-benar merepresentasikan perbedaan kelas kopi [19].

Konfigurasi sistem *E-Nose* pada penelitian ini ditunjukkan pada Gambar 2, yaitu aroma kopi dialirkan melalui sistem pompa udara menuju ruang sensor yang berisi *MQ-9*, *MQ-135*, dan *DHT22*, lalu sinyal dibaca *ESP32* dan dikirim ke komputer untuk disimpan sebagai *dataset*.



Gambar 2. Skema Konfigurasi Electronic Nose

2) Prosedur Eksperimen dan Akuisisi Data

Akuisisi data dilakukan pada malam hari dalam kondisi suhu lingkungan sekitar 26 °C berdasarkan pengamatan saat eksperimen. Sistem *Electronic Nose* berbasis *ESP32* terhubung dengan larik sensor *MQ-9*, *MQ-135*, dan *DHT22*. Sensor *MQ-9* dan *MQ-135* merekam respons sensor gas terhadap campuran senyawa volatil dari aroma kopi, sedangkan *DHT22* digunakan untuk memantau kondisi lingkungan pengukuran. Dalam dataset yang digunakan pada tahap pemodelan, fitur lingkungan yang dimanfaatkan adalah suhu (*Temp C*), sementara kelembapan berperan sebagai parameter pemantauan selama proses akuisisi. Perlu dicatat bahwa nilai suhu yang terekam oleh *DHT22* pada dataset berada pada kisaran sekitar 35 °C karena sensor ditempatkan pada sistem pengukuran yang dipengaruhi oleh kondisi lokal di sekitar *chamber* dan paparan uap panas dari sampel, sehingga nilainya tidak sepenuhnya merepresentasikan suhu lingkungan luar. Sampel kopi ditempatkan pada *chamber* uji berbahan akrilik, kemudian aliran aroma diatur melalui mekanisme sirkulasi udara menggunakan pompa agar paparan volatil ke permukaan sensor berlangsung lebih seragam antarsesi. Mekanisme pengaliran seperti ini umum diterapkan pada sistem *E-Nose* karena berperan penting dalam menjaga kestabilan respons sensor dan mengurangi variasi pengukuran yang tidak diinginkan [20].

Untuk menjamin konsistensi pengambilan data, setiap sesi pengukuran menggunakan 10 g bubuk kopi yang ditempatkan pada wadah gelas berkapasitas 100 mL, kemudian diberikan air bersuhu 100 °C guna membantu pelepasan senyawa volatil secara lebih stabil. Sinyal sensor direkam sebanyak 10 kali pengambilan per sesi sesuai konfigurasi sistem. Dalam konteks penelitian ini, sepuluh kali pengambilan tersebut diperlakukan sebagai pembacaan berulang dalam satu sesi akuisisi untuk menangkap variasi respons sensor terhadap paparan aroma yang sama secara lebih stabil. Setelah satu sesi pengukuran selesai, dilakukan fase pemulihan (*purging*) dengan mengalirkan udara bersih melalui jalur sirkulasi untuk menurunkan efek *carry-over* atau sisa aroma dari pengukuran sebelumnya, sehingga tidak memengaruhi sesi berikutnya.

Seluruh keluaran sensor dipaketkan oleh *ESP32* dan dikirimkan ke komputer melalui komunikasi serial untuk disimpan dalam format tabel (.xlsx) sebagai *dataset* mentah yang akan digunakan pada tahap pra-pemrosesan dan pemodelan [21].

Sebelum proses akuisisi data dimulai, sensor dinyalakan dan dibiarkan mencapai kondisi stabil (*warm-up*) guna meminimalkan *drift* awal, sehingga respons sensor yang direkam merepresentasikan kondisi pengukuran yang konsisten dan dapat dibandingkan antar sesi. Rincian prosedur akuisisi data ini dirancang untuk memastikan konsistensi antar sesi pengukuran serta meningkatkan reproduktibilitas penelitian, sehingga data aroma kopi yang diperoleh dapat dibandingkan secara objektif dan digunakan secara andal pada tahap pemodelan. Secara keseluruhan, proses akuisisi menghasilkan 1.503 sampel dari 10 kelas kopi, dengan proporsi sampel antarkelas yang relatif seimbang

sehingga layak digunakan untuk evaluasi klasifikasi multikelas.

Proses pengambilan *dataset* aroma kopi pada penelitian ini ditunjukkan pada Gambar 3, yang menggambarkan alur akuisisi data mulai dari penempatan sampel kopi pada *chamber* uji, pengaliran uap volatil melalui sistem sirkulasi udara, hingga perekaman respons sensor oleh ESP32. Pada tahap ini, senyawa volatil yang dilepaskan dari sampel kopi dialirkan secara terkontrol menuju ruang sensor sehingga respons MQ-9 dan MQ-135 serta parameter lingkungan dari DHT22 dapat direkam selama proses akuisisi. Dalam tahap pemodelan, fitur lingkungan yang digunakan adalah suhu (Temp C). Alur tersebut dirancang untuk memastikan bahwa setiap sesi pengukuran mengikuti prosedur yang sama, sehingga data yang dihasilkan bersifat konsisten dan layak digunakan sebagai masukan pada tahap pra-pemrosesan dan pemodelan klasifikasi.



Gambar 3. Proses Pengambilan Dataset Kopi

Karakteristik sensor yang digunakan dalam sistem *E-Nose* pada penelitian ini dirangkum pada Tabel 1, yang memuat jenis sensor beserta parameter fisik dan kimia yang dideteksi sebagai sumber fitur masukan pada proses pemodelan.

TABEL 1.
SENSOR DAN PARAMETER YANG DIDETEKSII

Sensor	Parameter yang Dideteksi
MQ-135	Karbon dioksida (CO ₂), amonia (NH ₃), VOC (etanol, dll.)
MQ-9	Karbon monoksida (CO), hidrokarbon aromatik
DHT22	Suhu udara (°C), kelembapan relatif (%)

Parameter yang terdeteksi oleh masing-masing sensor pada Tabel 1 selanjutnya digunakan sebagai fitur masukan pada tahap pra-pemrosesan dan pemodelan klasifikasi aroma kopi menggunakan model DNN, SVM, dan ANN.

Dataset penelitian merupakan hasil akuisisi aroma kopi menggunakan sistem *Electronic Nose* berbasis ESP32. Data yang digunakan sebagai fitur masukan pada proses klasifikasi terdiri atas tiga atribut numerik, yaitu MQ-9, MQ-135, dan suhu (Temp C), serta satu label kelas berupa jenis kopi. Meskipun sensor DHT22 secara umum mampu membaca suhu dan kelembapan, pada dataset yang digunakan dalam penelitian ini fitur lingkungan yang dimanfaatkan pada proses pemodelan adalah suhu (Temp C), sedangkan label "Jenis" digunakan sebagai target klasifikasi. Total data yang

diperoleh berjumlah 1.503 record yang mewakili 10 kelas kopi Arabika dan Robusta dari berbagai daerah di Indonesia.

3) Struktur Dataset Mentah Hasil Akuisisi E-Nose

Dataset mentah pada penelitian ini merupakan hasil perekaman langsung sistem *Electronic Nose* sebelum melalui tahapan pra-pemrosesan. Setiap *record* memuat nilai keluaran sensor MQ-9, MQ-135, suhu (Temp C) dari DHT22, serta label jenis kopi sebagai target klasifikasi. Pada berkas mentah, dataset juga memuat kolom timestamp sebagai penanda waktu perekaman. Namun, karena penelitian ini berfokus pada klasifikasi berdasarkan respons sensor, fitur yang digunakan pada tahap pemodelan dibatasi pada MQ-9, MQ-135, dan suhu (Temp C), sedangkan timestamp tidak digunakan sebagai variabel masukan. Penyajian struktur data mentah bertujuan menjaga transparansi eksperimen sekaligus menegaskan bahwa data awal masih berpotensi mengandung fluktuasi dan nilai ekstrem yang umum dijumpai pada data sensor, sehingga memerlukan pra-pemrosesan agar proses pembelajaran model berlangsung lebih stabil [9].

Seluruh data disimpan dalam berkas *Dataset E nose (3).xlsx* yang mencakup 1.503 *record* dari 10 kelas kopi, dengan contoh format *dataset* mentah disajikan pada Tabel 2 sebagai representasi struktur kolom sebelum tahap pra-pemrosesan.

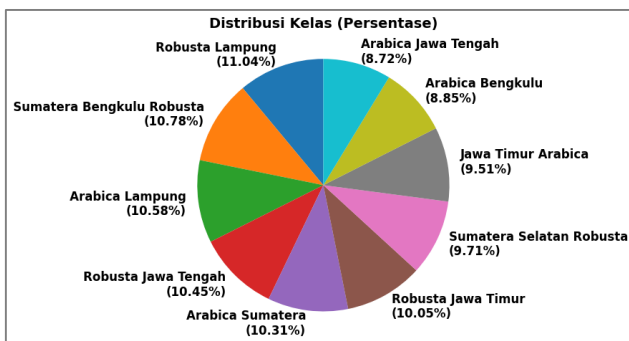
TABEL 2.
STRUKTUR DATASET MENTAH HASIL AKUISISI E-NOSE

No	MQ-9	MQ-135	Suhu (°C)	Jenis Kopi
1	149	1602	35.3	Robusta Jawa Timur
2	152	1650	35.2	Robusta Jawa Timur
...
...
153	111	426	35,5	Robusta Jawa Tengah
154	314	4095	35,4	Robusta Jawa Tengah
...
...
310	108	365	35,7	Robusta Lampung
311	252	3230	35,7	Robusta Lampung
...
...
476	95	256	35,6	Sumatera Bengkulu Robusta
477	201	2611	35,5	Sumatera Bengkulu Robusta
...
...
638	88	222	35,2	Sumatera Selatan Robusta
639	279	3801	35,2	Sumatera Selatan Robusta
...
...
784	118	311	30,6	Arabica Bengkulu
785	400	4095	30,6	Arabica Bengkulu
...

No	MQ-9	MQ-135	Suhu (°C)	Jenis Kopi
...
917	84	196	33,1	Arabica Jawa Tengah
918	239	3331	33,1	Arabica Jawa Tengah
...
...
1048	101	316	32	Arabica Sumatera
1049	260	3870	32,1	Arabica Sumatera
...
...
1203	99	318	33,8	Arabica Lampung
1204	329	4095	33,8	Arabica Lampung
...
...
1502	162	1330	35	Arabica Jawa Timur
1503	161	1329	35	Arabica Jawa Timur

Untuk memberikan gambaran mengenai proporsi data pada setiap kelas, distribusi sampel per kelas ditampilkan dalam bentuk diagram persentase pada Gambar 4.

Berdasarkan Gambar 4, distribusi sampel pada masing-masing kelas menunjukkan proporsi yang relatif merata, dengan rentang persentase antara 8,72% hingga 11,04%. Kelas dengan proporsi tertinggi adalah Robusta Lampung sebesar 11,04%, sedangkan proporsi terendah terdapat pada Arabica Jawa Tengah sebesar 8,72%, sehingga selisih antarkelas hanya sekitar 2,32%. Kondisi ini menunjukkan bahwa dataset tergolong relatif seimbang (*balanced*), sehingga dapat membantu meminimalkan bias model terhadap kelas tertentu dan mendukung evaluasi komparatif yang lebih proporsional pada seluruh kelas kopi.



Gambar 4. Distribusi kelas

B. Pra-pemrosesan Data

Penelitian ini menggunakan dua skenario eksperimen. Skenario pertama menerapkan pra-pemrosesan berurutan berupa pembersihan outlier menggunakan Interquartile Range (IQR), denoising sinyal menggunakan Discrete Wavelet Transform (DWT) basis Daubechies-4 (db4), dan normalisasi fitur menggunakan Min–Max scaling. Skenario kedua digunakan sebagai pembanding, yaitu tanpa penerapan IQR

dan DWT, namun tetap menggunakan normalisasi Min–Max agar seluruh model dibandingkan pada skala fitur yang setara. Kedua skenario tersebut kemudian dievaluasi menggunakan tiga model klasifikasi, yaitu DNN, SVM, dan ANN.

Pada implementasinya, kedua skenario pra-pemrosesan diterapkan setelah pemisahan data train–test untuk mencegah kebocoran informasi dari data uji ke proses pelatihan. Pada skenario pertama, data latih diproses secara berurutan melalui IQR, DWT, dan Min–Max scaling, kemudian parameter hasil proses tersebut diterapkan secara konsisten pada data validasi dan data uji. Pada skenario kedua, data hanya melalui normalisasi Min–Max dengan parameter yang diturunkan dari data latih. Dengan pengaturan ini, setiap model dievaluasi pada kondisi yang setara tanpa memanfaatkan informasi statistik dari data uji pada tahap pelatihan.

1) Pembersihan Outlier Menggunakan Interquartile Range (IQR)

Metode *Interquartile Range* (IQR) digunakan karena bersifat *robust*, tidak mensyaratkan distribusi data normal, serta efektif untuk mendeteksi nilai ekstrem pada data sensor [6]. Pada data *E-Nose*, *outlier* dapat muncul saat sensor mengalami saturasi, perubahan aliran mikro, atau gangguan listrik sesaat. Menghapus *outlier* membantu mencegah model meminimalkan *loss* dengan “mengikuti” nilai yang tidak wajar [9].

IQR didefinisikan sebagai selisih kuartil ketiga dan kuartil pertama:

$$IQR = Q_3 - Q_1 \tag{1}$$

Dimana:

Q_1 = kuartil pertama (persentil 25)

Q_3 = kuartil ketiga (persentil 75)

IQR = rentang antar-kuartil (*interquartile range*)

Batas bawah dan batas atas data “wajar”:

$$\begin{aligned} \text{Batas bawah} &= Q_1 - 1.5 \times IQR \\ \text{Batas atas} &= Q_3 + 1.5 \times IQR \end{aligned} \tag{2}$$

Dimana

Q_1, Q_3, IQR = sama seperti di atas

1.5 = konstanta standar untuk mendeteksi *outlier* moderat

Nilai $MQ-9, MQ-135$, dan $Temp C$ yang berada di luar batas (2) dieliminasi dari *dataset* sebelum tahap *denoising* dan normalisasi agar distribusi data lebih representatif.

2) Peredaman Derau (Denoising) Menggunakan Wavelet Transform (db4)

Setelah *outlier* dihapus, sinyal sensor masih berpotensi mengandung *noise* berfrekuensi tinggi. *Denoising* diperlukan agar model mempelajari pola *respons* utama (kenaikan–penurunan akibat paparan aroma), bukan fluktuasi kecil acak

yang tidak stabil [22]. *Discrete Wavelet Transform (DWT)* dipilih karena mampu merepresentasikan sinyal pada *domain* waktu dan frekuensi secara lokal, sehingga cocok untuk sinyal sensor yang bersifat *non-stasioner* [8].

Representasi sinyal dengan fungsi skala dan wavelet:

$$s(t) = \sum_k a_{j_0,k} \phi_{j_0,k}(t) + \sum_{j=j_0}^J \sum_k d_{j,k} \psi_{j,k}(t) \quad (3)$$

Dimana:

$s(t)$ = sinyal asli pada waktu t

$\phi_{j_0,k}(t)$ = fungsi skala (approximation)

$\psi_{j,k}(t)$ = fungsi wavelet (detail)

$a_{j_0,k}$ = koefisien aproksimasi pada level j_0

$d_{j,k}$ = koefisien detail pada level j dan indeks k

J = level dekomposisi maksimum

j_0 = level awal dekomposisi

Koefisien detail kemudian ditapis menggunakan *soft threshold*:

$$d'_{j,k} = \begin{cases} \text{sign}(d_{j,k})(|d_{j,k}| - \lambda), & |d_{j,k}| > \lambda \\ 0, & |d_{j,k}| \leq \lambda \end{cases} \quad (4)$$

Dimana:

$d_{j,k}$ = koefisien detail sebelum ditapis

$d'_{j,k}$ = koefisien detail setelah ditapis

λ = nilai ambang (*threshold*)

$\text{sign}(\cdot)$ = fungsi tanda (+/-)

$|\cdot|$ = nilai *absolut*

Pada tahap implementasi, DWT basis Daubechies-4 (db4) dengan level dekomposisi 2 digunakan untuk melakukan denoising pada sinyal setiap fitur sensor sebelum normalisasi. Koefisien detail ditapis menggunakan pendekatan *soft threshold*, kemudian sinyal direkonstruksi kembali untuk menghasilkan representasi sinyal yang lebih halus dan stabil sebagai masukan model klasifikasi.

3) Normalisasi Fitur Menggunakan Min–Max Scaling

Normalisasi diperlukan karena skala *MQ-135* umumnya jauh lebih besar dibanding *MQ-9* dan *Temp C*. Tanpa normalisasi, proses optimasi cenderung didominasi fitur berskala besar sehingga pembelajaran tidak seimbang [23]. Metode *Min–Max* dipilih karena sederhana dan mempertahankan proporsi antar nilai fitur pada rentang yang seragam [23].

Rumus Min–Max scaling:

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (5)$$

Dimana:

x = nilai fitur sebelum normalisasi

x' = nilai fitur setelah normalisasi

x_{\min} = nilai minimum fitur pada *dataset*

x_{\max} = nilai maksimum fitur pada *dataset*

Selain DNN sebagai model utama, penelitian ini juga menggunakan SVM dan ANN sebagai model pembanding. SVM dipilih karena dikenal efektif untuk klasifikasi data tabular dengan batas keputusan nonlinier melalui kernel Radial Basis Function (RBF). ANN digunakan sebagai pembanding dari kelompok jaringan saraf dengan arsitektur yang lebih sederhana dibandingkan DNN. Dengan demikian, perbandingan ketiga model tersebut dapat menunjukkan pengaruh kompleksitas model terhadap kemampuan klasifikasi aroma kopi multikelas pada data E-Nose.

C. Pembagian Data dan Pemodelan Klasifikasi

1) Pembagian Data Train–Test (80:20)

Dataset hasil pra-pemrosesan dibagi menjadi data latih dan data uji dengan rasio 80:20. Pembagian ini dilakukan untuk memastikan evaluasi akhir dilakukan pada data yang tidak pernah digunakan selama pelatihan, sehingga dapat menggambarkan kemampuan generalisasi model secara lebih objektif [17]. Untuk menjaga proporsi tiap kelas tetap terwakili pada kedua subset, pembagian data dilakukan secara stratified berdasarkan label kelas dengan `random_state = 42`. Jika N adalah jumlah sampel setelah pra-pemrosesan, maka:

$$\begin{aligned} N_{\text{train}} &= 0.8N \\ N_{\text{test}} &= 0.2N \end{aligned} \quad (6)$$

Dimana:

N = jumlah total sampel setelah pra-pemrosesan

N_{train} = jumlah sampel data latih

N_{test} = jumlah sampel data uji

2) K-Fold Cross Validation ($k = 5$)

Evaluasi model dilakukan menggunakan skema train–test split 80:20, di mana data uji disimpan terpisah dan hanya digunakan pada tahap evaluasi akhir. Selanjutnya, data latih dievaluasi menggunakan K-Fold Cross Validation dengan $k = 5$ untuk memperoleh estimasi performa yang lebih stabil terhadap variasi pembagian data. Pada penelitian ini, *K-Fold* diimplementasikan dengan `shuffle = True` dan `random_state = 42` agar pembagian fold dilakukan secara acak namun tetap reproduisibel. Pemilihan $k = 5$ dilakukan karena memberikan keseimbangan yang baik antara reliabilitas evaluasi dan kecukupan jumlah data pelatihan pada setiap fold.

Ukuran tiap *fold*:

$$N_{\text{fold}} = \frac{N_{\text{train}}}{k} \quad (7)$$

Dimana:

N_{train} = jumlah sampel data latih

k = jumlah *fold* (pada penelitian ini $k = 5$)

N_{fold} = jumlah sampel per *fold*

Akurasi pada *fold* ke- i :

$$Accuracy_i = \frac{TP_i + TN_i}{TP_i + TN_i + FP_i + FN_i} \quad (8)$$

Dimana:

TP_i = true positive pada fold ke- i

TN_i = true negative pada fold ke- i

FP_i = false positive pada fold ke- i

FN_i = false negative pada fold ke- i

$Accuracy_i$ = akurasi fold ke- i

Rata-rata akurasi K -Fold:

$$Accuracy = \frac{1}{k} \sum_{i=1}^k Accuracy_i \quad (9)$$

Dimana:

k = jumlah fold

$Accuracy_i$ = akurasi pada fold ke- i

$Accuracy$ = akurasi rata-rata

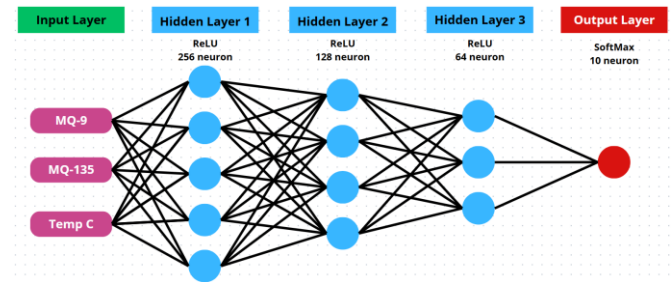
Untuk mencegah *data leakage*, pemisahan data dilakukan secara berjenjang. Pertama, data dibagi menjadi data latih (80%) dan data uji (20%), di mana data uji disimpan terpisah dan hanya digunakan pada evaluasi akhir. Kedua, *cross-validation* dilakukan hanya pada data latih, sehingga proses pelatihan dan validasi tidak pernah melibatkan data uji. Ketiga, setiap langkah pra-pemrosesan yang berpotensi “belajar” dari data khususnya penentuan parameter normalisasi ditetapkan berdasarkan data latih, kemudian diterapkan secara konsisten pada data validasi dan data uji. Dengan pengaturan ini, pelaporan akurasi uji merepresentasikan kemampuan model pada data yang benar-benar tidak teramati selama pelatihan.

3) Perancangan Model Deep Neural Network (DNN)

Pada penelitian ini, istilah *Deep Neural Network* (DNN) digunakan secara konsisten untuk merujuk pada jaringan saraf berlapis dengan lebih dari satu *hidden layer*, sedangkan istilah *Artificial Neural Network* (ANN) digunakan untuk model perbandingan dengan arsitektur yang lebih sederhana. Model utama yang diterapkan adalah DNN bertipe *multilayer perceptron* untuk klasifikasi multikelas. *Input layer* memuat tiga fitur numerik hasil pembacaan sensor, yaitu MQ-9, MQ-135, dan suhu (*Temp C*) dari DHT22.

Arsitektur jaringan terdiri atas tiga *hidden layer* berukuran 256, 128, dan 64 neuron dengan fungsi aktivasi *Rectified Linear Unit* (ReLU). Untuk meningkatkan kestabilan pelatihan dan membantu mengurangi risiko *overfitting*, pada setiap *hidden layer* diterapkan *Batch Normalization* dan *Dropout*. Nilai *dropout* yang digunakan adalah 0,3 pada dua lapisan tersembunyi pertama dan 0,2 pada lapisan tersembunyi ketiga. Pada *output layer* digunakan fungsi aktivasi *softmax* dengan jumlah neuron sesuai banyaknya kelas kopi, yaitu 10 neuron, untuk menghasilkan probabilitas prediksi pada klasifikasi multikelas.

Pemilihan arsitektur ini didasarkan pada kebutuhan untuk memodelkan hubungan nonlinier antara respons multisensor dan kelas kopi. Kombinasi beberapa *hidden layer*, *Batch Normalization*, dan *Dropout* diharapkan mampu membantu jaringan dalam mempelajari pola aroma yang kompleks secara lebih stabil, sekaligus menjaga kemampuan generalisasi model pada data yang tidak teramati selama pelatihan.



Gambar 5. Arsitektur Deep Neural Network

Arsitektur pada Gambar 5 menunjukkan alur pemrosesan dari lapisan masukan hingga lapisan keluaran. Vektor fitur masukan, yaitu MQ-9, MQ-135, dan *Temp C*, ditransformasikan secara bertahap melalui tiga *hidden layer* beraktivasi ReLU untuk mengekstraksi representasi fitur yang semakin abstrak. Selanjutnya, lapisan keluaran dengan fungsi aktivasi *softmax* mengubah hasil transformasi tersebut menjadi distribusi probabilitas terhadap 10 kelas kopi. Susunan ini dipilih untuk menjaga keseimbangan antara kapasitas representasi model dan kompleksitas komputasi, serta mendukung pelatihan yang stabil pada dataset penelitian. Gambar 5 disajikan sebagai ilustrasi skematik arsitektur jaringan, sedangkan jumlah neuron pada masing-masing lapisan ditunjukkan melalui label layer.

Secara matematis, vektor masukan:

$$x = \begin{bmatrix} x_{MQ9} \\ x_{MQ135} \\ x_{Temp} \end{bmatrix} \quad (10)$$

Dimana:

x : vektor fitur masukan

x_{MQ9} : fitur MQ-9

x_{MQ135} : fitur MQ-135

x_{Temp} : fitur suhu (*Temp C*)

$(\cdot)^T$: transpose

Transformasi pada neuron lapisan tersembunyi:

$$z_j^{(l)} = \mathbf{w}_j^{(l)} \cdot \mathbf{a}^{(l-1)} + b_j^{(l)} \quad (11)$$

$$a_j^{(l)} = \phi_{ReLU}(z_j^{(l)}) = \max(0, z_j^{(l)}) \quad (12)$$

Dimana:

l : indeks lapisan

j : indeks neuron pada lapisan l

$z_j^{(l)}$: nilai pra-aktivasi neuron j pada lapisan l

$w_j^{(l)}$: vektor bobot *neuron* j pada lapisan l

$a^{(l-1)}$: aktivasi dari lapisan sebelumnya

$b_j^{(l)}$: bias *neuron* j pada lapisan l

$\phi_{ReLU}(\cdot)$: fungsi aktivasi *ReLU*

$\max(\cdot)$: operator maksimum

Output layer menggunakan *Softmax*:

$$\text{Softmax}(z_i) = \frac{e^{z_i}}{\sum_{j=1}^K e^{z_j}}, K = 10 \quad (13)$$

Dimana:

z_i : logit untuk kelas ke- i

e : bilangan *Euler*

K : jumlah kelas (10 kelas kopi)

4) Konfigurasi Model Perbandingan SVM dan ANN

Sebagai perbandingan terhadap model utama DNN, penelitian ini juga mengimplementasikan dua model klasifikasi lain, yaitu *Support Vector Machine* (SVM) dan *Artificial Neural Network* (ANN). Penggunaan kedua model ini bertujuan untuk memberikan evaluasi komparatif terhadap pendekatan *machine learning* klasik maupun jaringan saraf dengan arsitektur yang lebih sederhana pada data aroma kopi berbentuk tabular.

Model SVM digunakan karena merupakan salah satu algoritma klasifikasi yang umum diterapkan pada data tabular dan dikenal memiliki kemampuan yang baik dalam memisahkan kelas pada ruang fitur nonlinier. Pada penelitian ini, SVM dikonfigurasi menggunakan *kernel radial basis function* (RBF) dengan parameter $C = 1.0$ dan parameter kernel $\gamma = \text{scale}$. Konfigurasi ini dipilih karena mampu membentuk batas keputusan nonlinier yang fleksibel pada data dengan jumlah fitur relatif terbatas, sekaligus tetap umum digunakan pada klasifikasi multikelas berbasis fitur numerik. Secara umum, fungsi keputusan SVM dapat dinyatakan sebagai:

$$f(x) = \text{sign} \left(\sum_{i=1}^n \alpha_i y_i K(x_i, x) + b \right) \quad (14)$$

Dengan:

$f(x)$ = hasil keputusan klasifikasi untuk data masukan x

α_i = koefisien atau bobot *support vector* ke- i

y_i = label kelas dari data latih ke- i

$K(x_i, x)$ = fungsi *kernel* antara data latih x_i dan data uji x

b = bias

n = jumlah *support vector*

Pada penelitian ini digunakan *kernel* RBF yang dirumuskan sebagai:

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|_2) \quad (15)$$

Dengan:

$K(x_i, x_j)$ = nilai kemiripan antara dua vektor fitur

γ = parameter *kernel* yang mengatur jangkauan pengaruh tiap data

$\|x_i - x_j\|_2^2$ = jarak Euclidean kuadrat antara dua sampel

Nilai $C = 1.0$ digunakan sebagai parameter penalti untuk mengontrol kompromi antara margin yang lebar dan kesalahan klasifikasi pada data latih, sedangkan $\gamma = \text{scale}$ digunakan untuk menyesuaikan parameter *kernel* secara otomatis terhadap distribusi data. Konfigurasi ini dipilih karena bersifat umum dan stabil untuk klasifikasi multikelas pada data dengan jumlah fitur yang relatif terbatas.

Sementara itu, model ANN digunakan sebagai perbandingan berbasis jaringan saraf pada skenario klasifikasi tabular yang sama. Pada penelitian ini, ANN diimplementasikan menggunakan arsitektur *multilayer perceptron* dengan tiga *hidden layer* berukuran 256, 128, dan 64 *neuron*, fungsi aktivasi *ReLU*, serta regularisasi berupa *Batch Normalization* dan *Dropout* berturut-turut sebesar 0,3, 0,3, dan 0,2. Pada lapisan keluaran digunakan fungsi aktivasi *softmax* dengan jumlah *neuron* sesuai banyaknya kelas kopi untuk menghasilkan probabilitas prediksi pada klasifikasi multikelas.

Secara matematis, transformasi pada *neuron* ANN dinyatakan sebagai:

$$\begin{aligned} z_j^{(l)} &= \sum_{i=1}^m w_{ij}^{(l)} a_i^{(l-1)} + b_j^{(l)} \\ a_j^{(l)} &= g(z_j^{(l)}) \end{aligned} \quad (16)$$

engan:

$z_j^{(l)}$ = nilai pra-aktivasi *neuron* ke- j pada lapisan ke- l

$w_{ij}^{(l)}$ = bobot dari *neuron* ke- i pada lapisan sebelumnya ke *neuron* ke- j pada lapisan ke- l

$a_i^{(l-1)}$ = keluaran *neuron* ke- i dari lapisan sebelumnya

$b_j^{(l)}$ = bias *neuron* ke- j

$g(\cdot)$ = fungsi aktivasi

m = jumlah *neuron* pada lapisan sebelumnya

Pada *hidden layer*, fungsi aktivasi yang digunakan adalah *ReLU*:

$$g(z) = \max(0, z) \quad (17)$$

sedangkan pada lapisan keluaran digunakan fungsi *softmax*:

$$\hat{y}_k = \frac{e^{z_k}}{\sum_{c=1}^K e^{z_c}} \quad (18)$$

Dengan:

\hat{y}_k = probabilitas prediksi untuk kelas ke- k

z_k = logit atau keluaran sebelum aktivasi pada kelas ke- k

K = jumlah kelas

Dengan konfigurasi ini, perbandingan antara DNN dan ANN dapat menunjukkan pengaruh kedalaman arsitektur jaringan terhadap performa klasifikasi pada data E-Nose. Ketiga model, yaitu DNN, SVM, dan ANN, diuji menggunakan skenario pra-pemrosesan yang sama, yaitu IQR, DWT, dan *Min-Max scaling*, serta pada skenario

perbandingan tanpa IQR dan DWT, sehingga perbandingan performa dilakukan pada kondisi eksperimen yang setara.

5) *Parameter Pelatihan dan Implementasi Model*

Pelatihan model DNN dan ANN dilakukan selama 100 *epoch*. Pada kedua model tersebut, optimisasi bobot dilakukan menggunakan *optimizer* Adam dengan *learning rate* sebesar 0,001. Untuk proses pelatihan jaringan, *batch size* yang digunakan adalah 16. Fungsi aktivasi pada lapisan tersembunyi menggunakan ReLU, sedangkan lapisan keluaran menggunakan *softmax* untuk klasifikasi multikelas. Pada implementasi DNN, konfigurasi pelatihan dipadukan dengan *Batch Normalization* dan *Dropout* untuk meningkatkan kestabilan proses belajar. Pada implementasi ANN, konfigurasi pelatihan juga dipadukan dengan *Batch Normalization* dan *Dropout* untuk membantu meningkatkan kestabilan proses belajar serta mengurangi risiko *overfitting*.

Dalam klasifikasi multikelas, fungsi objektif yang digunakan pada model jaringan saraf dinyatakan dengan *categorical cross-entropy*, yaitu:

$$L = - \sum_{k=1}^K y_k \log(\hat{y}_k) \tag{19}$$

Dengan:

L = nilai fungsi loss

y_k = label aktual untuk kelas ke- k

\hat{y}_k = probabilitas prediksi untuk kelas ke- k

K = jumlah kelas

Fungsi loss tersebut digunakan untuk mengukur selisih antara distribusi probabilitas hasil prediksi model dengan label kelas sebenarnya. Semakin kecil nilai loss, maka semakin dekat hasil prediksi model terhadap target yang diharapkan. Penggunaan fungsi *categorical cross-entropy* sesuai untuk kasus klasifikasi multikelas karena keluaran model dinyatakan dalam bentuk probabilitas melalui fungsi *softmax*.

Optimizer Adam digunakan karena mampu menggabungkan keunggulan *adaptive learning rate* dan momentum, sehingga proses konvergensi parameter berlangsung lebih stabil pada pelatihan jaringan saraf. Secara umum, pembaruan parameter pada Adam dapat dituliskan sebagai:

$$\theta_{t+1} = \theta_t - \eta \frac{\hat{m}_t}{\sqrt{\hat{v}_t + \epsilon}} \tag{20}$$

Dengan:

θ_t = parameter model pada iterasi ke- t

θ_{t+1} = parameter model setelah pembaruan

η = *learning rate*

\hat{m}_t = estimasi momen pertama gradien

\hat{v}_t = estimasi momen kedua gradien

ϵ = konstanta kecil untuk menjaga stabilitas numerik

Berbeda dengan DNN dan ANN yang memerlukan proses pelatihan berbasis iterasi, model SVM diimplementasikan menggunakan *kernel* RBF dengan parameter $C = 1.0$ dan $\gamma =$

scale. Model ini tidak menggunakan *epoch* maupun *batch size*, tetapi dilatih langsung pada data latih hasil pra-pemrosesan.

Seluruh model diimplementasikan pada lingkungan Google Colab menggunakan bahasa pemrograman Python. Proses pengolahan data memanfaatkan pustaka *pandas*, *NumPy*, *PyWavelets*, dan *scikit-learn*, sedangkan implementasi DNN dan ANN dilakukan menggunakan *TensorFlow/Keras*. Pada tahap evaluasi, pembagian data dilakukan dengan skema *train-test split* 80:20, di mana data latih selanjutnya dievaluasi menggunakan *K-Fold Cross Validation* dengan $k = 5$. Kinerja model diukur menggunakan *accuracy*, *precision*, *recall*, dan *F1-score*, serta divisualisasikan melalui *confusion matrix* untuk melihat pola klasifikasi pada masing-masing kelas kopi.

Untuk menjamin konsistensi eksperimen, evaluasi DNN, SVM, dan ANN dilakukan menggunakan subset data latih dan data uji yang sama pada masing-masing skenario pra-pemrosesan.

D. *Evaluasi Kinerja Model*

Evaluasi dilakukan menggunakan *confusion matrix* untuk menilai ketepatan klasifikasi multikelas dengan membandingkan hasil prediksi masing-masing model terhadap label sebenarnya pada data uji. Berdasarkan matriks tersebut dihitung metrik *accuracy*, *precision*, *recall*, dan *F1-score* untuk menggambarkan kinerja model secara keseluruhan maupun per kelas. Penggunaan metrik-metrik ini umum diterapkan pada penelitian klasifikasi berbasis E-Nose dan jaringan saraf karena mampu memberikan gambaran performa model secara ringkas dan informatif [24], [25].

Secara teoritis, metrik evaluasi yang diturunkan dari *confusion matrix* pada klasifikasi multikelas memiliki sifat invariansi tertentu dan banyak digunakan sebagai dasar analisis performa model, khususnya untuk pelaporan berbasis rata-rata antar kelas [26].

Dalam konteks aplikasi *Electronic Nose*, *confusion matrix* dan metrik turunannya juga digunakan untuk mengevaluasi kinerja sistem sensor secara komprehensif, sehingga perbandingan performa antar kelas dapat dilakukan secara objektif [27]. Struktur umum *confusion matrix* multi-kelas ditampilkan pada Tabel 3.

TABEL 3. STRUKTUR CONFUSION MATRIX MULTI-KELAS

Kelas Aktual	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	Jumlah
C1	n ₁₁	n ₁₂	Σ ₁
C2	n ₂₁	n ₂₂	...								Σ ₂
C3	n ₃₃								Σ ₃
...
Jumlah	ΣC	ΣC	ΣC	ΣC	ΣC	ΣC	ΣC	ΣC	ΣC	ΣC	ΣC
	1	2	3	4	5	6	7	8	9	10	N

Dimana:

n_{ij} : jumlah sampel kelas aktual i yang diprediksi sebagai kelas j

Σ : jumlah total pada baris atau kolom

N : total sampel evaluasi

Akurasi:

$$Accuracy = \frac{\sum_{i=1}^K n_{ii}}{N} \quad (21)$$

Keterangan:

n_{ii} : jumlah prediksi benar pada kelas ke- i

K : jumlah kelas

N : total sampel evaluasi

Precision dan *recall* per kelas ke- k :

$$Precision_k = \frac{TP_k}{TP_k + FP_k} \quad (22)$$

$$Recall_k = \frac{TP_k}{TP_k + FN_k} \quad (23)$$

Dimana:

TP_k : benar-positif untuk kelas k

FP_k : salah-positif untuk kelas k

FN_k : salah-negatif untuk kelas k

F1-score per kelas:

$$F1_k = 2 \cdot \frac{Precision_k \cdot Recall_k}{Precision_k + Recall_k} \quad (24)$$

Keterangan

$F1_k$: rata-rata harmonik antara *precision* dan *recall* untuk kelas ke- k

Untuk pelaporan multikelas, metrik juga dirata-ratakan secara makro (*macro-average*) agar setiap kelas memiliki bobot yang sama [28].

III. HASIL DAN PEMBAHASAN

A. Karakteristik Dataset dan Dampak Pra-Pemrosesan

Dataset aroma kopi diperoleh dari sistem Electronic Nose (E-Nose) berbasis ESP32 dengan sensor MQ-9, MQ-135, dan DHT22, yang menghasilkan 1.503 record dari 10 kelas kopi Arabika dan Robusta. Berdasarkan analisis awal, respons sensor, khususnya MQ-135, menunjukkan rentang nilai yang lebar dan cenderung mengandung nilai ekstrem, sehingga diperlukan pra-pemrosesan untuk meningkatkan kualitas representasi fitur sebelum klasifikasi.

Penelitian ini menggunakan dua skenario pra-pemrosesan. Skenario pertama menerapkan IQR, DWT, dan normalisasi Min-Max, sedangkan skenario kedua hanya menerapkan Min-Max sebagai pembanding. Rancangan ini digunakan untuk menilai sejauh mana pembersihan outlier dan denoising sinyal berkontribusi terhadap performa klasifikasi pada DNN, SVM, dan ANN.

B. Hasil Validasi K-Fold Cross Validation ($k = 5$)

Setelah tahap pra-pemrosesan dan pembagian data train-test 80:20, evaluasi pada data latih dilakukan

menggunakan K-Fold Cross Validation dengan $k = 5$ untuk menilai performa tiga model klasifikasi, yaitu DNN, SVM, dan ANN. Evaluasi dilakukan pada dua skenario, yaitu skenario pertama dengan pra-pemrosesan berurutan berupa IQR, DWT, dan Min-Max scaling, serta skenario kedua yang hanya menerapkan Min-Max scaling. Ringkasan hasil evaluasi ditunjukkan pada Tabel 4.

TABEL 4.
RINGKASAN HASIL EVALUASI K-FOLD CROSS VALIDATION
($K = 5$) PADA DUA SKENARIO PRA-PEMROSESAN

Skenario	DNN (Akurasi)	SVM (Akurasi)	ANN (Akurasi)
IQR + DWT + Min-Max	91,45 %	81,78 %	61,43 %
Min-Max	80,07 %	61,46 %	73,72 %

Berdasarkan Tabel 4, model DNN memberikan performa terbaik pada kedua skenario evaluasi. Pada skenario pertama, DNN mencapai akurasi 91,45%, lebih tinggi dibandingkan SVM sebesar 81,78% dan ANN sebesar 61,43%. Pada skenario kedua, DNN tetap menjadi model terbaik dengan akurasi 80,07%, diikuti ANN sebesar 73,72% dan SVM sebesar 61,46%. Hasil ini menunjukkan bahwa DNN paling efektif dalam memodelkan hubungan nonlinier antar fitur sensor pada data E-Nose berbentuk tabular.

Dari sisi pengaruh pra-pemrosesan, penerapan IQR dan DWT memberikan dampak yang berbeda pada masing-masing model. Pada DNN, akurasi meningkat dari 80,07% pada skenario kedua menjadi 91,45% pada skenario pertama. Pada SVM, peningkatan juga terlihat cukup besar, yaitu dari 61,46% menjadi 81,78%. Sebaliknya, ANN justru menunjukkan performa yang lebih baik pada skenario kedua, dengan akurasi 73,72%, dibandingkan skenario pertama yang hanya mencapai 61,43%. Temuan ini menunjukkan bahwa kombinasi IQR dan DWT sangat membantu DNN dan SVM dalam membentuk fitur yang lebih diskriminatif, tetapi pada konfigurasi ANN yang digunakan, transformasi tambahan tersebut belum memberikan keuntungan yang sama.

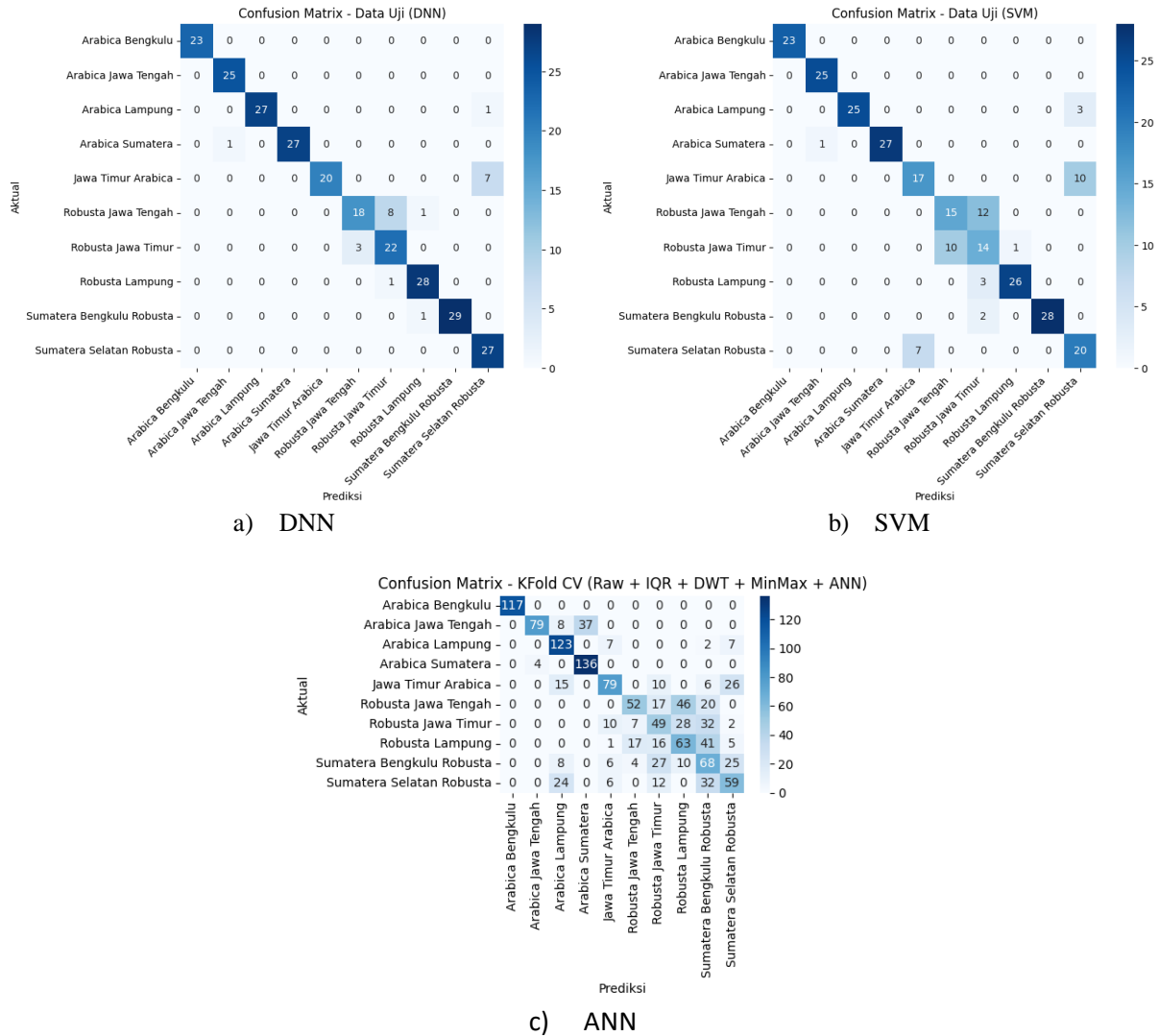
C. Hasil Pengujian dan Analisis Confusion Matrix

Evaluasi pada data uji dilakukan untuk menilai kemampuan generalisasi model terhadap data yang tidak pernah digunakan selama pelatihan maupun validasi. Pada penelitian ini, hasil klasifikasi divisualisasikan dalam bentuk confusion matrix untuk tiga model, yaitu DNN, SVM, dan ANN, pada dua skenario pra-pemrosesan. Visualisasi ini digunakan untuk melihat distribusi prediksi benar dan salah pada masing-masing kelas kopi serta untuk membandingkan pengaruh skenario pra-pemrosesan terhadap pola klasifikasi setiap model.

Berdasarkan Gambar 6, skenario pertama menunjukkan pola klasifikasi yang paling kuat pada model DNN, ditandai dengan dominasi nilai pada diagonal utama yang lebih jelas dibandingkan model lain. Hal ini menunjukkan bahwa sebagian besar sampel kopi berhasil diprediksi dengan benar

ketika data terlebih dahulu melalui pembersihan outlier, denoising sinyal, dan normalisasi. Pada model SVM, diagonal utama juga tampak cukup dominan, meskipun masih terlihat beberapa kekeliruan pada kelas-kelas tertentu, khususnya pada kelompok Robusta. Sementara itu, ANN pada skenario ini menunjukkan sebaran salah klasifikasi yang lebih luas,

terutama pada kelas-kelas dengan karakteristik aroma yang saling berdekatan, sehingga performanya lebih rendah dibandingkan DNN dan SVM. Temuan ini konsisten dengan hasil pada Tabel 4, bahwa skenario pertama paling efektif untuk DNN dan SVM, tetapi belum memberikan keuntungan yang sama pada ANN.

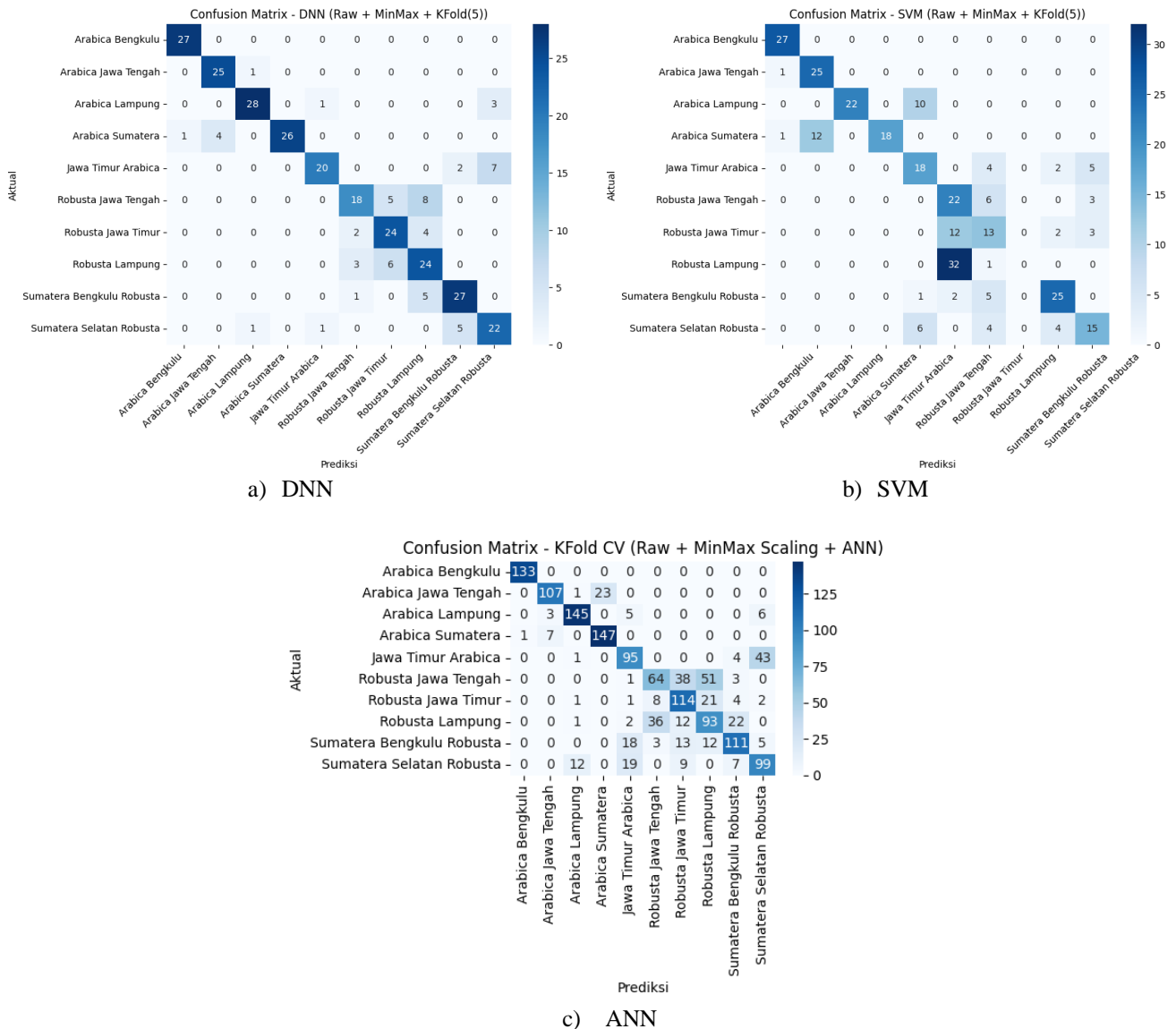


Gambar 6 Confusion matrix pada skenario 1

(data mentah + IQR + DWT + Min–Max + k = 5): (a) DNN, (b) SVM, (c) ANN.

Berdasarkan Gambar 7, pola confusion matrix pada skenario kedua menunjukkan hasil yang berbeda antar model. Pada DNN, diagonal utama masih terlihat dominan, tetapi jumlah salah klasifikasi meningkat dibandingkan skenario pertama, yang menunjukkan bahwa normalisasi Min–Max saja belum mampu menghasilkan representasi fitur sebaik ketika IQR dan DWT juga diterapkan. Pada SVM, penurunan kualitas klasifikasi terlihat lebih jelas, terutama pada beberapa kelas Arabika dan Robusta yang

saling tertukar, sehingga pola diagonal utama menjadi kurang kuat dibandingkan skenario pertama. Sebaliknya, ANN pada skenario kedua menunjukkan pola klasifikasi yang lebih baik dibandingkan ANN pada skenario pertama. Hal ini tampak dari diagonal utama yang lebih dominan dan sebaran salah klasifikasi yang relatif lebih terkendali, sehingga mendukung hasil evaluasi bahwa ANN justru memperoleh performa lebih baik pada skenario kedua.



Gambar 7. Confusion matrix pada skenario 2 (data mentah + Min-Max + k = 5): (a) DNN, (b) SVM, (c) ANN.

Jika dibandingkan secara langsung, confusion matrix pada kedua skenario memperlihatkan bahwa pengaruh pra-pemrosesan tidak identik pada semua model. DNN dan SVM menunjukkan pola diagonal utama yang lebih kuat pada skenario pertama, yang menandakan bahwa IQR dan DWT membantu memperjelas pemisahan antar kelas. Sebaliknya, ANN menunjukkan hasil yang lebih baik pada skenario kedua, sehingga mengindikasikan bahwa fitur hasil transformasi tambahan pada skenario pertama belum tentu paling sesuai untuk konfigurasi ANN yang digunakan. Dengan demikian, confusion matrix tidak hanya memperkuat hasil kuantitatif pada Tabel 4, tetapi juga menunjukkan bahwa efektivitas tahapan pra-pemrosesan sangat bergantung pada karakteristik model klasifikasi.

D. Analisis Hasil dan Error Analysis

Hasil evaluasi menunjukkan bahwa kombinasi model klasifikasi dan skenario pra-pemrosesan memberikan pengaruh nyata terhadap performa klasifikasi aroma kopi multikelas. Temuan utama penelitian ini adalah bahwa DNN merupakan model yang paling unggul dan paling konsisten pada kedua skenario. Pada skenario pertama, DNN mencapai akurasi 91,45%, sedangkan pada skenario kedua nilainya menurun menjadi 80,07%. Perbedaan ini menunjukkan bahwa DNN memperoleh manfaat yang besar dari pembersihan outlier dan denoising sinyal sebelum proses normalisasi.

SVM juga menunjukkan pola yang serupa, bahkan dengan peningkatan yang relatif besar dari 61,46% pada skenario

kedua menjadi 81,78% pada skenario pertama. Hal ini mengindikasikan bahwa model berbasis margin sangat sensitif terhadap kualitas separabilitas fitur. Ketika outlier telah dikurangi dan sinyal telah dibuat lebih halus melalui DWT, batas pemisah antar kelas menjadi lebih jelas sehingga SVM dapat bekerja lebih efektif. Dengan kata lain, skenario pertama memberikan representasi fitur yang lebih mendukung proses klasifikasi pada SVM.

Sebaliknya, ANN menunjukkan kecenderungan yang berbeda. Akurasi ANN pada skenario kedua mencapai 73,72%, lebih tinggi dibandingkan 61,43% pada skenario pertama. Hasil ini menunjukkan bahwa pra-pemrosesan tambahan berupa IQR dan DWT tidak selalu meningkatkan performa semua model. Pada konfigurasi ANN yang digunakan, transformasi fitur setelah IQR dan DWT diduga belum memberikan representasi yang paling sesuai untuk proses pembelajaran, sehingga performanya justru lebih rendah dibandingkan ketika hanya menggunakan normalisasi Min–Max.

Dari perspektif klasifikasi multikelas, kesalahan prediksi masih cenderung muncul pada kelas-kelas kopi yang memiliki respons sensor saling berdekatan, terutama pada beberapa kelas Robusta dan sebagian kelas Arabika dengan pola volatil yang mirip. Pola ini menunjukkan bahwa salah klasifikasi tidak terjadi secara acak, melainkan dipengaruhi oleh tumpang tindih karakteristik aroma yang tertangkap oleh sensor. Dalam konteks tersebut, DNN menunjukkan kemampuan terbaik untuk membentuk pemisahan nonlinier antar kelas, SVM sangat bergantung pada kualitas fitur hasil pra-pemrosesan, sedangkan ANN pada konfigurasi ini lebih sesuai pada fitur yang lebih sederhana.

Secara keseluruhan, hasil penelitian ini menegaskan bahwa skenario pertama, yaitu data mentah + IQR + DWT + Min–Max + K-Fold ($k = 5$), merupakan skenario terbaik untuk DNN dan SVM, sedangkan skenario kedua, yaitu data mentah + Min–Max + K-Fold ($k = 5$), memberikan hasil lebih baik pada ANN. Di antara ketiga model yang dibandingkan, DNN tetap menjadi model dengan performa terbaik dan paling konsisten untuk klasifikasi aroma kopi multikelas berbasis data E-Nose tabular.

IV. KESIMPULAN

Penelitian ini menunjukkan bahwa klasifikasi aroma kopi multikelas berbasis Electronic Nose dapat dilakukan secara objektif menggunakan data sensor MQ-9, MQ-135, dan DHT22 melalui evaluasi komparatif DNN, SVM, dan ANN pada dua skenario pra-pemrosesan. Skenario pertama, yaitu data mentah + IQR + DWT + Min–Max, memberikan hasil terbaik untuk DNN dan SVM, masing-masing dengan akurasi 91,45% dan 81,78%, sedangkan skenario kedua, yaitu data mentah + Min–Max, lebih sesuai untuk ANN dengan akurasi 73,72%. Secara keseluruhan, DNN merupakan model dengan performa terbaik dan paling konsisten, sehingga pendekatan ini berpotensi menjadi metode pendukung yang lebih objektif dan andal untuk identifikasi aroma kopi multikelas.

DAFTAR PUSTAKA

- [1] K. Agropolitan and K. Kudus, "Pengembangan Potensi Kopi Sebagai Komoditas Unggulan Kawasan Agropolitan Kabupaten Kudus," vol. 21, 2024.
- [2] F. Amalia *et al.*, "Identification of potential quality markers in Indonesia's Arabica specialty coffee using GC/MS-based metabolomics approach," *Metabolomics*, vol. 19, no. 11, pp. 1–11, 2023, doi: 10.1007/s11306-023-02051-5.
- [3] C. Gonzalez Viejo, E. Tongson, and S. Fuentes, "Integrating a low-cost electronic nose and machine learning modelling to assess coffee aroma profile and intensity," *Sensors*, vol. 21, no. 6, pp. 1–16, 2021, doi: 10.3390/s21062016.
- [4] C. H. Lee, "An AI-powered Electronic Nose System with Fingerprint Extraction for Aroma Recognition of Coffee Beans," *Sensors*, 2022.
- [5] M. Telaumbanua, D. D. Novita, S. Triyono, and C. Saragih, "Tipe Chamber Dan Posisi Sensor E-Nose Untuk Mendeteksi Aroma Biji Kopi Robusta Menggunakan Mikrokontroler," *J. Ilm. Rekayasa Pertanian dan Biosist.*, vol. 9, no. 1, pp. 84–95, 2021, doi: 10.29303/jrpb.v9i1.237.
- [6] D. P. Purbawa *et al.*, "Adaptive filter for detection outlier data on electronic nose signal," *Sens. Bio-Sensing Res.*, vol. 36, no. April, p. 100492, 2022, doi: 10.1016/j.sbsr.2022.100492.
- [7] A. S. AlSalehy and M. Bailey, "Improving Time Series Data Quality: Identifying Outliers and Handling Missing Values in a Multilocation Gas and Weather Dataset," *Smart Cities*, vol. 8, no. 3, pp. 1–39, 2025, doi: 10.3390/smartcities8030082.
- [8] F. G. Handayani, A. Rajaguguk, I. H. Rosma, and E. Ervianto, "Comparative Analysis Of Mother Wavelet For Voltage Sag And Swell Classification Using Discrete Wavelet Transform (Dwt) And Radial Basis Function Neural Network (Rbfn) Analisis Komparatif Mother Wavelet Untuk Klasifikasi Gangguan Voltage Sag Dan Swell," vol. 25, no. 2, pp. 132–140, 2025.
- [9] M. A. Barata *et al.*, "Hydrogen Sulfide Leak Detection Using The C4.5 Algorithm: Optimizing Feature Extraction For Enhanced Accuracy Mula," vol. 1, pp. 348–358, 2024.
- [10] I. Wisma, D. Prastya, and R. Sarno, "Coffee Aroma Classification Based On World Coffee Research Sensory Lexicon Using Electronic Nose," no. March, pp. 19–23, 2025, doi: 10.31602/v0i0.15389.
- [11] A. Syafi'i, M. A. Barata, and R. Rohmah, "Artificial Neural Network Algorithm on e-Nose Devices for Honey Classification," *J. Telemat.*, vol. 20, no. 1, pp. 28–48, 2025, doi: 10.61769/telematika.v20i1.722.
- [12] E. Rianty and K. Budi, "Perbandingan Kinerja Algoritma C4.5 dan Naive Bayes Dalam Klasifikasi Data Penjualan Buku PT. XYZ," vol. 12, no. 6, 2025, doi: 10.30865/jurikom.v12i6.9345.
- [13] Malikhah *et al.*, "Detection of Infectious Respiratory Disease Through Sweat From Axillary Using an E-Nose With Stacked Deep Neural Network," *IEEE Access*, vol. 10, pp. 51285–51298, 2022, doi: 10.1109/ACCESS.2022.3173736.
- [14] D. Erwanto *et al.*, "Development and Application of an Electronic Nose System for Classifying Coffee Varieties Based on Aromatic Profiles," *J. Intell. Syst. Control*, vol. 3, no. 3, pp. 186–200, 2024, doi: 10.56578/jisc030305.
- [15] S. D. Astuti *et al.*, "Electronic nose coupled with artificial neural network for classifying of coffee roasting profile," *Sens. Bio-Sensing Res.*, vol. 43, no. February, p. 100632, 2024, doi: 10.1016/j.sbsr.2024.100632.
- [16] V. Borisov, T. Leemann, K. Sebler, J. Haug, M. Pawelczyk, and G. Kasneci, "Deep Neural Networks and Tabular Data : A Survey," no. June, pp. 1–22, 2022.
- [17] F. Hasehzadeh, "E-nose equipped with Artificial Intelligence Technology for diagnosis of dairy cattle disease in veterinary," 2021.
- [18] D. B. Magfira, F. Yudianto, T. D. Wulan, T. Herlambang, R. P. N. Budiarti, and A. T. R. Siswanti, "Perancangan IoT Sederhana Untuk Sistem Pendeteksi Kemurnian Kopi Bubuk," *J. Tecnoscienza*, vol. 9, no. 1, pp. 86–96, 2024, doi:

- 10.51158/cykvay82.
- [19] B. Sumanto, D. R. Java, W. Wijaya, and J. Hendry, "Seleksi Fitur Terhadap Performa Kinerja Sistem E-Nose untuk Klasifikasi Aroma Kopi Gayo," *MATRIK J. Manajemen, Tek. Inform. dan Rekayasa Komput.*, vol. 21, no. 2, pp. 429–438, 2022, doi: 10.30812/matrik.v21i2.1495.
- [20] I. S. Nasution, D. P. Delima, Z. Zaidiyah, and R. Fadhil, "Mathematical Modelling of Engineering Problems A Low Cost Electronic Nose System for Classification of Gayo Arabica Coffee Roasting Levels Using Stepwise Linear Discriminant and K-Nearest Neighbor," vol. 9, no. 5, pp. 1271–1276, 2022.
- [21] N. Dzakwan, S. Gandhi, M. Y. Hariyawan, and H. Briantoro, "Stasiun Cuaca Mini Pintar Berbasis Deep Learning," vol. 12, no. 5, pp. 8026–8033, 2025.
- [22] M. J. Vikri, I. W. D. Prastya, U. P. Sanjaya, and M. A. Barata, "Rice Quality Identification for Indonesian Food Standards Based on Electronic Nose Identifikasi Kualitas Beras Berdasarkan Standar Pangan Indonesia Berbasis Electronic Nose," no. 1, 2025.
- [23] P. P. Allorerung, A. Erna, M. Bagussahrir, and S. Alam, "Analisis Performa Normalisasi Data untuk Klasifikasi K-Nearest Neighbor pada Dataset Penyakit," *JISKA (Jurnal Inform. Sunan Kalijaga)*, vol. 9, no. 3, pp. 178–191, 2024, doi: 10.14421/jiska.2024.9.3.178-191.
- [24] M. A. Barata, E. Noersasongko, Purwanto, and M. A. Soeleman, "Improving the Accuracy of C4.5 Algorithm with Chi-Square Method on Pure Tea Classification Using Electronic Nose," vol. 5, no. 158, pp. 226–235, 2026.
- [25] F. I. Silfana and M. A. Barata, "Using K-NN Algorithm for Evaluating Feature Selection on High," vol. 17, no. 2, 2024.
- [26] G. Zeng, "Invariance Properties and Evaluation Metrics Derived from the Confusion Matrix in Multiclass Classification," *Mathematics*, vol. 13, no. 16, 2025, doi: 10.3390/math13162609.
- [27] Z. Peng *et al.*, "A Comprehensive Evaluation Model for Optimizing the Sensor Array of Electronic Nose," *Appl. Sci.*, vol. 13, no. 4, 2023, doi: 10.3390/app13042338.
- [28] J. Tampubolon, "Comparative performance of LSTM and DNN in sentiment analysis," vol. 14, no. 1, pp. 1–11, 2025, [Online]. Available: www.ejournal.isha.or.id/index.php/Mandiri