

Klik disini untuk menuliskan kategori naskah

## Klasifikasi BAU Dalam Kulkas Menggunakan Sensor SGP-30 Dan Algoritma Random Forest Classifier

Amira Muhammad Salim Banem <sup>1</sup>, Halim Maulana <sup>2</sup>

<sup>1</sup> Fakultas Ilmu Komputer dan Teknologi Informasi, Teknologi Informasi, Universitas Muhammadiyah Sumatera Utara, Medan, Indonesia

<sup>2</sup> Fakultas Ilmu Komputer dan Teknologi Informasi, Sistem Informasi, Universitas Muhammadiyah Sumatera Utara, Medan, Indonesia

### INFORMASI ARTIKEL

Diterima Redaksi: 00 Januari 00  
Revisi Akhir: 00 Februari 00  
Diterbitkan *Online*: 00 Maret 00

### KATA KUNCI

Klasifikasi Bau; SGP-30; Random Forest Classifier; Kulkas

### KORESPONDENSI

Phone: +6283165383328  
E-mail: meeramodz9@gmail.com

### A B S T R A K

Penelitian ini bertujuan untuk mengembangkan sistem klasifikasi bau dalam kulkas menggunakan sensor SGP-30 dan algoritma Random Forest Classifier. Sistem ini dirancang untuk mendeteksi dan mengklasifikasikan kondisi bau makanan ke dalam tiga kategori, yaitu Segar, Sedang, dan Busuk, berdasarkan data sensor berupa nilai eCO<sub>2</sub>, TVOC, suhu, dan kelembapan. Data yang diperoleh kemudian diproses menggunakan algoritma Random Forest Classifier yang mampu menghasilkan akurasi sebesar 97,78% dalam melakukan klasifikasi bau. Hasil penelitian menunjukkan bahwa sistem yang dibangun dapat mengidentifikasi kondisi bau secara akurat dan efektif. Sistem ini berpotensi untuk dikembangkan lebih lanjut dengan memperluas variasi data dan mengeksplorasi algoritma klasifikasi lainnya guna meningkatkan performa sistem.

### PENDAHULUAN

Machine learning atau pembelajaran mesin merupakan salah satu cabang penting dalam bidang kecerdasan buatan (Artificial Intelligence/AI), yang memungkinkan komputer atau sistem untuk secara otomatis belajar dari data yang tersedia. Alih-alih hanya menjalankan instruksi yang telah diprogram secara eksplisit oleh manusia, machine learning dirancang untuk memungkinkan komputer mengenali pola, mempelajari struktur data, serta membuat keputusan atau prediksi berdasarkan hasil analisis terhadap data tersebut. Teknologi ini bekerja dengan membangun model matematis atau statistik dari kumpulan data, kemudian menggunakan model tersebut untuk menggeneralisasi atau mengantisipasi hasil pada data baru yang sebelumnya belum pernah dilihat. Artinya, dengan menggunakan machine learning, sistem dapat meningkatkan kemampuannya seiring waktu melalui proses pelatihan (training), tanpa perlu perintah yang secara langsung mendefinisikan setiap langkah atau logika keputusan. Kemampuan untuk menganalisis data secara otomatis ini menjadikan machine learning sangat efektif dalam berbagai aplikasi modern, seperti pengenalan suara, klasifikasi gambar, deteksi anomali, serta sistem rekomendasi. Dalam konteks penelitian ini, machine learning digunakan untuk melakukan klasifikasi terhadap jenis bau dalam kulkas berdasarkan data sensor yang dikumpulkan, sehingga sistem dapat menentukan apakah bau yang terdeteksi termasuk kategori segar, sedang, atau busuk, tanpa intervensi manual dari pengguna (Chyan et al., n.d.).

Kulkas merupakan salah satu perangkat elektronik yang hampir selalu ada di setiap rumah tangga. Fungsi utamanya adalah menjaga kesegaran bahan makanan dengan mengontrol suhu dan kelembapan udara di dalam ruang penyimpanan. Namun, permasalahan yang kerap muncul dan sering diabaikan adalah munculnya bau tidak sedap akibat akumulasi senyawa organik volatil (Volatile Organic Compounds, VOC) dari bahan makanan yang membusuk atau tidak tersimpan

dengan baik. Masalah ini tidak hanya mengganggu kenyamanan pengguna, tetapi juga berpotensi membahayakan kesehatan jika tidak segera ditangani. Senyawa-senyawa berbahaya seperti amonia, hidrogen sulfida, dan metana seringkali merupakan hasil dekomposisi bahan organik yang mengindikasikan potensi adanya kontaminasi bakteri atau jamur (Sensirion, Rüffer et al. 2018).

Menurut penelitian Macías-Quijas, keberadaan bau tidak sedap di lingkungan penyimpanan makanan dapat berfungsi sebagai indikator dini terhadap kualitas makanan. Dalam konteks ini, permasalahan bau dalam kulkas dapat menjadi lebih signifikan ketika pengguna tidak dapat secara langsung mengetahui sumber atau tingkat keparahan bau tersebut. Kondisi tertutup pada kulkas membuat deteksi manual menjadi tidak efisien, dan seringkali pengguna baru menyadari bau saat sudah terlalu terlambat. Selain itu, menurut Springer, banyak kasus pemborosan makanan disebabkan oleh tidak disadarinya kerusakan bahan makanan secara dini, yang akhirnya berdampak pada ekonomi rumah tangga dan isu keberlanjutan pangan (Palma et al., 2025).

Masalah ini semakin kompleks dengan meningkatnya kebutuhan akan pengelolaan makanan secara efisien, terutama di tengah gaya hidup modern. Oleh karena itu, dibutuhkan solusi yang mampu mendeteksi bau secara otomatis, akurat, dan real-time. Dalam hal ini, teknologi Internet of Things (IoT) menawarkan pendekatan yang sangat menjanjikan. Dengan memanfaatkan sensor gas pintar seperti SGP-30, sistem dapat secara langsung mengukur konsentrasi TVOC dan estimasi CO<sub>2</sub> sebagai indikator kualitas udara di dalam kulkas. Sensor ini berbasis metal-oxide dan telah banyak digunakan untuk aplikasi pemantauan kualitas udara dalam ruangan karena kepekaannya terhadap berbagai senyawa organik yang berhubungan dengan bau (Budiman, Wibowo 2020).

Namun, sensor saja tidak cukup. Data yang dihasilkan oleh sensor gas sering kali bersifat kompleks dan memiliki pola non-linear, sehingga memerlukan teknik pengolahan lanjutan untuk mengubah data mentah menjadi informasi yang berguna. Dalam hal ini, algoritma pembelajaran mesin (machine learning) memainkan peran penting. Salah satu algoritma yang terbukti sangat efektif dalam klasifikasi data sensor adalah Random Forest Classifier. Algoritma ini bekerja dengan membangun sejumlah pohon keputusan dan menggabungkan hasilnya untuk memperoleh klasifikasi akhir yang lebih akurat dan tahan terhadap overfitting (Aira et al., 2024) (Anderson et al., 2024).

Random Forest sangat cocok untuk mengatasi tantangan dalam klasifikasi bau karena mampu menangani banyak fitur input, serta memiliki toleransi terhadap noise pada data. Dalam studi yang dilakukan Maharani et al., Random Forest digunakan untuk klasifikasi kualitas udara berbasis IoT dan menunjukkan hasil akurasi tinggi dibandingkan algoritma lain seperti K-Nearest Neighbor dan Naïve Bayes. Sementara itu, Macías-Quijas juga menemukan bahwa Random Forest lebih unggul dalam mengklasifikasikan data sensor bau lingkungan dibandingkan metode klasifikasi tradisional.

Selain keunggulan teknisnya, alasan lain pemilihan Random Forest adalah kestabilan hasil klasifikasinya meskipun pada data pelatihan yang tidak seimbang atau fluktuatif. Hal ini sangat penting dalam konteks sensor bau, karena data yang dikumpulkan bisa sangat bervariasi tergantung jenis makanan, kondisi suhu, hingga umur penyimpanan. Keandalan ini membuat sistem klasifikasi berbasis Random Forest lebih adaptif terhadap berbagai kondisi lingkungan dalam kulkas (Pratama, 2022) (Gusti 2023).

Berdasarkan pendekatan ini, sistem klasifikasi bau dalam kulkas yang diusulkan terdiri dari tiga komponen utama: yang pertama Sensor SGP-30 sebagai perangkat akuisisi data bau, kedua Mikrokontroler ESP32 sebagai pemroses data dan penghubung ke jaringan IoT, serta Random Forest Classifier sebagai pengklasifikasi jenis bau berdasarkan dataset yang telah dilabeli sebelumnya. Sistem ini nantinya akan mampu mendeteksi bau-bau tertentu, seperti bau makanan basi, fermentasi, atau kimiawi, dan mengklasifikasinya menjadi 3 kategori bau yaitu bau segar, sedang, busuk .

Gambaran hasil yang diharapkan dari penelitian ini adalah terciptanya prototipe sistem IoT yang mampu melakukan klasifikasi bau secara otomatis dan real-time dengan tingkat akurasi yang tinggi. Sistem ini juga diharapkan dapat memberikan peringatan atau notifikasi kepada pengguna saat terdeteksi adanya bau tidak sedap, sehingga tindakan preventif dapat segera dilakukan. Dengan demikian, sistem ini tidak hanya meningkatkan kenyamanan pengguna, tetapi juga membantu dalam mengurangi pemborosan makanan, menjaga kesehatan, dan mendukung praktik penyimpanan makanan yang lebih higienis dan efisien.

Lebih jauh, sistem ini berpotensi diintegrasikan dengan teknologi smart home dan dikembangkan sebagai bagian dari sistem manajemen makanan cerdas yang lebih besar. Hal ini sejalan dengan tren industri 4.0 dan smart living, di mana

teknologi digital digunakan untuk mengoptimalkan berbagai aspek kehidupan rumah tangga. Dengan pendekatan yang komprehensif dan berbasis teknologi mutakhir, penelitian ini diharapkan dapat memberikan kontribusi nyata dalam pemecahan masalah sehari-hari yang selama ini kerap diabaikan.

## TINJAUAN PUSTAKA

### *Machine Learning (ML)*

Machine Learning (ML) atau dalam bahasa Indonesia disebut pembelajaran mesin, merupakan salah satu pendekatan utama dalam bidang Artificial Intelligence (AI) yang dirancang untuk meniru atau menggantikan perilaku manusia dalam menyelesaikan suatu masalah secara otomatis. Sejalan dengan namanya, machine learning bekerja dengan meniru cara manusia belajar dari pengalaman dan kemudian menggunakan pengetahuan tersebut untuk membuat keputusan atau menyelesaikan tugas yang serupa di masa depan. Konsep dasar dari machine learning terletak pada kemampuan sistem untuk belajar dari data yang diberikan tanpa harus diprogram secara eksplisit untuk setiap tugas tertentu. Proses ini dikenal sebagai training atau pelatihan, di mana sistem dipaparkan pada sejumlah besar data yang disebut data training untuk mengidentifikasi pola, hubungan, dan struktur dalam data tersebut. Secara umum, ada dua penerapan utama dalam machine learning, yaitu klasifikasi dan prediksi (regresi). Klasifikasi digunakan ketika sistem ditugaskan untuk mengelompokkan data ke dalam kategori tertentu berdasarkan ciri-ciri atau fitur yang dimilikinya, layaknya manusia yang mengenali dan membedakan objek berdasarkan pengamatan. Contohnya, dalam penelitian ini, sistem ML digunakan untuk mengklasifikasikan jenis bau dalam kulkas menjadi segar, asam, atau busuk. Sementara itu, prediksi digunakan untuk memperkirakan nilai keluaran dari suatu data input berdasarkan pola yang telah dipelajari selama proses pelatihan. Beberapa metode machine learning yang banyak digunakan dalam berbagai aplikasi antara lain adalah Decision Tree atau sistem pengambil keputusan, Support Vector Machine (SVM) yang unggul dalam klasifikasi data kompleks, dan Neural Network, yang mengadopsi cara kerja jaringan saraf manusia untuk menyelesaikan permasalahan pembelajaran yang lebih kompleks dan non-linear (Ahmad Hania, 2017).

### *Algoritma Random Forest Classifier*

Random Forest Classifier adalah sebuah metode klasifikasi yang terdiri dari sekumpulan pohon keputusan (decision trees) yang dibangun secara independen satu sama lain. Setiap pohon keputusan tersebut dilatih menggunakan data latih yang berbeda-beda serta fitur-fitur acak yang tidak saling bergantung, sehingga setiap pohon memiliki karakteristik dan perspektif yang unik dalam memproses data. Hasil dari setiap pohon keputusan ini kemudian digabungkan melalui mekanisme voting, di mana kelas akhir atau prediksi ditentukan berdasarkan mayoritas suara dari seluruh pohon yang ada dalam hutan (forest) tersebut. Proses pembuatan pohon keputusan dimulai dengan menentukan node akar (root node) yang menjadi titik awal pengambilan keputusan, dan kemudian berjalan melalui cabang-cabang hingga mencapai node daun (leaf nodes) yang berisi keputusan akhir atau hasil klasifikasi. Metode ini sangat efektif dalam mendeteksi pola seperti sarkasme, karena kemampuan Random Forest dalam menangani fitur yang bervariasi dan memanfaatkan keragaman pohon untuk meningkatkan akurasi dan ketahanan model terhadap kesalahan data. Sebagai gambaran, pohon keputusan diilustrasikan sebagai struktur bercabang yang berawal dari satu titik pusat dan berujung pada beberapa titik hasil yang mewakili keputusan akhir dari proses klasifikasi (Adriansyah et al., 2022).

### *Sensor SGP30*

SGP30 merupakan salah satu jenis sensor gas digital yang dirancang dan dikembangkan oleh perusahaan Sensirion, yang terkenal akan teknologi sensor berkinerja tinggi. Sensor ini secara khusus ditujukan untuk digunakan dalam berbagai aplikasi yang berhubungan dengan pemantauan kualitas udara, seperti perangkat pembersih udara otomatis, sistem ventilasi pintar, serta alat-alat berbasis Internet of Things (IoT). SGP30 menggunakan protokol komunikasi I2C yang memungkinkan integrasi yang sederhana dan cepat dengan berbagai jenis mikrokontroler, termasuk yang populer di kalangan pengembang seperti ESP32 dan Arduino. Selain mudah dalam implementasinya, sensor ini juga sudah dikalibrasi sejak dari pabrik, sehingga pengguna tidak perlu lagi melakukan kalibrasi manual. Kemampuan adaptasi otomatis terhadap kondisi lingkungan sekitar menjadikan SGP30 ideal untuk aplikasi yang memerlukan data akurat dan real-time dalam memantau kualitas udara di dalam ruangan maupun luar ruangan.

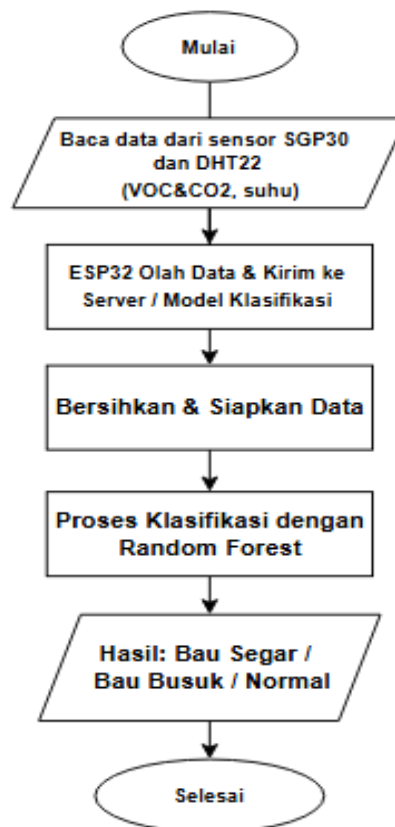
Fungsi utama dari SGP30 adalah untuk mendeteksi dan mengukur dua parameter penting dalam kualitas udara, yaitu Total Volatile Organic Compounds (TVOC) dan equivalent carbon dioxide (eCO<sub>2</sub>). TVOC merupakan kumpulan senyawa organik yang mudah menguap dan biasanya berasal dari bahan kimia rumah tangga, asap rokok, parfum, atau

zat pelarut lainnya yang berpotensi membahayakan kesehatan jika terakumulasi dalam ruangan tertutup. Sedangkan eCO<sub>2</sub> bukanlah pengukuran langsung dari karbon dioksida, melainkan estimasi kadar CO<sub>2</sub> berdasarkan konsentrasi gas VOC yang terdeteksi—sehingga disebut "equivalent CO<sub>2</sub>". Dalam kondisi normal, sensor ini dapat mengukur TVOC dalam rentang 0–60 ppm dan eCO<sub>2</sub> dalam kisaran 400–1000 ppm. Namun, dalam aplikasi khusus yang membutuhkan pemantauan intensif atau industri, jangkauan pengukuran dapat diperluas hingga 60.000 ppb untuk TVOC dan 60.000 ppm untuk eCO<sub>2</sub>. Sensor ini memiliki tingkat akurasi pengukuran sekitar  $\pm 15\%$  dari nilai aktual, yang dianggap cukup baik untuk aplikasi komersial dan penelitian lingkungan (Kusumah et al., 2018).

## METODOLOGI

Metodologi penelitian dimulai dengan studi literatur mengenai sensor gas SGP-30 dan algoritma Random Forest untuk klasifikasi bau. Kemudian dilakukan identifikasi kebutuhan sistem untuk mengklasifikasikan bau dalam kulkas berdasarkan data sensor. Setelah perancangan model klasifikasi, dilakukan pengumpulan data dan pelatihan algoritma Random Forest menggunakan data hasil sensor. Selanjutnya, sistem diuji secara real-time untuk mengevaluasi akurasi klasifikasi bau dalam kulkas berdasarkan sensor SGP-30. Hasil pengujian digunakan untuk menyempurnakan model agar klasifikasi bau lebih tepat dan dapat diandalkan.

### Flowchart Sistem Klasifikasi



Gambar 1. Flowchart Sistem Klasifikasi

Sistem Flowchart di atas menggambarkan alur sistem klasifikasi bau berbasis IoT yang dimulai dari pembacaan data oleh sensor SGP30 dan DHT22 (mengukur VOC, CO<sub>2</sub>, dan suhu), kemudian data dikirim dan diolah oleh ESP32 untuk diteruskan ke server atau model klasifikasi. Selanjutnya, data dibersihkan dan dipersiapkan sebelum diproses oleh algoritma Random Forest untuk klasifikasi bau (segar, sedang, busuk). Hasil klasifikasi ditampilkan dan sistem berakhir setelah proses selesai. Alur ini mencerminkan sistem pendeteksi bau otomatis yang efisien dan terintegrasi dari sensor hingga hasil.

### ***Klasifikasi dengan Menggunakan Algoritma Random Forest***

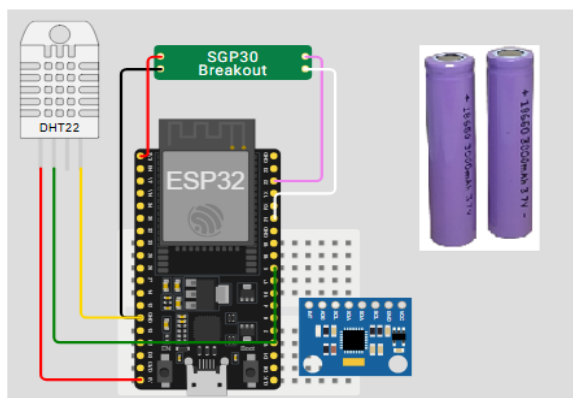
Sistem ini mengklasifikasikan bau dalam kulkas menjadi tiga jenis: bau segar (makanan masih baik), bau sedang (awal pembusukan), dan bau busuk (makanan sudah rusak). Algoritma Random Forest digunakan untuk memproses data dari banyak pohon keputusan. Setiap pohon memberi prediksi berdasarkan data yang dilatih secara acak, lalu hasil akhir ditentukan dari suara mayoritas. Fitur utama yang digunakan adalah nilai TVOC dan eCO<sub>2</sub> dari sensor SGP-30, serta suhu dan kelembaban jika tersedia. Contohnya, TVOC 400 ppb dan eCO<sub>2</sub> 1200 ppm biasanya berarti bau sedang, sedangkan nilai yang lebih tinggi seperti TVOC 1500 ppb dan eCO<sub>2</sub> 2500 ppm menunjukkan bau busuk. Data sensor diterima secara real-time melalui UDP dari mikrokontroler. Model yang sudah dilatih lalu memprediksi jenis bau dan hasilnya langsung disimpan serta divisualisasikan untuk memudahkan pemantauan.

## **HASIL DAN PEMBAHASAN**

### ***Gambaran Umum Sistem***

Sistem yang dikembangkan bertujuan untuk melakukan klasifikasi bau dalam kulkas secara otomatis dan real-time menggunakan sensor SGP-30 serta algoritma Random Forest Classifier. Sistem terdiri dari dua bagian utama:

1. Perangkat keras: ESP32, sensor SGP-30 untuk mendeteksi nilai eCO<sub>2</sub> dan TVOC, serta DHT22 untuk pengukuran suhu dan kelembaban.



Gambar 2. Rancangan Alat

Alat ini dirancang untuk mengumpulkan data bau di dalam kulkas secara real-time dengan bantuan sensor SGP30 yang mendeteksi konsentrasi gas eCO<sub>2</sub> dan TVOC sebagai indikator kualitas udara. Sensor DHT22 juga digunakan untuk mencatat suhu dan kelembaban yang memengaruhi pembentukan bau. Semua data yang dikumpulkan dikirim dan diproses oleh mikrokontroler ESP32, kemudian digunakan untuk pelatihan dan pengujian algoritma klasifikasi bau menggunakan machine learning. Dengan demikian, alat ini berfungsi sebagai sistem akuisisi data yang penting dalam proses klasifikasi bau kulkas.

2. Perangkat lunak: Skrip Python `main_realtime.py` yang bertugas: menerima data sensor via protokol UDP, melakukan prediksi klasifikasi bau menggunakan model Random Forest (`model_random_forest.pkl`), menyimpan hasil ke file `data_prediksi_realtime.csv`, dan menampilkan visualisasi real-time dengan Matplotlib.

### ***Dataset yang Digunakan***

Dataset yang digunakan dalam penelitian ini dikumpulkan secara real-time selama periode tertentu menggunakan sensor gas dan sensor lingkungan yang terhubung ke mikrokontroler ESP32. Data sensor dikirim melalui protokol UDP ke komputer yang menjalankan skrip Python untuk menerima, memproses, dan menyimpan data secara otomatis dalam format CSV. Setiap data yang direkam terdiri dari beberapa parameter penting, yaitu:

1. Waktu Pengambilan Data (WAKTU)  
Mencatat waktu pengambilan data sensor dalam format datetime untuk memudahkan analisis temporal.
2. Nilai eCO<sub>2</sub> (ppm)  
Konsentrasi karbon dioksida ekuivalen yang terdeteksi oleh sensor gas.
3. Nilai TVOC (ppb)  
Total Volatile Organic Compounds, diukur sebagai indikator senyawa organik yang mudah menguap.
4. Suhu (°C)  
Suhu lingkungan sekitar sensor yang diukur menggunakan sensor suhu.

## 5. Kelembapan (%)

Persentase kelembapan udara di sekitar sensor.

Data yang diterima oleh sistem secara otomatis disimpan dalam file CSV, seperti `data_bau.csv` atau `data_prediksi_realtime.csv`, yang menjadi sumber utama untuk pelatihan dan pengujian model klasifikasi bau. Contoh data pada CSV adalah sebagai berikut:

	eCO2	TVOC	SUHU	KELEMBAPAN
0	513.46	194.99	4.72	70.42
1	433.66	217.75	5.59	69.72
2	595.43	249.99	4.63	71.16
3	500.24	231.02	5.17	67.85
4	594.01	291.89	5.62	71.67

Gambar 3. Data Awal Sensor

Lima baris pertama dari data menunjukkan hasil pembacaan sensor SGP-30 yang terdiri dari empat parameter yaitu eCO2 (karbon dioksida ekuivalen), TVOC (Total Volatile Organic Compounds), suhu, dan kelembapan. Nilai-nilai ini merepresentasikan kondisi udara di dalam kulkas yang digunakan sebagai input dalam sistem klasifikasi bau. Data ini penting untuk mendeteksi perubahan kualitas udara yang mengindikasikan tingkat kesegaran atau kebusukan makanan di dalam kulkas.

### Deskripsi Data

Langkah awal eksplorasi dimulai dengan menampilkan lima data teratas dan deskripsi statistik dari dataset. Dataset terdiri dari empat fitur utama: eCO2 (karbon dioksida ekuivalen), TVOC (total senyawa organik volatil), SUHU, dan KELEMBAPAN. Keempat fitur ini merepresentasikan kondisi lingkungan dalam kulkas. Melalui fungsi `.describe()`, diperoleh informasi statistik seperti nilai minimum, maksimum, rata-rata (mean), dan standar deviasi untuk tiap fitur. Analisis ini penting untuk mengetahui skala data dan mendeteksi nilai ekstrem atau tidak wajar yang dapat mengganggu pemodelan.

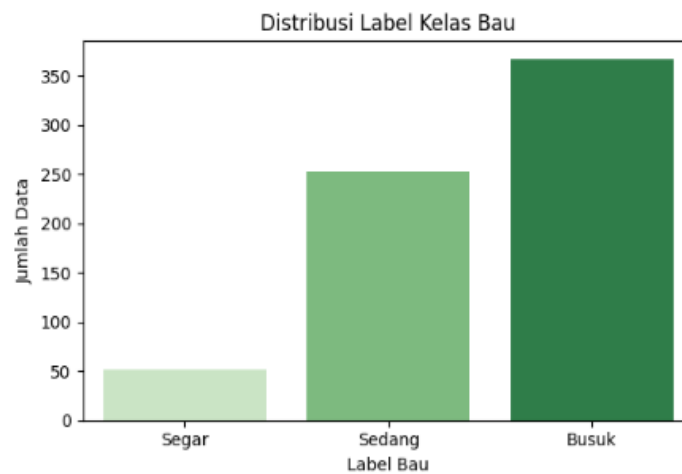
	eCO2	TVOC	SUHU	KELEMBAPAN
count	672.000000	672.000000	672.000000	672.000000
mean	1489.287723	1572.230580	6.870923	73.047292
std	630.304491	849.590182	1.546800	7.637067
min	331.190000	105.210000	2.900000	58.210000
25%	944.457500	841.910000	5.640000	66.982500
50%	1482.205000	1577.580000	6.820000	72.905000
75%	1975.015000	2267.460000	7.980000	77.925000
max	3045.860000	3344.920000	10.650000	93.430000

Gambar 4. Deskripsi Data

Ringkasan statistik di atas menunjukkan bahwa data sensor memiliki rentang nilai yang bervariasi dan tidak terdapat nilai kosong (missing value), sehingga data dapat langsung digunakan dalam proses pemodelan.

### Distribusi Label Kelas

Distribusi label kelas sangat penting untuk dipahami karena mempengaruhi performa model klasifikasi. Label kelas di sini adalah kategori bau (misalnya 'Segar', 'Sedang', 'Busuk') yang dihasilkan dari klasifikasi berdasarkan data sensor. Dengan mengetahui seberapa banyak data yang masuk ke tiap kategori, kita dapat mengevaluasi apakah dataset seimbang atau terjadi ketidakseimbangan kelas (class imbalance). Ketidakseimbangan data dapat menyebabkan model bias ke kelas mayoritas sehingga performa prediksi untuk kelas minoritas menjadi kurang baik. Visualisasi distribusi label dengan grafik batang memudahkan interpretasi dan memberikan gambaran cepat tentang proporsi tiap kelas dalam data.

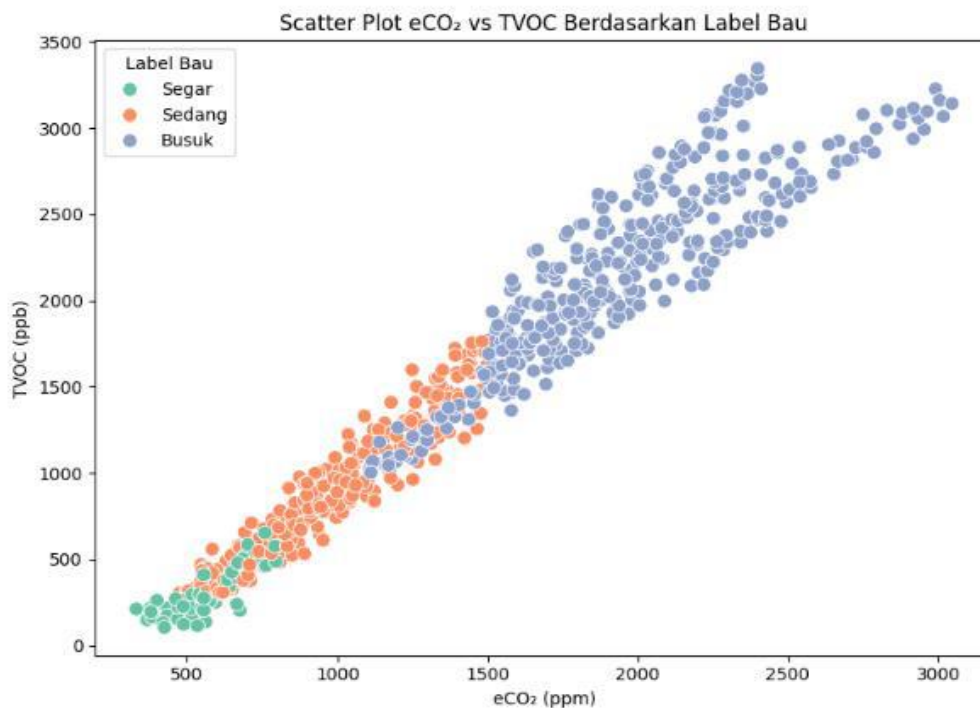


Gambar 5. Distribusi Label Kelas

Gambar di atas menampilkan distribusi jumlah data untuk masing-masing kategori label bau dalam bentuk diagram batang berwarna hijau. Grafik ini menunjukkan seberapa banyak data yang termasuk dalam kategori Segar, Sedang, dan Busuk, berdasarkan hasil klasifikasi aturan ambang batas sensor. Melalui grafik ini, kita dapat dengan mudah mengamati apakah data distribusinya seimbang atau didominasi oleh salah satu kategori bau tertentu, sehingga membantu dalam memahami karakteristik dataset serta memastikan tidak terjadi ketidakseimbangan kelas (class imbalance) yang bisa mempengaruhi performa model klasifikasi.

#### *Distribusi Fitur per Kelas*

Analisis distribusi fitur per kelas dilakukan untuk melihat bagaimana nilai sensor eCO<sub>2</sub>, TVOC, suhu, dan kelembapan tersebar di setiap kategori bau. Misalnya, apakah nilai eCO<sub>2</sub> cenderung lebih rendah pada kategori 'Segar' dibanding 'Busuk'. Informasi ini berguna untuk memahami pola dan perbedaan karakteristik data pada tiap kelas, yang akan membantu dalam proses pemodelan dan interpretasi hasil. Visualisasi menggunakan boxplot memperlihatkan sebaran data, median, kuartil, serta potensi outlier pada tiap fitur untuk setiap kelas, sehingga kita bisa menilai kualitas dan relevansi fitur dalam klasifikasi.

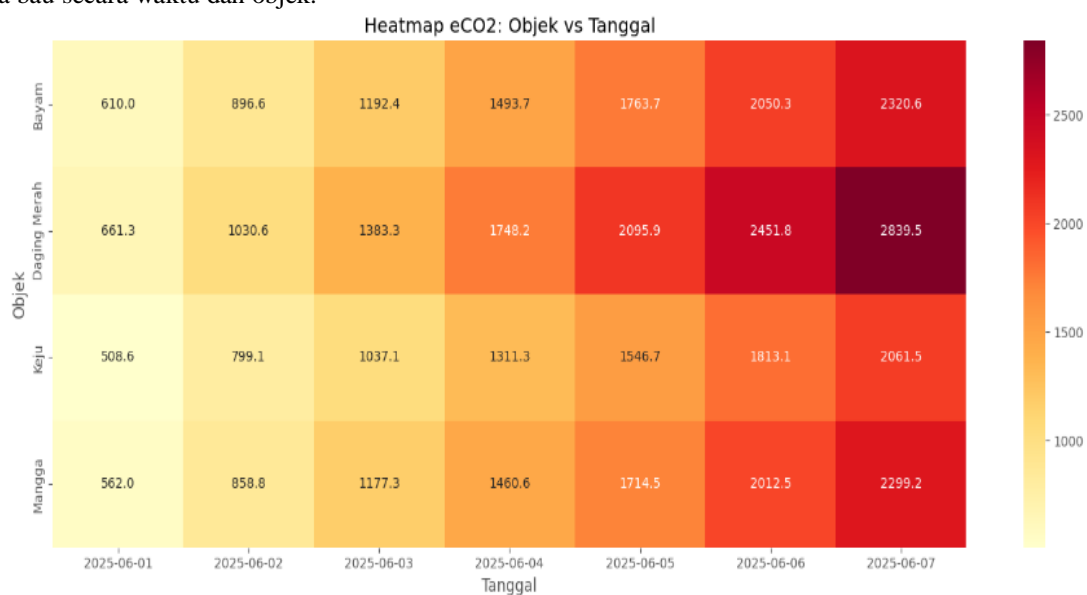


Gambar 6. Distribusi Fitur per Kelas

Scatter plot di atas menjelaskan hubungan antara nilai eCO<sub>2</sub> dan TVOC dalam data sensor, dengan titik-titik yang dikelompokkan berdasarkan kategori label bau yaitu Segar, Sedang, dan Busuk. Setiap titik mewakili satu sampel data, dengan posisi horizontal menunjukkan nilai eCO<sub>2</sub> dan posisi vertikal menunjukkan nilai TVOC. Warna titik-titik dibedakan berdasarkan label bau, sehingga memudahkan dalam melihat pola distribusi dan keterkaitan antara dua parameter sensor tersebut terhadap klasifikasi bau. Dari grafik ini dapat diamati apakah terdapat pemisahan visual antar kategori bau yang jelas, yang juga mencerminkan kekuatan fitur eCO<sub>2</sub> dan TVOC dalam membantu proses klasifikasi bau oleh model.

### Visualisasi Rata-rata eCO<sub>2</sub> terhadap Tanggal dan Objek

Visualisasi ini menggunakan heatmap untuk menunjukkan bagaimana nilai rata-rata eCO<sub>2</sub> berubah terhadap masing-masing objek makanan dari hari ke hari. Setiap sel menunjukkan rata-rata nilai eCO<sub>2</sub> untuk objek tertentu pada tanggal tertentu. Dengan visualisasi ini, pola-pola tren peningkatan atau penurunan kadar gas eCO<sub>2</sub> bisa terlihat, yang dapat dihubungkan dengan proses pembusukan atau degradasi kualitas bahan makanan. Pendekatan ini membantu melihat dinamika bau secara waktu dan objek.



Gambar 7. Visualisasi Rata-rata eCO<sub>2</sub> terhadap Tanggal dan Objek

Grafik heatmap menggambarkan rata-rata nilai eCO<sub>2</sub> untuk setiap objek pada masing-masing tanggal pengambilan data. Setiap sel dalam heatmap menunjukkan intensitas nilai eCO<sub>2</sub> dengan gradasi warna dari kuning muda hingga merah tua, di mana warna yang lebih gelap menunjukkan konsentrasi eCO<sub>2</sub> yang lebih tinggi. Grafik ini memungkinkan analisis visual terhadap perubahan tingkat eCO<sub>2</sub> dari waktu ke waktu untuk setiap objek yang diamati, sehingga dapat membantu dalam mengidentifikasi pola atau tren bau berdasarkan akumulasi gas yang diukur. Dengan menggunakan format tabel yang disusun dalam bentuk matriks objek dan tanggal, heatmap ini efektif untuk membandingkan performa kebersihan atau kesegaran antara berbagai sampel kulkas pada periode yang berbeda.

### Evaluasi Model

Evaluasi model dilakukan untuk mengukur performa algoritma Random Forest Classifier dalam mengklasifikasikan bau berdasarkan data sensor. Dataset dibagi menjadi dua bagian:

- 80% untuk pelatihan (training set)
- 20% untuk pengujian (testing set)

Setelah pelatihan, model dievaluasi menggunakan data uji (testing set) dan menghasilkan metrik performa sebagai berikut:

#### 1. Akurasi

Akurasi serupakan ukuran seberapa banyak prediksi model yang benar dibandingkan dengan seluruh jumlah data yang diuji.

Akurasi: 97.8 %  
 Mean Squared Error: 0.044  
 Cohen's Kappa Score: 0.96

Gambar 8. Akurasi

Akurasi 97.8% menunjukkan bahwa model mampu memprediksi bau dengan sangat tepat. Nilai MSE 0.044 berarti kesalahan prediksi model sangat kecil. Sedangkan Cohen's Kappa 0.96 menunjukkan bahwa prediksi model sangat konsisten dan hampir selalu sesuai dengan label sebenarnya, bahkan setelah memperhitungkan kemungkinan tebakan acak.

2. Precision, Recall, F1-Score

Precision, Recall, dan F1-Score merupakan metrik evaluasi untuk menilai kinerja model klasifikasi. Precision menunjukkan seberapa tepat model saat memprediksi suatu kelas, Recall menunjukkan seberapa baik model menangkap semua data dari kelas tersebut, dan F1-Score adalah rata-rata harmonik dari keduanya yang memberikan gambaran seimbang antara akurasi dan cakupan prediksi.

```

Classification Report:
              precision    recall  f1-score   support

   Busuk         1.00        0.97        0.99         74
   Sedang         0.93        1.00        0.96         38
   Segar         1.00        0.96        0.98         23

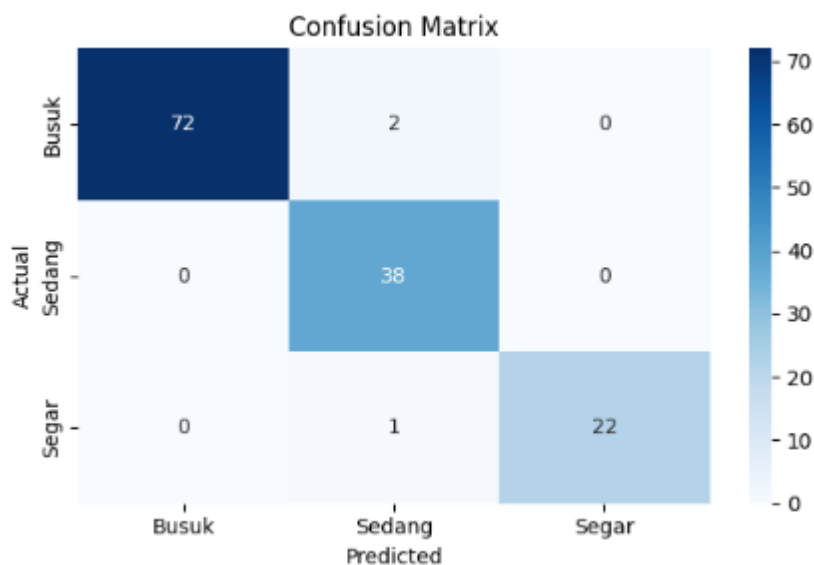
 accuracy          0.98
 macro avg         0.98        0.98        0.98         135
 weighted avg     0.98        0.98        0.98         135
    
```

Gambar 9. Precision, Recall, F1-Score

Berdasarkan classification report di atas, model Random Forest menunjukkan performa yang sangat baik dalam mengklasifikasikan bau ke dalam tiga kategori: Busuk, Sedang, dan Segar. Kategori Busuk memiliki precision dan recall hampir sempurna (1.00 dan 0.97), menunjukkan model sangat akurat dalam mengenali bau busuk. Kategori Sedang juga diklasifikasi dengan baik, dengan precision 0.93 dan recall 1.00, menandakan hampir semua data "Sedang" berhasil dikenali tanpa kesalahan. Sementara itu, kategori Segar memiliki precision sempurna (1.00), namun recall hanya 0.96, artinya sebagian data Segar masih salah diklasifikasi sebagai kelas lain. Secara keseluruhan, akurasi model mencapai 98%, dengan skor rata-rata tertimbang (weighted avg) precision, recall, dan F1-score juga tinggi di angka 0.98, menandakan model bekerja secara konsisten dan andal di semua kelas.

3. Confusion Matrix

Confusion matrix menggambarkan jumlah prediksi benar dan salah antar kelas, Contoh visualisasi Confusion Matrix:



Gambar 10. Confusion Matrix

Visualisasi confusion matrix pada Gambar 4.11 menunjukkan performa model Random Forest dalam mengklasifikasikan bau ke dalam tiga kategori: Segar, Sedang, dan Busuk. Setiap sel pada matriks merepresentasikan jumlah prediksi model terhadap label tertentu dibandingkan dengan label sebenarnya. Warna biru yang lebih gelap di diagonal utama (kiri atas ke kanan bawah) menunjukkan jumlah prediksi yang benar pada masing-masing kelas, sedangkan warna yang lebih terang di luar diagonal menunjukkan kesalahan prediksi. Semakin gelap warna pada diagonal, semakin baik performa model dalam mengenali kategori tersebut. Berdasarkan hasil visualisasi ini, dapat disimpulkan bahwa model memiliki tingkat akurasi yang tinggi dalam membedakan jenis bau, dengan sedikit kesalahan klasifikasi antar kategori.

## KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa: Sistem klasifikasi bau berbasis IoT menggunakan sensor SGP-30 dan algoritma Random Forest berhasil diimplementasikan dan berfungsi dengan baik. Algoritma Random Forest menunjukkan performa yang tinggi dengan akurasi mencapai 97.78 % dalam mengklasifikasi kondisi bau menjadi tiga kelas: Segar, Sedang, dan Busuk. Sistem ini efektif untuk mengidentifikasi kondisi bau dalam kulkas secara real time.

## DAFTAR PUSTAKA

### Buku

- [1] Indah Purnama Sari. Algoritma dan Pemrograman. Medan: UMSU Press, 2023, pp. 290.
- [2] Indah Purnama Sari. Buku Ajar Pemrograman Internet Dasar. Medan: UMSU Press, 2022, pp. 300.
- [3] Indah Purnama Sari. Buku Ajar Rekayasa Perangkat Lunak. Medan: UMSU Press, 2021, pp. 228.
- [4] Janner Simarmata Arsan Kumala Jaya, Syarifah Fitrah Ramadhani, Niel Ananto, Abdul Karim, Betrisandi, Muhammad Ilham Alhari, Cucut Susanto, Suardinata, Indah Purnama Sari, Edson Yahuda Putra. Komputer dan Masyarakat. Medan: Yayasan Kita Menulis, 2024, pp.162.
- [5] Mahdianta Pandia, Indah Purnama Sari, Alexander Wirapraja Fergie Joanda Kaunang, Syarifah Fitrah Ramadhani Stenly Richard Pungus, Sudirman, Suardinata Jimmy Herawan Moedjahedy, Elly Warni, Debby Erce Sondakh. Pengantar Bahasa Pemrograman Python. Medan : Yayasan Kita Menulis, 2024, pp.180
- [6] Zelvi Gustiana Arif Dwinanto, Indah Purnama Sari, Janner Simarmata Mahdianta Pandia, Supriadi Syam, Semmy Wellem Taju Fitrah Eka Susilawati, Asmah Akhriana, Rolly Junius Lontaan Fergie Joanda Kaunang. Perkembangan Teknologi Informatika. Medan: Yayasan Kita Menulis, 2024, pp.158

### Jurnal

- [7] Adriansyah, I., Mahendra, M. D., Rasywir, E., & Pratama, Y. (2022). Perbandingan Metode Random Forest Classifier dan SVM Pada Klasifikasi Kemampuan Level Beradaptasi Pembelajaran Jarak Jauh Siswa. *Bulletin of Informatics and Data Science*, 1(2), 98. <https://doi.org/10.61944/bids.v1i2.49>
- [8] Ahmad Hania, A. (2017). Mengenal Artificial Intelligence, Machine Learning, & Deep Learning. *Jurnal Teknologi Indonesia*, 1(June), 1–6. <https://amt-it.com/mengenal-perbedaan-artificial-intelligence-machine-learning-deep-learning/>
- [9] Sari, I.P., Al-Khowarizmi, A.K., Apdilah, D., Manurung, A.A., & Basri, M. (2023). Perancangan Sistem Pengaturan Suhu Ruangan Otomatis Berbasis Hardware Mikrokontroler Berbasis AVR. *sudo Jurnal Teknik Informatika* 2 (3), 131-142
- [10] Wardani., S, & Dewantoro., RW. (2024). Internet of Things: Home Security System based on Raspberry Pi and Telegram Messenger. *Indonesian Journal of Applied Technology, Computer and Science* 1 (1), 7-13
- [11] Aira, A., Maharani, P., Sakti, R. H., Fajar, M., Haq, I., & Ajis, M. (2024). Air Quality Classification System using Random Forest Algorithm using MQ-7 and MQ-135 Sensors with IoT-based. 1(2), 65–74.
- [12] Al-Okby, M. F. R., Neubert, S., Roddelkopf, T., Fleischer, H., & Thurow, K. (2022). Evaluating of IAQ-index and TVOC As Measurements Parameters for IoT-based Hazardous Gases Detection and Alarming Systems. *SAMI 2022 - IEEE 20th Jubilee World Symposium on Applied Machine Intelligence and Informatics, Proceedings*, 291–298. <https://doi.org/10.1109/SAMI54271.2022.9780751>
- [13] Sari, I.P., Al-Khowarizmi, A.K., Hariani, P.P., Perdana, A., & Manurung, A.A. (2023). Implementation And Design of Security System On Motorcycle Vehicles Using Raspberry Pi3-Based GPS Tracker And Facedetection. *Sinkron: jurnal dan penelitian teknik informatika* 8 (3), 2003-2007

- [14] Sari, I.P., Basri, M., Ramadhani, F., & Manurung, A.A. (2023). Penerapan Palang Pintu Otomatis Jarak Jauh Berbasis RFID di Perumahan. *Blend Sains Jurnal Teknik 2* (1), 16-25
- [15] Anderson, J., Johnson, E., & Brown, M. (2024). IoT , Anomaly Detection , Machine Learning , K-Nearest Neighbors , Random Forest , Real-Time Detection. *1*(1), 1–6.
- [16] Budiman, H. et al. (2021). Implementasi Sensor Gas untuk Smart Refrigerator. *Jurnal Teknologi.*, 12(2), 5–10.
- [17] Sari, I.P., & Batubara, I.H. (2020). Aplikasi Berbasis Teknologi Raspberry Pi Dalam Manajemen Kehadiran Siswa Berbasis Pengenalan Wajah. *JMP-DMT 1* (4), 6
- [18] M. Saleh and M. Haryanti, "Design of a Home Security System Using Relays", *Journal of Electrical Technology, Mercu Buana University*, Vol. 8 No. May 2, 2017
- [19] Sari, I.P., Novita, A., Al-Khowarizmi, A., Ramadhani, F., & Satria, A. (2024). Pemanfaatan Internet of Things (IoT) pada Bidang Pertanian Menggunakan Arduino UnoR3. *Blend Sains Jurnal Teknik 2* (4), 337-343
- [20] Husaini, A., & Sari, I.P. (2023). Konfigurasi dan Implementasi RB750Gr3 sebagai RT-RW Net pada Dusun V Suka Damai Desa Sei Meran. *sudo Jurnal Teknik Informatika 2* (4), 151-158
- [21] Chyan, P., Arni, S., & Thayf, M. S. S. (n.d.). *Pengantar Machine Learning PT . MIFANDI MANDIRI DIGITAL.*
- [22] Sari, I.P., Batubara, I.H., & Al-Khowarizmi, A. (2021). Sensitivity Of Obtaining Errors In The Combination Of Fuzzy And Neural Networks For Conducting Student Assessment On E-Learning. *International Journal of Economic, Technology and Social Sciences (Injects) 2* (1), 331-338
- [23] Sari, I.P., Al-Khowarizmi, A., & Batubara, I.H. (2021). Cluster Analysis Using K-Means Algorithm and Fuzzy C-Means Clustering For Grouping Students' Abilities In Online Learning Process. *Journal of Computer Science, Information Technology and Telecommunication Engineering 2* (1), 139-144
- [24] Erlangga, F., & Sari, I.P. (2024). Perancangan Sistem Untuk Merekomendasikan Produk Skincare Menggunakan Metode NLP. *Portal Riset dan Inovasi Sistem Perangkat Lunak 2* (4), 1-11
- [25] Bisono, A.T., & Zulherry, A. (2025). Analisis Sentimen Game Genshin Impact Untuk Mengetahui Reaksi Dan Harapan Pemain Menggunakan Metode Naïve Bayes. *Sudo Jurnal Teknik Informatika 4*(2), 183-193
- [26] Zulherry, A., Gunawan, T.S., & Wanayumini, W. (2021). Analisis Hasil Pendukung Keputusan Mendapatkan Rumah Dinas Perusahaan Menggunakan Metode Analytical Hierarcy Process (AHP) Dan Teknik For Order Referenci By Similarity (Topsis). *Media Informatika Budi Darma 5*(2), 695-704
- [27] Degeng, I. W., & Santoso, M. (2018). Smart Kulkas dengan Fitur SMS untuk Melaporkan Ketersediaan Bahan Makanan. *STRING (Satuan Tulisan Riset Dan Inovasi Teknologi)*, 3(1), 26. <https://doi.org/10.30998/string.v3i1.2724>