

BAB I PENDAHULUAN

1.1 Latar Belakang

Apel (*Malus domestica*) adalah buah yang memiliki tingkat popularitas tinggi di seluruh dunia (Nuples, 2022). Di banyak negara, apel menjadi pilihan utama konsumen karena rasanya yang manis, asam, dan segar, serta kandungan gizi yang beragam. Buah ini mengandung vitamin C, serat, kalium, dan berbagai senyawa antioksidan yang bermanfaat untuk meningkatkan sistem imun tubuh dan mencegah berbagai penyakit degeneratif seperti penyakit jantung, diabetes tipe 2, dan kanker. Selain itu, apel juga dapat membantu dalam menurunkan berat badan, meningkatkan kesehatan pencernaan, dan memperbaiki metabolisme tubuh. Oleh karena itu, apel sering dianggap sebagai simbol makanan sehat, yang menjadikannya buah yang banyak dibudidayakan, termasuk di Indonesia (Pamungkasih dkk., dan Panunggul dkk., 2023).

Di Indonesia, apel banyak ditanam di daerah dengan iklim sejuk, seperti Malang, Batu, dan wilayah lain di Jawa Timur, yang dikenal menghasilkan apel berkualitas tinggi. Apel yang tumbuh di daerah ini memiliki kualitas unggul karena kondisi iklim yang mendukung, seperti suhu rendah dan curah hujan yang cukup (Pamungkasih, dkk., 2023). Wilayah terbesar penghasil buah apel di Indonesia adalah Jawa Timur, dengan jumlah produksi mencapai 480.834 ton pada tahun 2019 dan terus meningkat ditahun-tahun selanjutnya (Badan Pusat Statistik 2019-2022). Namun, meskipun apel mudah ditemukan di pasar, kualitasnya sangat bervariasi, dipengaruhi oleh banyak faktor, baik selama pertumbuhan di kebun maupun pada proses pascapanen (Tahir, 2023).

Kualitas buah apel sering dikategorikan berdasarkan beberapa parameter fisik yang jelas terlihat oleh produsen atau pedagang. Beberapa faktor yang dipertimbangkan dalam penilaian kualitas apel meliputi ukuran, warna, bentuk, tekstur kulit, kekerasan buah, serta adanya cacat pada kulit buah akibat kerusakan alami atau kerusakan mekanis selama transportasi. Apel berkualitas baik biasanya memiliki ukuran yang seragam, kulit yang cerah dan merata, serta bentuk yang utuh tanpa cacat. Buah apel yang bebas dari kerusakan atau penyakit serta memiliki rasa

yang enak umumnya dihargai lebih tinggi (Brahmantyo, dkk., 2022).

Sebaliknya, apel dengan cacat fisik seperti goresan, bekas benturan, atau warna kulit yang tidak merata biasanya dianggap sebagai apel kualitas rendah. Apel yang terlalu matang atau memiliki tekstur lembek juga dinilai rendah dan tidak cocok untuk konsumsi langsung. Oleh karena itu, proses pengklasifikasian kualitas apel menjadi penting bagi produsen dan pedagang dalam menetapkan harga jual serta memenuhi harapan konsumen (Indrasari, 2019).

Proses pengklasifikasian kualitas apel biasanya dilakukan secara manual, di mana pekerja memeriksa setiap buah apel secara visual untuk menentukan apakah buah tersebut memenuhi standar kualitas tertentu. Meskipun metode ini sederhana dan telah digunakan lama, terdapat beberapa kelemahan yang dapat mempengaruhi akurasi dan efektivitas proses klasifikasi (Amrullah, 2017).

Salah satu kelemahan utama adalah ketergantungan pada tenaga manusia yang banyak dan memakan waktu. Sebuah fasilitas pengolahan apel besar mungkin membutuhkan banyak pekerja untuk memeriksa dan mengklasifikasikan apel, yang tentu saja menambah biaya operasional dan waktu yang dibutuhkan. Selain itu, kesalahan manusia yang terjadi karena kelelahan atau kurangnya konsentrasi bisa mengarah pada kesalahan dalam penilaian kualitas, yang berdampak pada distribusi yang tidak tepat, misalnya, apel kualitas rendah yang seharusnya dibuang malah dijual dengan harga lebih tinggi (Finda & Utomo, 2024).

Seiring dengan meningkatnya permintaan pasar dan ekspektasi konsumen yang semakin tinggi, diperlukan sistem otomatis yang dapat mengklasifikasikan kualitas apel dengan lebih cepat, akurat, dan efisien. Kemajuan dalam teknologi kecerdasan buatan (AI) dan pembelajaran mesin (*machine learning*) dapat menjadi solusi untuk mengatasi masalah yang timbul dari sistem klasifikasi manual (Dwi, 2023 dan Agustyaningru, dkk., 2020). Dengan menggunakan algoritma machine learning, kualitas apel dapat diprediksi secara otomatis dengan menganalisis berbagai fitur fisik yang dimilikinya.

Pembelajaran mesin adalah cabang dari kecerdasan buatan yang memungkinkan sistem untuk belajar dari data, menemukan pola tersembunyi, dan membuat keputusan berdasarkan informasi tersebut. Salah satu keunggulannya

adalah kemampuan untuk memproses data dalam jumlah besar secara cepat dan efisien, serta menghasilkan hasil yang akurat meski data tersebut sangat kompleks (Mardiani, dkk., 2023). Dalam pengklasifikasian kualitas apel, berbagai algoritma pembelajaran mesin dapat digunakan untuk mengevaluasi karakteristik fisik apel seperti ukuran, bentuk, tekstur, dan warna kulit, guna menentukan kualitasnya (Purnamawati, dkk., 2020).

Salah satu algoritma pembelajaran mesin yang populer dalam klasifikasi adalah *Decision Tree*. Algoritma ini bekerja dengan membagi data menjadi subset lebih kecil berdasarkan atribut yang dianggap penting untuk memisahkan data ke dalam kelas yang berbeda. Model ini menghasilkan struktur pohon keputusan, di mana setiap simpul mewakili pengujian terhadap fitur tertentu, dan cabangnya menggambarkan hasil dari pengujian tersebut (Permana, dkk., 2021).

Kelebihan utama dari *Decision Tree* adalah kemudahan pemahaman dan interpretasi model. Struktur pohon memungkinkan pengguna untuk melihat langkah-langkah yang diambil oleh model dalam membuat keputusan. Namun, salah satu kelemahannya adalah kecenderungan untuk mengalami *overfitting*, yaitu ketika pohon keputusan menjadi terlalu rumit dan menyesuaikan diri terlalu banyak dengan data pelatihan, sehingga kurang mampu menggeneralisasi data baru (Lishania, dkk., 2020).

Random Forest menggabungkan banyak pohon keputusan untuk menghasilkan model yang lebih kuat dan stabil. Dalam algoritma ini, setiap pohon keputusan dibangun menggunakan subset acak dari data pelatihan, dan hasil akhirnya diperoleh dengan menggabungkan hasil dari setiap pohon (Adi & Wintarti, 2022 dan Agustyaningrum, dkk., 2020).

Keunggulan utama dari *Random Forest* adalah kemampuannya untuk mengurangi risiko *overfitting* dan lebih akurat dalam menghadapi masalah yang lebih kompleks. Dengan banyaknya pohon keputusan yang digunakan, *Random Forest* juga lebih tahan terhadap fluktuasi data dan lebih mampu menggeneralisasi data baru. Selain itu, *Random Forest* juga dapat menangani data yang hilang dan masalah data tidak seimbang (Syukron & Subekti, 2018; Jali, dkk., 2024).

Meskipun keduanya memiliki tujuan yang sama dalam mengklasifikasi data,

Decision Tree dan *Random Forest* memiliki perbedaan yang signifikan. *Decision Tree* lebih sederhana, cepat dilatih, dan mudah dipahami, tetapi rentan terhadap *overfitting*. Sementara itu, *Random Forest* lebih tahan terhadap *overfitting* dan lebih stabil, namun memerlukan lebih banyak waktu dan sumber daya komputasi (Noviant, dkk., 2023).

Oleh karena itu, penting untuk melakukan analisis perbandingan antara kedua algoritma ini dalam konteks klasifikasi kualitas buah apel, guna menentukan algoritma mana yang lebih efektif dan efisien dalam menghasilkan prediksi yang akurat (Kanugrahan, dkk., 2024).

Penelitian ini akan mengevaluasi dan membandingkan kinerja algoritma *Decision Tree* dan *Random Forest* dalam klasifikasi kualitas buah apel, serta menilai kedua algoritma berdasarkan metrik evaluasi seperti akurasi, presisi, *recall*, dan *F1-score*. Penelitian ini juga akan mengeksplorasi bagaimana pemilihan fitur, teknik pemrosesan data, dan pembagian data dengan rasio. (Dwi, 2023 dan Permana, dkk., 2021).

Penelitian ini diharapkan dapat memberikan kontribusi signifikan bagi industri pertanian, terutama dalam meningkatkan efisiensi dan akurasi proses klasifikasi kualitas buah apel. Dengan memanfaatkan algoritma pembelajaran mesin, proses klasifikasi dapat diotomatisasi, mengurangi ketergantungan pada tenaga kerja manual, dan meminimalkan kesalahan yang disebabkan oleh faktor manusia. Penelitian ini juga dapat menjadi acuan bagi industri pertanian untuk mengadopsi teknologi canggih guna meningkatkan daya saing dan keberlanjutan dalam industri pertanian global (Panunggul, dkk., 2023 dan Lintang & Layuk, 2020).

1.2 Rumusan Masalah

1. Bagaimana perbandingan kinerja algoritma *Decision Tree* dan *Random Forest* dalam klasifikasi kualitas buah apel?
2. Algoritma mana yang memberikan akurasi lebih tinggi dalam klasifikasi kualitas buah apel?
3. Rasio pembagian data manakah yang menghasilkan performa paling tinggi pada masing – masing algoritma ?

1.3 Tujuan Penelitian

1. Untuk mengetahui perbandingan kinerja algoritma *Decision Tree* dan *Random Forest* dalam klasifikasi kualitas buah apel.
2. Untuk mengetahui dan mengidentifikasi algoritma mana yang memberikan akurasi lebih tinggi dalam klasifikasi kualitas buah apel.
3. Untuk mengetahui rasio pembagian data yang menghasilkan performa paling tinggi pada masih – masing algoritma

1.4 Manfaat Penelitian

Manfaat yang diambil dari penelitian ini sebagai berikut :

1. Memberikan pengetahuan tentang algoritma *Decision Tree* dan *Random Forest* dalam konteks pengolahan citra dan data untuk klasifikasi kualitas buah, serta mengungkap kekuatan dan kelemahan kedua algoritma tersebut.
2. Dapat memberikan solusi otomatis berbasis machine learning untuk klasifikasi buah apel menggantikan proses manual yang memakan waktu dan tenaga, sehingga lebih cepat.
3. Dengan adanya model klasifikasi, dapat menjadi dasar pengembangan sistem otomatis dalam klasifikasi buah apel untuk meningkatkan efisiensi sortir.
4. Dengan akurasi klasifikasi dapat membantu memastikan buah apel yang dipasarkan memenuhi standar kualitas, meningkatkan kepuasan konsumen.

1.5 Batasan Penelitian

Batasan yang dapat diambil dari penelitian ini sebagai berikut

1. Data yang digunakan dalam penelitian ini bersumber dari dataset publik *apple Quality* yang tersedia di Kaggle <https://www.kaggle.com/datasets/nelgiriyewithana/apple-quality> yang diambil pada bulan Mei 2025.
2. Dataset yang digunakan berjumlah 4001 baris dan 9 kolom.
3. Fitur yang digunakan, seperti warna, tekstur, ukuran, dan bentuk buah apel.
4. Analisis data dilakukan menggunakan algoritma *machine learning*, yaitu *decision Tree* dan *Random Forest* dengan Python.
5. Pengukuran kinerja model dilakukan dengan menggunakan matrik akurasi, presisi, recall, dan F1-score.

