E-ISSN : 3025-1311
https://journal.cib.institute/index.php/kohesi

Kohesi: Jurnal Multidisiplin Saintek
Volume 11 No 3 Tahun 2026

KORELASI DATA FAKTOR PRODUKSI TERHADAP CAPAIAN PRODUKSI KELAPA SAWIT MENGGUNAKAN METODE APRIORI (STUDI KASUS: PT PERKEBUNAN NUSANTARA IV REGIONAL II KEBUN TANJUNG JATI)

Dini Wahyuni [1*], Yani Maulita [2], Rusmin Saragih [3]
[1,2,3] Program Studi Sistem Informasi, STMIK Kaputama Binjai
E-mail: dini16442@gmail.com[1] yanimaulita26@gmail.com [2] evitha12014@gmail.com[3]

ABSTRACT

Oil palm is one of Indonesia's important plantation commodities, with production influenced by various factors, including land area, plant age, rainfall, rainy days, fertilizer achievement, and harvest rotation. PT Perkebunan Nusantara IV Regional II Kebun Tanjung Jati has historical data on production factors and production achievement that have not been optimally utilized to identify relationships among variables. This study aims to apply the Apriori algorithm to identify association rules between production factors and oil palm production achievement. The data consisted of 1,692 instances covering the 2023-2025 period. Data processing was performed using WEKA with a minimum support of 5% and a minimum confidence of 90%. The results produced 585 association rules, which were subsequently screened based on the seven research variables, resulting in 12 rules with 7-itemsets. Further selection based on the structure of production factors as antecedents and production achievement as the consequent produced two candidate rules. The best rule was Rule No. 147, consisting of narrow land area, immature plant age, high rainfall, moderate rainy days, adequate fertilizer achievement, and frequent harvest rotation leading to low production achievement. This rule obtained 6.74% support and 100% confidence. The results can provide supporting information for evaluating production conditions and data-driven decision-making.

Keywords: Apriori Algorithm, Association Rule, Production Factors, Oil Palm, Production Achievement

ABSTRAK

Kelapa sawit merupakan salah satu komoditas perkebunan penting yang produksinya dipengaruhi oleh berbagai faktor, seperti luas lahan, usia tanaman, curah hujan, hari hujan, capaian pemupukan, dan rotasi panen. PT Perkebunan Nusantara IV Regional II Kebun Tanjung Jati memiliki data historis faktor produksi dan capaian

Article History
Received: Agustus 2026
Reviewed:Agustus 2026
Published: Agustus 2026

Plagiarism Checker No 237.u873247.fh8374

Copyright : Author
Publish by : Kohesi

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License

produksi yang belum dimanfaatkan secara optimal untuk menemukan pola hubungan antarvariabel. Penelitian ini bertujuan menerapkan algoritma Apriori untuk mengidentifikasi aturan asosiasi antara faktor produksi dengan capaian produksi kelapa sawit. Data yang digunakan merupakan data periode 2023–2025 sebanyak 1.692 instances. Pengolahan dilakukan menggunakan perangkat lunak WEKA dengan minimum support 5% dan minimum confidence 90%. Hasil pengujian menghasilkan 585 association rules, kemudian diseleksi berdasarkan kesesuaian tujuh variabel penelitian dan diperoleh 12 rule dengan 7-itemset. Selanjutnya, seleksi berdasarkan struktur faktor produksi sebagai antecedent dan capaian produksi sebagai consequent menghasilkan dua kandidat rule. Best rule diperoleh pada Rule No. 147 dengan kombinasi luas lahan sempit, usia tanaman remaja, curah hujan tinggi, hari hujan sedang, capaian pemupukan cukup, dan rotasi panen banyak menuju capaian produksi rendah. Rule tersebut memiliki support 6,74% dan confidence 100%. Hasil penelitian dapat menjadi informasi pendukung dalam evaluasi kondisi produksi dan pengambilan keputusan berbasis data.

Kata Kunci: ***Algoritma Apriori, Association Rule, Faktor Produksi, Kelapa Sawit, Capaian Produksi***

1. PENDAHULUAN

Kelapa sawit merupakan salah satu komoditas perkebunan strategis di Indonesia yang memiliki peran penting dalam perekonomian dan industri berbasis minyak nabati. Perkembangan sektor kelapa sawit tidak hanya ditentukan oleh luas areal perkebunan, tetapi juga oleh kemampuan pengelola dalam mempertahankan produktivitas tanaman. Data pemerintah menunjukkan bahwa luas areal kelapa sawit Indonesia mencapai sekitar 16,83 juta hektare pada tahun 2024 dengan produksi minyak sawit mentah sebesar 45,44 juta ton. Kondisi tersebut menunjukkan pentingnya pengelolaan faktor produksi secara optimal untuk mempertahankan produktivitas perkebunan kelapa sawit secara berkelanjutan [1].

Capaian produksi kelapa sawit dipengaruhi oleh berbagai faktor yang saling berkaitan, baik faktor lingkungan maupun faktor budidaya. Faktor tersebut antara lain luas lahan, usia tanaman, curah hujan, hari hujan, pemupukan, dan aktivitas panen. Faktor lingkungan dan budidaya memiliki keterkaitan dengan pertumbuhan serta produktivitas tanaman, sehingga pengelolaan setiap faktor produksi perlu dilakukan secara terukur. Direktorat Jenderal Perkebunan juga menempatkan aspek lingkungan, iklim, serta pengelolaan budidaya sebagai bagian penting dalam upaya peningkatan produktivitas perkebunan [2].

PT Perkebunan Nusantara IV Regional II Kebun Tanjung Jati memiliki data operasional yang memuat informasi mengenai faktor-faktor produksi dan capaian produksi kelapa sawit. Data

penelitian mencakup Luas Lahan, Usia Tanaman, Curah Hujan, Hari Hujan, Capaian Pemupukan, Rotasi Panen, dan Capaian Produksi. Ketersediaan data historis tersebut memberikan peluang untuk dilakukan pengolahan lebih lanjut sehingga tidak hanya digunakan sebagai dokumentasi atau bahan evaluasi operasional, tetapi juga dapat dimanfaatkan untuk menemukan pola keterkaitan antarfaktor produksi. Penelitian ini menggunakan data produksi kelapa sawit periode 2023–2025 sebagai objek analisis.

Permasalahan yang muncul tidak hanya berkaitan dengan bagaimana mengetahui kondisi masing-masing faktor produksi, tetapi juga bagaimana mengidentifikasi kombinasi beberapa faktor yang muncul secara bersamaan dan berkaitan dengan kategori capaian produksi tertentu. Analisis terhadap setiap faktor secara terpisah belum tentu dapat menggambarkan pola kombinasi yang terdapat dalam data historis. Oleh karena itu, diperlukan pendekatan data mining yang mampu menemukan keterkaitan antarfaktor dan menyajikannya dalam bentuk aturan yang mudah dipahami. Salah satu pendekatan yang dapat digunakan adalah Association Rule Mining, yaitu teknik untuk menemukan hubungan atau pola kemunculan bersama antaritem. Algoritma Apriori merupakan salah satu metode yang dapat digunakan untuk membentuk frequent itemset dan association rule dengan mempertimbangkan nilai support dan confidence [2] [3].

Penerapan algoritma Apriori pada bidang pertanian telah dilakukan dalam beberapa penelitian. [2] menerapkan algoritma Apriori untuk menentukan pola penjualan produk pertanian, sedangkan [3] menerapkan Association Rule menggunakan algoritma Apriori untuk menganalisis pola yang berkaitan dengan peningkatan hasil pertanian. Penelitian tersebut menunjukkan bahwa algoritma Apriori dapat digunakan untuk menemukan pola keterkaitan dalam data pertanian berdasarkan nilai support dan confidence. Selain itu, [4] menerapkan algoritma Apriori dalam menentukan kombinasi produk pertanian sebagai pendukung pengambilan keputusan. Hasil penelitian terdahulu tersebut memperlihatkan bahwa Apriori memiliki potensi dalam menggali pola dari data pertanian yang memiliki hubungan antarkomponen.

Meskipun demikian, penelitian terdahulu yang menerapkan algoritma Apriori pada bidang pertanian masih lebih banyak berfokus pada pola penjualan, rekomendasi pertanian, kombinasi produk, atau hubungan umum pada data pertanian. Belum banyak penelitian yang secara khusus mengidentifikasi pola asosiasi antara kombinasi faktor produksi kelapa sawit yang mencakup kondisi lahan, usia tanaman, curah hujan, hari hujan, capaian pemupukan, dan rotasi panen terhadap kategori capaian produksi pada data operasional perkebunan. Selain itu, analisis dalam penelitian ini tidak diarahkan untuk memprediksi nilai produksi, tetapi untuk menemukan aturan asosiasi yang menggambarkan kombinasi kondisi faktor produksi yang berkaitan dengan kategori capaian produksi tertentu.

Berdasarkan research gap tersebut, penelitian ini menerapkan algoritma Apriori menggunakan WEKA untuk menganalisis hubungan antara faktor-faktor produksi dengan capaian produksi kelapa sawit pada PT Perkebunan Nusantara IV Regional II Kebun Tanjung Jati. Data faktor produksi dikategorikan berdasarkan kriteria yang telah ditetapkan, kemudian dianalisis melalui proses pembentukan frequent itemset dan association rule menggunakan nilai support

dan confidence. Pendekatan ini digunakan untuk memperoleh pola hubungan yang menempatkan faktor-faktor produksi sebagai antecedent dan capaian produksi sebagai consequent, sehingga hasil analisis dapat memberikan gambaran mengenai kombinasi kondisi produksi yang berkaitan dengan capaian produksi.

Dengan demikian, penelitian ini bertujuan menerapkan algoritma Apriori menggunakan WEKA untuk mengidentifikasi pola hubungan antara faktor-faktor produksi dengan capaian produksi kelapa sawit serta menentukan Best Rule Association berdasarkan nilai support dan confidence. Penelitian ini diharapkan dapat memberikan kontribusi berupa pola asosiasi dari data historis produksi yang dapat digunakan sebagai informasi pendukung dalam proses evaluasi kondisi faktor produksi pada PT Perkebunan Nusantara IV Regional II Kebun Tanjung Jati.

2. TINJAUAN PUSTAKA

2.1. Data Mining

Data Mining merupakan proses untuk menggali dan menemukan informasi, pola, hubungan, atau pengetahuan yang bernilai dari sekumpulan data dengan memanfaatkan teknik statistik, kecerdasan buatan, dan machine learning. Data Mining merupakan bagian dari proses Knowledge Discovery in Database (KDD) yang bertujuan menghasilkan pengetahuan yang valid, baru, dan bermanfaat dari data yang tersedia. Melalui proses tersebut, data yang sebelumnya hanya berfungsi sebagai arsip dapat diolah menjadi informasi yang mendukung pengambilan keputusan [5].

Dalam penelitian ini, Data Mining digunakan untuk mengolah data historis faktor produksi kelapa sawit sehingga dapat ditemukan pola hubungan antara beberapa faktor produksi dengan kategori capaian produksi. Pendekatan tersebut sesuai dengan karakteristik penelitian yang berorientasi pada pencarian pola, bukan melakukan prediksi nilai produksi.

2.2. Association Rule

Association Rule atau aturan asosiasi merupakan salah satu teknik Data Mining yang digunakan untuk menemukan pola hubungan atau keterkaitan antaritem dalam suatu kumpulan data. Teknik ini mencari kombinasi item yang sering muncul secara bersamaan dan menyajikannya dalam bentuk aturan $A \rightarrow B$, yang menunjukkan hubungan antara antecedent dan consequent. Association Rule termasuk metode deskriptif karena berfokus pada penemuan pola yang terdapat dalam data tanpa melakukan prediksi terhadap data masa mendatang [6].

Pada penelitian ini, Association Rule digunakan untuk menemukan hubungan antara kombinasi faktor produksi sebagai antecedent, yaitu Luas Lahan, Usia Tanaman, Curah Hujan, Hari Hujan, Capaian Pemupukan, dan Rotasi Panen, dengan Capaian Produksi sebagai consequent. Dengan demikian, aturan yang terbentuk dapat menggambarkan kombinasi kondisi faktor produksi yang berkaitan dengan kategori capaian produksi tertentu.

2.3. Algoritma Apriori

Algoritma Apriori merupakan algoritma Data Mining yang digunakan untuk menemukan frequent itemset dan membentuk Association Rule dari sekumpulan data. Proses Apriori dilakukan dengan mencari kombinasi item yang memenuhi nilai minimum support dan kemudian membentuk aturan berdasarkan nilai minimum confidence. Apriori diperkenalkan oleh Agrawal dan Srikant pada tahun 1994 dan menggunakan prinsip Apriori Property dalam proses pencarian kombinasi item [6].

Prinsip Apriori Property menyatakan bahwa apabila suatu itemset memenuhi minimum support, maka seluruh subset dari itemset tersebut juga harus memenuhi minimum support. Sebaliknya, apabila suatu itemset tidak memenuhi minimum support, maka kombinasi yang lebih besar yang mengandung itemset tersebut tidak perlu diproses lebih lanjut. Prinsip ini digunakan untuk mengurangi jumlah kandidat yang harus diperiksa selama proses pencarian frequent itemset. Secara umum, penerapan Apriori meliputi penentuan minimum support dan confidence, pencarian frequent itemset, pembentukan kandidat itemset secara bertingkat, pembentukan Association Rule, serta evaluasi aturan berdasarkan nilai confidence.

Dalam penelitian ini, algoritma Apriori diterapkan pada data faktor produksi kelapa sawit untuk menemukan kombinasi faktor produksi yang memiliki keterkaitan dengan Capaian Produksi. Proses pencarian dilakukan secara bertingkat mulai dari L1 hingga kombinasi itemset yang memenuhi batas minimum support.

2.4. Support

Support merupakan ukuran yang menunjukkan tingkat frekuensi kemunculan suatu item atau kombinasi item dalam keseluruhan dataset. Nilai support menunjukkan seberapa sering suatu pola ditemukan dibandingkan dengan seluruh transaksi yang dianalisis. Semakin tinggi nilai support, semakin sering pola tersebut muncul dalam dataset. Secara umum, support suatu item dapat dirumuskan sebagai:

$$Support\ (A) = \frac{\text{Jumlah transaksi yang mengandung A}}{\text{Total transaksi}} \text{ x } 100\% \quad (1)$$

Untuk kombinasi dua item, support menunjukkan proporsi transaksi yang mengandung kedua item tersebut secara bersamaan. Dalam penelitian ini, nilai minimum support digunakan sebagai batas untuk menentukan frequent itemset. Itemset yang tidak memenuhi batas tersebut tidak dilanjutkan ke tahap pembentukan aturan asosiasi.

2.5. Confidence

Confidence merupakan ukuran yang menunjukkan tingkat keyakinan suatu aturan asosiasi. Nilai tersebut menggambarkan seberapa besar kemungkinan consequent terjadi ketika kondisi pada antecedent telah terpenuhi. Dengan demikian, confidence digunakan untuk mengukur kekuatan keterkaitan antara kondisi A dan kondisi B dalam aturan A → B. Secara umum, confidence dapat dirumuskan sebagai:

$$Confidence\ (A \rightarrow B) = \frac{\text{Jumlah transaksi } yang \text{ mengandung A } dan\ B}{Jumlah \text{ transaksi yang mengandung A}} \text{ x } 100\% \quad (2)$$

Dalam penelitian ini, confidence digunakan sebagai salah satu dasar dalam menyeleksi Association Rule yang memenuhi kriteria penelitian. Nilai confidence yang semakin tinggi menunjukkan semakin besar konsistensi consequent muncul ketika antecedent terpenuhi.

2.6. WEKA

WEKA (Waikato Environment for Knowledge Analysis) merupakan perangkat lunak yang digunakan untuk penerapan dan analisis berbagai metode Data Mining dan machine learning. Dalam penelitian ini, WEKA digunakan sebagai perangkat implementasi algoritma Apriori untuk mengolah dataset faktor produksi kelapa sawit dan menghasilkan frequent itemset serta Association Rule [7].

Penggunaan WEKA memungkinkan proses pencarian itemset dilakukan secara sistematis berdasarkan parameter yang telah ditentukan. Hasil yang diperoleh selanjutnya digunakan untuk melakukan seleksi aturan berdasarkan struktur variabel penelitian serta nilai support dan confidence.

2.7. Faktor Produksi Kelapa Sawit

Faktor produksi dalam penelitian ini merupakan variabel yang digunakan untuk menggambarkan kondisi yang berkaitan dengan capaian produksi kelapa sawit. Berdasarkan objek penelitian, faktor tersebut terdiri atas Luas Lahan, Usia Tanaman, Curah Hujan, Hari Hujan, Capaian Pemupukan, dan Rotasi Panen, sedangkan Capaian Produksi digunakan sebagai hasil atau consequent dalam aturan asosiasi [8].

Ketujuh atribut tersebut dikategorikan berdasarkan kriteria penelitian sehingga dapat direpresentasikan sebagai item dalam proses Apriori. Pendekatan kategorisasi memungkinkan kondisi masing-masing faktor produksi dianalisis dalam bentuk kombinasi item dan selanjutnya dicari pola kemunculannya terhadap kategori Capaian Produksi.

3. METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan metode Data Mining untuk mengidentifikasi pola hubungan antara faktor-faktor produksi dan capaian produksi kelapa sawit. Objek penelitian adalah data faktor produksi kelapa sawit pada PT Perkebunan Nusantara IV Regional II Kebun Tanjung Jati periode 2023–2025. Data yang dianalisis merupakan data historis operasional perkebunan yang mencakup faktor produksi dan capaian produksi. Dataset penelitian terdiri atas 1.692 data dengan tujuh atribut, yaitu Luas Lahan (Ha), Usia Tanaman (Tahun), Curah Hujan (MM), Hari Hujan (HH), Capaian Pemupukan (%), Rotasi Panen, dan Capaian Produksi (Kg).

3.1. Data Penelitian

Data penelitian diperoleh melalui dokumentasi data operasional yang berkaitan dengan faktor produksi dan capaian produksi kelapa sawit pada objek penelitian. Data yang diperoleh

kemudian disusun dan dikategorikan berdasarkan kriteria yang telah ditetapkan sehingga setiap variabel dapat direpresentasikan sebagai item dalam proses Association Rule Mining.

Tabel 1. Variabel dan Kategori Data Penelitian

No.	Variabel	Kategori
1	Luas Lahan (Ha)	Sempit, Sedang, Luas
2	Usia Tanaman (Tahun)	Tua, Dewasa, Remaja, Muda
3	Curah Hujan (MM)	Rendah, Sedang, Tinggi
4	Hari Hujan (HH)	Sedikit, Sedang, Banyak
5	Capaian Pemupukan (%)	Kurang, Cukup, Baik
6	Rotasi Panen	Sedikit, Normal, Banyak
7	Capaian Produksi (Kg)	Rendah, Sedang, Banyak

Data penelitian merupakan data operasional produksi kelapa sawit pada PT Perkebunan Nusantara IV Regional II Kebun Tanjung Jati periode 2023–2025. Dataset yang digunakan dalam penelitian berjumlah 1.692 data. Data tersebut memuat informasi mengenai tahun tanam, luas lahan, usia tanaman, kondisi curah hujan, hari hujan, capaian pemupukan, rotasi panen, dan capaian produksi. Data tersebut kemudian dikategorikan sesuai dengan kriteria penelitian untuk dapat digunakan dalam proses Association Rule Mining menggunakan algoritma Apriori. Struktur data penelitian dapat dilihat pada Tabel 2.

Tabel 2. Sampel Data Penelitian

No.	Luas Lahan (Ha)	Usia Tanaman (Tahun)	Curah Hujan (MM)	Hari Hujan (HH)	Capaian Pemupukan (%)	Rotasi Panen	Capaian Produksi (Kg)
1	315,06	10	372	16	57,22	5	806.340
2	164,43	10	372	16	54,79	5	606.870
3	23,16	17	146	8	87,44	4	40.460
4	87,61	16	146	8	70,36	4	170.030
5	51,31	14	146	8	90,58	5	60.340
6	30,00	13	146	8	49,26	5	47.320
7	166,70	11	146	8	88,56	5	284.680
8	315,06	10	146	8	83,03	5	419.350
9	7,50	9	146	8	73,90	5	9.840
10	6,50	8	146	8	89,74	5	7.550
11	32,29	6	146	8	73,92	5	32.990
12	143,09	25	146	8	88,00	4	117.380
13	48,99	24	146	8	71,09	4	32.430
14	24,68	16	146	8	70,69	4	53.500
15	25,10	14	146	8	71,46	5	29.970
16	42,74	13	146	8	76,03	5	83.720
17	53,22	12	146	8	81,51	5	82.290
18	33,84	11	146	8	73,67	5	59.110
19	164,43	10	146	8	91,34	5	432.270
20	229,56	22	247	10	83,10	4	307.680
21	37,05	19	247	10	90,56	4	45.400
22	1,15	18	247	10	90,36	5	1.870
23	15,36	15	247	10	89,16	5	18.720

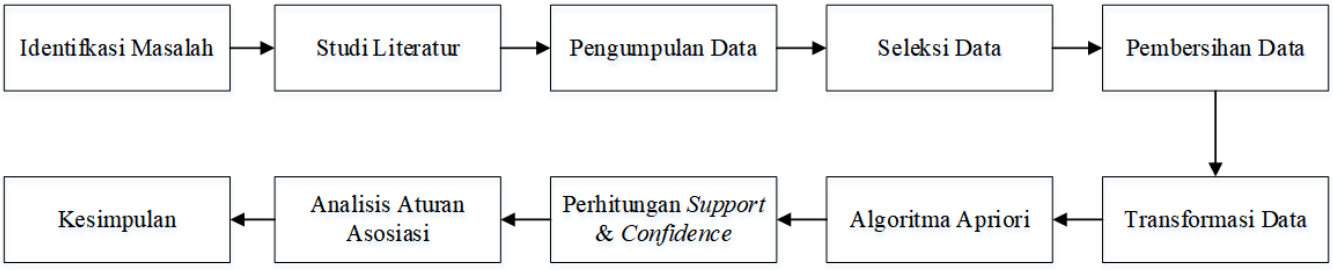
No.	Luas Lahan (Ha)	Usia Tanaman (Tahun)	Curah Hujan (MM)	Hari Hujan (HH)	Capaian Pemupukan (%)	Rotasi Panen	Capaian Produksi (Kg)
24	46,98	14	247	10	80,12	5	53.800
25	59,04	11	247	10	82,42	5	114.000
26	7,30	10	247	10	91,08	5	12.520
27	30,71	9	247	10	81,11	5	45.490
28	84,00	8	247	10	78,79	5	101.140
29	37,17	7	247	10	82,35	5	50.230
30	17,10	6	247	10	120,30	5	17.500

Berdasarkan Tabel 2, setiap data terdiri atas informasi tahun tanam dan enam faktor yang digunakan untuk menggambarkan kondisi produksi, yaitu luas lahan, usia tanaman, curah hujan, hari hujan, capaian pemupukan, dan rotasi panen. Capaian Produksi (Kg) digunakan sebagai variabel keluaran yang selanjutnya dikategorikan menjadi Rendah, Sedang, dan Tinggi sesuai dengan kriteria penelitian. Data numerik tersebut selanjutnya melalui tahap preprocessing dan kategorisasi sebelum digunakan dalam algoritma Apriori. Kategorisasi diperlukan karena algoritma Apriori bekerja dengan item atau kategori yang dapat dikombinasikan menjadi frequent itemset. Dengan demikian, nilai numerik pada data awal tidak langsung digunakan sebagai item, tetapi terlebih dahulu direpresentasikan ke dalam kategori yang telah ditentukan.

Dalam penelitian ini, Tahun Tanam digunakan sebagai informasi pendukung untuk memperoleh Usia Tanaman, sedangkan enam variabel lainnya digunakan dalam pembentukan Association Rule. Struktur analisis menempatkan Luas Lahan, Usia Tanaman, Curah Hujan, Hari Hujan, Capaian Pemupukan, dan Rotasi Panen sebagai faktor produksi, sedangkan Capaian Produksi digunakan sebagai consequent. Struktur tersebut kemudian menjadi dasar dalam proses seleksi Association Rule hingga diperoleh Best Rule Association. Struktur variabel ini juga sesuai dengan data penelitian pada skripsi yang mencantumkan ketujuh atribut tersebut.

3.2. Tahapan Analisis

Analisis data dilakukan menggunakan algoritma Apriori dengan perangkat lunak WEKA. Tahapan penelitian dimulai dari pengumpulan dan persiapan data, kategorisasi variabel, import dataset ke WEKA, pemilihan atribut yang digunakan, proses pembentukan frequent itemset, pembentukan Association Rule, seleksi aturan berdasarkan kriteria penelitian, dan penentuan Best Rule Association. Tahapan analisis dapat dirumuskan sebagai berikut:



Gambar 1. Alur Penelitian

Pada tahap awal, dataset disiapkan dalam format yang dapat dibaca oleh WEKA. Selanjutnya dilakukan pemeriksaan atribut dan data untuk memastikan seluruh variabel yang digunakan dalam penelitian dapat diproses. Data yang telah dikategorikan kemudian digunakan dalam proses pencarian kombinasi item menggunakan algoritma Apriori.

3.3. Penerapan Algoritma Apriori

Algoritma Apriori digunakan untuk menemukan frequent itemset dan Association Rule dari data faktor produksi kelapa sawit. Proses pencarian dilakukan secara bertahap mulai dari L1, kemudian dilanjutkan ke kombinasi itemset yang lebih besar hingga tidak terdapat lagi kombinasi yang memenuhi batas minimum support. Parameter algoritma Apriori ditentukan sebelum proses analisis dilakukan. Penelitian menggunakan minimum support sebesar 5% dan minimum confidence sebesar 90%. Minimum support digunakan untuk menentukan frekuensi minimum kemunculan suatu kombinasi item dalam dataset, sedangkan minimum confidence digunakan untuk menentukan tingkat keyakinan minimum suatu aturan asosiasi.

Nilai support digunakan untuk mengukur frekuensi kemunculan suatu pola dalam keseluruhan dataset, sedangkan confidence digunakan untuk mengukur tingkat kepastian bahwa consequent muncul ketika kondisi pada antecedent terpenuhi. Secara matematis, nilai support dan confidence dihitung dengan persamaan 1 dan 2. Parameter tersebut digunakan sebagai batas dalam proses pembentukan dan penyaringan Association Rule.

3.4. Implementasi Menggunakan WEKA

Implementasi dilakukan menggunakan perangkat lunak WEKA melalui fasilitas Associate dengan algoritma Apriori. Dataset yang telah dipersiapkan kemudian dimasukkan ke dalam WEKA dan digunakan sebagai input dalam proses pencarian frequent itemset serta Association Rule. Pengaturan algoritma dilakukan dengan menetapkan minimum support sebesar 0,05, minimum confidence sebesar 0,90, dan jumlah rule yang dapat ditampilkan sebanyak 1.000 rule. Parameter outputItemSets diaktifkan agar frequent itemset yang terbentuk dapat ditampilkan sebagai bagian dari proses analisis.

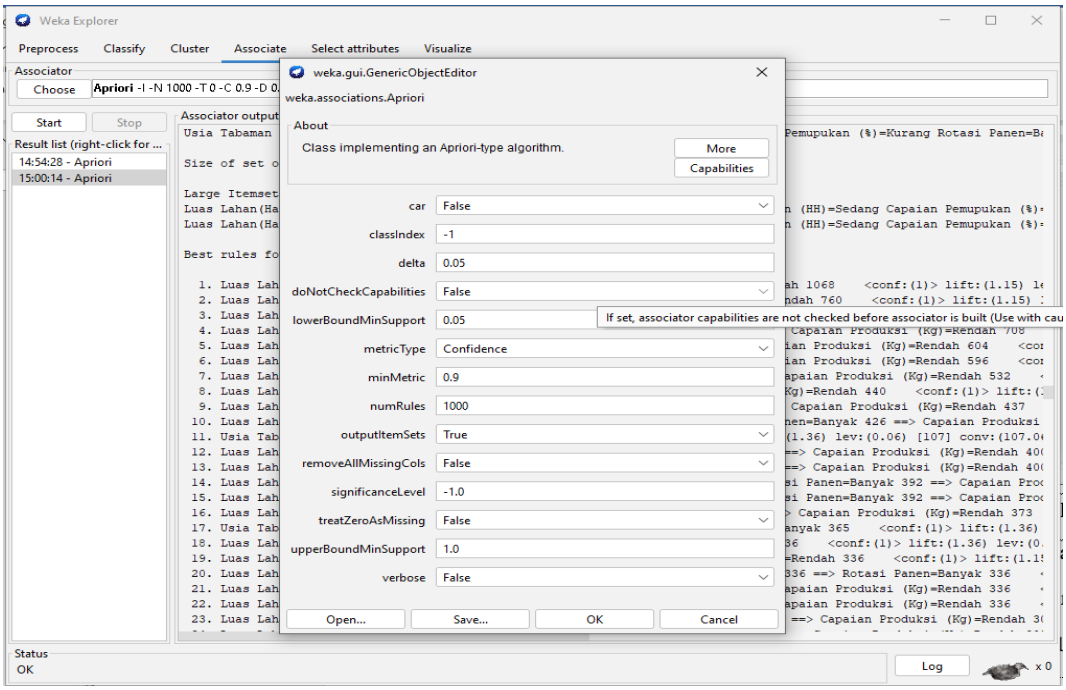
3.5. Seleksi Association Rule

Setelah Association Rule terbentuk, dilakukan seleksi berdasarkan tujuan penelitian. Seleksi dilakukan dengan mempertahankan aturan yang memiliki struktur sesuai dengan variabel penelitian, yaitu enam faktor produksi sebagai antecedent dan Capaian Produksi sebagai consequent. Aturan yang digunakan sebagai kandidat Best Rule Association harus memuat seluruh tujuh atribut penelitian, yaitu Luas Lahan, Usia Tanaman, Curah Hujan, Hari Hujan, Capaian Pemupukan, Rotasi Panen, dan Capaian Produksi. Selanjutnya, kandidat dibandingkan berdasarkan nilai confidence dan support. Confidence digunakan sebagai ukuran utama tingkat kepastian aturan, sedangkan support digunakan sebagai pertimbangan berikutnya apabila terdapat aturan dengan nilai confidence yang sama. Tahap seleksi tersebut dilakukan untuk memastikan bahwa aturan yang dipilih benar-benar merepresentasikan hubungan kombinasi faktor produksi dengan capaian produksi sesuai dengan fokus penelitian.

4. HASIL DAN PEMBAHASAN

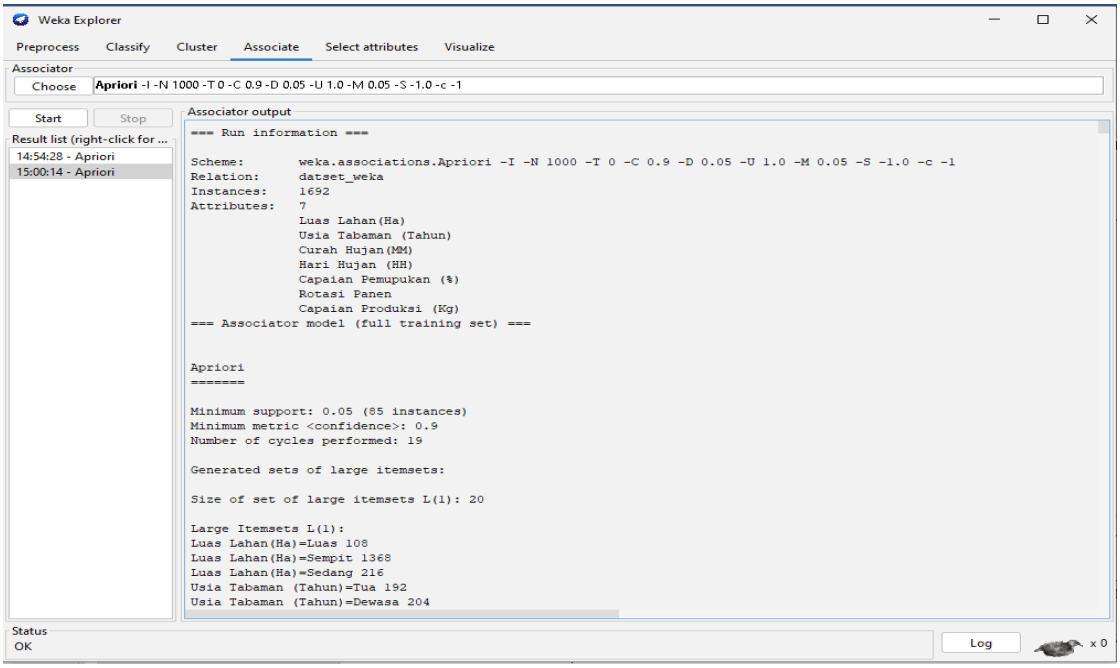
4.1. Implementasi dan Pembentukan Frequent Itemset

Implementasi algoritma Apriori dilakukan menggunakan perangkat lunak WEKA terhadap 1.692 data faktor produksi kelapa sawit periode 2023–2025. Dataset terdiri atas tujuh atribut penelitian, yaitu Luas Lahan, Usia Tanaman, Curah Hujan, Hari Hujan, Capaian Pemupukan, Rotasi Panen, dan Capaian Produksi. Data yang telah dikategorikan kemudian diproses menggunakan algoritma Apriori untuk membentuk frequent itemset dan Association Rule. Pada pengujian awal digunakan minimum support sebesar 10% dan minimum confidence sebesar 90%. Berdasarkan proses pencarian menggunakan WEKA, kombinasi frequent itemset yang terbentuk hanya mencapai tingkat tertentu dan belum menghasilkan kombinasi yang sesuai dengan struktur tujuh variabel penelitian. Oleh karena itu, dilakukan pengujian lanjutan dengan menurunkan minimum support menjadi 5%, sedangkan minimum confidence tetap sebesar 90%.



Gambar 2. Pengaturan Parameter Algoritma Apriori pada WEKA

Pada pengaturan akhir tersebut, nilai minimum support sebesar 5% setara dengan sekitar 85 transaksi dari 1.692 data. Artinya, suatu kombinasi item harus muncul sekurang-kurangnya pada 85 transaksi agar dapat dipertahankan dalam proses pembentukan frequent itemset. Selanjutnya, WEKA melakukan pencarian kombinasi item secara bertahap mulai dari L1 hingga tingkat kombinasi yang memenuhi batas minimum support. Hasil proses menunjukkan bahwa pembentukan frequent itemset berlangsung hingga L7. Pada tahap L7 diperoleh 585 instances sebagai bagian dari hasil proses pencarian kombinasi item yang selanjutnya digunakan dalam pembentukan aturan asosiasi.



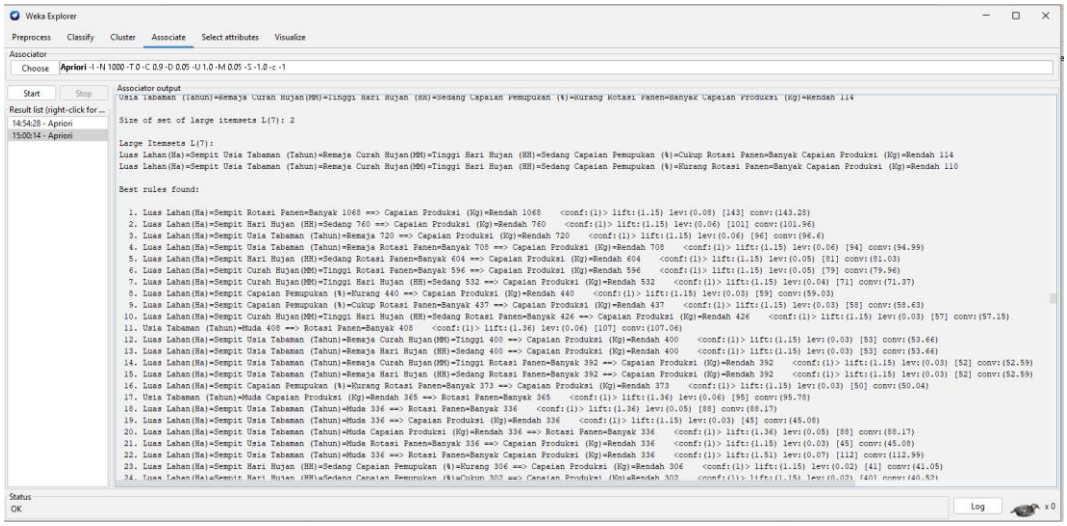
Gambar 3. Hasil Pembentukan Frequent Itemset pada WEKA

Proses tersebut kemudian dilanjutkan dengan pembentukan Association Rule. WEKA menghasilkan aturan berdasarkan parameter minimum support dan minimum confidence yang telah ditetapkan. Seluruh aturan yang terbentuk kemudian diperiksa berdasarkan tingkat itemset untuk memperoleh aturan yang sesuai dengan struktur variabel penelitian.

4.2. Seleksi Association Rule

Tahap selanjutnya adalah melakukan seleksi terhadap aturan asosiasi yang terbentuk. Seleksi dilakukan karena tidak seluruh aturan yang dihasilkan oleh WEKA secara langsung sesuai dengan tujuan penelitian. Penelitian ini secara khusus mengidentifikasi hubungan antara enam faktor produksi dengan Capaian Produksi. Kriteria aturan yang digunakan adalah Luas Lahan, Usia Tanaman, Curah Hujan, Hari Hujan, Capaian Pemupukan, dan Rotasi Panen sebagai antecedent, sedangkan Capaian Produksi sebagai consequent. Dengan demikian, aturan yang dipertahankan harus memuat seluruh tujuh atribut penelitian dalam satu kombinasi aturan.

Berdasarkan hasil penyaringan terhadap aturan yang terbentuk, diperoleh 12 aturan dengan 7-itemset. Selanjutnya, dari 12 aturan tersebut dilakukan seleksi kembali berdasarkan kesesuaian struktur variabel penelitian. Proses seleksi menghasilkan dua kandidat Association Rule yang memenuhi seluruh kriteria tersebut.



Gambar 4. Hasil Seleksi Association Rule 7-Itemset pada WEKA

Selanjutnya hasil seleksi Seleksi Association Rule 7-Itemset pada WEKA disajikan pada tabel berikut.

Tabel 3. Kandidat Best Rule Association

No.	*Association Rule*	*Support*	*Confidence*
1	Luas Lahan = Sempit, Usia Tanaman = Remaja, Curah Hujan = Tinggi, Hari Hujan = Sedang, Capaian Pemupukan = Cukup, Rotasi Panen = Banyak → Capaian Produksi = Rendah	6,74%	100,00%
2	Luas Lahan = Sempit, Usia Tanaman = Remaja, Curah Hujan = Tinggi, Hari Hujan = Sedang, Capaian Pemupukan = Kurang, Rotasi Panen = Banyak → Capaian Produksi = Rendah	6,50%	100,00%

Berdasarkan Tabel 3, kedua kandidat rule memiliki confidence sebesar 100%. Nilai tersebut menunjukkan bahwa seluruh data yang memenuhi kondisi pada bagian antecedent masing-masing aturan juga memiliki Capaian Produksi dalam kategori Rendah. Perbedaan kedua kandidat terdapat pada variabel Capaian Pemupukan. Kandidat pertama memiliki kategori Cukup, sedangkan kandidat kedua memiliki kategori Kurang. Dari sisi support, kandidat pertama memiliki nilai 6,74%, sedangkan kandidat kedua sebesar 6,50%. Dengan demikian, kandidat pertama memiliki frekuensi kemunculan yang lebih tinggi dalam keseluruhan dataset.

4.3. Penentuan Best Rule Association

Penentuan Best Rule Association dilakukan setelah seluruh kandidat yang sesuai dengan struktur penelitian diperoleh. Kriteria utama yang digunakan adalah nilai confidence, kemudian apabila terdapat beberapa aturan dengan nilai confidence yang sama, dilakukan perbandingan berdasarkan nilai support. Kedua kandidat memiliki confidence yang sama, yaitu 100%. Oleh karena itu, proses seleksi dilanjutkan dengan membandingkan nilai support. Kandidat pertama memiliki support sebesar 6,74%, lebih tinggi daripada kandidat kedua yang memiliki support sebesar 6,50%. Berdasarkan kriteria tersebut, kandidat pertama ditetapkan sebagai Best Rule Association.

Tabel 4. Best Rule Association

Parameter	Hasil
Best Rule	Luas Lahan = Sempit, Usia Tanaman = Remaja, Curah Hujan = Tinggi, Hari Hujan = Sedang, Capaian Pemupukan = Cukup, Rotasi Panen = Banyak → Capaian Produksi = Rendah
Support	6,74%
Confidence	100,00%
Jumlah kemunculan	114 instances
Capaian Produksi	Rendah

Berdasarkan Tabel 4, Best Rule Association yang diperoleh menunjukkan bahwa kombinasi Luas Lahan = Sempit, Usia Tanaman = Remaja, Curah Hujan = Tinggi, Hari Hujan = Sedang, Capaian Pemupukan = Cukup, dan Rotasi Panen = Banyak berasosiasi dengan Capaian Produksi = Rendah. Aturan tersebut memiliki support sebesar 6,74%, yang menunjukkan bahwa pola tersebut muncul pada 114 data dari 1.692 data penelitian. Sementara itu, confidence sebesar 100% menunjukkan bahwa seluruh transaksi yang memenuhi kombinasi keenam kondisi pada antecedent juga memiliki Capaian Produksi dalam kategori Rendah. Pemilihan aturan tersebut sebagai Best Rule bukan karena memiliki confidence lebih tinggi dibandingkan kandidat lainnya, karena kedua kandidat sama-sama memiliki confidence 100%. Aturan tersebut dipilih karena memiliki nilai support yang lebih tinggi, sehingga kombinasi faktor produksi tersebut muncul lebih banyak dalam dataset dibandingkan kandidat lainnya.

5. KESIMPULAN

Berdasarkan hasil penelitian, penerapan algoritma Apriori menggunakan WEKA berhasil digunakan untuk mengidentifikasi pola asosiasi antara faktor-faktor produksi dengan capaian produksi kelapa sawit pada PT Perkebunan Nusantara IV Regional II Kebun Tanjung Jati. Data sebanyak 1.692 periode 2023–2025 yang terdiri atas Luas Lahan, Usia Tanaman, Curah Hujan, Hari Hujan, Capaian Pemupukan, Rotasi Panen, dan Capaian Produksi dapat diolah melalui pembentukan frequent itemset dan Association Rule secara bertahap dari L1 hingga L7. Dengan minimum support 5% dan minimum confidence 90%, diperoleh pola asosiasi yang selanjutnya diseleksi berdasarkan struktur variabel penelitian. Hasil seleksi menghasilkan dua kandidat rule yang memenuhi struktur tujuh variabel, dengan confidence masing-masing 100%. Berdasarkan perbandingan support, diperoleh Best Rule Association dengan kombinasi Luas Lahan = Sempit, Usia Tanaman = Remaja, Curah Hujan = Tinggi, Hari Hujan = Sedang, Capaian Pemupukan = Cukup, dan Rotasi Panen = Banyak → Capaian Produksi = Rendah. Aturan tersebut memiliki support sebesar 6,74%, confidence 100%, dan muncul pada 114 data. Hasil ini menunjukkan bahwa algoritma Apriori dapat memberikan pola asosiasi yang dapat dimanfaatkan sebagai informasi pendukung dalam evaluasi kondisi produksi kelapa sawit.

Berdasarkan hasil penelitian, PT Perkebunan Nusantara IV Regional II Kebun Tanjung Jati dapat memanfaatkan pola asosiasi yang dihasilkan sebagai salah satu informasi pendukung dalam melakukan evaluasi terhadap kondisi faktor produksi yang berkaitan dengan capaian produksi rendah. Penelitian selanjutnya disarankan menggunakan data dengan periode pengamatan yang lebih panjang dan jumlah data yang lebih besar agar pola asosiasi yang diperoleh dapat

menggambarkan kondisi produksi secara lebih luas. Selain itu, penelitian berikutnya dapat menambahkan faktor lain yang relevan, seperti jenis tanah, pemeliharaan tanaman, serangan hama dan penyakit, tenaga kerja, jenis pupuk, serta faktor lingkungan lainnya. Penggunaan algoritma lain seperti FP-Growth juga dapat dipertimbangkan sebagai pembanding untuk mengetahui perbedaan pola dan efisiensi proses pencarian Association Rule

DAFTAR PUSTAKA

[1] B. P. Statistik, "Statistik Tanaman Perkebunan Tahunan Indonesia 2024: Kelapa Sawit, Kopi, Kakao, Karet, Teh, dan Komoditas Perkebunan Unggulan," Badan Pusat Statistik, 2025.

[2] D. J. Perkebunan, "Laporan Kinerja Direktorat Jenderal Perkebunan Tahun 2025," Kementerian Pertanian Republik Indonesia, 2026.

[3] Y. Anatesya, A. Fauzi, and R. Saragih, "Implementasi Association Rule pada Sistem Rekomendasi Peningkatan Hasil Pertanian Menggunakan Metode Apriori," no. 4, 2024.

[4] N. F. Hasan, R. Hammad, D. E. Profesi, and K. Kusrini, "Sistem Pendukung Keputusan Penentuan Kombinasi Paket Produk Pertanian Menggunakan Algoritma Apriori," *Eksplora Inform.*, vol. 9, no. 1, pp. 38–49, 2019, doi: 10.30864/eksplora.v9i1.261.

[5] A. Sapitri, N. Elisya, N. M. Mustafa, and M. Badrul, "Penerapan Data Mining Menggunakan Algoritma Apriori Untuk Menentukan Minat Customer Parfume Dari Riwayat Data Penjualan," *Inf. Syst. Educ. Prof.*, vol. 7, no. 1, pp. 11–20, 2022.

[6] N. Q. Rahmawati and Z. Fatah, "IMPLEMENTASI DATA MINING TERHADAP DATA MINAT MAHASISWA MENGGUNAKAN METODE APRIORI," vol. 2, no. November, pp. 215–228, 2024, doi: https://doi.org/10.29103/jtku.v13i2.19563 1.

[7] I. E. Mustika, Yunita Ardilla, Abraham Manuhutu, Nazaruddin Ahmad, Imanuddin Hasbi, Guntoro, Melda Agnes Manuhutu, Mohamad Ridwan, Hozairi, Anindya Khrisna Wardhani, Syariful Alim, Ikhsan Romli, Yoga Religia, D Tri Octafian, Unggul Utan Sufandi, *Data Mining dan Aplikasinya*. Penerbit Widina Bhakti Persada, Bandung, 2021.

[8] T. Ningsih, I. O. Y. Sitompul, and S. D. Siahaan, "Analisa Faktor-Faktor Yang Mempengaruhi Produksi Kelapa Sawit Di Kebun Tanah Raja PT . Bakrie Sumatera Plantations," no. 2, pp. 166–174, 2023, doi: https://doi.org/10.30596/jasc.v7i2.16628.