

Data Mining untuk Memprediksi Jumlah Penjualan Hasil Pertanian Menggunakan Algoritma Forecasting (Studi Kasus : Dinas Pertanian Kabupaten Banggai)

Dewi Puspa Lamondjong¹, Mardi Hardjianto²

^{1,2}Magister Ilmu Komputer, Universitas Budi Luhur

e-mail: ¹lamondjongpuspa@gmail.com, ²mardi.hardjianto@budiluhur.ac.id

INTISARI

Teknologi data mining sangat berguna dalam membantu memprediksi jumlah penjualan dalam menemukan informasi yang sangat penting dari data mereka. Data mining mengeksplorasi basis data untuk menemukan pola-pola yang tersembunyi, meramalkan jumlah penjualan yang sangat berguna untuk mendukung pengambilan keputusan, pencarian informasi dalam memprediksi yang mungkin saja tidak dilihat atau dilupakan oleh pemerintahan daerah Kabupaten Banggai, karena di luar dari ekspektasi mereka. Analisis yang otomatisasi dilakukan oleh data mining menjawab pertanyaan-pertanyaan jumlah penjualan hasil panen tanaman pangan yaitu padi, jagung, kedelai dan ubi kayu yang terjual dengan baik dan jika dilakukan dengan cara tradisional memerlukan cukup banyak waktu dalam menjawabnya. Hasil dari model ini adalah menggali informasi dari data hasil panen dan jumlah penjualan tanaman pangan tersebut, sehingga dapat memprediksi jumlah penjuala hasil panen tanaman pangan tersebut yang terjual dan pemerintah daerah bisa memprediksi hasil panen tanaman pangan apa yang lebih serius untuk di tingkatkan kualitasnya. Pada penelitian ini untuk memprediksi jumlah penjualan pertanian menggunakan metode forecasting yaitu regresi linier, untuk kemudian dapat menjadi keluaran angka yang dapat di tentukan berapa banyak jumlah penjualan hasil panen tanaman pangan tersebut, dan akan menjadi bahan evaluasi bagi pemerintah daerah untuk setiap jumlah penjualan hasil panen tanaman pangan tersebut. Diharapkan model yang dibuat dalam memprediksi jumlah penjualan untuk setiap hasil panen tanaman pangan tersebut sesuai data yang ada.

Kata Kunci: jumlah penjualan, forecasting, prediksi, data mining.

ABSTRACT

Data mining technology is very useful in helping predict the number of sales in finding very important information from their data. Data mining explored databases to find hidden patterns, forecast the number of sales which is very useful to support decision making, search for information in predicting what the regional government of Banggai Regency may not see or forget, because it is beyond their expectations. Automated analysis carried out by data mining answers questions on the number of food crops sales, wnamely rice, maize, soybeans and cassava which are sold well and if done in the traditional way requires a lot of time to answer them. The result of this model is to find information from the data on yields and the number of these food plants sales, then it can predict the amount of these crops sales that are sold and the local government can predict the yields of more serious food crops to improve their quality. In this study, to predict the amount of agricultural sales using a calculation method, namely linear regression, then the number output can determine how much the these food crops sales are sold, and will be used as evaluation material for local governments for each total sales of crop yields. It is expected that the model made in predicting the number of sales for each crop yield is in accordance with existing data.

Keywords: number of sales, prediction, prediction, data mining

I. PENDAHULUAN

Kabupaten Banggai merupakan salah satu wilayah yang terletak pada kawasan dengan potensi iklim dan kondisi lahan yang beragam karakteristiknya. Adanya keragaman sifat fisik lahan ini dapat dijadikan modal dasar yang dapat di pakai dalam melakukan pertimbangan

dalam menentukan perwilayahan komoditas pertanian. Keragaman sifat fisik lahan akan menentukan jenis komoditas yang dapat diusahakan serta akan berpengaruh terhadap tingkat produktifitasnya. Subsektor pertanian sebagai bagian dari sektor pertanian dapat berperan aktif dalam rangka peningkatan

perekonomian wilayah. Untuk lebih meningkatkan kemakmuran yang lebih seimbang dan merata antar daerah diperlukan perencanaan pembangunan dengan kebijakan-kebijakan yang dapat merangsang perkembangan wilayah sesuai dengan potensinya masing-masing.

Sektor pertanian di Kabupaten Banggai merupakan sektor yang mampu menopang perekonomian masyarakat di kabupaten Banggai, Karena sebagian besar penduduk bermata pencaharian dengan bercocok tanam, hal ini sesuai dengan data dari keseluruhan Desa/Kelurahan di Kabupaten Banggai sekitar 93,98% merupakan potensi sector tanaman pangan, Perkebunan dan Perikanan (Banggai dalam Angka, 2018). Selain itu sektor pertanian di Kabupaten Banggai juga memberikan Kontribusi yang cukup besar terhadap perekonomian di Kabupaten Banggai. (Banggai, 2018)

Tabel 1.1 Grafik Tanaman Pangan Dinas Pertanian Kabupaten Banggai



Berdasarkan data dari laporan potensi pertanian di daerah Banggai hasil dari produksi pertanian dalam hal ini tanaman pangan yaitu padi, jagung, kedelai dan ubi kayu bermacam-macam dan sesuai musim panen masing-masing tanaman di masing-masing daerah, kualitas panen dan sedikit banyaknya stok hasil tanaman pangan. Harga yang selalu berubah-ubah setiap musim panennya mengakibatkan terjadi beberapa kendala dalam hal ini, kurang terorganisirnya jumlah penjualan hasil pangan dengan baik dan mengakibatkan harga bisa menjadi rendah atau menjadi sangat tinggi dikarenakan jumlah stok padi, jagung, kedelai dan ubi kayu setiap panennya. Di sektor pertanian, setiap aktivitas proses produksi selalu dihadapkan dengan situasi resiko (risk) dan ketidakpastian (uncertainty). Sumber tidak

tetapnya kepastian yang paling penting di sektor pertanian dalam hal ini tanaman pangan adalah tidak tetapnya hasil panen dan jumlah

penjualan hasil tanaman pangan. tidak tetapnya hasil panen akan mengakibatkan ikut turun naiknya pendapatan yang diterima oleh petani dari hasil produksi tanaman pangan mereka. Sehingga perlu memprediksi jumlah penjualan tanaman pangan yaitu padi, jagung, kedelai dan ubi kayu setiap kali panen.. (Willmen TB Panjaitan, 2018)

Salah satu upaya untuk mengantisipasi terjadinya fluktuasi jumlah penjualan tanaman pangan yaitu padi, jagung, kedelai dan ubi kayu adalah dengan melakukan prediksi jumlah penjualan setiap produk tanaman pangan tersebut. Dimana metode Algoritma Regresi linear adalah pilihan yang dapat digunakan untuk melakukan prakiraan/prediksi jumlah penjualan produk pangan tersebut yang terjual di masa depan dalam kurun waktu tertentu. Prediksi penjualan hasil tanaman pangan tersebut ditujukan untuk mendapatkan gambaran tentang keadaan yang akan terjadi di masa yang akan datang agar dapat digunakan untuk melakukan pengambilan keputusan terkait, guna meminimalisasi resiko dan memaksimalkan potensi keuntungan yang dapat diraih untuk menunjang perekonomian daerah dan pemerintah daerah dapat memprioritaskan hasil panen apa saja yang jumlah penjualannya meningkat..

Dalam penelitian ini juga di jabarkan penelitian terdahulu: Pradana, Andhika Mulya. Prediksi pemesanan menu pada report point of sales sambapos v4 menggunakan metode asosiasi.(studi kasus: pt. planetmas adidaya boga). Dengan permasalahan Penulis membuat sistem penelitian pada data transaksi penjualan tersebut dapat disimpulkan bahwa aplikasi pengelolaan data penjualan pada sebuah perusahaan atau toko sangat dibutuhkan untuk mengetahui produk yang sering dibeli oleh konsumen. Dengan menggunakan suatu algoritma yang dapat mengelola data penjualan yang menghasilkan informasi berupa produk terlaris. Algoritma apriori termasuk jenis aturan asosiasi pada data mining, Algoritma Apriori yang bertujuan untuk menemukan frequent itemsets dijalankan pada sekumpulan data. Analisis Apriori didefinisikan suatu proses untuk menemukan semua aturan apriori yang memenuhi syarat minimum untuk support dan syarat minimum untuk confidence. Hasil 1. Telah dihasilkan rancangan sistem data mining pada transaksi penjualan bibit buah menggunakan algoritma apriori di UD Buah Asri. 2. Telah dihasilkan program

aplikasi data mining pada transaksi penjualan bibit buah menggunakan algoritma apriori di UD Buah Asri yang berbasis web menggunakan Bahasa pemrograman PHP. Penelitian selanjutnya Nurhid ayati dan Alimud din (2019). Klasifikasi Penjualan Obat Pertanian Laris Dan Kurang Laris Pada UD Cahaya Tani Menggunakan Metode Decision Tree. Solusi Penulis melakukan penelitian Untuk mengatasi permasalahan tersebut digunakan metode data mining dengan mengimplementasikan algoritma decision tree yang merupakan salah satu teknik klasifikasi dengan membentuk sebuah pohon keputusan. Setelah dilakukan pengujian dengan 100 data pasien diperoleh hasil dengan nilai akurasi sebesar 90% menggunakan Confusion Matrix. Hasil Obat pertanian yang laris atau yang paling dibutuhkan pada Musim Tanam I adalah obat pertanian dengan jumlah terjual di atas 11 adalah sidafur dengan ukuran 1 kg, virtakoukuran 100 ml, dll. Kemudian dengan jenis Fungisida dan Pupuk Daun dengan harga ≤ 87500 dan jumlah terjual ≤ 8 adalah Folicur Gold, Dithane, Nativo, Gandasil B dan D, dll. Sedangkan obat pertanian yang kurang laris atau yang kurang dibutuhkan oleh petani adalah obat pertanian dengan jenis Herbisida, Insektisida dan Perekat dengan harga ≤ 87500 dan dengan jumlah terjual ≤ 8 adalah alika, curacron, goal, besmor, dll. Penelitian selanjutnya Mohamad Efendi Lasulika (2017). Solusi Penulis melakukan penelitian dengan mengumpulkan data- data. data yang diperoleh adalah data harian time series univariate, data tersebut berupa data harga komoditi jagung/kg dari tahun 2010-2015 dengan jumlah record 1435. Data tersebut untuk hari sabtu dan minggu tidak terjadi transaksi serta tanggal merah setiap tahunnya kosong dan terdiri dari variabel harga. Preprocessing data yang dilakukan pada penelitian ini menggunakan normalisasi data. Normalisasi data digunakan untuk mengelompokkan data ke dalam skala atau jangkauan tertentu sehingga mempermudah dalam mengolah data. Hasil Hasil dari semua percobaan yang telah dilakukan untuk menentukan model terbaik dari KNearest Neighbor, dari periode 1-8 periode dimana uji coba dilakukan dengan nilai K-Fold cross validation = 10, serta menentukan nilai dari Parameter K antara 1,3,5 dan 7. Berikut adalah rangkuman dari hasil uji coba model terbaik dari pada setiap periodenya.

II. METODOLOGI PENELITIAN

Metode yang digunakan dalam melakukan penelitian ini menggunakan dataset history jumlah penjualan dan stok setiap hasil panen tanaman pangan yaitu padi, jagung, kedelai dan ubi kayu pada waktu tertentu. dengan menggunakan dataset history jumlah penjualan dan stok setiap hasil panen tanaman pangan serta perkiraan cuaca di daerah kabupaten banggai pada tahun-tahun sebelumnya. Sumber data diambil dari Dinas Pertanian Kabupaten Banggai dan BMKG kabupaten banggai. Penelitian yang dilakukan yaitu memberikan prediksi jumlah penjualan tanaman pangan tersebut, dengan memprediksi jumlah dari setiap hasil panen tanaman pangan tersebut pada waktu tertentu,

A. Metode Pengumpulan Data

Pengumpulan data dilakukan untuk mendapatkan informasi sehingga data yang diperoleh berhubungan dengan penelitian ini. Metode pengumpulan data yang digunakan oleh peneliti, antara lain:

1. Wawancara

Wawancara merupakan teknik pengumpulan data yang dilakukan dengan mewawancarai bidang yang terkait dengan penjualan dan hasil panen tanaman pangan di dinas Pertanian Banggai serta perkiraan cuaca di BMKG kabupaten Banggai. Kemudian penulis melakukan observasi lapangan atau tinjauan langsung untuk mengamati dan mengetahui karakteristik dari hasil panen tanaman pangan yaitu padi, jagung, kedelai dan ubi kayu. Dari hasil wawancara ini diharapkan dapat menambah data yang diperoleh dari hasil pengamatan.

2. Studi Pustaka

Pengumpulan data yang diperoleh dengan mempelajari, meneliti dan membaca buku informasi dari internet, jurnal dan tesis.

B. Instrumen Penelitian

Berikut rincian instrumentasi yang digunakan atau dibutuhkan dalam penelitian antara lain:

1. Perangkat Keras:

- Processor Intel Core i5 6500U
- Memory 8 GB DDR4
- Graphics AMD Radeon 3 GB
- Storage Internal 250GB SSD

2. Perangkat Lunak:

- Android Studio 3.1
- IntelliJ IDEA 2016.3.8
- pgAdmin 4

d) Command Promt Windows

C. Langkah-Langkah Penelitian

Langkah-langkah metodologi penelitian dapat digambarkan sebagai berikut ini :



Gambar 1 Langkah-Langkah Penelitian

a. Identifikasi masalah

Langkah awal dari penelitian ini adalah dengan mengidentifikasi masalah yang terdapat pada jumlah penjualan hasil panen pertanian Kabupaten Banggai, dimana untuk memprediksi jumlah penjualan hasil panen dengan melakukan wawancara langsung ke dinas pertanian kabupaten Banggai untuk mendapatkan gambaran yang jelas terkait keadaan pertanian di kabupaten Banggai.

b. Studi pustaka

Peneliti melakukan pencarian jurnal-jurnal dan searching tentang masalah yang ada dan bagaimana cara menyelesaikan masalah yang ada, dan bisa diselesaikan menggunakan teknik datamining .

c. Pengumpulan data

Pada tahap ini peneliti melakukan wawancara untuk mendapatkan data-data tentang penjualan, hasil panen dan harga panen di Dinas Pertanian kabupaten Banggai 3 tahun sebelumnya dan data tersebut akan digunakan dalam metode yang digunakan.

d. Seleksi data

Pada tahap ini Peneliti mengambil data yang digunakan yaitu hanya data stok dan data penjualan tanaman pangan yaitu padi, jagung, kedelai, ubi

kayu, data data seperti harga dan lainnya tidak dimasukan karena tidak sesuai dengan penelitian terkait.

e. Tranformasi data

Pada tahap transformasi data ini peneliti menggunakan fungsi aktivasi model persamaan Regresi Linear sederhana untuk melakukan proses normalisasi dalam persamaan

$$a = \frac{(y)(x^2) - (x)(xy)}{n(x^2) - (x)^2}$$

$$b = \frac{n(xy) - (x)(y)}{n(x^2) - (x)^2}$$

ket.

Y= variable response atau variable akibat (Dependent)

X= variabelprediktor atau variable factor (Independent) a = konstanta

b = koefisien regresi (kemiringan); besaran Responden yang ditimbulkan

f. Proses Mining

Peneliti melakukan penambangan data atau proses mining dari data yang sudah di transformasi sehingga data dapat dilakukan penambangan yang sesuai dengan metode Regresi Linear.

g. Analisis

Pada tahap analisis ini peneliti mengamati data yang didapatkan serta pola jumlah penjualan produksi hasil tanaman pangan yaitu padi, jagung, kedelai dan ubi kayu. kumpulan data hasil survey terkait dengan jumlah penjualan produksi hasil tanaman pangan tersebut. Survey dilakukan di area pertanian di wilayah Kabupaten Banggai. Data yang diambil adalah hasil penjualan dari tahun 2017-2019

h. Pengujian metode

Pada tahapan ini penulis memprediksi jumlah penjualan hasil tanaman pangan yaitu padi, jagung, kedelai dan ubi kayu tahun ini dengan menguji algoritma Regresi Linear menggunakan Confusion Matrix untuk akurasi terbaik dari prediksi.

Parameter yang digunakan dalam melakukan pengujian adalah dengan menghitung nilai True Negative (TN), False Positive (FP), False Negative (FN), dan True Positive (TP) dapat diperoleh nilai akurasi, presisi dan recall. Nilai akurasi menggambarkan seberapa akurat sistem dapat mengklasifikasikan data secara benar.

nilai akurasi merupakan perbandingan antara data yang terklasifikasi benar dengan keseluruhan data. Setelah itu peneliti akan melihat apakah tingkat akurasi prediksi sesuai yang diinginkan atau diperbaiki variable pendukung penelitian ini.

i. Presentasi Pengetahuan

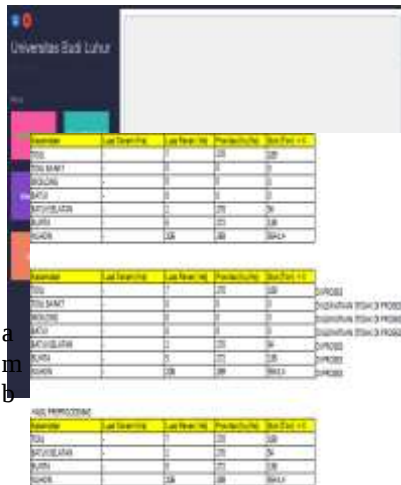
Dalam tahap ini peneliti melakukan visualisasi agar data dapat dibaca oleh dinas pertanian kabupaten banggai dan dapat ditampilkan secara menarik..

III. HASIL DAN PEMBAHASAN

A. Implementasi

Implementasi merupakan salah satu tahapan pembuatan aplikasi. Tujuan dari pengimplementasian ini adalah untuk mengukur sejauh mana program aplikasi ini dapat menyelesaikan masalah. Dengan adanya pengimplementasian serta uji coba tersebut, dapat mempermudah untuk melihat adanya kekurangan pada program yang telah dibuat. Hasil implementasi adalah sebagai berikut:

1) Tampilan Form Utama



ar 2. Tampilan Layar Utama

Pada gambar 2 adalah halaman untuk user melakukan import dan pemilihan berkas dataset dan data uji, serta melakukan proses prediksi jumlah penjualan dengan memilih metode Regresi Linear.

2) Tampilan Import Berkas

Pada gambar 3 adalah halaman untuk user memilih berkas dan lokasi berkas dataset atau data uji. Berkas yang di

pilih hanya diperbolehkan untuk tipe Microsoft Excel (.xls atau .xlsx).

Gambar 3: Tampilan Import dan Pemilihan Berkas Excel

3) Tampilan Hasil Prediksi

Pada gambar 4. adalah halaman hasil prediksi jumlah penjual setelah user melakukan prediksi dan menekan tombol Start Predictions. Pada halaman tersebut dapat dilihat kolom predictions di setiap data uji. Selain itu juga dapat dilihat validasi Confusion Matrix berupa akurasi, presisi dan recall.



Gambar 4. Tampilan Layar Hasil Prediksi

4) Preprocessing

Proses Preprocessing dilakukan untuk memilih data yang dapat diproses oleh aplikasi. Data yang dapat diproses adalah data yang memiliki nilai X dan Y tidak 0. Jika nilai X atau Y bernilai 0, maka data tersebut tidak akan digunakan. Untuk contoh proses preprocessing dapat dilihat pada gambar 5.

Gambar 5. Hasil Preprocessing Dataset

Berikut adalah dataset pangan jagung yang di pakai untuk pengujian ini, dengan jumlah 69 data.

Tabel 1. Tabel Dataset



Kecamatan	Luas Tanam (Ha)	Luas Panen (Ha)	Provitas (Ku/Ha)	Stok (Ton)	Jumlah Yang Berhasil Terjual (Ton)
TOILI	-	1616	42	6787.2	6777.2
TOILI BARAT	-	1	40	4	3.5
MOILONG	-	60	45	270	265
BATUI	-	1577	33.25	5243.53	5238.53
BATUI SELATAN	-	2211	46	10170.6	10160.6
BUNTA	-	3246	54.71	17758.87	17748.87
NUHON	-	3483	53.78	18731.57	18726.57
SIMPANG RAYA	-	3329	54.83	18252.91	18242.91
KINTOM	-	226	33.4	754.84	749.84
LUWUK	-	664	36.79	2442.86	2432.86
LUWUK TIMUR	-	1025	35	3587.5	3577.5
LUWUK UTARA	-	540	24	1296	1291
LUWUK SELATAN	-	685	46	3151	3141
NAMBO	-	96	42	403.2	398.2
PAGIMANA	-	1598	54.31	8678.74	8668.74
BUALEMO	-	5619	57.25	32168.78	32158.78
LOBU	-	385	43	1655.5	1650.5
LAMALA	-	957	38	3636.6	3626.6
MASAMA	-	567	44	2494.8	2484.8
MANTOH	-	1598	47.5	7590.5	7580.5
BALANTAK	-	1024	49.43	5061.63	5056.63
BALANTAK SELATAN	-	1895	53.47	10132.57	10122.57
BALANTAK UTARA	-	383	43	1646.9	1641.9
Toili	423,00	190	50.26	954.94	950.94
Toili Barat	13,00	13	39.66	51.558	49.1
Moilong	70,00	26	46.42	120.692	113.12
Batui	85,00	79	50.64	400.056	391.06
Batui Selatan	892,00	603	51.92	3130.776	3121.9
Bunta	1.169,00	949	48.58	4610.242	4602.5
Nuhon	1.561,00	1578	49.72	7845.816	7841.11
Simpang Raya	1.024,00	857	48.03	4116.171	4111.17
Kintom	78,00	27	47.61	128.547	122.22
Luwuk	88,00	41	47.61	195.201	190.11

Dari seluruh dataset pangan jagung, akan digunakan 2 attribute untuk kebutuhan prediksi Regresi Linear, yaitu attribute Stok sebagai X dan Jumlah Penjualan sebagai Y.

B. Data Uji

Berikut adalah data uji yang di pakai untuk pengujian ini, dengan jumlah 23 data uji.

Tabel 2. Tabel Data Uji

Kecamatan	Luas Tanam (Ha)	Luas Panen (Ha)	Provitas (Ku/Ha)	Stok (Ton)	Jumlah Yang Berhasil Terjual (Ton)
TOILI	-	1616	42	6787.2	6777.2
TOILI BARAT	-	1	40	4	3.5
MOILONG	-	60	45	270	265
BATUI	-	1577	33.25	5243.525	5238.53
BATUI SELATAN	-	2211	46	10170.6	10160.6
BUNTA	-	3246	54.71	17758.866	17748.87
NUHON	-	3483	53.78	18731.574	18726.57
SIMPANG RAYA	-	3329	54.83	18252.907	18242.91
KINTOM	-	226	33.4	754.84	749.84
LUWUK	-	664	36.79	2442.856	2432.86
LUWUK TIMUR	-	1025	35	3587.5	3577.5
LUWUK UTARA	-	540	24	1296	1291
LUWUK SELATAN	-	685	46	3151	3141
NAMBO	-	96	42	403.2	398.2
PAGIMANA	-	1598	54.31	8678.738	8668.74
BUALEMO	-	5619	57.25	32168.775	32158.78
LOBU	-	385	43	1655.5	1650.5
LAMALA	-	957	38	3636.6	3626.6
MASAMA	-	567	44	2494.8	2484.8
MANTOH	-	1598	47.5	7590.5	7580.5
BALANTAK	-	1024	49.43	5061.632	5056.63
BALANTAK SELATAN	-	1895	53.47	10132.565	10122.57
BALANTAK UTARA	-	383	43	1646.9	1641.9

Dari seluruh data uji pangan jagung, akan digunakan 1 attribute untuk kebutuhan prediksi Regresi Linear,

yaitu attribute Jumlah Penjualan sebagai Y. Dan untuk melakukan validasi menggunakan Confusion Matrix, digunakan attribute Stok dari data uji.

C. Regresi Linear

1) Pelatihan

Pada proses prediksi menggunakan metode Regresi Linear, Langkah pertama adalah mencari nilai X, Y, X², Y², dan XY pada dataset dengan rumus dibawah ini:

$$X = \text{Attribute 1}$$

$$Y = \text{Attribute 2}$$

$$X^2 = X * X$$

$$Y^2 = Y * Y$$

$$XY = X * Y$$

Maka didapat hasil dari Langkah pertama dapat dilihat pada table 4.3.

Tabel 3. Dataset Hasil Langkah Pertama

Kecamatan	Stok (X)	Jumlah Penjualan (Y)	X ²	Y ²	XY
TOILI	6787.2	6777.2	46066083,84	45930439,84	45998211,84
TOILI BARAT	4	3.5	16	12,25	14
MOILONG	270	265	72900	70225	71550
BATUI	5243.53	5238.53	27494606,86	27442196,56	27468389,21
BATUI SELATAN	10170.6	10160.6	103441104,4	103237792,4	103339398,4
BUNTA	17758.87	17748.87	315377463,7	315022386,3	315199875
NUHON	18731.57	18726.57	350871714,7	350684424	350778056,8
SIMPANG RAYA	18252.91	18242.91	333168723,5	332803765,3	332986194,4
KINTOM	754.84	749.84	569783,4256	562260,0256	566009,2256
LUWUK	2442.86	2432.86	5967564,98	5918807,78	5943136,38
LUWUK TIMUR	3587.5	3577.5	12870156,25	12798506,25	12834281,25
LUWUK UTARA	1296	1291	1679616	1666681	1673136
LUWUK SELATAN	3151	3141	9928801	9865881	9897291
NAMBO	403.2	398.2	162570,24	158563,24	160554,24
PAGIMANA	8678.74	8668.74	75320527,99	75147053,19	75233740,59
BUALEMO	32168.78	32158.78	1034830407	1034187131	1034508719

Dari table 4.3 telah didapat nilai X, Y, X², Y², dan XY masing-masing dataset. Selanjutnya adalah melakukan summary pada nilai X, Y, X², Y², dan XY. Rumus summary dapat dilihat dibawah ini:

$$\begin{aligned} \text{Summary}(X) &= \sum_{i=0}^1 X \\ \text{Summary}(Y) &= \sum_{i=0}^1 Y \\ \text{Summary}(X^2) &= \sum_{i=0}^1 X^2 \\ \text{Summary}(Y^2) &= \sum_{i=0}^1 Y^2 \\ \text{Summary}(XY) &= \sum_{i=0}^1 XY \end{aligned}$$

Dari rumus diatas maka didapat hasil perhitungan summary yang dapat dilihat pada table 4.

Tabel 4. Hasil Summary

Stok (X)	Jumlah Penjualan (Y)	X ²	Y ²	XY
286047,6	285557,4	3601706571	3596714600	3599208605

Sekarang nilai X, Y, X², Y², dan XY telah didapat. Selanjutnya adalah mencari nilai A dan B yang nantinya digunakan untuk proses prediksi. Rumus yang digunakan dapat dilihat dibawah ini:

$$\begin{aligned} A &= \frac{(\text{summary}Y * \text{summary}X^2) - (\text{summary}X * \text{summary}XY)}{(\text{Jumlah dataset} * \text{summary}X^2) - (\text{summary}X * \text{summary}X)} \\ B &= \frac{(\text{Jumlah dataset} * \text{summary}XY) - (\text{summary}X * \text{summary}Y)}{(\text{Jumlah dataset} * \text{summary}X^2) - (\text{summary}X * \text{summary}X)} \end{aligned}$$

Dari rumus diatas, maka didapat nilai A dan B sebagai berikut:

$$\begin{aligned} A &= -6.3 \\ B &= 1.0 \end{aligned}$$

D. Pengujian

Pada tahap pengujian akan dilakukan prediksi pada data uji menggunakan nilai A dan B yang sebelumnya telah didapat pada proses pelatihan dataset. Untuk contoh penelitian ini akan memprediksi data uji yang dapat dilihat pada table 5.

Tabel 5. Tabel Data Uji

Kecamatan	Luas Tanam (Ha)	Luas Panen (Ha)	Provitas (Ku/Ha)	Stok (Ton)	Jumlah Yang Berhasil Terjual (Ton)
TOILI	-	1616	42	6787.2	6777.2

Dari table 4.5 dapat dilihat nilai yang akan adalah dengan attribute stok bernilai 6787,2. Perhitungan dapat dilihat sebagai berikut:

$$Y = A + (B * \text{Attribute Stok})$$

$$Y = -6,3 + (1,0 * 6787,2) = 6780,9$$

Bisa dilihat bahwa hasil prediksi yang didapat adalah 6780,9 dari perhitungan

Y. Untuk data uji yang lain dapat dilakukan berulang sesuai nilai attribute stok nya. Hasil prediksi seluruh data uji dapat dilihat pada tabel 6.

Tabel 6. Hasil Prediksi Seluruh Data Uji

Kecamatan	Stok	Jumlah Penjual asli	Hasil Prediksi Jumlah Jual	Akurasi
TOILI	6787.2	6777.2	6780.9	99.95
TOILI BARAT	4	3.5	-2.3	-65.71
MOILONG	270	265	263.7	99.51
BATUI	5243.525	5238.53	5237.22	99.97
BATUI SELATAN	10170.6	10160.6	10164.3	99.96
BUNTA	17758.866	17748.87	17752.57	99.98
NUHON	18731.574	18726.57	18725.27	99.99
SIMPANG RAYA	18252.907	18242.91	18246.61	99.98
KINTOM	754.84	749.84	748.54	99.83
LUWUK	2442.856	2432.86	2436.56	99.85
LUWUK TIMUR	3587.5	3577.5	3581.2	99.9
LUWUK UTARA	1296	1291	1289.7	99.9
LUWUK SELATAN	3151	3141	3144.7	99.88
NAMBO	403.2	398.2	396.9	99.67
PAGIMANA	8678.738	8668.74	8672.44	99.96
BUALEMO	32168.775	32158.78	32162.48	99.99
LOBU	1655.5	1650.5	1649.2	99.92
LAMALA	3636.6	3626.6	3630.3	99.9
MASAMA	2494.8	2484.8	2488.5	99.85
MANTOH	7590.5	7580.5	7584.2	99.95
BALANTAK	5061.632	5056.63	5055.33	99.97
BALANTAK SELATAN	10132.565	10122.57	10126.27	99.96
BALANTAK UTARA	1646.9	1641.9	1640.6	99.92

E. Validasi

Dari hasil prediksi pada tabel 6. maka dilakukan proses validasi, yaitu mencari akurasi, presisi, dan recall pada hasil

prediksi menggunakan metode Confusion Matrix.

Beberapa rules yang digunakan penulis untuk mencari true positive, true negative, false positive, dan false negative. Rules dapat dilihat sebagai berikut:

True Positive : Akurasi ≥ 70 dan Hasil Prediksi $\leq$ Stok

True Negative : Akurasi ≥ 70 dan Hasil Prediksi $\geq$ Stok

False Positive : Akurasi < 70 dan Hasil Prediksi $\leq$ Stok

False Negative : Akurasi < 70 dan Hasil Prediksi $\geq$ Stok

Dari rules diatas maka didapat 22 true positive, 0 true negative, 1 false positive, dan 0 false negative.

Selanjutnya adalah menghitung akurasi, presisi, dan recall dari data true dan false dengan rumus sebagai berikut:

$$\begin{aligned} \text{Akurasi} &= \frac{\text{True Pos} + \text{True Neg}}{\text{True Pos} + \text{True Neg} + \text{False Pos} + \text{False Ne}} \\ \text{Presisi} &= \frac{\text{True Pos}}{\text{True Pos} + \text{False Pos}} \\ \text{Recall} &= \frac{\text{True Pos}}{\text{True Pos} + \text{False Neg}} \end{aligned}$$

Dengan menggunakan rumus diatas maka didapat hasil validasi menggunakan Confusion Matrix yang dapat dilihat pada tabel 7.

Tabel 7. Hasil Validasi Confusion Matrix

Confusion Matrix	Akurasi	Presisi	Recall
	96%	96%	100%

IV. KESIMPULAN

Permasalahan yang ada yaitu sulitnya memprediksi jumlah penjualan hasil panen tanaman pangan yaitu padi, jagung, kedelai dan ubi kayu yang terjual setiap kali panen ditahun mendatang. berapa banyak jumlah penjualan hasil panen tanaman pangan tersebut yang terjual dan juga sebagai bahan pertimbangan pemerintah daerah memprioritaskan hasil tanaman pangan apa yang penjualannya baik. maka penulis mencoba untuk memprediksi jumlah penjualan hasil tanaman pangan setiap kali panen dengan menggunakan data mining dengan metode algoritma regresi linear dan

hasil dari prediksi data mining dengan menggunakan confusion matrix Akurasi 96 % , Presisi 90%, Recall 100%. Diharapkan dengan model data mining yang ditawarkan tersebut Dinas Pertanian Kabupaten Banggai bisa mengetahui berapa banyak penjualan hasil tanaman pangan tersebut di daerah.

UCAPAN TERIMA KASIH

Terima kasih untuk semua pihak yang telah membantu untuk pembuatan jurnal ini.

REFERENSI

- [1] www.arduino.cc
- [2] aziz., M. R. (2017). Implementasi data mining pada transaksi penjualan bibit buah dengan Algoritma apriori (studi kasus UD Buah Asri). Jurnal Program studi Teknik informatika fakultas Teknik informatika fakultas Teknik universitas nusantara PGRI kediri .
- [3] Banggai, D. P. (2018). Laporan Produksi Pertanian Kabupaten Banggai.
- [4] E.Prasetyo. (2012). Data Mining: Konsep dan Aplikasi menggunakan Matlab, 1 ed.Yogyakarta: Andi Offset, 2012. yogyakarta.
- [5] Effrida Manalu, F. A. (2017). penerapan algoritma naive bayes untuk memprediksi jumlah produksi barang berdasarkan data persediaan dan jumlah pemesanan pada cv. papadan mama pastries.
- [6] Ghebyla Najla Ayuni, D. F. (t.thn.). Penerapan Metode Regresi Linear Untuk Prediksi Penjualan Properti pada PT XYZ. Jurnal Telematika, vol. 14 no. 2, Institut Teknologi Harapan Bangsa, Bandung p-ISSN: 1858-2516 e-ISSN: 2579-377.
- [7] Haerul Fatah, A. S. (2018). prediksi harga cryptocurrency dengan metode k-nearest neighbours. Jurnal PILAR Nusa Mandiri Vol. 14, No. 2 September 2018.
- [8] Han, J. a. (2006). Data Mining Concepts and Techniques (2nd Edition).
- [9] Hayuningtyas, R. Y. (2019). Penerapan Algoritma Naïve Bayes Untuk Rekomendasi Pakaian Wanita. Jurnal Informatika 6 (1): 18–22.
- [10] John, R. a. (2017). Rancang Bangun Aplikasi Prediksi Penjualan Menggunakan Algoritma Double Exponential Smoothing Berbasis Web (Studi Kasus: Pt. Sanpak Unggul). Jurnal Informatika 14 (1): 28–35.

- <https://doi.org/10.9744/informatika.14.1.28-35>.
- [11] Maesaroh siti, K. (2017). Sistem Prediksi Produktifitas Pertanian Padi Menggunakan Data Mining.
- [12] Nurhidayati, A. (2019). Klasifikasi Penjualan Obat Pertanian Laris Dan Kurang Laris Pada UD Cahaya Tani Menggunakan Metode Decission Tree. *Jurnal Informatika dan Teknologi* Vol. 2 No. 2, Juli 2019, hal 101 - 108 .
- [13] Raharja, M. A. (2018). Prediksi Ketinggian Gelombang Laut Menggunakan Metode Backpropagation Pada Pantai Lebih Gianyar. *Jurnal Ilmiah Ilmu Komputer* XI (1): 19–26.
- [14] Rezza Hary Dwi Satriya, E. S. (2018). Implementasi Metode Ensemble K-Nearest Neighbor untuk Prediksi Nilai Tukar Rupiah Terhadap Dollar Amerika. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer* e-ISSN: 2548-964X Vol. 2, No. 4, April 2018, hlm. 1718-1725 <http://j-ptiik.ub.ac.id>.
- [15] S. Visa, B. R. (2011). "Confusion matrix-based feature selection," *Midwest Artificial Intelligence and Cognitive Science Conference*, vol. 710, April 2011. *Midwest Artificial Intelligence and Cognitive Science Conference*, vol. 710, April 2011. http://www2.cs.uregina.ca/~dbd/cs831/notes/confusion_matrix/confusion_matrix.html.
- [16] Syarfi Aziz, A. S. (2017). Penerapan Metode ARIMA untuk Peramalan Pengunjung Perpustakaan UIN Suska Riau. *Seminar Nasional Teknologi Informasi, Komunikasi dan Industri (SNTIKI) 9* ISSN (Printed) : 2579-7271. Riau.
- [17] Taghsya Izmi Andini, W. W. (2016). Prediksi Potensi Pemasaran Produk Baru dengan Metode Naïve Bayes Classifier dan Regresi Linear. *journal Uii*, <https://journal.uui.ac.id/Snati/article/view/6231>.
- [18] Taghsya Izmi Andini, W. W. (2016). Prediksi Potensi Pemasaran Produk Baru dengan Metode Naïve Bayes Classifier dan Regresi Linear. *Journal UII*, <https://journal.uui.ac.id/Snati/article/view/6231>.
- [19] Turban, E. J.-P. (2007). *Decision Support Systems and Intelligent Systems*. *Decision Support Systems and Intelligent Systems*. Vol. 7 <http://www.amazon.co.uk/dp/0131230131>.
- [20] Willmen TB Panjaitan, E. U. (2018). prediksi panen padi menggunakan algoritma k-nearest neighbour. *Prosiding SNATIF Ke -5 Tahun 2018* ISBN: 978-602-1180-86-0.
- [21] Yustanti, W. (2012). Algoritma K-Nearest Neighbour untuk Memprediksi Harga Jual Tanah. *Jurnal matematika, statistic dan komputasi* <http://journal.unhas.ac.id/index.php/jmsk> ..