

PENERAPAN DATA MINING DALAM PREDIKSI PRODUKSI BERAS MENGUNAKAN METODE REGRESI LINEAR

Adinata ¹, Ade Irma Purnamasari ², Irfan Ali ³

^{1,2} Program Studi Teknik Informatika, STMIK IKMI Cirebon

³ Program Studi Rekayasa Perangkat Lunak, STMIK IKMI Cirebon

Jl Perjuangan No.10B Majasem Kec. Kesambi Kota Cirebon

Adinata706@gmail.com

ABSTRAK

Kondisi geografis Indonesia yang mendukung serta memiliki iklim tropis sangat cocok untuk membudidayakan padi. Sehingga pada tahun 2018, Indonesia berhasil menjadi negara penghasil beras terbanyak di dunia. Hal ini menjadikan Indonesia sebagai negara penghasil beras terbanyak di dunia pada tahun 2018. Analisis regresi digunakan untuk memprediksi nilai dari suatu variabel response (y) dengan menggunakan nilai dari variabel prediktor/independent (x). Tujuan penelitian ini adalah untuk Mendapatkan gambaran daerah-daerah penghasil beras di Kabupaten Cirebon, Mengklasifikasikan daerah-daerah penghasil beras di Kabupaten Cirebon berdasarkan kuantitas hasil panen per tahun, Mengimplementasikan data mining dengan metode regresi linear untuk memprediksi produksi beras di daerah-daerah prioritas produksi beras di Kabupaten Cirebon. Hasil penelitian menunjukkan bahwa RMSE dan Relative Error, menunjukkan jika metode prediksi yang dipergunakan baik. Hasil penelitian nilai evaluasi yang di dapat dari perhitungan prediki menggunakan RMSE 285.590 dan Relative Error 0.000 %. Dan prediksi produk beras yang disimulasikan dengan rapidminer mempunyai nilai min 140 dan nilai max 41060 dengan nilai rata-rata 9037.220 dan standard deviation 7837.949, Hasil evaluasi nilai RMSE pada produksi beras adalah 285.590, sedangkan nilai Relative Error adalah 0.000%. Nilai RMSE dan Relative Error yang mendekati 0 menunjukkan bahwa prediksi mendekati nilai aktual.

Kata kunci : *produksi beras, prediksi, regresi linear*

1. PENDAHULUAN

Salah satu sumber makanan pokok yang sebagian besar dikonsumsi oleh masyarakat Indonesia adalah beras salah satu makanan pokok yang dikonsumsi oleh sebagian besar masyarakat Indonesia. Tanaman padi (*Oryza sativa* L.) merupakan tanaman penghasil beras Padi adalah tanaman yang menghasilkan beras. Padi dapat tumbuh di daerah yang memiliki iklim tropis atau subtropis. Kondisi geografis Indonesia yang mendukung, terutama iklim tropisnya, sangat cocok untuk membudidayakan padi. Indonesia berhasil menduduki peringkat pertama sebagai negara penghasil beras terbanyak di dunia pada tahun 2018. Produksi beras sangat penting bagi kehidupan masyarakat. Beras adalah makanan pokok yang memberikan energi bagi tubuh untuk melakukan berbagai aktivitas. Populasi masyarakat yang terus bertambah menyebabkan kebutuhan akan beras juga meningkat dari waktu ke waktu. Kekurangan pasokan beras dapat menyebabkan kelangkaan bahan pangan, yang dapat berdampak pada kegiatan masyarakat. Pemerintah harus berupaya untuk menjaga ketersediaan pasokan beras yang berkualitas dan mencukupi untuk memenuhi kebutuhan pangan masyarakat. [1]

Beras merupakan salah satu kebutuhan penting yang harus dipenuhi oleh setiap orang yang tinggal di Indonesia. Beras merupakan kebutuhan pokok yang tidak dapat digantikan dengan barang lain. Beras merupakan kebutuhan pokok yang penting bagi masyarakat, sehingga kenaikan atau penurunan harga tidak akan mempengaruhi pola konsumsi masyarakat.

Oleh karena itu, perhatian terhadap harga beras sangat penting karena harga beras cenderung berubah-ubah.

Menganalisa faktor-faktor yang mempengaruhi produksi beras dan memprediksi jumlah produksi beras tersebut dengan 2 metode selanjutnya akan membandingkan hasil yang lebih akurat dengan cara menghitung nilai rata-rata kesalahan relatif dari peramalan menggunakan Logika Fuzzy dan Analisis Regresi Berganda. Tujuannya agar membantu pemerintah mengetahui faktor-faktor penentu dalam memaksimalkan produksi beras sehingga Indonesia mampu untuk berswasembada. [2]

Saat ini pertanian menjadi sebuah unsur penting dalam menyokong ketahanan pangan. Berdasarkan data dari BPS Jawa Timur menunjukkan bahwa jumlah produksi padi mengalami peningkatan dan penurunan dari tahun ke tahun. Berapa penelitian yang dilakukan sebelumnya diantaranya metode Simple Regresi Linear memiliki MAPE diatas 10% dan metode Single Exponential Smoothing memiliki MAPE dibawah 10%. [3].

Meskipun Indonesia menduduki peringkat ketiga sebagai produsen padi di Asia, bahkan di dunia; mempertimbangkan jumlah penduduk Indonesia berada pada peringkat keempat dunia setelah Tiongkok, India, dan Amerika Serikat dengan jumlah penduduk Indonesia tahun 2019 tercatat sekitar 270 juta jiwa (Badan Pusat Statistik, 2020), Indonesia juga merupakan negara pengimpor beras dominan. BPS Indonesia mencatat pada periode tahun 2010–2018, Indonesia melakukan impor beras rata-rata per tahun sebesar 1,2 juta ton dengan impor pada tahun

2018 tercatat 2,2 juta ton. Bila dibandingkan dengan ekspor beras Indonesia pada tahun 2018 sebesar 3 196 ton, maka secara empirik terlihat meskipun menduduki peringkat keti- ga sebagai produsen beras Indonesia merupakan salah satu negara pengimpor beras terbesar.[4][5]

Semakin meningkat. Kebutuhan yang tinggi harus diikuti dengan peningkatan jumlah produksi agar tidak timbul permasalahan. Pemerintah sejak awal telah memberikan perhatian khusus tentang kebijakan beras nasional.[6] Banyak kalangan dari masyarakat yang memilih beras tipe ini sebagai bahan pokok di tempatnya masing-masing. Sebagai negara yang memproduksi beras serta berada di lingkungan negara-negara tetangga yang juga memproduksi beras, Indonesia melalui bagian pemerintah terkait juga turut mengontrol harga dari beras tersebut khususnya beras premium. Hal ini dilakukan agar harga beras premium yang ada di masyarakat sesuai dengan kondisi perekonomian di wilayah sekitar. Peluang para pedagang untuk berbuat curang dalam memainkan harga juga dapat direduksi dengan adanya kebijakan dari pemerintah terkait harga tersebut.[7]

Sebagai salah satu pilar ekonomi negara, sektor pertanian diharapkan dapat meningkatkan pendapatan terutama dari penduduk pedesaan yang masih di bawah garis kemiskinan. Untuk itu, berbagai investasi dan kebijakan telah dilakukan pemerintah untuk mendorong pertumbuhan di sektor pertanian. Investasi di sektor pertanian seringkali sangat mahal, ditambah lagi tingkat pengembaliannya sangat rendah dan waktu investasinya juga panjang sehingga tidak terlalu menarik swasta. pembangunan irigasi, penyuluhan pertanian dan berbagai bentuk investasi dalam bentuk subsidi dan lainnya pada umumnya harus dilakukan oleh pemerintah.[8]

Objek penelitian ini adalah petani padi sawah di Desa Pamotan Kecamatan Kalipucang Kabupaten Pangandaran. Data yang dipergunakan merupakan data primer dan data sekunder. Metode penarikan sampel menggunakan purposive sampling dan pemilihan daerah penelitian ditentukan secara sengaja. Analisis data dengan regresi berganda.[9]

2. TINJAUAN PUSTAKA

2.1. Beras

Beras merupakan salah satu bahan makanan pokok yang menjadi komoditas penting terutama di negara Indonesia dikarenakan beras menjadi bahan konsumsi utama bagi masyarakat untuk memperoleh asupan karbohidrat. [10]

2.2. Regresi Linear

Regresi linear adalah metode statistika yang digunakan untuk menganalisis hubungan sebab-akibat antara variabel, serta dapat digunakan untuk memprediksi nilai variabel tertentu. Variabel yang digunakan pada penelitian ini adalah luas panen, curah hujan, kelembaban, dan suhu rata-rata.

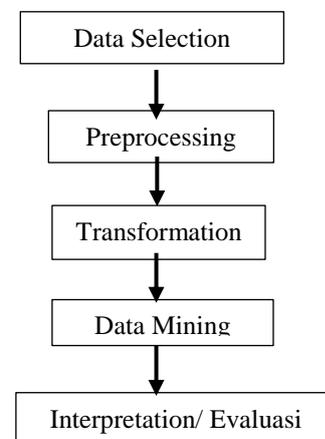
2.3. Data Mining

Data mining adalah teknologi yang dapat digunakan untuk menemukan informasi penting dari kumpulan data besar. Beberapa aplikasi data mining digunakan untuk memprediksi apa yang akan terjadi di masa depan. Analisis regresi adalah metode statistika yang digunakan untuk mempelajari hubungan antara dua variabel atau lebih. Hubungan antara dua variabel tersebut bersifat linier, yaitu hubungan yang dapat digambarkan oleh sebuah garis lurus.

2.4. Rapidminer

RapidMiner adalah perangkat lunak yang dapat digunakan secara mandiri untuk analisis data, serta dapat diintegrasikan ke dalam produk lain sebagai mesin data mining.

2.5. Tahapan Perencanaan



a. Data Selection

Pemilihan data dilakukan untuk memilih data operasional yang akan digunakan dalam tahap ekstraksi data pada regresi linear. Hasil pemilihan data ditampilkan dalam satu halaman berdasarkan data operasional.

b. Preprocessing

Sebelum tahap selanjutnya dilakukan, perlu dilakukan pembersihan data untuk memastikan kualitas data yang akan digunakan dalam regresi linear. Pembersihan data meliputi penghapusan data duplikat, pengecekan data yang tidak konsisten, serta koreksi kesalahan data, seperti kekeliruan dalam cetak.

c. Transformation

Data perlu diubah atau ditransformasikan agar sesuai dengan langkah penambangan data. Prediksi merupakan langkah yang kreatif dan sangat bergantung pada jenis data atau model yang digunakan.

d. Data Mining

Pada tahap ini, teknik atau metode digunakan untuk menemukan pola atau hasil yang menarik dari data. Terdapat berbagai teknik, metode, atau algoritma penambangan data yang dapat digunakan. Pemilihan teknik atau metode yang

tepat sangat bergantung pada tujuan dan keseluruhan tahapan dalam regresi linear.

e. Evaluasi

Model data yang diperoleh dari analisis data dapat diinterpretasikan agar mudah dimengerti oleh pihak yang berkepentingan. Tujuan interpretasi model data adalah untuk menentukan apakah kebijakan atau informasi yang diperoleh sesuai dengan informasi yang ada.

2.6. Sumber Data

Sumber Data yang digunakan dalam penelitian adalah data primer dan data sekunder. Sumber data primer merupakan sumber data yang diperoleh secara langsung dari sumber asli dan tidak melalui media perantara. Data historis transaksi hasil panen yang digunakan diperoleh secara langsung dari objek penelitian melalui wawancara dan dokumentasi. Sedangkan data sekunder merupakan sumber data penelitian yang diperoleh secara tidak langsung melalui media perantara diperoleh dan dicatat oleh pihak lain. Data sekunder pada umumnya berupa bukti catatan atau laporan historis yang dipublikasikan. Data sekunder yang dimaksud dalam penelitian ini adalah sumber data yang digunakan untuk menunjang kelengkapan teori data primer. Adapun data yang didapatkan melalui hasil studi pendahuluan bahwasanya perkembangan produksi beras.

2.7. Teknik Pengumpulan Data

2.7.1. Observasi

Observasi atau pengamatan adalah teknik atau cara mengumpulkan data dengan cara mengamati kegiatan yang sedang berlangsung. Observasi partisipasif adalah jenis observasi di mana peneliti terlibat langsung dengan kegiatan sehari-hari orang yang sedang diamati. Hal ini dilakukan untuk melihat dari dekat kegiatan yang dilakukan. Penelitian melakukan kunjungan ke Dinas Pertanian dan Peternakan Kabupaten Cirebon untuk mendapatkan informasi secara langsung tentang produksi beras di kabupaten tersebut pada periode 2020 hingga 2022.

2.7.2. Wawancara (Interview)

Wawancara adalah kegiatan percakapan antara dua orang untuk bertukar informasi dan ide. Percakapan tersebut dilakukan dengan cara saling bertanya dan menjawab, sehingga dapat menghasilkan makna baru dalam suatu topik tertentu. Dalam penelitian ini, wawancara dilakukan kepada salah satu staf IT dan pengguna sistem di dinas terkait di Kabupaten Cirebon. Wawancara dilakukan peneliti untuk memastikan bahwa data yang dicari berasal dari orang yang kompeten di bidangnya.

2.7.3. Studi Dokumentasi

Studi dokumentasi adalah teknik pengumpulan data dengan mengumpulkan dan menganalisis dokumen-dokumen, baik dokumen tertulis, gambar, maupun elektronik. Data diperoleh dengan mencari

referensi melalui jurnal dan buku yang berkaitan dengan permasalahan yang diteliti. Peneliti akan melakukan pencatatan secara lengkap dan cepat setelah data terkumpul untuk menghindari kemungkinan hilangnya data. Dalam studi dokumentasi, peneliti dapat melakukan kegiatan seperti mendokumentasikan atau meminta rekapitulasi hasil produksi beras.

2.8. Contoh Persamaan Matematika

Persamaan regresi linier sederhana adalah model matematika yang menggambarkan hubungan antara satu variabel bebas (X) dengan satu variabel tak bebas (Y). Persamaan regresi linier sederhana secara matematik diekspresikan oleh : [5]

$$Y^{\wedge} = a + bX$$

Yang mana :

Y ^ = garis regresi/ variable response

a = konstanta (intersep), perpotongan dengan sumbu vertikal

b = konstanta regresi (slope)

X = variabel bebas/ predictor

Besarnya konstanta a dan b dapat ditentukan menggunakan persamaan :

$$a = \frac{(\sum Y_1)(\sum X_1^2) - (\sum X_1)(\sum X_1 Y_1)}{n(\sum X_1^2) - (\sum X_1)^2}$$

$$b = \frac{n(\sum X_1 Y_1) - (\sum X_1)(\sum Y_1)}{n\sum X_1^2 - (\sum X_1)^2}$$

yang mana n = jumlah data

Langkah-langkah Analisis dan Uji Regresi Linier Sederhana Adapun langkah-langkah yang perlu dilakukan untuk melakukan analisis dan uji regresi linier sederhana adalah sebagai berikut

- a. Menentukan tujuan dari Analisis Regresi Linear Sederhana.
- b. Mengidentifikasi variabel predictor dan variabel response.
- c. Melakukan pengumpulan data dalam bentuk table
- d. Menghitung X², XY dan total dari masing-masingnya
- e. Menghitung a dan b menggunakan rumus yang telah ditentukan.
- f. Membuat model Persamaan Garis Regresi
- g. Melakukan prediksi terhadap variabel predictor atau response
- h. Uji signifikansi menggunakan Uji-t dan menentukan Taraf Signifikan
- i. Koefisien regresi b ditentukan dengan menggunakan rumus yang telah diberikan

3. METODE PENELITIAN

Penelitian ini menggunakan metode data sampel. Metode ini dipilih karena data yang digunakan dalam

penelitian ini merupakan data yang diperoleh dari kumpulan data terbaik. Metode ini sangat sesuai diterapkan dalam penelitian ini karena sebelum uji coba data dilakukan, data tersebut terlebih dahulu diproses untuk menghasilkan data yang valid. Data yang telah diproses kemudian diolah menggunakan program RapidMiner. RapidMiner adalah tools data mining yang memiliki antarmuka yang nyaman. Dalam RapidMiner, analisis data dikonfigurasi dalam sebuah proses view. Setiap langkah analisis digambarkan dengan sebuah operator dalam proses analisis.

Operator-operator ini memiliki port untuk input dan output, sehingga dapat berkomunikasi dengan operator lain untuk mendapatkan input data atau mengirim data yang telah diubah dan menggenerasi model. Davies-Bouldin Index adalah salah satu metode yang digunakan untuk mengukur validitas cluster pada suatu metode pengelompokan. Kohesi didefinisikan sebagai jumlah dari kedekatan data terhadap titik pusat cluster dari cluster yang diikuti. Evaluasi menggunakan Davies-Bouldin Index memiliki skema evaluasi internal, yaitu evaluasi yang dilakukan dengan melihat kuantitas dan kedekatan antar data hasil cluster.

4. HASIL DAN PEMBAHASAN

Pembahasan tentang cara menerapkan data mining untuk meningkatkan produksi beras menggunakan metode regresi linier.

4.1. Data Selection

Pemilihan data adalah proses memilih atribut yang diinginkan dari kumpulan data asli untuk digunakan dalam proses data mining selanjutnya. Proses pemilihan data dalam Rapidminer dilakukan dengan menggunakan operator pemilihan atribut. Operator ini dapat digunakan untuk memfilter atribut yang diinginkan berdasarkan kriteria tertentu. Jika metadata diketahui, maka semua atribut akan ditampilkan dan atribut yang diperlukan dapat dipilih dengan mudah. Hasil pemilihan data adalah kumpulan atribut yang akan digunakan dalam proses data mining selanjutnya.



Gambar 1. Hasil seleksi atribut

Berdasarkan gambar di atas, dapat diketahui bahwa hasil seleksi atribut adalah kolom nama kecamatan, tahun produksi 2021, dan produksi 2020.

Kolom-kolom tersebut akan digunakan sebagai data pelatihan dalam proses analisis data. Hasil seleksi atribut untuk jumlah kolom kecamatan adalah 40.

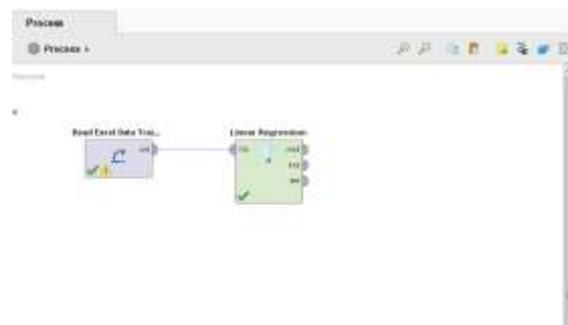
Row No.	nama daerah	hasil produksi	percobaan
1	Kecamatan	2023	2022
2	Waled	7	7761
3	Pasipaman	7	148
4	Cleatag	7	3481
5	Pabuatan	7	3566
6	Lozan	7	5831
7	Pabedian	7	7716
8	Bukan	7	7573
9	Getang	7	7622
10	Karangsemb	7	9125
11	Karangwang	7	6359
12	Lemahabang	7	5944
13	Sukat Lel	7	8572
14	Sedang	7	12381
15	Arangpura	7	5984
16	Pangeran	7	4525
17	Mudu	7	9996
18	Deber	7	6756

Gambar 2. Hasil seleksi atribut (data prediksi)

Berdasarkan gambar di atas, dapat diketahui bahwa hasil prediksi jumlah produksi beras pada tahun 2023 adalah berdasarkan nilai percobaan produksi beras pada tahun 2022. Hasil prediksi tersebut akan digunakan sebagai data pendukung dalam proses analisis data. Jumlah kolom kecamatan yang didapatkan dari hasil prediksi adalah 40.

4.2. Preprocessing

Proses prapemrosesan data meliputi penghapusan duplikat data, pemeriksaan data yang tidak konsisten, dan perbaikan kesalahan dalam data. Dalam langkah ini, menggunakan operator filter sampel RapidMiner untuk menghilangkan missing value. Data missing value adalah data yang tidak memiliki nilai pada salah satu atributnya. Data missing value yang dihilangkan dalam langkah ini adalah data yang terdapat pada dua atribut, yaitu jumlah produksi beras tahun prediksi dan jumlah produksi beras tahun sebelumnya. Berikut adalah proses penghapusan missing value di RapidMiner menggunakan operator filter examples.

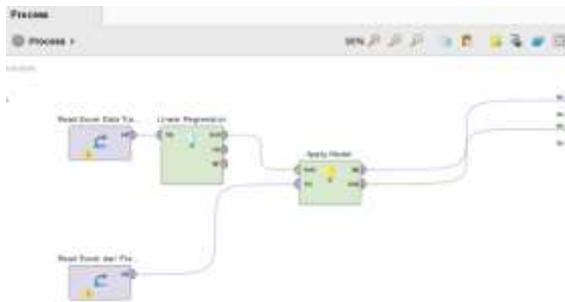


Gambar 3. Preprocessing

Pada proses prapemrosesan data, data diolah menggunakan metode regresi linear. Oleh karena itu, atribut yang diteliti sebagai data training dihubungkan dengan metode regresi linear.

4.3. Transformation

Pada tahap transformasi, data diubah menjadi bentuk yang sesuai dengan metode data mining yang akan digunakan. Tujuan transformasi adalah untuk mengubah data yang dipilih menjadi bentuk yang dapat diproses oleh prosedur penambangan. Dalam proses ini, kode kecamatan diubah menjadi ID, dan kode produksi diubah menjadi label. Atribut yang diberi label bertindak sebagai label, dan atribut dalam peran pelabelan bertindak sebagai operator pembelajaran. Label sering kali dianggap sebagai variabel atau kelas. Data berikut diproses untuk data mining.



Gambar 4. Penggunaan *linear regression* dan *apply model*

Gambar di atas menunjukkan bahwa untuk menentukan suatu prediksi, diperlukan data pelatihan dan metode yang digunakan, serta data yang akan diprediksi. Data-data tersebut kemudian dihubungkan dengan model yang telah dilatih.

No	Nama	Produksi (kg)	Label
1	Blora	17020.00	1
2	Pekalongan	141.840	1
3	Cirebon	1407000	1
4	Pekalongan	1406500	1
5	Lidah	1387120	1
6	Pekalongan	1378000	1
7	Bekasi	1376000	1
8	Cirebon	1322000	1
9	Pekalongan	1314000	1
10	Pekalongan	1300000	1
11	Pekalongan	1284000	1
12	Kediri	1271000	1
13	Cirebon	1250000	1
14	Pekalongan	1240000	1
15	Pekalongan	1232000	1
16	Blora	1221000	1
17	Blora	1217000	1
18	Blora	1212000	1
19	Blora	1208000	1
20	Cirebon	1199000	1
21	Pekalongan	1188000	1
22	Pekalongan	1184000	1
23	Blora	1181000	1
24	Blora	1177000	1
25	Blora	1173000	1
26	Blora	1169000	1
27	Blora	1165000	1
28	Blora	1161000	1
29	Blora	1157000	1
30	Blora	1153000	1
31	Pekalongan	1149000	1
32	Pekalongan	1145000	1
33	Blora	1141000	1
34	Blora	1137000	1
35	Blora	1133000	1
36	Blora	1129000	1
37	Blora	1125000	1
38	Blora	1121000	1
39	Blora	1117000	1
40	Blora	1113000	1
41	Blora	1109000	1
42	Blora	1105000	1
43	Blora	1101000	1
44	Blora	1097000	1
45	Blora	1093000	1
46	Blora	1089000	1
47	Blora	1085000	1
48	Blora	1081000	1
49	Blora	1077000	1
50	Blora	1073000	1
51	Blora	1069000	1
52	Blora	1065000	1
53	Blora	1061000	1
54	Blora	1057000	1
55	Blora	1053000	1
56	Blora	1049000	1
57	Blora	1045000	1
58	Blora	1041000	1
59	Blora	1037000	1
60	Blora	1033000	1
61	Blora	1029000	1
62	Blora	1025000	1
63	Blora	1021000	1
64	Blora	1017000	1
65	Blora	1013000	1
66	Blora	1009000	1
67	Blora	1005000	1
68	Blora	1001000	1
69	Blora	997000	1
70	Blora	993000	1
71	Blora	989000	1
72	Blora	985000	1
73	Blora	981000	1
74	Blora	977000	1
75	Blora	973000	1
76	Blora	969000	1
77	Blora	965000	1
78	Blora	961000	1
79	Blora	957000	1
80	Blora	953000	1
81	Blora	949000	1
82	Blora	945000	1
83	Blora	941000	1
84	Blora	937000	1
85	Blora	933000	1
86	Blora	929000	1
87	Blora	925000	1
88	Blora	921000	1
89	Blora	917000	1
90	Blora	913000	1
91	Blora	909000	1
92	Blora	905000	1
93	Blora	901000	1
94	Blora	897000	1
95	Blora	893000	1
96	Blora	889000	1
97	Blora	885000	1
98	Blora	881000	1
99	Blora	877000	1
100	Blora	873000	1

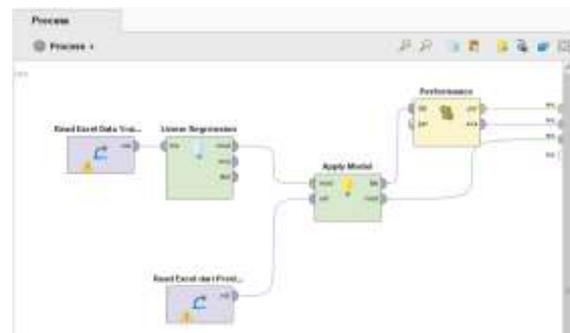
Gambar 5. Hasil Transformasi data prediksi

Data prediksi menggambarkan perkiraan hasil produksi beras yang akan dihasilkan pada tahun 2023. Perkiraan ini diperoleh dari hasil pengolahan data pelatihan dengan menggunakan metode regresi linear.

4.4. Data Mining

Data mining Proses data mining yang diterapkan dalam penelitian ini yaitu prediksi menggunakan regresi linear. Pada penelitian ini, akan dimuat perhitungan prediksi dari beberapa kecamatan yang ada di kab Cirebon secara spesifik dan nilai evaluasi dari hasil prediksi diukur dengan RMSE dan Relative Error. Entry Kategori entry adalah tahun berjalan produksi beras yang memiliki jumlah dalam tonase, pada sesi ini dilakukan data mining dengan metode regresi linear menggunakan aplikasi rapid miner, untuk memprediksi penjualan kedepannya. Pada proses ini dilakukan pemodelan algoritma regresi linear dengan mengubah nama atribut pada data Entry menjadi variabel X (tahun) dan Y (jumlah produksi), dilakukan perhitungan untuk mencari nilai a dan b untuk menghitung persamaan regresi linear.

Dalam proses ini, beberapa operator digunakan, yaitu operator Read Excel untuk membaca data produksi beras dari file Microsoft Excel. Operator Multiply digunakan untuk menyalin data dari operator Read Excel dan menghubungkannya ke operator Linear Regression. Operator Linear Regression digunakan untuk menghitung model regresi linear dari data input. Operator Apply Model digunakan untuk menerapkan model regresi linear yang telah dilatih pada data testing. Hasil prediksi kemudian ditampilkan di port Result. Desain pemodelan dapat dilihat pada gambar 6.



Gambar 6. Desain proses pemodelan

Berdasarkan gambar di atas, kinerja model dapat dievaluasi secara otomatis dengan menampilkan daftar nilai kriteria produksi beras yang disediakan dalam tugas tertentu.

4.5. Evaluasi Accuracy

Pada tahap ini, untuk memudahkan pengumpulan data, dilakukan proses evaluasi hasil record mining, salah satunya dengan menggunakan matriks konfusi. Matriks konfusi adalah alat yang digunakan untuk menilai akurasi klasifikasi. Benar menguntungkan dan benar negatif menunjukkan bahwa klasifikasi dilakukan dengan benar, sedangkan salah

menguntungkan dan salah negatif menunjukkan bahwa klasifikasi dilakukan dengan salah.

Model	Coefficient	Max Error	Min Coefficient	Tolerance	F-Stat	p-Value	Coll.
Intercept	1.000	0.075	0.999	1	30.244	0	---
lnarea2022	-68.234	100.000	-	-	0.760	0.384	---

Gambar 7. Tahapan Evaluasi

```

root_mean_squared_error
root_mean_squared_error: 285.590 (1)
    
```

Gambar 8. RSME linear regresi

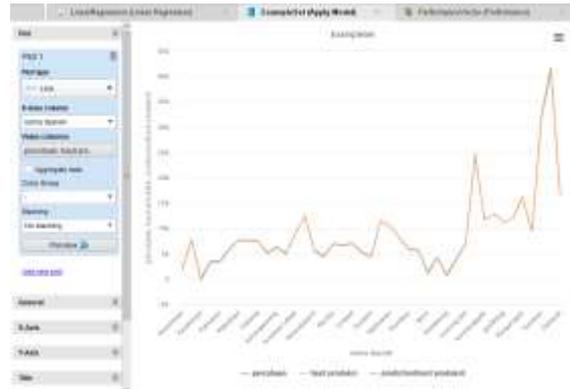
Berdasarkan hasil evaluasi, nilai RMSE pada produksi beras adalah 285.590. Nilai Relative Error adalah 0%. Nilai RMSE dan Relative Error yang mendekati 0 atau prediksi yang mendekati nilai aktual menunjukkan hasil prediksi yang sangat baik.

Berikut adalah prediksi jumlah produksi beras pada tahun 2023:

Row No.	wilayah daerah	hasil produksi	prediksi...	persentase
1	Beccamatari	2003	1737.410	3022
2	Waled		7643.306	7701
3	Pasolemas		139.267	140
4	Cheakuj		3238.814	3401
5	Pasuaran		3329.371	3509
6	Lusat		3760.019	5801
7	Papedian		7588.800	7716
8	Sakelaki		7449.742	7573
9	Cetang		7530.190	7622
10	Ranangpans		4830.594	5105
11	Ranangwang		4230.405	4309
12	Lerangwang		4847.240	5044
13	Sekolah Loo		8982.308	9072
14	Sedang		12337.478	12361
15	Asteropans		2020.041	2004
16	Pangemas		4313.190	4525
17	Sekada		6852.972	6936
18	Sekar		6889.602	6708

Gambar 9. Hasil Prediksi Jumlah produksi beras tahun 2023

Berdasarkan gambar di atas, dapat dilihat bahwa prediksi jumlah produksi beras rata-rata mengalami kenaikan di berbagai wilayah kecamatan di Kabupaten Cirebon. Detail produksi beras dapat dilihat pada tabel di atas.



Gambar 10. Visualisasi Prediksi Produksi Beras tahun 2023

Gambaran prediksi produksi beras di Kabupaten Cirebon dapat disimulasikan dalam bentuk grafik. Grafik tersebut menunjukkan bahwa produksi beras di setiap wilayah di Kabupaten Cirebon sangat bervariasi. Variasi tersebut dipengaruhi oleh luas lahan dan jumlah curah hujan di setiap wilayah.



Gambar 11. Data Mean, Min dan Max Prediksi produksi beras

Berdasarkan hasil simulasi dengan Rapidminer, prediksi produksi beras memiliki nilai minimum 140 ton, nilai maksimum 41.060 ton, nilai rata-rata 9.037,22 ton, dan simpangan baku 7.837,95 ton. Dengan menggunakan perintah yang ada di ms excel didapatkan hasil.

SUMMARY OUTPUT						
Regression Statistics						
Multiple R	0.99992413					
R Square	0.99984826					
Adjusted R Square	0.99979667					
Standard Error	648.878900					
Observations	46					
ANOVA						
	df	SS	MS	F	Significance F	
Regression	1	258831260	2.58831E+08	6131.681	1.86E-43	
Residual	44	15061185.6	342299.7			
Total	45	260000000				
Coefficients						
		Standard Error	t Stat	P-value	Lower 95% Lower	Upper 95% Upper
Intercept	402.218018	151.421769	2.65429	0.01207	93.23743	711.1986
X Variable 1	0.96502402	0.011294791	78.80723	1.08E-49	0.942524	0.987524

Gambar 12. data produksi beras

Dari data diatas dapat diinterpretasikan bahwa
 $Y = b_0 + b_1 X$
 $Y = 402.2 + 0.965 X$

Setiap kecamatan di Kabupaten Cirebon diperkirakan akan menghasilkan produksi beras sekitar 402,2 ton per tahun. Produksi beras tersebut akan mengalami kenaikan sebesar 0,96 ton per tahun.

5. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian tersebut, dapat disimpulkan bahwa setelah dilakukan proses pemodelan, diperoleh hasil prediksi produksi beras selama 3 bulan di kecamatan-kecamatan di Kabupaten Cirebon menggunakan algoritma Regresi Linear. RMSE dan Relative Error, Semakin kecil nilai error hasil prediksi dari kondisi sebenarnya, menunjukkan jika metode prediksi yang dipergunakan baik. Hasil nilai evaluasi yang di dapat dari perhitungan prediksi menggunakan RMSE 285.590 dan Relative Error 0.000 %. Dan prediksi produk beras yang disimulasikan dengan rapidminer mempunyai nilai min 140 dan nilai max 41060 dengan nilai rata-rata 9037.220 dan standard deviation 7837.949. Hasil evaluasi nilai RMSE pada produksi beras adalah 285.590, sedangkan nilai Relative Error adalah 0.000%. Nilai RMSE dan Relative Error yang mendekati 0 menunjukkan bahwa prediksi mendekati nilai aktual.

Berikut adalah beberapa saran yang dapat dipertimbangkan untuk pengembangan sistem peningkatan produksi beras menggunakan metode regresi linier: Bagi Peneliti Sistem ini diharapkan dapat ditingkatkan dengan metode lain agar lebih akurat dalam memprediksi produksi beras dan menyediakan layanan yang lebih baik. Pengembangan sistem ini dapat dilakukan dengan menambahkan fitur-fitur baru sesuai kebutuhan dan memperluas cakupannya dengan menggunakan metode lainnya. Bagi Dinas Pertanian Kab Cirebon Hasil penelitian ini diharapkan dapat digunakan sebagai rujukan untuk menyusun strategi ke depan agar dapat mengoptimalkan potensi pertanian di setiap kecamatan. Bagi Kampus Hasil penelitian ini diharapkan dapat menambah wawasan ilmiah, khususnya di bidang data mining dengan metode regresi linear. Selain itu, hasil penelitian ini juga diharapkan dapat menjadi dasar untuk pengembangan metode prediksi lainnya di masa depan.

DAFTAR PUSTAKA

- [1] S. Wijayanto and Y. Fathoni, M, "Pengelompokan Produktivitas Tanaman Padi di Jawa Tengah Menggunakan Metode Clustering K-Means," *J. JUPITER*, vol. 13, no. 2, pp. 212–219, 2021.
- [2] S. Yulia Retno, "Perbandingan Analisis Logika Fuzzy dan Regresi Linier Berganda dalam Menentukan Produksi Beras Nasional," *J. KomtekInfo*, vol. 8, no. 4, pp. 239–248, 2021, doi: 10.35134/komtekinfo.v8i4.186.
- [3] J. Terapan, S. & Teknologi, R. Novanda Putra, A. Aziz, and A. Zaini, "Implementasi Metode Simple Regresi Linear dan Single Exponential Smoothing untuk Memprediksi Produksi Padi Jawa Timur," *Fak. Sains dan Teknol. PGRI Kanjuruhan Malang*, vol. 5, no. 2, p. 2023, 2023, [Online]. Available: <https://doi.org/10.21067/jtst.v5i2.8545>
- [4] E. N. Kencana, D. Arnawa, and K. Jayanegara, "Memodelkan Impor Beras Menggunakan Regresi Data Panel," *J. Mat.*, vol. 10, no. 2, p. 135, 2021, doi: 10.24843/jmat.2020.v10.i02.p130.
- [5] A. N. Hanna, J. S. Mcdonald, C. H. Miller, and D. Couri, "Pretreatment with paracetamol inhibits metabolism of enflurane in rats," *Br. J. Anaesth.*, vol. 62, no. 4, pp. 429–433, 1989, doi: 10.1093/bja/62.4.429.
- [6] S. Sulistyorini and L. T. Sunaryanto, "Dampak Efisiensi Usahatani Padi Terhadap Peningkatan Produktivitas," *Jambura Agribus. J.*, vol. 1, no. 2, pp. 43–51, 2020, doi: 10.37046/jaj.v1i2.2680.
- [7] R. E. Putra and A. S. Sinaga, "Perkiraan Harga Beras Premium DKI Jakarta Menggunakan Regresi Linier," *J. Inf. Eng. Educ. Technol.*, vol. 6, no. 2, pp. 80–85, 2022, doi: 10.26740/jieet.v6n2.p80-85.
- [8] U. Seplida, S. Tan, and Y. Yulmardi, "Strategi peningkatan pendapatan petani padi di Kecamatan Gunung Tujuh Kabupaten Kerinci," *J. Paradig. Ekon.*, vol. 15, no. 2, pp. 213–228, 2020, doi: 10.22437/paradigma.v15i2.10324.
- [9] N. R. Walis, B. Setia, and A. Y. Isyanto, "Faktor-Faktor Yang Berpengaruh Terhadap Produksi Padi Di Desa Pamotan Kecamatan Kalipucang Kabupaten Pangandaran," *J. Ilm. Mhs. Agroinfo Galuh*, vol. 8, no. 3, p. 648, 2021, doi: 10.25157/jimag.v8i3.5419.
- [10] I. Febriani, M. Safii, and O. Alfina, "Implementasi Data Mining Peningkatan Produksi Beras Menggunakan Metode K-Means Clustering," *Maj. Ilm. METHODODA*, vol. 12, no. 3, pp. 258–268, 2022, doi: 10.46880/methoda.vol12no3.pp258-268.