

ESTIMASI LUAS PANEN PADI DI KABUPATEN REMBANG MENGUNAKAN ALGORITMA LINEAR REGRESSION

Anis Fitriya¹, Suprayogi²

^{1,2}Teknik Informatika – S1, Fakultas Ilmu Komputer, Universitas Dian Nuswantoro

Jln. Nakula 1 No 5-11 Semarang 50131 INDONESIA

111201105902@mhs.dinus.ac.id

Abstrak

Potensi pertanian yang dimiliki kabupaten Rembang juga tak kalah dengan kabupaten lain, salah satunya adalah tanaman padi. Secara umum, padi dapat dipanen pada umur antara 110-115 hari setelah tanam. Akibat musim penghujan yang mundur dan berlangsungnya musim kemarau berkepanjangan yang melanda Kabupaten Rembang, menyebabkan berkurangnya hasil panen padi dari tahun sebelumnya. Untuk memenuhi kebutuhan tanaman pangan padi, dinas Pertanian dan Kehutanan Kabupaten Rembang berupaya untuk mengoptimalkan hasil pertanian dengan mengestimasi luas panen padi di Kabupaten Rembang sebagai langkah antisipasi kemungkinan buruk pada hasil panen padi sehingga nantinya dapat membantu para petani dalam mengatasi masalah tersebut serta meningkatkan produksi hasil pertaniannya. Untuk mengetahui hasil luas panen padi, penulis menggunakan metode estimasi dengan algoritma Linear Regresi. Dari hasil perhitungan yang telah dilakukan mendapatkan hasil persamaan regresi $Y = -23,839 + 0,73 x_1 + 5,135 x_2$. Nilai RMSE (Root Mean Squared Error) atau nilai performansi sebesar 185,52. Nilai yang dihasilkan masih terbilang cukup besar. Karena data testing yang digunakan berjumlah 36 data. Semakin banyak data yang digunakan maka akan semakin kecil RMSE yang dihasilkan.

Kata Kunci : Linear Regresi, Estimasi, Luas Panen Padi.

Abstract

Agricultural potentials in Rembang are not lost than other districts, one of them is a rice plant. In general, the rice can be harvested between 110-115 days after planting. Due to the retreat of the rainy season and the course of the dry season that hit Rembang regency, cause a reduction in the harvest from the previous year. To meet the needs of food crops of rice, Department of Agriculture and Forestry Rembang seeks to optimize crops by estimating the rice harvested area in Rembang to anticipate the bad possibility in the harvest so that later it can help farmers to overcome these problems and to improve the production of their crops. To know the results of rice harvested area, the authors use the method of estimation with Linear Regression algorithm. From the results of the calculations that have been done to get the regression equation $Y = -23.839 + 0.73 x_1 + 5.135 x_2$. The value of RMSE (Root Mean Squared Error) or the value of performance of 185.52. This resulting value is still quite large. Because the data used were 36 testing data. The more data is used it will be smaller RMSE produced.

Keywords: Linear Regression, Estimating, Rice Harvest Area.

1. PENDAHULUAN

Potensi pertanian yang dimiliki kabupaten Rembang juga tak kalah dengan kabupaten lain, salah satunya adalah tanaman padi. Secara umum, padi dapat dipanen pada umur antara 110-115 hari setelah tanam. Akibat musim penghujan yang mundur dan berlangsungnya musin kemarau berkepanjangan yang melanda Kabupaten Rembang, menyebabkan berkurangnya hasil panen padi dari tahun sebelumnya. Untuk memenuhi kebutuhan tanaman pangan padi, dinas Pertanian dan Kehutanan Kabupaten Rembang berupaya untuk mengoptimalkan hasil pertanian dengan mengestimasi luas panen padi di Kabupaten Rembang sebagai langkah antisipasi kemungkinan buruk pada hasil panen padi sehingga nantinya dapat membantu para petani dalam mengatasi masalah tersebut serta meningkatkan produksi hasil pertaniannya. Dari penelitian yang dilakukan oleh Bin Hariyati dan Sutikno [5] dengan judul Pemodelan Luas Panen Padi Di Kabupaten Lamongan Dengan Indikator El Nino Southern Oscillation Melalui Pendekatan Robust Bootstrap Least Trimmed Square dengan menggunakan data dari periode tahun 1990-2009, menyimpulkan pemodelan luas panen padi dengan menggunakan indikator *El Nino Southern Oscillation* yaitu Sea Surface Temperature mampu menangani permasalahan outlier dan data kecil. Pemodelan luas panen padi melibatkan indikator SST (Sea Surface Temperature) atau Tekanan Permukaan Laut serta distribusi curah hujan. Peramalan luas panen padi menunjukkan ketepatan prediksi untuk periode 1 adalah sebesar 95,32%,

periode 2 sebesar 66,71%, dan periode 3 sebesar 36,82%.

Berdasarkan penelitian sebelumnya yang sama-sama melakukan pemodelan estimasi dengan data yang diperoleh dari tahun-tahun terdahulu yaitu periode 1990-2009 menggunakan *El Nino Southern Oscillation* atau tekanan permukaan laut dengan indikator iklim dan menunjukkan pendekatan metode regresi *robust* dengan indikator curah hujan. Maka muncul gagasan untuk membuat model baru luas panen padi yang melibatkan indikator / variabel area tanam dan curah hujan. Hal ini menjadi latar belakang penggunaan metode estimasi menggunakan algoritma Linear Regression dengan data yang diperoleh dari tahun sebelumnya yaitu periode 2010-2014. Dengan metode tersebut diharapkan dapat memodelkan luas panen yang akan dihasilkan nantinya. Metode algoritma regresi merupakan sebuah metode penelitian yang berfungsi sebagai pengolah data untuk memprediksi keadaan data pada masa yang akan datang berdasarkan adanya perubahan pola hubungan variabel data pada masa lalu dengan asumsi bahwa pola hubungan variabel data yang diolah dapat berkelanjutan. Oleh karena itu, metode algoritma regresi ini cocok untuk diimplementasikan ke dalam simulasi forecasting. Pengimplementasian algoritma regresi dapat dibentuk dengan melihat dan menganalisis terlebih dahulu adanya pola hubungan data pada masa lalu yang selalu berubah kemudian dilanjutkan dengan melakukan sebuah keputusan untuk menentukan jenis algoritma regresi yang akan digunakan untuk dijadikan sebuah metode pengolah data [10]. Dengan metode tersebut

diharapkan dapat memodelkan luas panen yang akan dihasilkan nantinya.

2. METODE

2.1 Data

Data diperoleh secara langsung dari bagian pusat data dinas Pertanian dan Kehutanan Kabupaten Rembang, yaitu data hasil pertanian padi tahun 2010-2014. Di kabupaten Rembang terdapat 14 kecamatan, yaitu : Rembang, Kaliori, Sulang, Sumber, Bulu, Lasem, Pancur, Sluke, Kragan, Sarang, Sedan, Pamotan, Gunem, Sale. Jumlah data yang digunakan yakni 176 data.

2.2 Instrumen Penelitian

- a. Kebutuhan Software
 1. Sistem Operasi Windows 7
 2. NetBeans IDE 7.3
 3. Ms. Word
- b. Kebutuhan Hardware
 1. Laptop dengan spesifikasi
 1. Processor : Pentium Dual Core
 2. Sistem Operasi : Windows 7
 3. RAM : 2GB

2.3 Teknik Pengumpulan Data

- a. Wawancara

Hal yang ditanyakan mengenai : komoditi tanaman yang baik di kabupaten rembang, pada bagian ini dijelaskan tanaman padi mempunyai komoditi baik yang tiap kecamatan mempunyai tanaman tersebut. Bagaimana hasil panen padi di kabupaten Rembang, Pada bagian ini dijelaskan mengenai data hasil pertanian padi di setiap kecamatan di Kabupaten Rembang, periode panen

padi, serta ditunjukkan data - data hasil luas panen padi dari tahun 2010-2014.

- b. Studi Dokumentasi

Dengan mempelajari dokumentasi data – data yang berkaitan dengan hasil pertanian padi di bagian pusat data dinas Pertanian dan Kehutanan Kabupaten Rembang. Peneliti memperoleh informasi tentang tanaman padi dan hasil panen padi di setiap daerahnya dan model pengolahan data yang ada di dinas tersebut.

2.4 Variabel Penelitian

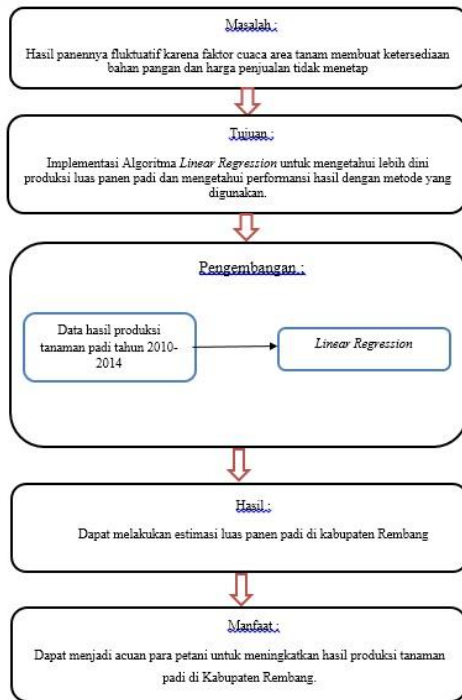
Variabel yang digunakan dalam penelitian ini ada 3, yaitu : tambah tanam, curah hujan, dan hasil panen. Penentuan variabel ini dibedakan menjadi 2 jenis, yaitu *variable independen* dan *variable dependen*. Variabel *independen* atau variabel bebas merupakan variabel yang mempengaruhi variabel lain atau menjadi sebab terjadinya perubahan variabel lain. Dalam hal ini yang termasuk dalam *variable independen* adalah variabel tambah tanam dan curah hujan. Sedangkan *variable dependen* atau variabel terikat merupakan variabel yang dipengaruhi variabel lain yaitu variabel hasil panen.

2.5 Teknik Analisis Data

- a. Menyeleksi Data mentah yang diperoleh menjadi data primer yang akan digunakan dalam penelitian.
- a. Melakukan perhitungan menggunakan perhitungan algoritma Linear Regression dan di dalam metode Estimasi.

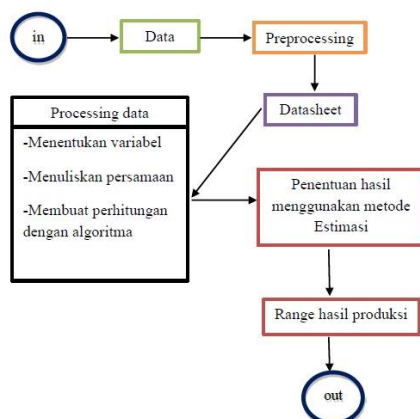
2.6 Kerangka Pemikiran

Kerangka pemikiran yang terbentuk digambarkan dalam bentuk diagram sebagai berikut :



Gambar 1. Kerangka Pemikiran

2.7 Prosedur Penyelesaian



Gambar 2. Prosedur Penyelesaian

Dari gambar diatas dapat dijelaskan langkah demi langkah sebagai berikut :

1. Seperti yang sudah dijelaskan diatas, data yang didapat kemudian dikelompokkan menjadi data-data yang diperlukan.
2. Dari data yang ada dilakukan preprocessing, pada tahap ini akan dilakukan proses data cleaning untuk menghasilkan dataset yang bersih yaitu dengan menghilangkan data yang tidak relevan, *missing value* (data yang tidak berisi nilai atau kosong) harus dibersihkan.
3. Setelah dilakukan pemrosesan data kemudian menentukan hasil produksi padi yang ingin dicapai dengan menggunakan metode estimasi.
4. Berdasarkan perhitungan dengan range tahun yang telah ditentukan didapat *output* yang menghasilkan nilai yang didapat dalam penentuan estimasi luas panen padi.

3. HASIL DAN PEMBAHASAN

3.1 Menuliskan Persamaan Regresi

Data hasil luas panen padi di Kabupaten Rembang yang digunakan berjumlah 176 data. Dari data keseluruhan dibagi menjadi dua kelompok yaitu *data training* dan *data testing*. *Data training* yang digunakan berjumlah 140 data sedangkan *data testing* yang digunakan berjumlah 36 data. Variable yang digunakan yaitu : area/tambah tanam yang diinisialisasikan sebagai x_1 dan curah hujan yang diinisialisasikan sebagai x_2 . Keduanya disebut dengan variable *independen*. Kemudian variable luas panen diinisialisasikan sebagai y ,

dimana y merupakan variable *dependen*. Setelah diketahui beberapa variable yang digunakan maka langkah selanjutnya adalah menuliskan persamaan linear regresi yang nantinya akan digunakan pada proses perhitungan. X_1 dan x_2 merupakan bobot-bobot yang diperlukan, sedangkan y merupakan *class* atau variable *dependen*.

nilai-nilai yang diperlukan dimasukkan ke dalam persamaan untuk mendapatkan nilai a_0 , a_1 , dan a_2 , dimana :

$$\begin{aligned} \sum x_1 &= 91746 \\ \sum x_2 &= 2234.913 \\ \sum y &= 75134 \\ \sum x_1y &= 90574311 \\ \sum x_2y &= 1248738.34 \\ \sum x_1x_2 &= 1512609.297 \\ \sum x_1^2 &= 116395942 \\ \sum x_2^2 &= 38455.83488 \end{aligned}$$

Rumus awal :

$$\begin{aligned} n\alpha_0 + \alpha_1 \sum_{i=1}^n x_{i1} + \alpha_2 \sum_{i=1}^n x_{i2} + \dots \\ + \alpha_k \sum_{i=1}^n x_{ik} = \sum_{i=1}^n y_i \end{aligned} \quad (1)$$

$$\begin{aligned} \alpha_0 \sum_{i=1}^n x_{i1} + \alpha_1 \sum_{i=1}^n x_{i1}^2 + \alpha_2 \sum_{i=1}^n x_{i1} x_{i2} \\ + \dots + \alpha_k \sum_{i=1}^n x_{i1} x_{ik} = \sum_{i=1}^n x_{i1} y_i \end{aligned} \quad (2)$$

:

$$\begin{aligned} \alpha_0 \sum_{i=1}^n x_{ik} + \alpha_1 \sum_{i=1}^n x_{ik} x_{i1} + \\ \alpha_2 \sum_{i=1}^n x_{ik} x_{i2} + \dots + \alpha_k \sum_{i=1}^n x_{ik}^2 = \\ \sum_{i=1}^n x_{ik} y_i \end{aligned} \quad (3)$$

Dari hasil perhitungan menggunakan metode substitusi rumus diatas, maka menghasilkan :

$$\begin{aligned} a &= -23.83945757 \\ a_1 &= 0.730206489 \\ a_2 &= 5.135770152 \end{aligned}$$

maka didapat Persamaan Regresi :

$$Y = -23.839 + 0.73 x_1 + 5.13 x_2 \quad (4)$$

3.2 Pengujian Persamaan

Data testing / pengujian yang akan digunakan berjumlah 36 data . Dari data testing tersebut kemudian dilakukan pengujian terhadap persamaan yang telah terbentuk dengan memasukkan masing-masing nilai X_1 atau variabel area/tambah tanam dan memasukkan nilai X_2 atau variable curah hujan ke dalam persamaan regresi yang telah dihasilkan.

3.3 Perhitungan RMSE

$$RMSE = \frac{\sum_{i=1}^n \sqrt{(x_i - f_i)^2}}{n}$$

Keterangan :

x_i = data akhir (data hasil prediksi)

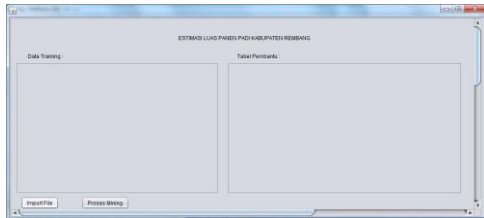
f_i = data awal (data sebenarnya)

n = jumlah data

Hasil perhitungan *Root Mean Squared Error* mendapatkan hasil sebesar 185.5258. Sesuai dengan teori RMSE yaitu semakin kecil RMSE yang dihasilkan (mendekati nilai 0) akan menghasilkan prediksi *output* yang lebih baik. Maka nilai RMSE yang dihasilkan termasuk dalam kategori kurang baik. Karena data uji / data testing yang digunakan berjumlah 36 data sehingga RMSE yang dihasilkan masih terbilang cukup besar. Karena semakin banyak

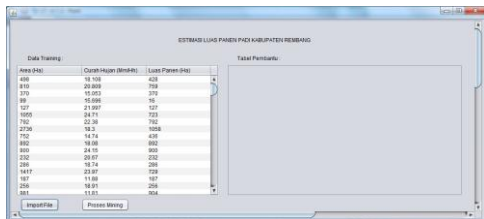
data yang digunakan maka akan semakin kecil RMSE yang dihasilkan.

3.4 Implementasi Hasil Penelitian



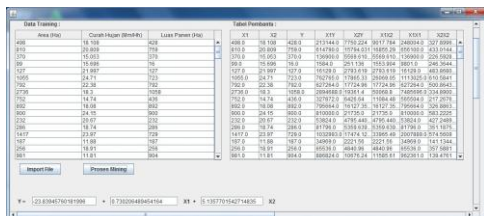
Gambar 3. Halaman Aplikasi

Di dalam halaman aplikasi terdapat 2 tabel, yaitu table training dan table pembantu dan 2 tombol yang terletak di bagian bawah table, yaitu tombol import file dan tombol proses mining.



Gambar 4. Tampilan Data Training

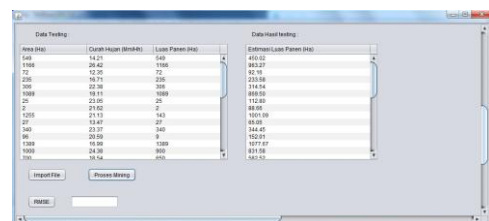
Dimana table training merupakan data hasil luas panen padi yang diimport dari file excel yang berekstensi .xls.



Gambar 5. Tampilan Hasil Regresi

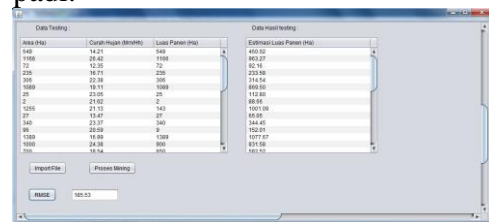
Tabel pembantu merupakan table perhitungan yang digunakan untuk membantu proses perhitungan persamaan regresi. Tombol proses mining berfungsi untuk menghitung

persamaan yang sudah diinputkan. Serta terdapat fungsi output yang menampilkan hasil dari proses perhitungan dari table training, yaitu menghasilkan persamaan regresi.



Gambar 6. Tampilan Data Testing

Setelah data berhasil diimport kemudian dilakukan proses mining dengan memasukkan data testing ke dalam persamaan regresi sehingga menghasilkan estimasi luas panen padi.



Gambar 7. Tampilan Hasil RMSE

Setelah dilakukan proses pembentukan persamaan regresi dan menguji persamaan dengan data testing yang berjumlah 36 data, maka dihasilkan proses perhitungan error sebesar 185.5258.

4. KESIMPULAN DAN SARAN

4.1 Kesimpulan

Berdasarkan dari penelitian yang dilakukan dan sudah diuraikan, maka dapat disimpulkan sebagai berikut :

1. Implementasi algoritma Linear Regresi ke dalam sebuah aplikasi sederhana menghasilkan sebuah informasi mengenai prediksi hasil

luas panen padi, sehingga dapat mengetahui lebih dini luas panen padi di kabupaten Rembang.

2. RMSE (*Root Mean Squared Error*) yang dihasilkan adalah 185.5258 . Performansi yang dihasilkan masih terbilang besar karena data uji yang digunakan berjumlah 36 yang berasal dari 20% dari data keseluruhan. Semakin besar jumlah data yang dipakai maka standar eror yang dihasilkan akan semakin kecil.

4.2 Saran

Ada beberapa saran yang ingin disampaikan untuk pengembangan penelitian lebih lanjut, yaitu :

1. Pada penelitian selanjutnya dapat mencoba menggunakan dataset yang berbeda dan dengan jumlah data yang lebih besar lagi sehingga nilai data selanjutnya yang dihasilkan dapat menghasilkan tingkat eror yang lebih kecil.
2. Selain penerapan secara teoritis dan aplikatif pada penelitian berikutnya dapat dicoba untuk membuat suatu aplikasi dengan teknik dan algoritma data mining yang berbeda sehingga dapat menghasilkan informasi yang variatif.
3. Penelitian ini disarankan dapat menjadi bahan referensi yang dipergunakan dan dikembangkan untuk penelitian selanjutnya.

DAFTAR PUSTAKA

- [1] A. A. Ilham, N. Harun dan Arwansyah, "Simulasi Penerapan Teknologi Data Mining Untuk Menghasilkan Model Pola Tanam Berkelanjutan," *Hasanuddin University*, pp. 124 – 132, April 2013.
- [2] Prasetyo, E., *DATA MINING : Konsep dan Aplikasi Menggunakan Matlab*, Yogyakarta, CV. Andi Offset, 2012.
- [3] A. Ardiyanto, "Dinas Pertanian dan Kehutanan Kabupaten Rembang," 20 Oktober 2014. [Online]. Available: <http://dintanhut.rembangkab.go.id/index.php/berita/201-produktivitas-padi-turun-akibat-kemarau-panjang>. [Accessed 28 Oktober 2014].
- [4] Kosasi S., "Penerapan Metode Jaringan Saraf Tiruan Backpropagation Untuk Memprediksi Nilai Ujian Sekolah," vol. 7, no. 1, pp. 20 – 28, Juni 2014.
- [5] Bin Hariyati dan Sutikno, "Pemodelan Luas Panen Padi Di Kabupaten Lamongan Dengan Indikator El Nino Southern Oscillation Melalui Pendekatan Robust Bootstrap Least Trimmed Square," Institut Teknologi Sepuluh Nopember, Surabaya, 2010.
- [6] Prasetyo, E., *DATA MINING : Mengolah Data Menjadi Informasi Menggunakan Matlab*, Yogyakarta, CV. Andi Offset, 2014.
- [7] F.A. Hermawati, *Data Mining*, Yogyakarta: CV. Andi Offset, 2013.
- [8] D.M. Silaban, DR. D. Saepudin dan DR. Adiwijaya, "Prediksi Time Series Harga Saham Menggunakan Metode

- Learning Support Vector Machine (SVM) dengan Studi Kasus PT. Aneka Tambang (ANTAM) Tbk,” *Telkom University*, 2013.
- [9] D.P. Sari, “Analisis Performansi Support Vector Regression Dalam Memprediksi Bonus Tahunan Karyawan,” *Universitas Diponegoro Semarang*, 2009.
- [10] D. Rosmala, J. Pardede, dan Baehaqi, "Sistem Simulasi Forecasting Potensi Penerimaan Pajak Kendaraan Bermotor (PKB) di Jawa Barat dengan Mengimplementasikan Algoritma Regresi," *Informatika*, vol. 3, no. 1, pp. 8-17, Januari-April 2012.
- [11] F.O. Musalim, L.W. Santoso, dan A. Setiawan, “Pembuatan Aplikasi Analisa Sistem Pendukung Keputusan Dalam Penentuan Harga Penjualan,” *Universitas Kristen Petra*, 2012.
- [12] Anton Azwar Ardywinata, "Estimasi Pelanggan dan Kebutuhan Daya Listrik Kalimantan Selatan kategori Rumah Tangga," Universitas Lambung Mangkurat, Banjarbaru, Bachelor Thesis 2013.
- [13] Ian H. Witten and Eibe Frank, *Data Mining Practical Machine Learning Tools and Techniques*, 2nd ed. San Francisco, United States of America: Diane Cerra, 2005.
- [14] N. Samosir. “USU Institutional Repository”, 2011, Home page on-line. Available from repository.usu.ac.id/bitstream./4/Chapter%20II.pdf;
- Internet ; accessed 12 Desember 2014.*
- [15] Susanto, S. dan Suryadi, D., Pengantar Data Mining: Menggali Pengetahuan dari Bongkahan Data, Yogyakarta, CV. Andi Offset, 2010.
- [16] “Pengujian dan Analisa Data,” 2015. [Online]. Available: <http://digilib.its.ac.id/public/ITS-paper-24639-1308100085-Paper.pdf>. [Accessed 28 Mei 2015].

