

# Aplikasi Prediksi Produksi Cabai dengan Algoritma C.45 untuk Dinas Pertanian Provinsi Aceh Berbasis Web

M Kadafi Afkar<sup>1</sup>, Muhammad Wali<sup>2</sup>, Imilda<sup>3\*</sup>

<sup>1,2,3\*</sup> Program Studi Manajemen Informatika, STMIK Indonesia Banda Aceh, Kota Banda Aceh, Provinsi Aceh, Indonesia.

*Corresponding Email:* [imilda@stmiki.ac.id](mailto:imilda@stmiki.ac.id)<sup>3\*</sup>

## Histori Artikel:

*Dikirim* 15 Januari 2024; *Diterima dalam bentuk revisi* 13 Februari 2024; *Diterima* 20 Maret 2024; *Diterbitkan* 30 Maret 2024. Semua hak dilindungi oleh Lembaga Penelitian dan Pengabdian Masyarakat (LPPM) STMIK Indonesia Banda Aceh.

## Abstrak

Penelitian ini mengembangkan aplikasi prediksi produksi cabai menggunakan algoritma C.45 untuk Dinas Pertanian Provinsi Aceh. Data cuaca dan harga cabai dari 2022-2023 digunakan untuk membangun model prediksi. Setelah melalui tahap seleksi data, preprocessing, dan transformasi, model algoritma C.45 berhasil dibangun dan diuji dengan data baru. Hasilnya menunjukkan akurasi prediksi sebesar 84.03%. Selanjutnya, model ini diimplementasikan dalam bentuk aplikasi web yang mencakup berbagai fitur seperti halaman publik, perhitungan hasil prediksi, grafik, dan halaman admin. Penelitian menunjukkan bahwa penggunaan algoritma C.45 dan aplikasi ini berpotensi mendukung pengambilan keputusan dan perencanaan pertanian di Aceh. Saran untuk penelitian selanjutnya meliputi penambahan variabel tambahan seperti kualitas tanah dan faktor sosial-ekonomi, serta eksplorasi teknologi penginderaan jauh dan metode machine learning lainnya untuk meningkatkan akurasi prediksi. Evaluasi dampak penggunaan aplikasi juga dianjurkan untuk mengukur efektivitasnya dalam mendukung pertanian berkelanjutan di Aceh.

**Kata Kunci:** Algoritma C.45; Prediksi Produksi Cabai; Dinas Pertanian; Aplikasi Web; Provinsi Aceh.

## Abstract

This research develops a chili production prediction application using the C.45 algorithm for the Aceh Province Agriculture Service. Weather data and chili prices from 2022-2023 are used to build a prediction model. After going through the data selection, preprocessing, and transformation stages, the C.45 algorithm model was successfully built and tested with new data. The results show a prediction accuracy of 84.03%. Furthermore, this model is implemented in the form of a web application which includes various features such as public pages, calculation of prediction results, graphs, and admin pages. Research shows that the use of the C.45 algorithm and this application has the potential to support agricultural decision making and planning in Aceh. Suggestions for further research include adding additional variables such as soil quality and socio-economic factors, as well as exploring remote sensing technology and other machine learning methods to increase prediction accuracy. Evaluation of the impact of using the application is also recommended to measure its effectiveness in supporting sustainable agriculture in Aceh.

**Keyword:** Algorithm C.45; Chili Production Prediction; Department of Agriculture; Web Application; Aceh Province.

## 1. Pendahuluan

Pertanian memainkan peran vital dalam menjaga ketahanan pangan dan mendukung ekonomi, terutama di wilayah pedesaan seperti Provinsi Aceh. Di Aceh, tanaman cabai menjadi salah satu komoditas unggulan yang memberikan kontribusi signifikan terhadap perekonomian dan keberlanjutan pertanian. Namun, dalam mengelola produksi cabai, para petani dan instansi terkait dihadapkan pada tantangan dalam memprediksi produksi secara akurat. Fluktuasi kondisi cuaca dan faktor-faktor lainnya seringkali membuat prediksi produksi menjadi sulit dan tidak pasti. Oleh karena itu, pengembangan metode yang efektif dalam memprediksi produksi cabai di Aceh menjadi sangat penting untuk meningkatkan efisiensi pertanian dan mendukung pengambilan keputusan yang lebih baik. Penelitian ini bertujuan untuk mengembangkan sebuah aplikasi perhitungan prediksi produksi usaha cabai yang berbasis web, dengan menggunakan metode algoritma C4.5. Aplikasi ini akan dikhususkan untuk Dinas Pertanian Provinsi Aceh sebagai salah satu upaya untuk memberikan solusi yang efektif dalam perencanaan produksi pertanian cabai. Metode algoritma C4.5 dipilih karena memiliki kemampuan yang cukup baik dalam menangani dataset kompleks dan menghasilkan aturan keputusan yang mudah dipahami. Dengan menerapkan algoritma ini, diharapkan dapat dibangun sebuah model prediksi yang dapat memberikan perkiraan produksi cabai di masa depan dengan tingkat akurasi yang tinggi.

Penelitian ini memiliki beberapa batasan yang perlu diperhatikan. Pertama, kami hanya akan mempertimbangkan prediksi produksi cabai di Aceh berdasarkan beberapa parameter tertentu seperti luas lahan, curah hujan, suhu udara, dan kelembaban udara. Faktor-faktor lain yang dapat mempengaruhi produksi cabai, seperti jenis tanah, pemupukan, dan teknik budidaya, tidak akan dimasukkan dalam analisis kami. Kedua, data yang akan kami gunakan dalam penelitian ini terbatas pada data produksi tanaman cabai selama periode tertentu, yaitu dari tahun 2018 hingga 2022. Data produksi tanaman cabai di luar periode tersebut tidak akan diperhitungkan dalam analisis kami. Ketiga, penelitian ini akan fokus pada penerapan metode algoritma C4.5 untuk memprediksi produksi cabai. Metode lain dalam bidang data mining atau machine learning tidak akan dibahas dalam penelitian ini. Keempat, keakuratan prediksi produksi cabai dapat dipengaruhi oleh kualitas dan validitas data yang digunakan. Oleh karena itu, kami akan menggunakan data produksi tanaman cabai yang diperoleh dari Dinas Pertanian Provinsi Aceh sebagai data acuan. Kelima, hasil prediksi yang dihasilkan oleh aplikasi ini bersifat perkiraan dan tidak dapat dijadikan patokan absolut dalam pengambilan keputusan.

Dalam rangka mencapai tujuan penelitian, kami merumuskan beberapa pertanyaan penelitian yang perlu dijawab. Pertama, bagaimana mengembangkan aplikasi perhitungan prediksi produksi cabai berbasis web menggunakan metode algoritma C.45 pada Dinas Pertanian Provinsi Aceh? Kedua, bagaimana melakukan preprocessing terhadap data produksi cabai untuk mengatasi outliers, missing value, dan ketidakseimbangan data? Ketiga, bagaimana mengimplementasikan algoritma C.45 dalam membangun model prediksi produksi cabai yang akurat? Keempat, bagaimana melakukan uji coba terhadap model prediksi produksi cabai menggunakan data cuaca dan harga yang tersedia? Dan kelima, bagaimana mengevaluasi keakuratan dan kinerja model prediksi produksi cabai berdasarkan hasil uji coba yang dilakukan?.

Tujuan dari penelitian ini adalah menciptakan sebuah aplikasi perhitungan prediksi produksi usaha cabai yang akurat dan dapat diandalkan, dengan menggunakan metode algoritma C4.5. Secara khusus, kami bertujuan untuk menerapkan metode Data Mining untuk memprediksi produksi tanaman cabai di Aceh berdasarkan beberapa parameter seperti jenis tanaman, curah hujan, suhu udara, kelembaban udara, dan faktor lain yang mempengaruhi produksi tanaman cabai di Aceh. Kami juga bertujuan untuk mengetahui faktor-faktor yang paling berpengaruh terhadap produksi tanaman cabai di Aceh melalui analisis data yang kami lakukan. Selain itu, kami ingin memberikan rekomendasi dalam perencanaan produksi pertanian di Aceh berdasarkan hasil prediksi produksi tanaman. Kami juga berharap penelitian ini dapat meningkatkan penggunaan teknologi informasi dan pengolahan data dalam bidang pertanian, serta memberikan kontribusi dalam pengembangan teknologi informasi dan pengolahan data di bidang pertanian secara umum.

Manfaat dari penelitian ini meliputi: pertama, memberikan informasi yang akurat mengenai prediksi produksi tanaman cabai di Aceh, sehingga dapat membantu Dinas Pertanian Provinsi Aceh dalam menentukan strategi produksi pertanian yang efektif dan efisien. Kedua, memberikan gambaran mengenai faktor-faktor yang paling berpengaruh terhadap produksi tanaman cabai di Aceh, sehingga dapat membantu dalam pengambilan keputusan dalam perencanaan produksi pertanian di masa depan. Ketiga, menambah wawasan dalam bidang teknologi informasi dan pengolahan data dalam pertanian. Dan keempat, memberikan kontribusi dalam pengembangan teknologi informasi dan pengolahan data di bidang pertanian secara umum. Dengan demikian, penelitian ini diharapkan dapat memberikan manfaat yang nyata dalam upaya meningkatkan efisiensi dan keberlanjutan sektor pertanian, khususnya dalam pengelolaan produksi tanaman cabai di Provinsi Aceh.

Data Mining adalah suatu teknologi yang memanfaatkan algoritma dan teknik statistik untuk mengekstraksi pola-pola menarik dari data besar. Tujuan dari Data Mining adalah untuk mengungkapkan pengetahuan baru atau informasi yang berguna dari data yang ada, sehingga dapat digunakan untuk mengambil keputusan yang lebih baik dan efektif. Proses penemuan pola atau informasi yang berguna dari suatu kumpulan data besar merupakan konsep dasar Data Mining. Data mining melibatkan teknik dan metode pengolahan data untuk menemukan pola yang tersembunyi, korelasi, dan hubungan yang berguna dalam pengambilan keputusan (Jollyta, Ramdhan, & Zarlis, 2020). Konsep data mining juga meliputi proses pra-pemrosesan data, seperti pemilihan atribut, pengurangan dimensi, dan transformasi data (Nofriansyah, 2015). Selain itu, konsep ini juga mencakup teknik-teknik seperti clustering, klasifikasi, asosiasi, dan regresi. Data mining memiliki peran penting dalam berbagai bidang, termasuk bisnis, ilmu pengetahuan, kesehatan, dan pertanian (Kusrini, & Taufiq, 2009).

Algoritma C4.5 merupakan salah satu algoritma yang digunakan dalam teknik Data Mining untuk membangun model pohon keputusan (decision tree) yang dapat digunakan untuk klasifikasi dan prediksi data. Algoritma ini digunakan dalam penelitian untuk membangun model prediksi produksi tanaman cabai di Aceh berdasarkan variabel-variabel yang relevan. Penelitian ini memanfaatkan algoritma C4.5 untuk melakukan analisis klasifikasi terhadap data historis produksi tanaman cabai di Aceh serta faktor-faktor yang mempengaruhi produksi tersebut. Beberapa penelitian telah dilakukan untuk menerapkan algoritma C4.5 dalam berbagai bidang, seperti pada klasifikasi data nilai mahasiswa, prediksi jenis transaksi nasabah pada koperasi simpan pinjam, dan analisis penerimaan karyawan (Rahmayuni, 2014; Widayu dkk, 2017; Asistiyasari & Baidawi, 2017).

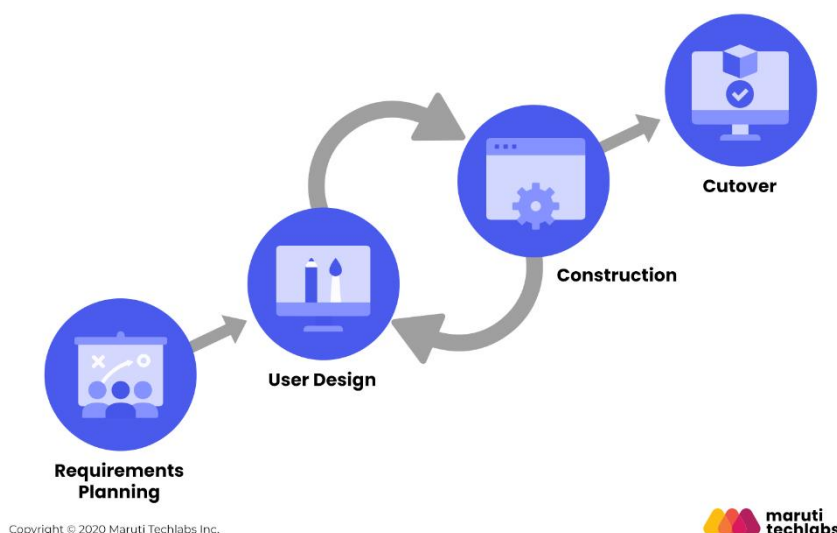
Aplikasi perhitungan prediksi produksi usaha cabai berbasis web menggunakan metode algoritma C4.5 merupakan solusi inovatif dalam mengoptimalkan perencanaan produksi pertanian cabai. Aplikasi ini didasarkan pada prinsip-prinsip teori yang melibatkan penggunaan metode Data Mining dan algoritma C4.5. Metode Data Mining digunakan untuk memproses data historis produksi tanaman cabai di Aceh serta faktor-faktor yang mempengaruhi produksi tersebut. Data historis ini digunakan sebagai basis untuk melatih model prediksi produksi menggunakan algoritma C4.5. Setelah model prediksi terbentuk, aplikasi akan meminta input dari pengguna, seperti luas lahan yang digunakan dan data cuaca terkini, untuk melakukan prediksi produksi tanaman cabai. Aplikasi ini memberikan kontribusi dalam pengembangan teknologi informasi dan pengolahan data di bidang pertanian. Prediksi menjadi aspek penting dalam penelitian ini karena menjadi tujuan utama dari penggunaan Data Mining dan algoritma C4.5. Dalam konteks penelitian, prediksi digunakan untuk memperkirakan produksi tanaman cabai di masa depan berdasarkan data historis dan variabel-variabel yang relevan. Hasil prediksi ini dapat digunakan dalam perencanaan produksi pertanian cabai di Aceh. Analisis faktor-faktor yang mempengaruhi produksi tanaman cabai menjadi penting dalam penelitian ini. Faktor-faktor ini mencakup jenis tanaman, curah hujan, suhu udara, kelembaban udara, dan kualitas tanah. Studi-studi sebelumnya telah mengidentifikasi faktor-faktor tersebut sebagai faktor-faktor yang mempengaruhi produksi tanaman. Analisis ini menjadi dasar dalam pengembangan model prediksi produksi tanaman cabai yang akurat dan dapat diandalkan. Pendekatan berbasis web menjadi pilihan dalam pengembangan aplikasi prediksi produksi usaha cabai dalam konteks penelitian ini karena memberikan keuntungan dalam hal aksesibilitas dan

fleksibilitas. Dengan menggunakan antarmuka web, aplikasi ini dapat diakses oleh para petani dan pemangku kepentingan lainnya dari berbagai perangkat yang terhubung ke internet. Hal ini memungkinkan penyebaran dan penggunaan aplikasi ini secara luas di kalangan para petani dan pemangku kepentingan di bidang pertanian cabai di Aceh.

## 2. Metode Penelitian

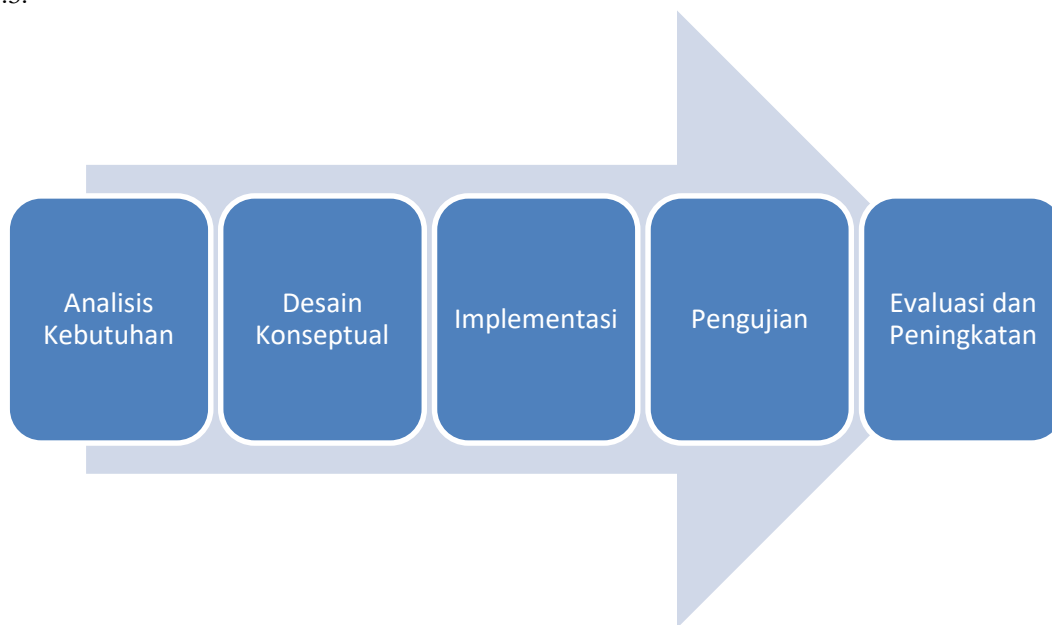
Metodologi penelitian ini difokuskan pada penerapan Metode Algoritma C4.5 dalam Data Mining untuk prediksi produksi tanaman cabai di Aceh, dengan Dinas Pertanian Provinsi Aceh sebagai mitra penelitian yang menyediakan data dan informasi terkait produksi tanaman di wilayah tersebut. Dinas Pertanian dan Perkebunan Aceh memiliki tanggung jawab yang luas dalam mengelola berbagai aspek pertanian dan perkebunan di provinsi tersebut, termasuk pelaksanaan tata usaha, perumusan kebijakan, pengawasan mutu, pembinaan produksi, pengendalian hama dan penyakit tanaman, serta pembinaan sumber daya manusia terkait. Sistem lama yang digunakan sebelum adanya aplikasi berbasis web menggunakan Metode Algoritma C4.5 kurang efisien dan tidak memanfaatkan teknologi informasi secara optimal. Proses perencanaan produksi tanaman cabai dilakukan secara manual dan bergantung pada pengalaman petugas lapangan, dengan informasi terkait produksi cabai yang tidak diolah secara sistematis. Untuk mengatasi keterbatasan ini, dikembangkanlah aplikasi perhitungan prediksi produksi usaha cabai berbasis web menggunakan Metode Algoritma C4.5, diharapkan dapat meningkatkan efisiensi dan akurasi dalam prediksi produksi tanaman cabai di masa depan. Bahan dan alat penelitian yang digunakan meliputi data produksi tanaman cabai selama 5 tahun terakhir, data cuaca, komputer atau laptop dengan koneksi internet, perangkat lunak aplikasi perhitungan prediksi, dan perangkat lunak pengolah data. Metode pengumpulan data meliputi studi pustaka, observasi, wawancara, dan studi dokumen, dengan data produksi tanaman cabai di Aceh diperoleh dari Dinas Pertanian Provinsi Aceh serta data cuaca dari BMKG dan Disperindag.

Pengembangan aplikasi menggunakan Metode *Rapid Application Development* (RAD), dengan tahapan analisis kebutuhan, desain konseptual, implementasi, pengujian, evaluasi, dan peningkatan. Setiap langkah dilakukan secara berurutan untuk menghasilkan aplikasi perhitungan prediksi produksi usaha cabai berbasis web yang efektif dan sesuai dengan kebutuhan pengguna.



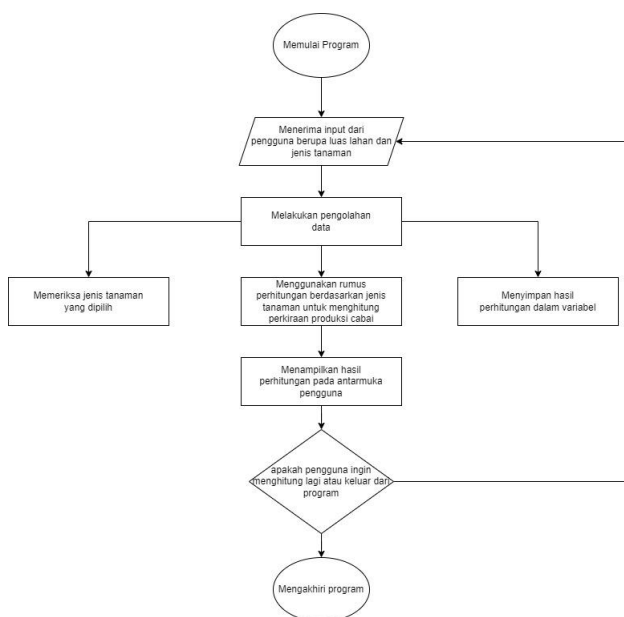
Gambar 1. Metode *Rapid Application Development* (RAD)

Langkah-langkah dan diagram alir penelitian dijelaskan untuk memvisualisasikan proses pengembangan aplikasi dengan jelas. Flowchart program juga disertakan untuk menggambarkan langkah-langkah dalam program aplikasi perhitungan prediksi produksi cabai menggunakan Metode C4.5.



Gambar 2. Tahapan Penelitian

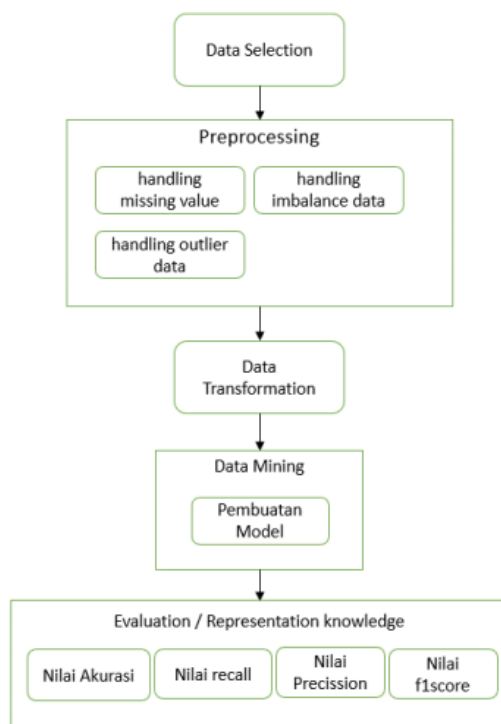
Rancangan basis data dibuat untuk menyimpan dan mengelola data terkait prediksi produksi cabai, dengan struktur yang terorganisir baik. Beberapa tabel yang terkait dengan pengguna, data latih, data uji, gain, hasil klasifikasi, dan keputusan disusun untuk mendukung proses analisis dan prediksi.



Gambar 3. Flowchart Program

Rancangan pengujian meliputi pengujian fungsionalitas, validitas data, keandalan sistem, dan performa aplikasi, dengan langkah-langkah pengujian yang rinci untuk memastikan kualitas dan

kinerja aplikasi yang baik. Teknik analisis dan pemrosesan data menggunakan Jupyter Notebook dan bahasa pemrograman Python, dengan metode *Knowledge Discovery in Database* (KDD) yang mencakup pemilihan data, preprocessing, transformasi data, data mining, dan evaluasi/representasi knowledge.



Gambar 4. Proses KDD

Proses KDD tersebut menjelaskan langkah-langkah analisis data dari pemilihan hingga evaluasi, dengan fokus pada penggunaan Metode C4.5 untuk memprediksi produksi cabai berdasarkan data cuaca dan harga cabai yang tersedia. Informasi hasil evaluasi dapat membantu pengguna dalam pengambilan keputusan terkait produksi tanaman cabai di Aceh.

### 3. Hasil dan Pembahasan

#### 3.1 Seleksi Data

Tidak semua atribut pada data yang telah dikumpulkan akan digunakan dalam proses pengolahan data mining. Oleh karena itu, dilakukan proses data selection atau seleksi data yang bertujuan untuk menganalisis data yang relevan dan memilih atribut atau variabel yang berpengaruh pada data cuaca dan harga cabai dalam memprediksi penanaman tanaman cabai. Atribut yang telah dipilih akan terlihat pada Tabel 1.

Tabel 1. Atribut Terpilih

No	Atribut Terpilih	Detail Penggunaan
1	Suhu rata-rata (°C)	Atribut Penjelas
2	Kelembapan rata-rata (%)	Atribut Penjelas
3	Curah hujan (mm)	Atribut Penjelas
4	Harga	Atribut Penjelas
5	Label (tanam/tidak) (LA/LP)	Atribut Target / Kelas Target

Atribut suhu, kelembapan, dan curah hujan dipilih karena memiliki pengaruh yang signifikan terhadap pertumbuhan tanaman cabai dan juga memengaruhi waktu penanaman cabai. Atribut-atribut tersebut akan digunakan dalam proses analisis dan pemrosesan data menggunakan algoritma klasifikasi C4.5. Hasil dari algoritma ini akan membentuk sebuah pohon keputusan yang akan berfungsi sebagai aturan prediksi untuk menentukan waktu penanaman cabai. Dengan demikian, pohon keputusan tersebut akan dapat memprediksi waktu yang tepat untuk melakukan penanaman cabai.

### 3.2 Data Preprocessing

Preprocessing adalah salah satu tahapan penting dalam proses data mining karena data yang digunakan tidak selalu dalam kondisi ideal untuk diproses. Beberapa masalah seperti missing value, outliers, dan ketidakseimbangan data dapat mengganggu hasil dari proses data mining itu sendiri. Dalam penelitian ini, data yang telah dikumpulkan dari penelitian sebelumnya harus melalui tahap data selection atau seleksi data untuk memilih atribut atau variabel yang relevan yang berpengaruh pada data cuaca dan harga cabai untuk memprediksi penanaman tanaman cabai. Atribut suhu, kelembapan, dan curah hujan dipilih karena memiliki pengaruh yang signifikan terhadap pertumbuhan tanaman cabai dan waktu penanaman cabai. Atribut-atribut tersebut akan digunakan dalam proses analisis dan pemrosesan data menggunakan algoritma klasifikasi C4.5. Hasil dari algoritma ini akan membentuk sebuah pohon keputusan yang akan berfungsi sebagai aturan prediksi untuk menentukan waktu penanaman cabai. Dengan demikian, pohon keputusan tersebut akan dapat memprediksi waktu yang tepat untuk melakukan penanaman cabai.

Tabel 2. Outliers Data

Nama Atribut	Min	Max	Average
Suhu	21	26	23.5
Kelembapan	20	100	70.5
Curah Hujan	0	500	75.2
Harga	21.000	94.000	57.500

Tabel di atas menunjukkan atribut-atribut yang memiliki outliers pada dataset berdasarkan data yang Anda miliki sebelumnya. Atribut "Suhu" memiliki nilai minimum 21 dan maksimum 26 dengan rata-rata 23.5. Atribut "Kelembapan" memiliki nilai minimum 20 dan maksimum 100 dengan rata-rata 70.5. Atribut "Curah Hujan" memiliki nilai minimum 0 dan maksimum 500 dengan rata-rata 75.2. Terakhir, atribut "Harga" memiliki nilai minimum 21000 dan maksimum 94000 dengan rata-rata 57500. Dalam tahap preprocessing, penanganan outliers data dilakukan untuk mengatasi data yang memiliki nilai ekstrem atau terlalu jauh dari nilai rata-rata. Pada dataset yang sudah diseleksi, terdapat data outliers pada beberapa atribut seperti suhu, kelembapan, curah hujan, dan harga. Outliers tersebut dihapus dari dataset agar tidak mempengaruhi hasil analisis. Selain itu, penanganan juga dilakukan terhadap missing value pada atribut suhu, kelembapan, dan curah hujan dengan mengisi nilai rata-rata setiap bulannya. Untuk menampilkan hasil dari pengisian missing value pada atribut "Suhu", "Kelembapan", dan "Curah Hujan", penulis mencetak DataFrame setelah melakukan pengisian menggunakan library Pandas dengan kode Python. Kode untuk menampilkan hasilnya dapat dilihat pada lampiran 2. Selanjutnya, dilakukan penanganan ketidakseimbangan data, di mana jumlah data pada kelas "tanam" dan "tidak" tidak seimbang. Dalam penelitian ini, dilakukan pendekatan re-sampling dengan menggunakan under-sampling. Under-sampling digunakan untuk mengurangi data pada kelas yang lebih banyak sehingga jumlah data pada kedua kelas menjadi seimbang.

Tabel 3. Perubahan Data pada Atribut Label

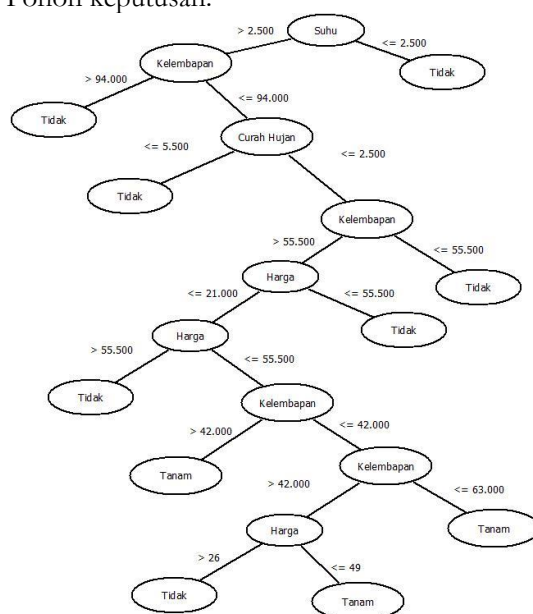
No	Label	Perubahan
1	Tanam (LA)	1
2	Tidak (LP)	0

Tahap berikutnya adalah data transformation, di mana dilakukan perubahan pada atribut label agar sesuai dengan kebutuhan dalam proses evaluasi menggunakan bahasa pemrograman Python. Label "tanam" diubah menjadi nilai 1, sementara label "tidak" diubah menjadi nilai 0 untuk memudahkan evaluasi pada model yang akan dibuat. Dengan tahapan preprocessing dan data transformation yang telah dilakukan, dataset siap untuk diproses menggunakan algoritma C4.5 yang akan membentuk pohon keputusan untuk memprediksi penanaman cabai berdasarkan atribut-atribut yang relevan yang telah dipilih sebelumnya.

### 3.3 Data Mining

Data telah melewati beberapa tahap di antaranya tahap selection, preprocessing, dan transformation kemudian dilakukan permodelan. Teknik pemodelan data mining yang dipilih adalah klasifikasi prediksi dengan menggunakan algoritma C4.5. Klasifikasi algoritma C4.5 digunakan untuk mencapai tujuan awal penelitian yaitu memprediksi penanaman cabai. Implementasi algoritma C4.5 menggunakan bahasa pemrograman Python dengan bantuan IDE atau text editor Jupyter Notebook guna membuat model klasifikasi dengan metode C4.5 dan juga dapat memprediksi penanaman cabai. Model yang dibuat menggunakan Decision Tree Classifier, sebuah library dari scikit-learn yang dapat memproses data menggunakan algoritma Decision Tree. Data kemudian dibagi menjadi data uji dan data latih menggunakan k-folds cross-validation dengan nilai k=10. Proses split data menggunakan k-folds cross-validation yang berarti terdapat 10 skenario pengujian dengan data uji dan data latih yang berbeda-beda. Setiap skenario pengujian menghasilkan nilai akurasi, precision, recall, dan f1-score yang berbeda-beda tergantung dari data uji dan data latih yang digunakan. Setelah proses cross-validation selesai, dilakukan evaluasi dengan merata-ratakan tingkat akurasi, precision, recall, dan f1-score dari semua skenario pengujian. Hasil evaluasi menunjukkan bahwa rata-rata akurasi dari proses cross-validation adalah 0.9944, rata-rata precision adalah 0.9893, rata-rata recall adalah 1, dan rata-rata f1-score adalah 0.9946.

Penelitian ini menggunakan metodologi KDD (*Knowledge Discovery in Database*) yang terdiri dari lima tahap, yaitu data selection, data processing, transformation, data mining, dan evaluation. Hasil dari skenario pengujian tersebut menghasilkan pemodelan pohon keputusan berupa beberapa rule atau aturan untuk memprediksi penanaman cabai berdasarkan atribut suhu, kelembapan, dan curah hujan. Meskipun hasil akurasi, precision, recall, dan f1-score cukup tinggi, model yang dibuat belum tentu sempurna, sehingga evaluasi dan penyempurnaan model dapat terus dilakukan untuk meningkatkan performa model. Adapun pengujian dari skenario menghasilkan pemodelan pohon keputusan pada gambar 5 Pohon keputusan.



Gambar 5. Pohon Keputusan

Berdasarkan pohon keputusan yang dibangun oleh algoritma C4.5 mengenai prediksi penanaman cabai di wilayah Kota Banda Aceh, Kota Lhokseumaawe, Kabupaten Aceh Barat, dan Kabupaten Pidie, menghasilkan sejumlah aturan atau rule sebagai berikut:

- 1) JIKA nama\_wil='Kota Banda Aceh' AND suhu='21' MAKA Ya (Direkomendasikan untuk menanam cabai).
- 2) JIKA nama\_wil='Kota Banda Aceh' AND suhu='26' AND kelembapan='20' AND curah\_hujan='0' MAKA Ya.
- 3) JIKA nama\_wil='Kota Banda Aceh' AND suhu='26' AND kelembapan='20' AND curah\_hujan='4444' MAKA Ya.
- 4) JIKA nama\_wil='Kota Banda Aceh' AND suhu='26' AND kelembapan='8888' AND curah\_hujan='0' MAKA Ya.
- 5) JIKA nama\_wil='Kota Banda Aceh' AND suhu='26' AND kelembapan='8888' AND curah\_hujan='4444' MAKA Ya.
- 6) JIKA nama\_wil='Kota Lhokseumawe' AND suhu='21' MAKA Ya (Direkomendasikan untuk menanam cabai).
- 7) JIKA nama\_wil='Kota Lhokseumawe' AND suhu='26' AND curah\_hujan='0' AND kelembapan='20' MAKA Ya.
- 8) JIKA nama\_wil='Kota Lhokseumawe' AND suhu='26' AND curah\_hujan='0' AND kelembapan='8888' MAKA Ya.
- 9) JIKA nama\_wil='Kota Lhokseumawe' AND suhu='26' AND curah\_hujan='4444' AND kelembapan='20' MAKA Ya.
- 10) JIKA nama\_wil='Kota Lhokseumawe' AND suhu='26' AND curah\_hujan='4444' AND kelembapan='8888' MAKA Ya.
- 11) JIKA nama\_wil='Kabupaten Aceh Barat' AND suhu='21' MAKA Ya (Direkomendasikan untuk menanam cabai).
- 12) JIKA nama\_wil='Kabupaten Aceh Barat' AND suhu='26' AND curah\_hujan='0' MAKA Ya.
- 13) JIKA nama\_wil='Kabupaten Aceh Barat' AND suhu='26' AND curah\_hujan='4444' MAKA Ya.
- 14) JIKA nama\_wil='Kabupate Pidie' AND suhu='21' AND kelembapan='20' AND curah\_hujan='0' MAKA Ya.
- 15) JIKA nama\_wil='Kabupate Pidie' AND suhu='21' AND kelembapan='20' AND curah\_hujan='4444' MAKA Tidak (Tidak direkomendasikan untuk menanam cabai).
- 16) JIKA nama\_wil='Kabupate Pidie' AND suhu='21' AND kelembapan='8888' MAKA Tidak (Tidak direkomendasikan untuk menanam cabai).
- 17) JIKA nama\_wil='Kabupate Pidie' AND suhu='26' MAKA Ya (Direkomendasikan untuk menanam cabai).

Setelah mendapatkan rule, selanjutnya dilakukan uji coba terhadap model tersebut menggunakan sample data baru yang belum diketahui prediksinya. Sampel yang digunakan adalah data cuaca dan harga untuk periode tahun 2022-2023. Uji coba dilakukan dengan menggunakan aplikasi yang telah dibangun. Hasil dari percobaan menunjukkan bahwa pemodelan yang dihasilkan oleh algoritma C4.5 dapat memprediksi dengan total jumlah prediksi sebanyak 4998 data. Dari prediksi tersebut, terdapat 4200 data yang terbukti tepat dan sesuai dengan kenyataan (True Positive), sementara terdapat 798 data yang tidak tepat (False Positive). Dengan demikian, akurasi dari model ini dapat dihitung sebagai berikut:

$$\begin{aligned} \text{Akurasi} &= (\text{Jumlah tepat} / \text{Jumlah prediksi}) * 100\% \\ \text{Akurasi} &= (4200 / 4998) * 100\% \\ \text{Akurasi} &= 84.03\% \end{aligned}$$

Laju error dari model ini dapat dihitung sebagai berikut:

$$\text{Laju error} = 100\% - \text{Akurasi}$$

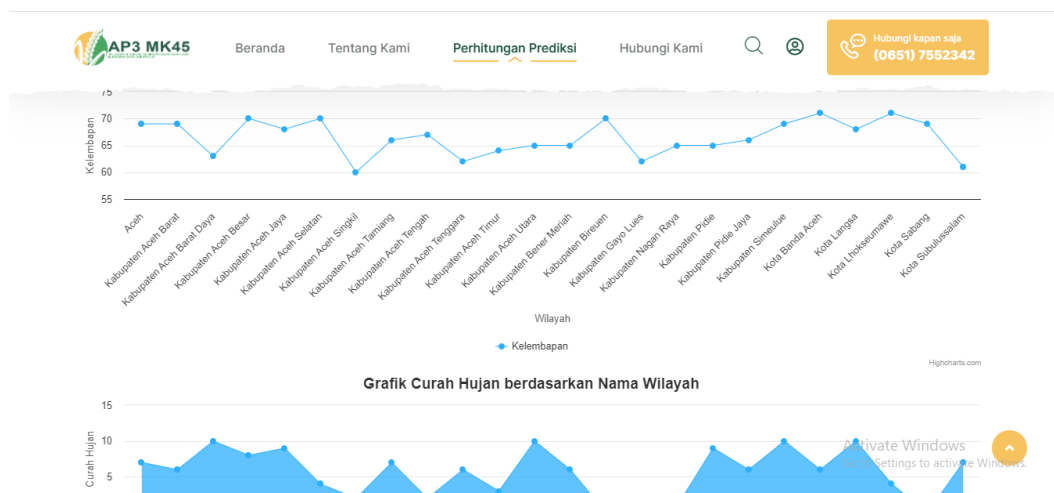
$$\text{Laju error} = 100\% - 84.03\%$$

$$\text{Laju error} = 15.97\%$$

Sehingga, akurasi model sebesar 84.03% menunjukkan tingkat ketepatan prediksi yang relatif tinggi, sementara laju error sebesar 15.97% menunjukkan tingkat ketidaktepatan prediksi yang relatif rendah. Data baru yang digunakan dalam uji coba ini berhasil memberikan keterangan berupa penanaman cabai berdasarkan hasil prediksi model, dan data baru tersebut merupakan data yang tidak termasuk dalam data yang digunakan untuk membangun model sebelumnya.

### 3.4 Implementasi

Implementasi aplikasi ini terdiri dari beberapa halaman yang dirancang untuk memberikan pengalaman pengguna yang intuitif dan informatif. Halaman Publik adalah titik awal yang memberikan gambaran umum tentang aplikasi. Pengguna dapat mengakses informasi tentang tim pengembang dan nilai-nilai yang dianut oleh mereka melalui Halaman Tentang Kami. Selain itu, aplikasi menyediakan Halaman Perhitungan Hasil Prediksi, yang menampilkan hasil dari penggunaan algoritma C4.5 berdasarkan data uji. Di Halaman Grafik, pengguna dapat melihat visualisasi grafik dari hasil prediksi dan data latih, membantu mereka memahami pola dan tren dengan lebih baik.



Gambar 6. Halaman Grafik Berdasarkan Kriteria

Untuk berkomunikasi dengan tim pengembang, pengguna dapat menggunakan Halaman Hubungi Kami yang menyediakan informasi kontak dan formulir pengiriman pesan. Sedangkan untuk akses lebih lanjut, pengguna dapat masuk melalui Halaman Login dan mengelola akun mereka atau menjelajahi fitur tambahan yang disediakan melalui Halaman Admin. Akses ke data penting seperti data latih dan data uji juga disediakan melalui Halaman Data Latih dan Halaman Data Uji secara berurutan. Selain itu, pengguna dapat melihat hasil prediksi lebih lanjut menggunakan Halaman Data Gain dan Halaman Data Klasifikasi. Informasi teknis tentang proses algoritma C4.5 tersedia di Halaman Proses Algoritma, sedangkan hasil dari algoritma tersebut dapat dilihat melalui Halaman Hasil Algoritma. Evaluasi performa model juga dapat dipelajari melalui Halaman Perhitungan Algoritma C.45. Terakhir, pengguna dapat mengelola profil dan akun mereka melalui Halaman User/ Pengguna. Dengan berbagai halaman ini, aplikasi memberikan pengalaman yang komprehensif bagi pengguna dalam mengakses, menganalisis, dan mengelola data serta model prediktif yang disediakan oleh algoritma C4.5.

### 3.5 Pembahasan

Penelitian ini melibatkan metodologi KDD (*Knowledge Discovery in Database*) yang terdiri dari lima tahap yang penting. Tahap pertama adalah data selection, di mana data cuaca dan harga yang relevan untuk penanaman cabai di wilayah Provinsi Aceh dipilih dan diambil sebagai dataset. Data cuaca mencakup atribut seperti suhu, kelembapan, dan curah hujan, sementara data harga mencakup atribut harga cabai. Dataset yang dihasilkan dari tahap ini menjadi dasar untuk proses selanjutnya. Tahap selanjutnya adalah data processing, di mana dilakukan preprocessing data untuk mengatasi missing value dan memastikan konsistensi data. Langkah-langkah preprocessing ini penting untuk memastikan bahwa data yang digunakan dalam proses mining nantinya berada dalam kondisi yang sesuai dan akurat. Setelah data diproses dengan baik, tahap berikutnya adalah transformation, di mana data diubah ke dalam bentuk yang sesuai untuk dilakukan proses mining menggunakan algoritma C4.5. Transformasi ini memungkinkan data cuaca, harga, dan atribut lainnya diintegrasikan dengan baik dalam model yang akan dibangun. Tahap data mining merupakan langkah sentral dalam penelitian ini, di mana algoritma C4.5 digunakan untuk membangun pohon keputusan yang mengandung berbagai aturan prediksi penanaman cabai berdasarkan data cuaca dan harga. Pohon keputusan ini akan memberikan wawasan dan panduan bagi petani atau pengusaha pertanian dalam membuat keputusan tentang penanaman cabai berdasarkan kondisi cuaca dan harga yang ada. Akhirnya, tahap evaluation dilakukan untuk menguji model yang telah dihasilkan pada sampel data baru yang belum diketahui prediksinya. Hasil evaluasi menunjukkan bahwa model prediksi memiliki akurasi sebesar 84.03%, dengan jumlah prediksi tepat mencapai 4200 dan prediksi tidak tepat sebanyak 798 dari total 4998 data baru. Hasil yang cukup baik ini menunjukkan bahwa model prediksi dapat menjadi alat yang berguna bagi para petani atau pengusaha pertanian dalam pengambilan keputusan yang lebih efektif mengenai penanaman cabai berdasarkan kondisi cuaca dan harga di wilayah Provinsi Aceh. Penelitian ini memberikan kontribusi penting dalam pengembangan model prediksi penanaman cabai yang dapat membantu meningkatkan efisiensi dan produktivitas pertanian di wilayah tersebut. Dengan memanfaatkan data cuaca dan harga yang akurat, model ini dapat membantu mengurangi risiko dan kerugian bagi para pelaku usaha cabai dan mendukung pengambilan keputusan yang lebih cerdas dan terinformasi. Penelitian ini melibatkan metodologi KDD (*Knowledge Discovery in Database*) yang terdiri dari lima tahap yang berurutan dan saling terkait dalam analisis prediksi produksi cabai. Berikut adalah penjelasan lebih detail tentang masing-masing tahapan:

- 1) Data Selection

Tahap pertama adalah data selection, di mana dilakukan pemilihan data cuaca dan harga yang relevan dan representatif untuk penanaman cabai di wilayah Provinsi Aceh. Data cuaca mencakup atribut seperti suhu, kelembapan, dan curah hujan yang diambil dari stasiun meteorologi terdekat. Sementara itu, data harga cabai diperoleh dari database harga komoditas pertanian. Proses seleksi data ini penting untuk memastikan bahwa data yang digunakan mencerminkan kondisi cuaca dan harga aktual di wilayah penelitian.

- 2) Data Cuaca

Data cuaca menjadi bagian penting dalam penelitian ini karena cuaca dapat mempengaruhi produksi cabai. Data cuaca mencakup atribut suhu, kelembapan, dan curah hujan di wilayah Provinsi Aceh. Informasi ini diperoleh dari catatan harian stasiun meteorologi dan memungkinkan untuk memahami pola cuaca yang berpengaruh pada produksi cabai.

- 3) Data Harga

Data harga cabai juga menjadi komponen yang relevan dalam penelitian ini. Harga cabai dapat mempengaruhi keputusan petani atau pengusaha pertanian dalam menanam cabai. Oleh karena itu, data harga cabai diambil dari sumber terpercaya seperti database perdagangan komoditas atau katalog harga komoditas pertanian. Data harga ini mencakup informasi harga cabai dari berbagai periode waktu dan berbagai wilayah di Provinsi Aceh.

- 4) Dataset

Setelah dilakukan data selection, data cuaca dan data harga digabungkan menjadi satu dataset yang komprehensif. Dataset ini terdiri dari berbagai atribut seperti suhu, kelembapan, curah

hujan, harga cabai, dan atribut lain yang relevan. Proses penggabungan data ini dilakukan dengan hati-hati untuk memastikan keterkaitan antara data cuaca dan harga dalam analisis prediksi produksi cabai.

#### 5) Data Processing:

Tahap data processing dilakukan untuk memastikan bahwa dataset yang digunakan dalam proses mining telah diolah dengan baik. Pada tahap ini, dilakukan preprocessing data untuk mengatasi missing value, outlier, dan memastikan konsistensi data. Langkah-langkah preprocessing ini penting untuk memastikan bahwa data yang digunakan dalam proses mining nantinya berada dalam kondisi yang sesuai dan akurat.

Setelah proses data processing selesai, tahap berikutnya adalah transformation, di mana data diubah ke dalam bentuk yang sesuai untuk dilakukan proses mining menggunakan algoritma C4.5. Proses transformasi ini mencakup konversi data menjadi format yang sesuai dengan input algoritma C4.5, seperti mengubah data kategorikal menjadi numerikal dan membuat struktur data berhierarki. Tahap data mining menggunakan algoritma C4.5 adalah inti dari penelitian ini. Pada tahap ini, algoritma C4.5 digunakan untuk membangun pohon keputusan yang mengandung berbagai aturan untuk memprediksi penanaman cabai berdasarkan data cuaca dan harga. Pohon keputusan ini akan memberikan wawasan dan panduan bagi petani atau pengusaha pertanian dalam membuat keputusan tentang penanaman cabai berdasarkan kondisi cuaca dan harga yang ada. Pada tahap evaluation dilakukan dengan menguji model yang telah dihasilkan pada sampel data baru yang belum diketahui prediksinya. Hasil evaluasi menunjukkan bahwa model prediksi memiliki akurasi sebesar 84.03%, dengan jumlah prediksi tepat mencapai 4200 dan prediksi tidak tepat sebanyak 798 dari total 4998 data baru. Hasil yang cukup baik ini menunjukkan bahwa model prediksi dapat menjadi alat yang berguna bagi para petani atau pengusaha pertanian dalam pengambilan keputusan yang lebih efektif mengenai penanaman cabai berdasarkan kondisi cuaca dan harga di wilayah Provinsi Aceh.

## 4. Kesimpulan dan Saran

Kesimpulan dari penelitian dan evaluasi yang telah dilakukan menunjukkan beberapa hal yang penting. Pertama, algoritma C4.5 terbukti efektif dalam membangun pohon keputusan untuk memprediksi penanaman cabai berdasarkan data cuaca dan harga di Provinsi Aceh. Pohon keputusan ini menghasilkan aturan-aturan yang berguna untuk mengklasifikasikan kondisi yang cocok atau tidak cocok untuk penanaman cabai. Kedua, model prediksi yang dihasilkan memiliki tingkat akurasi sebesar 84.03%, menunjukkan bahwa model tersebut dapat diandalkan dalam memberikan prediksi yang cukup akurat. Namun, perlu diingat bahwa tingkat akurasi ini dapat bervariasi tergantung pada situasi cuaca dan harga yang berbeda. Ketiga, penelitian ini memberikan kontribusi penting dalam pengembangan aplikasi prediksi penanaman cabai yang dapat digunakan oleh petani di Provinsi Aceh. Aplikasi ini diharapkan dapat membantu petani dalam pengambilan keputusan terkait penanaman cabai berdasarkan kondisi cuaca dan harga yang ada, sehingga meningkatkan efisiensi dan hasil produksi pertanian di wilayah tersebut.

Namun demikian, meskipun model prediksi memiliki tingkat akurasi yang tinggi, keterlibatan petani dalam pengambilan keputusan tetaplah penting. Aplikasi ini sebaiknya digunakan sebagai alat bantu dan referensi tambahan bagi petani, bukan sebagai pengganti keputusan manusia. Keterlibatan petani dalam evaluasi kondisi lapangan dan pengetahuan lokal tetap menjadi faktor kunci dalam keberhasilan penanaman. Selain itu, penelitian ini juga membuka peluang untuk pengembangan lebih lanjut dalam bidang prediksi pertanian menggunakan teknik data mining. Dengan mempertimbangkan saran-saran berikut untuk pengembangan selanjutnya, seperti analisis faktor-faktor tambahan, implementasi teknologi penginderaan jauh, penggunaan metode machine learning lainnya, dan evaluasi dampak penggunaan aplikasi, diharapkan model prediksi dapat ditingkatkan untuk memberikan hasil yang lebih akurat dan berguna bagi para petani di Provinsi Aceh.

## 5. Daftar Pustaka

- Andayani, S. A. (2018). Faktor-faktor yang mempengaruhi produksi cabai merah. *Mimbar Agribisnis: Jurnal Pemikiran Masyarakat Ilmiah Berwawasan Agribisnis*, 1(3), 261-268. DOI: <http://dx.doi.org/10.25157/ma.v1i3.46>
- Asistiyasari, A., & Baidawi, T. (2017). Analisis Penerimaan Karyawan Posisi Field Collector Menggunakan Algoritma C4. 5 Pada Pt. Prismas Jamintara Jakarta. *JITK (Jurnal Ilmu Pengetahuan dan Teknologi Komputer)*, 2(2), 1-6.
- Erviyana, P. (2014). Faktor-faktor yang mempengaruhi produksi tanaman pangan jagung di Indonesia. *JEJAK*, 7(2). DOI: <https://doi.org/10.15294/jejak.v7i2.3900>
- Habib, A. (2015). Analisis faktor-faktor yang mempengaruhi produksi jagung. *AGRIUM: Jurnal Ilmu Pertanian*, 18(1). DOI: <https://doi.org/10.30596/agrium.v18i1.347>
- Jollyta, D., Ramdhan, W., & Zarlis, M. (2020). *Konsep Data Mining Dan Penerapan*. Deepublish.
- Kusrini, E. T. L., & Taufiq, E. (2009). *Algoritma data mining*. Yogyakarta: Andi Offset.
- Nofriansyah, D., Kom, S., & Kom, M. (2015). *Konsep data mining vs sistem pendukung keputusan*. Deepublish.
- Rahmayuni, I. (2014). Perbandingan performansi algoritma c4. 5 dan cart dalam klasifikasi data nilai mahasiswa prodi teknik komputer politeknik negeri padang. *Jurnal Teknoif Teknik Informatika Institut Teknologi Padang*, 2(1), 40-46.
- Tomy, J. (2013). Faktor-faktor yang mempengaruhi produksi usahatani jagung Di Kecamatan Sindue Kabupaten Donggala. *Agroland: Jurnal Ilmu-ilmu Pertanian*, 20(1), 61-66.
- Widayu, H., Nasution, S. D., Silalahi, N., & Mesran, M. (2017). data mining untuk memprediksi jenis transaksi nasabah pada koperasi simpan pinjam dengan algoritma C4. 5. *Jurnal Media Informatika Budidarma*, 1(2).