

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pertanian memiliki peran strategis dalam meningkatkan kesejahteraan masyarakat, khususnya di pedesaan. Dengan menyediakan lapangan kerja yang cukup luas, sektor pertanian membantu mengurangi tingkat pengangguran dan kemiskinan. Selain itu, pertanian juga berkontribusi dalam menjaga ketahanan pangan nasional, memastikan ketersediaan pangan yang cukup dan berkualitas bagi seluruh penduduk. Lebih jauh lagi, pertanian yang berkelanjutan dapat melestarikan lingkungan dan sumber daya alam, sehingga memberikan manfaat jangka panjang bagi generasi mendatang. Perkebunan apel merupakan salah satu sektor pertanian yang memiliki nilai ekonomi tinggi dan berkontribusi besar terhadap sektor agribisnis di banyak negara.[1]

Apel adalah salah satu sumber pangan yang penting di seluruh dunia. Apel dapat digunakan sebagai bahan baku untuk berbagai jenis makanan, seperti saus, kue, jus, dan camilan. Selain itu, apel kaya akan nutrisi yang bermanfaat bagi kesehatan, seperti serat, vitamin, mineral, dan antioksidan. Jenis apel yang akan dibahas memiliki beragam varietas, seperti Granny Smith, Fuji, Honeycrisp, Gala, Pink Lady, dan Red Delicious. Buah apel (*Malus domestica*) termasuk dalam keluarga tanaman Rosaceae, yang menghasilkan buah-buahan berbiji, dan menjadi salah satu komoditas buah dengan produksi tertinggi di dunia.

Apel merupakan sumber senyawa fenolik yang berperan penting dalam berbagai proses fisiologis dan metabolik pada manusia. Selain itu, apel memiliki peran signifikan dalam sektor pertanian di Indonesia. Apel dikenal kaya akan nutrisi yang bermanfaat bagi kesehatan, seperti serat, vitamin C, kalium, dan antioksidan. Buah ini juga dapat diolah menjadi berbagai produk makanan, seperti jus, selai, keripik, dan bahan tambahan pada kue. Namun, banyak petani dan pedagang di Indonesia yang kesulitan menentukan apakah apel yang dihasilkan memiliki kualitas yang baik atau kurang layak untuk dijual. Penelitian ini bertujuan untuk membantu mengklasifikasikan apel berdasarkan kualitasnya, baik atau buruk untuk

dijual. Dengan begitu, para petani dan pedagang dapat lebih mudah memastikan kualitas apel dan meningkatkan nilai pasar, sehingga mendukung pertumbuhan sektor pertanian dan perdagangan apel di Indonesia.

Klasifikasi adalah pengelompokan data atau objek baru ke dalam kelas atau label berdasarkan atribut-atribut tertentu[2], Penerapan teknik klasifikasi dalam industri apel memungkinkan untuk membangun sistem otomatis yang mampu memilah apel berdasarkan kualitasnya secara akurat dan efisien. Sistem ini akan menganalisis berbagai karakteristik fisik apel, seperti warna, ukuran, dan bentuk, untuk mengklasifikasikannya menjadi beberapa kategori kualitas. Dengan demikian, produsen dapat memastikan bahwa produk yang dipasarkan kepada konsumen memiliki kualitas yang konsisten dan memenuhi standar yang ditetapkan. Selain itu, data yang dihasilkan dari proses klasifikasi dapat digunakan untuk mengoptimalkan proses produksi, meningkatkan efisiensi, dan mengambil keputusan bisnis yang lebih baik.

Dalam penelitian ini data yang diperoleh memiliki kelas atau label yang cukup banyak lebih tepatnya 7 label sehingga dibutuhkan metode yang dapat menghitung multiclass data. Salah Jika Decision Tree dapat diibaratkan sebagai satu pohon yang mengambil keputusan, maka Random Forest adalah sekelompok pohon yang bekerja sama untuk mencapai keputusan yang lebih akurat. Setiap pohon dalam hutan ini dibangun berdasarkan sampel data yang berbeda secara acak dari dataset asli. Dengan cara ini, setiap pohon akan memiliki karakteristik yang unik dan mampu menangkap pola yang berbeda-beda dalam data. Proses klasifikasi dengan Random Forest dimulai dengan membagi dataset menjadi beberapa bagian yang lebih kecil. Setiap bagian ini kemudian digunakan untuk membangun satu pohon keputusan. Setelah semua pohon selesai dibangun, setiap pohon akan memberikan suara atau prediksi terhadap data yang baru. Keputusan akhir diambil berdasarkan mayoritas suara dari seluruh pohon dalam hutan. Dengan menggabungkan banyak pohon keputusan, Random Forest mampu mengurangi risiko overfitting (model terlalu spesifik pada data latih) dan meningkatkan generalisasi model[3].

Salah satu alasan utama pemilihan metode ini adalah kemampuannya menghasilkan keputusan yang dapat diandalkan dan dijelaskan. Model yang dibangun melalui metode ini tidak hanya memberikan prediksi, tetapi juga menjelaskan alasan di balik prediksi tersebut. Hal ini sangat penting dalam berbagai aplikasi, terutama ketika keputusan yang diambil memiliki konsekuensi yang signifikan. Selain itu, transparansi model memungkinkan untuk mengidentifikasi potensi bias atau kesalahan dalam data dan model itu sendiri, klasifikasi apel ini terdiri dari 7 atribut dan 4000 data, dikarenakan atribut banyak maka akan dilakukan seleksi fitur yang paling mempengaruhi korelasi dengan atribut lainnya.

Dalam penelitian tahun 2024 menerapkan algoritma C4.5 untuk membangun model klasifikasi jenis kacang yang berasal dari negara Turkey. Dengan menggunakan teknik SMOTE untuk mengatasi ketidakseimbangan data, data yang diteliti memiliki total data berjumlah 11361 dan memiliki atribut 16 kolom serta 1 sebagai label, label atau target berjumlah 7 buah label sehingga cocok menggunakan decision tree c45 sebagai klasifikasi data multiclass, beliau berhasil mencapai tingkat akurasi, presisi, recall, dan F1-score sebesar 94%. Hasil ini menunjukkan kinerja model yang sangat baik dalam mengklasifikasikan berbagai jenis kacang[4].

Beberapa penelitian sebelumnya yaitu pada tahun 2020 penelitian tentang gizi bayi dengan klasifikasi, data yang digunakan berisi data 3 atribut dan memiliki 2 kelas, penelitian ini digunakan untuk dapat melakukan pemeriksaan dan pengukuran antropometri agar dapat mengetahui status gizi balita dapat mendatangi tempat pelayanan kesehatan masyarakat dan menghasilkan perhitungan software RapidMiner menunjukkan akurasi menggunakan tools performance adalah 100% [5]. dan juga pada penelitian tahun 2021 yang bertema tentang penerapan *decision tree* pada produk terlaris, penelitian ini menggunakan data dengan atribut dan 2 kelas. Dari hasil penelitian yang telah dilakukan untuk data penjualan dan stok barang di UD. Cipta Karya Gorontalo maka diperoleh akurasi sebesar 90% dan nilai AUC 0.709 [6].

1.2 Rumusan masalah

1.2.1 Permasalahan

Dengan memanfaatkan 8 ciri khas dari 4000 data apel, sistem berusaha membuat aturan untuk membedakan jenis apel. Sebelumnya, data perlu disiapkan dengan baik agar sistem dapat mengenali pola di dalamnya. Proses persiapan ini meliputi pemahaman data dan normalisasi data, sehingga mencapai hasil yang optimal.

1.2.2 Metode Usulan

Dari permasalahan yang telah dijelaskan, penelitian ini menggunakan metode Random Forest karena kemampuannya dalam menangani data yang kompleks dan memberikan hasil yang akurat dengan tingkat overfitting yang lebih rendah. Metode ini bekerja dengan membangun banyak pohon keputusan (decision trees) dari dataset yang sama, namun dengan menggunakan teknik bootstrap sampling dan seleksi fitur secara acak untuk setiap pohon. Hasil akhir diperoleh dengan menggabungkan prediksi dari semua pohon, baik melalui voting untuk klasifikasi maupun rata-rata untuk regresi. Proses ini meningkatkan stabilitas model sekaligus memastikan interpretasi yang tetap dapat dimengerti.

1.2.3 Pertanyaan Penelitian

1. Bagaimana proses klasifikasi data apel menggunakan metode Random Forest?
2. Berapa hasil dari evaluasi dengan menggunakan Confusion Matrix terhadap hasil performa (akurasi, presisi, *recall*, *f1-score*) sistem?

1.3 Tujuan dan manfaat

1.3.1 Tujuan Penelitian

1. Penelitian ini bertujuan untuk mengevaluasi kinerja algoritma Random Forest dalam mengklasifikasikan kualitas buah apel. Dengan menggunakan sejumlah besar data apel sebagai sampel, kemudian akan menganalisis seberapa akurat algoritma ini dapat mengidentifikasi kualitas berdasarkan karakteristik tertentu.
2. Dengan menggunakan sejumlah besar data apel sebagai sampel, akan menganalisis seberapa akurat algoritma ini dapat mengidentifikasi kualitas apel berdasarkan karakteristik tertentu. Selain itu, penelitian ini juga akan menguji pengaruh pemilihan fitur terhadap hasil klasifikasi, seperti akurasi, presisi, dan recall.

1.3.2 Manfaat Penelitian

Penelitian ini diharapkan dapat menjadi acuan bagi yang ingin mengklasifikasi data apel. Dengan menggunakan algoritma Random Forest, penelitian ini menunjukkan potensi besar dalam mengidentifikasi kualitas apel secara otomatis. Hasil penelitian ini dapat menjadi dasar untuk pengembangan sistem klasifikasi apel yang lebih canggih dan efisien, serta memberikan wawasan baru tentang penerapan algoritma Random Forest dalam bidang pertanian.

1.4 Batasan-Batasan

Batasan masalah didalam penelitian ini adalah sebagai berikut:

1. Data yang digunakan pada penelitian ini di upload oleh Nidula Elgiriye withana pada tanggal 11 January 2024 dan berikut adalah link download.

(<https://www.kaggle.com/datasets/nelgiriye withana/apple-quality>)

2. Data yang digunakan berjumlah 4000 data dengan 2 *class* dan 7 atribut data.
3. Data tiap *class* yaitu *good* berisi 2004 data, *Bad* memiliki total 1996 data.
4. Metode yang digunakan random forest dan confusion matrix

1.5 Sistematika Skripsi

Sistematika penulisan Skripsi ini terdiri dari lima bab, yaitu:

BAB I PENDAHULUAN

Pada bagian ini dibahas secara rinci mengenai latar belakang masalah yang diteliti, pertanyaan penelitian yang ingin dijawab, tujuan penelitian, manfaat yang diharapkan, serta bagaimana penelitian ini dilakukan dan dilaporkan.

BAB II KAJIAN PUSTAKA

Pada bagian ini, dijelaskan secara mendalam tentang teori-teori yang mendukung penelitian ini. Khususnya, dibahas bagaimana metode klasifikasi dapat digunakan untuk membedakan berbagai jenis apel berdasarkan karakteristik yang dimilikinya.

BAB III METODOLOGI PENELITIAN

Bab ini membahas kerangka kerja penelitian yang mencakup perencanaan, pelaksanaan, dan evaluasi. Di sini dijelaskan secara detail tentang metode yang digunakan, data yang diperlukan, serta langkah-langkah yang akan dilakukan untuk mencapai tujuan penelitian.

BAB IV HASIL DAN PEMBAHASAN

Bab ini menyajikan metodologi penelitian secara detail, mulai dari proses pengumpulan dan pengolahan data, pemilihan metode analisis, hingga interpretasi hasil. Ditekankan pada langkah-langkah perhitungan yang dilakukan untuk memastikan akurasi dan validitas hasil penelitian.

BAB V PENUTUP

Bab ini menyajikan hasil akhir dari penelitian, termasuk interpretasi terhadap data yang telah dianalisis. Selain itu, dibahas pula poin-poin penting yang dapat dijadikan landasan untuk penelitian lebih lanjut, serta saran bagi pembaca yang ingin menerapkan metode serupa dalam penelitiannya.

DAFTAR PUSTAKA

Bab ini menyajikan daftar lengkap publikasi ilmiah yang menjadi rujukan dalam penelitian ini. Daftar pustaka ini mencakup berbagai jurnal ilmiah, buku, dan sumber relevan lainnya yang telah memberikan landasan teori dan metodologi penelitian.