

# PERANCANGAN SISTEM *ELECTRONIC NOSE* BERBASIS MIKROKONTROLLER SEBAGAI ALAT PENGKLASIFIKASI JENIS TEH MURNI

Mula Agung Barata<sup>1</sup>, Muhammad Jauhar Vikri<sup>2</sup>, Teguh Pribadi<sup>3</sup>

<sup>1,2,3</sup> Program Studi Teknik Informatika, Fakultas Sains dan Teknologi, Universitas Nahdlatul Ulama Sunan Giri  
<sup>1</sup>mula.ab26@gmail.com, <sup>2</sup>vikri@unugiri.ac.id, <sup>3</sup>pribadi.teguh90@gmail.com

## Abstrak

Klasifikasi jenis teh murni secara manual sering kali memerlukan waktu yang lama dan bergantung pada kemampuan indera manusia, karena kemampuan ketajaman indera penciuman manusia yang berbeda penciuman anggapan terhadap obyek bersifat subyektif dan rentan terhadap kesalahan. Penelitian ini bertujuan untuk merancang sebuah sistem *electronic nose* berbasis mikrokontroler yang mampu mengklasifikasikan jenis teh murni secara otomatis dan konsisten. Sistem yang dikembangkan terdiri dari perangkat keras, termasuk mikrokontroler dan sejumlah sensor gas untuk mendeteksi senyawa volatil yang menjadi ciri khas setiap jenis teh. Data dari sensor diolah menggunakan algoritma machine learning untuk menghasilkan model klasifikasi yang akurat. Proses pengembangan melibatkan pengumpulan data aroma dari lima jenis teh murni yang diuji dalam berbagai kondisi lingkungan untuk memastikan robustitas sistem. Data yang diperoleh kemudian dianalisis dan diolah menggunakan algoritma supervised learning, yaitu algoritma *Decision Tree*. Sistem prototipe yang dihasilkan mampu mencapai akurasi klasifikasi sebesar 93,7%, menunjukkan keandalannya dalam mengenali pola aroma khas dari setiap jenis teh. Selain membahas hasil, penelitian ini juga mengidentifikasi tantangan seperti pengaruh variasi suhu dan kelembaban terhadap performa sensor, serta kebutuhan kalibrasi berkala untuk menjaga konsistensi sistem. Dengan hasil yang menjanjikan, sistem ini menawarkan solusi inovatif untuk mendukung industri teh dalam mengotomatisasi proses pengklasifikasian produk secara lebih efisien dan objektif.

**Kata kunci** : *Decision Tree, Electronic nose, E-nose, Teh murni*

## 1. Pendahuluan

Teh adalah salah satu hasil perkebunan di lingkungan Kementerian Pertanian Republik Indonesia yang berperan penting sebagai komoditas andalan yang mendorong ekonomi Indonesia (Susanti & Akbar, 2019). Pada tahun 2018, tercatat di Badan Pusat Statistik bahwa teh dapat mendorong devisa ekspor sebesar 1,5% dari PDB sektor pertanian atau sebesar 108,5 juta USD. Hal tersebut merupakan penghasil bahan baku sebagai bentuk pengembangan agroindustri sekaligus sumber pendapatan pertanian (*Statistik Teh Indonesia*, 2018). Selain menjadi salah satu komoditas penting dalam perekonomian Indonesia, teh juga merupakan produk nasional yang diekspor ke luar negeri mengingat Indonesia tercatat di urutan ketujuh sebagai negara penghasil teh dunia. Setiap jenis teh memiliki khasiat yang berbeda dan kualitas teh dapat diukur dari aroma masing-masing aromanya. Saat ini banyak jenis teh yang dimodifikasi oleh beberapa pelaku bisnis dengan menambahkan aroma pada teh sehingga kemurnian dari teh berkurang begitu juga khasiat yang dikandung. Indra penciuman manusia dirasa masih sangat terbatas dalam mengklasifikasikan jenis teh murni, sehingga masih sulit untuk membedakan mana jenis teh hitam, hijau, oolong ataupun teh putih (Putra et al., 2016). Oleh karena itu, dibutuhkan metode

untuk membantu dalam menangkap aroma teh dengan suatu alat instrumentasi berupa *electronic nose*. Alat tersebut membantu manusia dalam mengambil data dari aroma teh murni yang kemudian akan diklasifikasikan dengan algoritma data mining.

Klasifikasi merupakan salah satu algoritma data mining yang memiliki konsep pengelompokan suatu data ke dalam kriteria tertentu dengan membaca data yang sebelumnya sudah ada. Karakteristik dari algoritma klasifikasi adalah sudah adanya proses belajar berdasarkan nilai dari variabel prediktor, karena algoritma ini tergolong dalam supervised learning (Abena, P., Hadjinian, P., Stadler, R., Verhees, J., & Zanasi, 1998).

Penelitian terkait dengan pemanfaatan perangkat *electronic nose* dan teh pernah dilakukan oleh Runu Banerjee, dkk dengan topik penggunaan *electronic nose* dalam pengukuran kualitas teh hitam (Banerjee et al., 2016). Dataset teh hitam yang digunakan dalam penelitian ini diambil menggunakan perangkat *electronic nose*. Dataset diproses dengan algoritma Principal Component Analysis dan Neural Network dengan hasil 79,43% yang jelas dalam menetapkan ketepatan perangkat *electronic nose* dalam pengambilan data teh hitam.

Penelitian lain yang terkait dengan teh dan perangkat *electronic nose* dilakukan oleh Ozil Afindra Putra, dkk dengan topik pemanfaatan e-nose

untuk mengidentifikasi aroma teh dengan metode backpropagation (Putra et al., 2016). Pada penelitian ini menggunakan dataset teh dengan kelas black tea, green tea, white tea, jasmine tea dan oolong tea. Dataset yang diambil oleh perangkat E-Nose kemudian diproses menggunakan metode Backpropagation dengan hasil akurasi rata-rata sebesar 87,5%.

Penelitian lainnya dengan topik implementasi electronic nose dalam evaluasi kualitas teh yang dikombinasikan dengan metode kemometrik oleh Min X, Jun Wang dan Luyi Zhu (Xu et al., 2021). Pada penelitian ini peneliti menggunakan dataset yang diambil dari penerapan perangkat electronic nose. Hasil yang didapatkan dari penelitian ini memuaskan dengan akurasi 100% dari kinerja algoritma Logistic Regression dan Support Vector Machine dengan optimasi metode reduksi data antara lain Principal Component Analysis (PCA), Multidimensional Scalling (MDS) dan linear discriminant analysis (LDA).

Penelitian selanjutnya mengenai teh dengan topik klasifikasi teh komersial menurut asal dan jenisnya menggunakan kandungan unsur dengan spektroskopi fluoresensi sinar-x (xrf) yang dilakukan oleh Cia Min Lim, dkk (Lim et al., 2021). Hasil dari penelitian yang dilakukan bahwa tingkat akurasi tertinggi dalam klasifikasi teh ini adalah metode Partial Least Squares-Discriminant Analysis dengan akurasi sebesar 85%. Sedangkan metode Principal Component Analysis (PCA) dan Soft Independent Modelling of Class Analogies (SIMCA) dan mendapatkan nilai akurasi di bawahnya.

Penerapan electronic nose dalam suatu penelitian pada tahun 2019 yang digunakan dalam mengklasifikasikan dataset kopi luwak dan bukan kopi luwak dilakukan oleh Sulaiman Wakhid, dkk(Wakhid et al., 2020). Penelitian ini menggunakan algoritma *Decision Tree* dan parameter statistik standar deviasi dan menghasilkan akurasi prediksi sebesar 97%.

Pemanfaatan electronic nose untuk mendapatkan dataset daging sapi untuk bahan dalam pengujian algoritma machine learning (Wijaya et al., 2018). Penelitian ini dilakukan oleh Dedy Rahman Wijaya pada tahun 2018. Harapan dalam penelitian selanjutnya adalah penyempurnaan dalam menangani noise data pada dataset yang dihasilkan oleh perangkat electronic nose.

Penerapan algoritma deep neural network untuk menguji daging ayam segar dan tidak segar menghasilkan akurasi pengujian 98,70% dalam penelitian ini. Penelitian ini dilakukan oleh Achmad Ilham Fanany Al Isyrafie, dkk pada tahun 2022 dengan dataset daging ayam segar dan tidak segar (Al Isyrafie et al., 2022).

Selain pemanfaatan dalam pengambilan dataset teh dan kopi, perangkat electronic nose juga dimanfaatkan dalam pengujian mutu beras aromatik(Wijayanti et al., 2013). Penelitian yang

dilakukan oleh Sari Wijayanti, dkk pada tahun 2013 menghasilkan akurasi 84% dengan metode Principal Component Analys dan metode Support Vector Machine.

Penelitian lain yang berhubungan dengan algoritma C4.5 sebagai metode klasifikasi yang diusulkan pada penelitian ini pernah dilakukan oleh Pratik Gite, dkk dengan topik penelitian skema deteksi intrusi berbasis machine learning untuk berbagai jenis serangan di wireless sensor network menggunakan algoritma klasifikasi C4.5 dan CART(Gite et al., 2021). Penelitian ini menggunakan dataset serangan dalam jaringan yang dideteksi oleh WSN. Hasil penelitian ini menunjukkan bahwa akurasi dari algoritma C4.5 lebih tinggi dari CART ketika ukuran sampel pelatihan dan pengujian terus meningkat, kedua algoritma tersebut memiliki akurasi rata-rata antara 63% dan 68%.

Penelitian lain mengenai algoritma C4.5 dengan topik analisis kerentanan longsor menggunakan Random Forest, C4.5 dan C5.0 dengan dataset seimbang dan tidak seimbang yang dilakukan oleh Burak F. Tanyu, dkk (Tanyu et al., 2021). Hasil penelitian ini menyimpulkan bahwa algoritma Random Forest memiliki tingkat keefektifan paling rendah dan eksperimen tertinggi dari nilai akurasi dihasilkan oleh C4.5 dengan prosentase 99,8% dan nilai kappa 0,93 untuk data raster, sedangkan nilai akurasi dari data vector dengan prosentase 96% dan nilai kappa 0,92.

Penelitian selanjutnya dengan topik konstruksi Decision Tree berdasarkan algoritma C4.5 untuk penilaian stabilitas tegangan online oleh Xiangfei Menga, dkk(Meng et al., 2020). Penelitian ini menggunakan dataset penilaian stabilitas tegangan listrik online. Proses dalam penanganan dataset sebelum masuk pada model Decision Tree dan C4.5 terlebih dahulu melalui proses akuisisi sampel, seleksi atribut dan konstruksi Decision Tree. Penelitian ini memaparkan bahwa algoritma C4.5 lebih mampu menangani dataset dengan atribut kontinu dalam penelitian ini adalah dataset penilaian stabilitas tegangan listrik online.

Penelitian dengan topik optimasi Ahybrid Grey Wolf dan Particle Swarm Optimization dengan pendekatan C4.5 untuk prediksi Rheumatoid Arthritis dilakukan oleh Shanmugam Sundaramurthy dan Preethi Jayavel(Sundaramurthy & Jayavel, 2020). Dataset penyakit Rheumathoid Arthritis menjadi objek dalam penelitian ini. Hasil dari penelitian menyebutkan bahwa Grey Wolf Optimization dapat meningkatkan akurasi prediksi dari C4.5 sebesar 86,36% di mana prosentase tersebut adalah prosentase paling tinggi dibandingkan dengan akurasi algoritma C4.5 secara individu.

Penelitian dengan topik teh electronic nose sudah banyak ditemukan di beberapa tahun ini. Namun, penelitian yang mengerucut dengan topik klasifikasi jenis teh murni yang memanfaatkan penerapan electronic nose dengan pemrosesan algoritma

machine learning belum banyak ditemukan. Hal itu menjadi suatu pemikiran peneliti untuk mengembangkan penelitian dengan topik penerapan electronic nose untuk klasifikasi jenis teh murni dengan pemrosesan algoritma machine learning.

Berdasarkan latar belakang tersebut, maka peneliti menggunakan algoritma C4.5 mendapatkan akurasi tinggi dalam pengklasifikasian jenis teh murni.

**2. Metode**

**2.1 Data Mining**

Menurut Cabena P et.al. data mining merupakan disiplin ilmu yang memadukan teknik dari kelimuan database, statistic, pattern recognition, machine learning dan visualisasi dalam memecahkan masalah ekstraksi informasi dari basisdata berukuran besar (abena, P., Hadjinian, P., Stadler, R., Verhees, J., & Zanasi, 1998). Penjabaran mengenai data mining menurut Larose berdasarkan tugasnya dibagi menjadi beberapa kelompok, antara lain:

a. Estimation

Metode estimasi merupakan metode data mining yang mampu menyelesaikan permasalahan yang berkaitan dengan perkiraan hasil produksi berupa data time series dari beberapa tahun yang telah berlalu kemudian diolah sehingga mendapatkan hasil berupa persentase. Adapun algoritma data mining yang digunakan dalam *estimation*, yaitu *Support Vector Machine*, *Neural Network* dan *Linear Regression*.

b. Prediction

Metode prediksi atau bisa disebut dengan forecasting adalah metode data mining yang digunakan dalam melakukan prediksi nilai/hasil di masa depan. Adapun algoritma *data mining* yang digunakan dalam prediction, yaitu *Support Vector Machine*, *Neural Network* dan *Linear Regression*.

c. Classification

Metode klasifikasi adalah salah satu metode dalam mengelompokkan data sesuai dengan ciri-ciri atau karakteristik data tersebut. Adapun algoritma *data mining* yang digunakan dalam klasifikasi, yaitu *Logistic Regression*, *K-Nearest Neighbor*, *Decision Tree*, *Naive Bayes*.

d. Clustering

Metode *clustering* adalah metode yang digunakan untuk mengelompokkan data ke dalam satu cluster yang memiliki tingkat kemiripan nilai maksimum dan data antar cluster memiliki tingkat kemiripan yang minimum. Adapun algoritma data mining yang digunakan dalam klastering, yaitu *Fuzzy C-Means*, *K-Means*, *K-Medoids* dan *Self-Organizing Map (SOM)*.

e. Association

Metode asosiasi dalam data mining memiliki peran untuk mengukur hubungan antara 2 atau lebih atribut suatu data. Asosiasi banyak diterapkan pada dunia bisnis dengan analisis yang terkenal bernama afinitas atau cart analysis pasar. Adapun algoritma

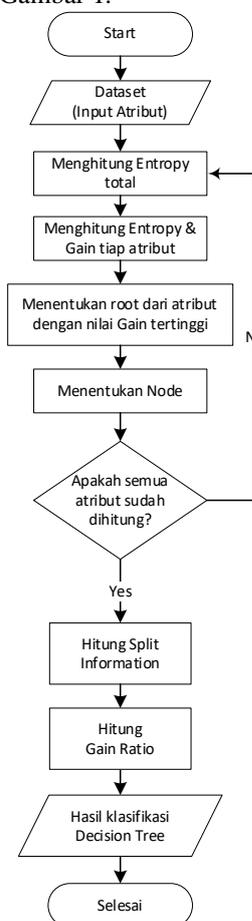
data mining yang digunakan dalam association, yaitu *FP-Growth*, *A-Priori* dan *Coefficient of Correlation*.

**2.2 Algoritma C4.5**

Algoritma C4.5 merupakan pengembangan dari algoritma ID3 yang masih satu rumpun dengan *Decision Tree*. C4.5 merupakan model klasifikasi yang berbentuk seperti pohon terbalik di mana C4.5 mudah dimengerti oleh orang awam. C4.5 dapat bekerja pada atribut bertipe kategorikal atau numerik, mengatasi *missing value*, melakukan proses *pruning* dan kinerjanya lebih baik dibandingkan metode pendahulunya seperti *CART* dan *ID3*. Proses atau tahapan *Decision Tree* dibagi menjadi 3 tahap yaitu:

- a. Pembentukan pohon keputusan,
- b. Pruning,
- c. Mengekstrak aturan dari pohon keputusan yang terbentuk.

Ekstraksi hasil model *Decision Tree* digambarkan seperti pohon yang terbalik yang mana sebuah pohon memiliki cabang, akar, dan daun atau *node*. Cabang menggambarkan keputusan sebagai pertanyaan klasifikasi, daun menggambarkan kelas atau label, dan akar menggambarkan *node* pertama atau *node* awal di atas (abena, P., Hadjinian, P., Stadler, R., Verhees, J., & Zanasi, 1998). Adapun langkah-langkah dalam perhitungan algoritma C4.5 dapat dilihat pada Gambar 1.



Gambar 1. Langkah Perhitungan Algoritma C4.5

Setelah memahami bagaimana tahapan yang harus dilakukan dalam melakukan perhitungan algoritma C4.5 maka dapat diuraikan bagaimana formula yang dirumuskan dalam perhitungan algoritma C4.5 ini dalam menguji suatu dataset. Adapun formula dalam perhitungan nilai Entropy ditunjukkan pada formula 1.

$$Entropy(S_1, S_2, \dots, S_n) = \sum_{i=1}^n P_i * \log_2(P_i) \quad (1)$$

Dimana, S = Himpunan kasus  
 n = Jumlah sampel  
 Pi = Proporsi kelas

Sedangkan formula 2 digunakan dalam menghitung nilai Gain.

$$Gain(A) = Entropy(S) - \sum_{i=1}^n \left( \frac{|S_i|}{|S|} \right) * Entropy(S_i) \quad (2)$$

Dimana, S = Himpunan kasus  
 A = Atribut  
 n = Jumlah sampel  
 |Si| = Jumlah Kasus pada partisi ke-i  
 |S| = Jumlah kasus dalam S

Kemudian untuk menghitung nilai Split Info digunakan formula seperti ditunjukkan pada formula 3.

$$Split\ Info(S, A) = \sum_{i=1}^0 \left( \frac{|S_i|}{|S|} \right) * \log_2 \left( \frac{|S_i|}{|S|} \right) \quad (3)$$

Dimana, S = Himpunan kasus  
 A = Atribut  
 |Si| = Jumlah Kasus pada partisi ke-i

Setelah nilai Gain dan Split Info ditemukan selanjutnya menghitung nilai Gain Ratio dengan formula 4.

$$Gain\ Ratio(S, A) = \frac{Gain\ Ratio(S, A)}{Split\ Information(S, A)} \quad (4)$$

### 2.3 Produksi Teh Indonesia

Teh adalah hasil olahan dari daun tanaman *camellia sinensis* yang dapat tumbuh secara subur dan ideal di daerah pegunungan dan memiliki kelembaban konstan. Daerah yang beriklim tropis dan subtropis di beberapa dataran tinggi Asia merupakan tempat yang cocok untuk memproduksi teh berkualitas tinggi. Teh banyak dimanfaatkan utamanya sebagai minuman, selain minuman teh juga dimanfaatkan sebagai pengawet dan obat seperti memperhalus kulit, melangsingkan tubuh, menekan kolesterol dalam darah, memperlancar sirkulasi darah, mencegah kanker dan penyakit jantung, perbaikan sel yang rusak dan sebagai antioksidan. Kekayaan manfaat dari teh tersebut didapatkan dari hasil bahan baku teh, proses olahan yang berbeda sehingga jenis teh yang dihasilkan juga berbeda. Jenis teh menurut cara pengolahannya diklasifikasikan menjadi 4 jenis yaitu *black tea*, *green tea*, *oolong tea* dan *white tea*

(Wijayanti et al., 2013). Ada beberapa pendapat dari *paper* yang telah dibaca oleh peneliti mengenai jenis teh murni

Menurut penuturan Bapak Ir. Purwanto Wahyu Priyono selaku Direktur Utama PT. Candi Loka bahwa 4 jenis teh murni ini berbahan dasar sama yaitu daun *camellia sinensis*. Keempat jenis teh murni tersebut menjalani proses pengolahan berbeda, sehingga jenis teh yang dihasilkan juga berbeda serta khasiat dari masing-masing jenisnya dalam penyembuhan penyakit. Di bawah ini merupakan uraian dari proses pengolahan dan khasiat dari masing-masing 4 jenis teh murni tersebut. antara lain:  
 a. *Black Tea*

*Black Tea* merupakan hasil fermentasi daun *camellia sinensis* dengan kafein tertinggi dalam kandungannya sehingga memiliki rasa dasar yang khas. Menurut penelitian yang sudah dilakukan bahwa teh hitam memiliki khasiat untuk menekan risiko terjadinya penyakit stroke, kandungan antioksidan yang dapat menekan kadar kolesterol, mencegah penyakit paru-paru dari akibat asap rokok. Gambar 2 memberikan ilustrasi dari the hitam.



Gambar 2. *Black Tea*

#### b. *Green Tea*

Pengukusan daun *camellia cinensis* menghasilkan teh hijau yang berkhasiat untuk menekan risiko stroke, menekan pertumbuhan kanker, menekan penyumbatan pembuluh darah, meningkatkan pembakaran lemak, menetralkan jumlah radikal bebas dalam jumlah besar dan meminimalisir gangguan saraf. Semua khasiat tersebut didapatkan karena kandungan antioksidan dalam teh hijau. Menurut penelitian, penyajian secangkir teh hijau setiap hari dapat dengan ukuran ± 200ml dapat menekan risiko penyakit jantung sebesar 10%. Gambar 3 menunjukkan ilustrasi dari the hijau.



Gambar 3. *Green Tea*

#### c. *Oolong Tea*

Proses fermentasi yang terjadi pada teh oolong berbeda dengan teh hitam yang mana teh oolong melalui proses fermentasi sebagian. Teh oolong dibuat dari bahan dasar daun, tunas dan batang

tanaman *camellia sinensis*. Ilustrasi dari tampilan the oolong ditunjukkan pada Gambar 4.



Gambar 4. Oolong Tea

d. *White Tea*

Jenis teh yang paling berbeda dari jenis teh yang berbeda adalah jenis teh putih, seperti diilustrasikan pada Gambar 5. Proses pengolahannya tidak melalui fermentasi dan tidak diawetkan. Penelitian yang telah dilakukan menunjukkan bahwa anti kanker paling tinggi dimiliki oleh jenis teh putih karena jenis teh lain lebih banyak melalui proses pengolahan. Polifenol yang tinggi terkandung di dalam teh putih dan memiliki sifat antibakteri, anti jamur dan antivirus. Bahan yang digunakan dalam pembuatan teh putih sangat berbeda dari jenis teh yang lain. Jenis teh lain dibuat dari bahan dasar daun, tunas dan batang tumbuhan *camellia sinensis*, sedangkan teh putih berbahan dasar dari kuncup tumbuhan *camellian sinensis* yang belum mengembang, sehingga harga di pasaran untuk teh putih hingga 100 kali lipat dari jenis teh yang lain.



Gambar 5. *White Tea*

2.4 *Electronic Nose*

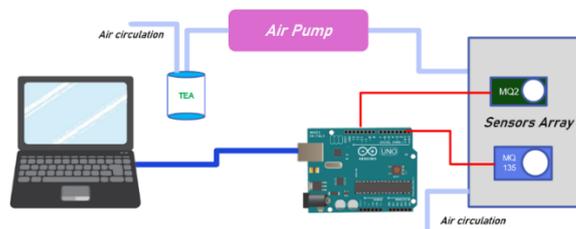
*Electronic nose* merupakan hidung buatan yang sengaja dirancang oleh manusia dengan prinsip kerja sama dengan sistem penciuman secara biologis. Perangkat ini dirancang dengan tujuan untuk dapat mendeteksi dan mengenali bau yang disajikan di sirkuit sensor gas. Perangkat ini dapat dikembangkan dan diterapkan dalam berbagai makanan dan minuman agar selalu menjaga konsistensi dalam aroma produk (Magfira & Sarno, 2018).

Perangkat *electronic nose* yang digunakan untuk mengambil data teh murni dalam penelitian ini disusun oleh beberapa komponen, di antaranya, *Arduino Uno AT Mega 256 microcontroller* dengan beberapa sensor MQ Series (Wakhid et al., 2020). Adapun sensor gas yang digunakan dalam perancangan *electronic nose* dapat dilihat pada Tabel 1.

Tabel 1. Tabel Sensor Gas

Sensor	Gas yang Dideteksi
MQ-135	Carbondioxide (CO <sub>2</sub> ), Carbonmonixide (CO)
MQ-8	Hydrogen (H <sub>2</sub> )
MQ-4	Methane (CH <sub>4</sub> )

Adapun skema konfigurasi yang digunakan sebagai perancangan dalam penelitian ini sesuai dengan mikrokontroler dan sensor-sensor yang dipasang dapat dilihat pada Gambar 6 di bawah ini.



Gambar 6. Skema konfigurasi *electronic nose*

2.5 *Dataset Penelitian*

Objek dalam penelitian ini adalah dataset dari 4 jenis teh murni yang merupakan hasil olahan dari tumbuhan daun *camellia sinensis* yang diambil dan diolah oleh Kebun Teh Jamus di bawah manajemen PT. Candi Loka yang berada di wilayah Kabupaten Ngawi, Jawa Timur yang tepatnya berada di Kawasan Desa Girikerto, Kecamatan Sine. Teh murni yang digunakan dalam penelitian ini terdiri dari black tea, green tea, oolong tea dan white tea.

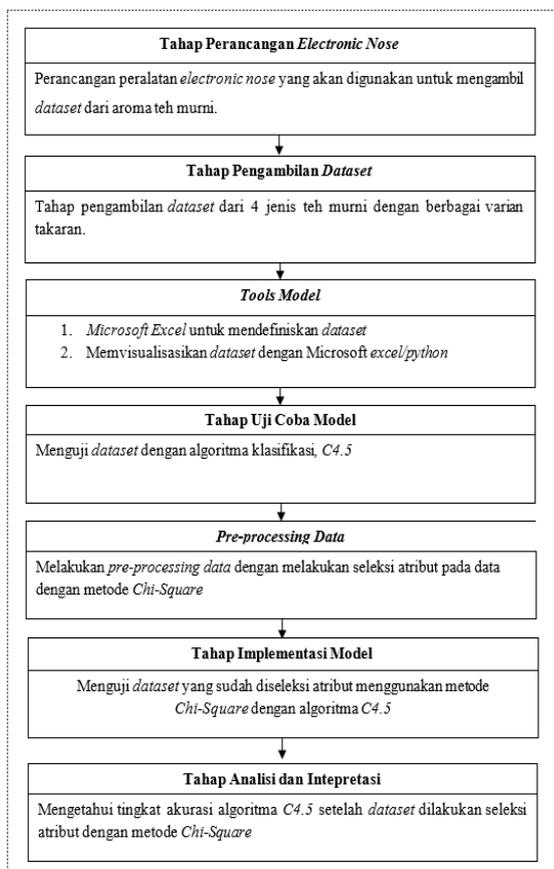
Dataset teh murni yang digunakan dalam penelitian ini terdiri dari 2756 records data dari 4 jenis teh murni asli Indonesia dan 4 atribut data pada masing-masing kelas. Atribut CO<sub>2</sub> mewakili uap gas karbondioksida, CO merupakan uap gas karbonmomoksida, H<sub>2</sub> merupakan hidrogen, CH<sub>4</sub> merupakan uap gas metana. Atribut data tersebut merupakan gas-gas yang terkandung dalam uap teh murni yang diambil oleh perangkat *electronic nose*. Penentuan atribut *dataset* yang digunakan merupakan hasil dari observasi dan wawancara dengan pakar dan produsen teh. Kemudian ditentukan sensor-senor yang memiliki kepekaan terhadap gas-gas yang ditentukan sebagai atribut dalam penelitian ini. Berikutnya data yang didapatkan akan dinormalisasi dan dilakukan proses pre-processing agar memaksimalkan kinerja algoritma digunakan dalam proses penelitian ini.

Tabel 2. Dataset Teh Murni

No.	CO <sub>2</sub>	CO	H <sub>2</sub>	CH <sub>4</sub>
1.	302	428	25	67
2.	302	428	25	67
3.	302	428	25	67
4.	302	428	25	67
5.	302	428	25	67
...	...	...	...	...
691.	276	377	25	97
692.	276	377	25	98
693.	276	377	25	97
694.	276	377	25	97
...	...	...	...	...
...	...	...	...	...
1379.	283	396	25	69
1380.	288	396	25	69
1381.	277	387	25	69
1382.	283	396	25	69
...	...	...	...	...
...	.....	.....	.....	.....
2755.	324	471	27	72
2756.	324	471	27	72

## 2.6 Kerangka Pemikiran Penelitian

Adapun kerangka pemikiran yang digunakan dalam penelitian dapat dilihat pada Gambar 7.



Gambar 7. Kerangka Pemikiran Penelitian

## 3. Hasil dan Pembahasan

### 3.1 Perancangan *Electronic Nose*

*Electronic nose* merupakan perangkat yang memiliki fungsi sebagai hidung tiruan manusia yang terdiri dari sensor-sensor gas yang disusun dan dikontrol dengan mikrokontroler untuk merespon aroma tertentu sesuai pengembangannya. Perancangan *electronic nose* dirancang mulai dari penentuan aroma gas yang akan diambil sehingga menentukan sensor yang akan digunakan dalam implementasinya. Perancangan perangkat *electronic nose* pada penelitian ini merupakan hasil rujukan dari pengembangan perangkat *electronic nose* pada penelitian-penelitian terdahulu yang memiliki fungsi berbeda. Beberapa penelitian terkait penerapan *electronic nose* menyebutkan bahwa *electronic nose* difungsikan sebagai pengukur kualitas kopi, pengukur kualitas beras dan pembanding perbedaan antara daging sapi dengan babi. Hasil dari perancangan perangkat *electronic nose* pada penelitian ini dapat dilihat pada Gambar 8.

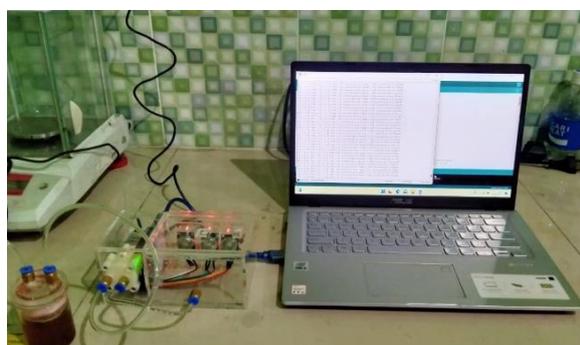
### 3.2 Pengujian Perangkat dan Pengambilan Data

Pengambilan *dataset* dari 4 jenis teh murni dilakukan dalam waktu 300 detik atau selama 5 menit, dimulai dari selesainya proses kalibrasi sensor

di dalam *electronic nose* kemudian larutan teh mulai diserap uapnya oleh pompa udara setelah katup terpasang dalam gelas kimia kemudian dialirkan ke dalam ruang udara *electronic nose*. Proses implementasi dalam pengambilan data larutan teh murni dapat dilihat pada Gambar 9.



Gambar 8. Hasil perancangan *electronic nose*



Gambar 9. Proses Pengambilan Data

### 3.3 Visualisasi Data

Penyajian data atau informasi lebih mudah dipahami melalui suatu tampilan grafis atau visual. Elemen visual yang ditampilkan untuk menyajikan suatu data atau informasi bertujuan untuk memberikan kemudahan dalam memahami tren data, *outlier*, *noise* dan *missing value* pada data. Sehingga peneliti lebih mudah untuk melihat kondisi data dan memutuskan metode dalam penanganan data sebelum dimasukkan ke dalam sebuah model.

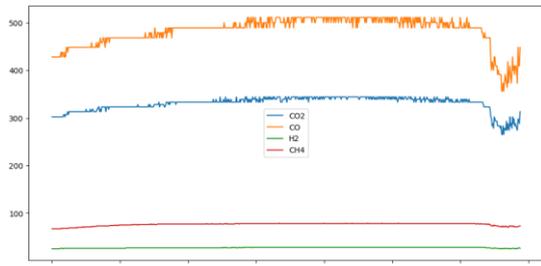
*Dataset* yang didapatkan dari perangkat *electronic nose* kemudian disajikan dalam bentuk visual grafis dengan pembagian visualisasi data perkelas. Visualisasi data disajikan dari kelas *black tea*, *green tea*, *oolong tea* dan *white tea*. Penyajian *dataset* perkelas dalam bentuk grafis seperti ini membantu peneliti dalam memahami kondisi data, sehingga peneliti mudah dalam menangani data pada tahap *preprocessing* ataupun penentuan model algoritma yang digunakan dalam menguji data.

Adapun penyajian visualisasi data dari *dataset* perkelasnya adalah sebagai berikut:

#### a. Visualisasi Data *Black Tea*

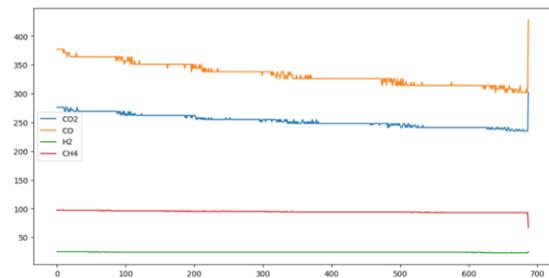
Pada Gambar 10 memvisualkan data *black tea* dengan 4 atributnya dari sampel data ke-1 sampai dengan ke-689. Tampilan visual pada data *black tea* memperlihatkan kondisi data yang mengalami ketidak stabilan di awal pengambilan data oleh sensor

hingga saat menjelang sensor berhenti mengambil data. Hal tersebut terlihat pada atribut untuk  $CO_2$  dan  $CO$ . Atribut  $H_2$  dan  $CH_4$  terlihat stabil dari awal pengambilan data hingga akhir sensor berhenti mengambil data. Ketidakstabilan sensor menghasilkan data yang kurang stabil dengan dampak akan mempengaruhi akurasi dari prediksi algoritma dalam pengujian.



Gambar 10. Visualisasi Data Black Tea

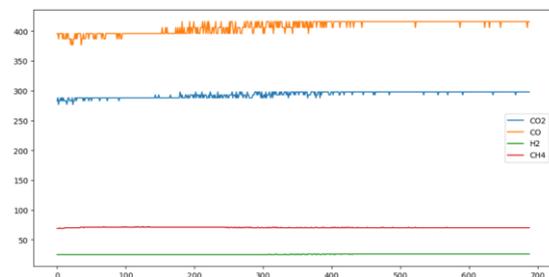
b. Visualisasi Data Green Tea



Gambar 11. Visualisasi Data Green Tea

Pada Gambar 11 memvisualkan data *green tea* dengan 4 atributnya dari sampel data ke-1 sampai dengan ke-689. Tampilan visual pada data *green tea* memperlihatkan kondisi data yang cukup stabil di awal pengambilan data oleh sensor, namun terjadi kenaikan nilai grafik saat menjelang sensor berhenti mengambil data. Hal tersebut terlihat pada atribut untuk  $CO_2$  dan  $CO$ . Atribut  $H_2$  dan  $CH_4$  terlihat stabil dari awal pengambilan data hingga akhir sensor berhenti mengambil data.

c. Visualisasi Data Oolong Tea

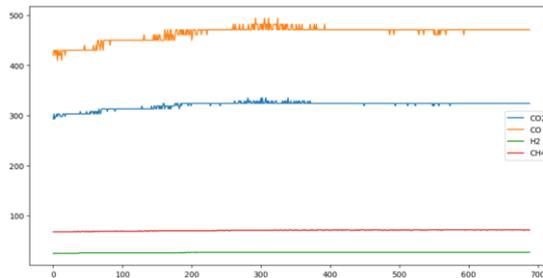


Gambar 12. Visualisasi Data Oolong Tea

Pada Gambar 12 memvisualkan data *oolong tea* dengan 4 atributnya dari sampel data ke-1 sampai dengan ke-689. Tampilan visual pada data *oolong tea* memperlihatkan kondisi data yang cukup stabil di

awal pengambilan data oleh sensor hingga menjelang sensor berhenti mengambil data. Hal tersebut terlihat pada atribut untuk  $CO_2$  dan  $CO$ . Atribut  $H_2$  dan  $CH_4$  terlihat stabil dari awal pengambilan data hingga akhir sensor berhenti mengambil data.

d. Visualisasi Data White Tea

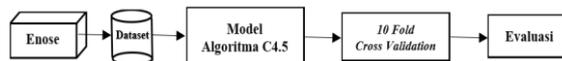


Gambar 13. Visualisasi Data White Tea

Pada Gambar 13 memvisualkan data *white tea* dengan 4 atributnya dari sampel data ke-1 sampai dengan ke-689. Tampilan visual pada data *white tea* memperlihatkan kondisi data yang cukup stabil di awal pengambilan data oleh sensor hingga menjelang sensor berhenti mengambil data. Hal tersebut terlihat pada atribut  $CO$ . Namun terlihat sedikit *noise* di awal pengambilan data atribut  $CO_2$ . Hal tersebut terlihat pada atribut untuk  $CO_2$  dan  $CO$ . Atribut  $H_2$  dan  $CH_4$  terlihat stabil dari awal pengambilan data hingga akhir sensor berhenti mengambil data.

3.4 Pengujian Dataset dengan Algoritma C4.5

Berdasarkan metode yang diusulkan dapat disajikan melalui skema metode pada Gambar 14 secara gambaran mendasar alur dari pengujian *dataset* dengan algoritma *C4.5*.



Gambar 14. Skenario penelitian

Pengujian *dataset* dengan algoritma *C4.5* menggunakan evaluasi *10fold cross validation* menghasilkan akurasi dari 10 iterasi. Hasil pengujian akhir berupa akurasi dihitung dari nilai rata-rata yang dihasilkan dari perhitungan 10 iterasi oleh algoritma *C4.5*. Metode ini lebih valid dalam perhitungan akurasi suatu hasil pengujian *dataset*, karena sistem perhitungan melakukan *cross check* secara merata terhadap sebaran data. Adapun hasil dari akurasi *10fold cross validation* dapat dilihat pada Tabel 3 di bawah ini.

Tabel 3. Pengujian C4.5 dengan 10fold Cross Validation

<i>k-fold</i>	<i>Accuracy Result</i>
1	55,07%
2	96,73%
3	100%
4	100%
5	100%
6	99,63%

7	100%
8	100%
9	100%
10	85,09%

Hasil perhitungan *recall*, *precision*, *accuracy* dan *F1-score* dari 2 kelas pada *dataset* teh disajikan pada Tabel 4 untuk dilakukan perhitungan total dan mengukur level kebenaran pada kinerja metode *Chi-square* dan algoritma *C4.5*.

Tabel 4. Hasil Evaluasi Algoritma *C4.5* dengan *Confusion matrix*

Class	Recall	Precision	Accuracy	F1-Score
Black Tea	0,942	0,85	0,95	0,896
Green Tea	0,933	0,999	0,981	0,964
Oolong Tea	0,942	0,977	0,98	0,96
White Tea	0,928	0,917	0,961	0,922
Rata-rata	0,936	0,935	0,935	0,935

Hasil akhir menunjukkan bahwa perhitungan algoritma *C4.5* menghasilkan akurasi yang tinggi yaitu sebesar 93,7% seperti pada Gambar 15 di bawah ini sehingga kinerja *electronic nose* menghasilkan data yang valid dalam melakukan klasifikasi jenis teh murni.

```
print("=====")
cv=cross_val_score(dt,data, y, cv=10)
print("Hasil Akurasi C45 Dataset Tea = ")
print(cv.mean())
#print(cv)

=====
Hasil Akurasi C45 Dataset Tea =
0.9365388669301712
```

Gambar 15. Hasil pengujian *dataset* teh murni

#### 4. Kesimpulan

Hasil analisis dari eksperimen yang dilakukan dalam penelitian ini dapat disimpulkan bahwa algoritma *C4.5* pada klasifikasi jenis teh murni menggunakan perangkat *electronic nose* menghasilkan kesimpulan sebagai berikut: Pengembangan *electronic nose* sebagai perangkat yang berfungsi untuk mengambil *dataset* teh murni dari uap teh menghasilkan *dataset* yang cukup stabil. Hal itu dibuktikan dengan pengujian *dataset* teh murni yang dimasukkan ke dalam model algoritma *C4.5* dengan hasil akurasi prediksi klasifikasi sebesar 93%,

Pada eksperimen yang telah dilakukan dalam penelitian ini masih terdapat beberapa kekurangan yang dapat disempurnakan untuk penelitian selanjutnya. Adapun beberapa kekurangan yang dapat diperbaiki adalah sebagai berikut:

- Electronic nose* untuk uap teh masih dapat dikembangkan lagi dengan tujuan klasifikasi jenis teh murni dengan menambah sensor-sensor selain *MQ-135*, *MQ-8*, *MQ-4* yang bertujuan mendapatkan atribut *dataset* lebih banyak dari penelitian ini.
- Pengembangan *electronic nose* masih bisa dilakukan untuk pengukuran kualitas teh dengan

jenis tertentu, klasifikasi jenis kopi Indonesia, kualitas daging dan obyek lain yang dapat diukur melalui gas yang ada dalam objek tertentu.

#### Daftar Pustaka:

Abena, P., Hadjinian, P., Stadler, R., Verhees, J., & Zanasi, A. (1998). *Discovering Data Mining. From Concept to Implementation*. Prentice Hall.

Al Isyrofie, A. I. F., Kashif, M., Aji, A. K., Aidatuzzahro, N., Rahmatillah, A., Winarno, Susilo, Y., Syahrom, A., & Astuti, S. D. (2022). Odor clustering using a gas sensor array system of chicken meat based on temperature variations and storage time. *Sensing and Bio-Sensing Research*, 37(July), 100508. <https://doi.org/10.1016/j.sbsr.2022.100508>

Banerjee, R., Bandyopadhyay, R., Tudu, B., & Bhattacharyya, N. (2016). Tea and the Use of the Electronic Nose. In *Electronic Noses and Tongues in Food Science* (pp. 125–135). Elsevier Inc. <https://doi.org/10.1016/B978-0-12-800243-8.00013-5>

Gite, P., Chouhan, K., Murali Krishna, K., Kumar Nayak, C., Soni, M., & Shrivastava, A. (2021). ML Based Intrusion Detection Scheme for various types of attacks in a WSN using C4.5 and CART classifiers. *Materials Today: Proceedings*, xxxx. <https://doi.org/10.1016/j.matpr.2021.07.378>

Lim, C. M., Carey, M., Williams, P. N., & Koidis, A. (2021). Rapid classification of commercial teas according to their origin and type using elemental content with X-ray fluorescence (XRF) spectroscopy. *Current Research in Food Science*, 4, 45–52. <https://doi.org/10.1016/j.crfs.2021.02.002>

Magfira, D. B., & Sarno, R. (2018). Classification of Arabica and Robusta coffee using electronic nose. *2018 International Conference on Information and Communications Technology, ICOIACT 2018, 2018-Janua*, 645–650. <https://doi.org/10.1109/ICOIACT.2018.8350725>

Meng, X., Zhang, P., Xu, Y., & Xie, H. (2020). Construction of decision tree based on C4.5 algorithm for online voltage stability assessment. *International Journal of Electrical Power and Energy Systems*, 118. <https://doi.org/10.1016/j.ijepes.2019.105793>

Putra, O. A., Firdaus, & Heryansyah, M. H. (2016). Identifikasi Aroma Teh dengan E-nose Menggunakan Metode Backpropagation. *Seminar Nasional Sains Dan Teknologi*, 34(1), 1–8.

Statistik Teh Indonesia. (2018). ©BPS RI/BPS – Statistics Indonesia Pencetak/Printed.

Sundaramurthy, S., & Jayavel, P. (2020). A hybrid Grey Wolf Optimization and Particle Swarm

- Optimization with C4.5 approach for prediction of Rheumatoid Arthritis. *Applied Soft Computing Journal*, 94. <https://doi.org/10.1016/j.asoc.2020.106500>
- Susanti, A. A., & Akbar. (2019). Buku Outlook Komoditas Perkebunan Teh. In *Pusat Data dan Sistem Informasi Pertanian*.
- Tanyu, B. F., Abbaspour, A., Alimohammadlou, Y., & Tecuci, G. (2021). Landslide susceptibility analyses using Random Forest, C4.5, and C5.0 with balanced and unbalanced datasets. *Catena*, 203. <https://doi.org/10.1016/j.catena.2021.105355>
- Wakhid, S., Sarno, R., Sabilla, S. I., & Maghfira, D. B. (2020). Detection and classification of Indonesian civet and non-civet coffee based on statistical analysis comparison using E-Nose. *International Journal of Intelligent Engineering and Systems*, 13(4), 56–65. <https://doi.org/10.22266/IJIES2020.0831.06>
- Wijaya, D. R., Sarno, R., & Zulaika, E. (2018). Electronic nose dataset for beef quality monitoring in uncontrolled ambient conditions. *Data in Brief*, 21, 2414–2420. <https://doi.org/10.1016/j.dib.2018.11.091>
- Wijayanti, S., Kartikadarma, E., & Wulandari, S. A. (2013). Perancangan Enose sebagai Alat Uji Cepat Mutu Beras Aromatik. *Seminar Nasional Teknologi Informasi & Komunikasi Terapan 2013 (Semantik 2013)*, 2013(November), 340–344.
- Xu, M., Wang, J., & Zhu, L. (2021). Tea quality evaluation by applying E-nose combined with chemometrics methods. *Journal of Food Science and Technology*, 58(4), 1549–1561. <https://doi.org/10.1007/s13197-020-04667-0>

*Halaman ini sengaja dikosongkan*