

PENERAPAN DATA MINING UNTUK MENENTUKAN ESTIMASI PRODUKTIVITAS TANAMAN TEBU DENGAN MENGGUNAKAN ALGORITMA LINIER REGRESI BERGANDA DI KABUPATEN REMBANG

Eggy Inaidi Andana Warih¹, Yuniarsi Rahayu²

Teknik Informatika, Fakultas Ilmu Komputer, Universitas Dian Nuswantoro

Jl. Nakula No. 5-11 Semarang-50131

E-mail : eggyiaw9999@gmail.com¹, yuniarsi_r@dsn.dinus.ac.id²

Abstrak

Dinas Pertanian dan Kehutanan Kabupaten Rembang adalah sebuah instansi yang bergerak dibidang pertanian, peternakan dan hasil hutan. Salah satu komoditi tanaman pangan di Kabupaten Rembang adalah tebu. Dilihat dari hasil yang didapatkan setiap tahunnya, produksi tanaman tebu tidak konsisten perkembangannya atau tergolong naik turun disetiap tahunnya, dengan memiliki luas lahan 147313 hektar. Kabupaten Rembang sendiri seharusnya menjadi salah satu penghasil tebu yang banyak. Dengan adanya faktor luas, produksi, jumlah petani dan produktivitas maka dapat dilakukan estimasi untuk mengetahui produksi tanaman tebu di Kabupaten Rembang. Oleh sebab itu, dengan digunakannya algoritma Linier Regresi Berganda diharapkan akan membantu untuk memperoleh hasil seberapa besar produksi tanaman tebu di Kabupaten Rembang sebagai acuan petani untuk menambah hasil panen tebu di tiap tahunnya.

Kata kunci : *Data Mining, Linier Regresi Berganda, Produktivitas, Pertanian, Tebu, Estimasi*

Abstract

Department of Agriculture and Forestry Rembang is an institution that is engaged in agriculture, livestock and forest products. One commodity crops in Rembang is sugarcane. Judging from the results obtained each year, crop production is not consistent development or classified up and down every year, with a land area of 147 313 hectares. Rembang itself should be one of the many sugar cane producer. With the large factors, the production, the number of farmers and productivity can be performed to determine the estimates of crop production in Rembang. Therefore, with the use of Linear Regression algorithm will hopefully help to obtain the results of how big the production of sugarcane in Rembang as a reference for farmers to increase crop yields of sugar cane in each year.

Keyword : *Mining, Linear Regression, Productivity, Agriculture, Sugar, Estimate.*

1. Pendahuluan

Gula merupakan salah satu komoditas pangan dimana selama hampir satu dekade terakhir ini kondisinya menunjukkan kecenderungan semakin terpuruk. Produksi gula mengalami penurunan, beberapa diantaranya adalah disebabkan oleh menurunnya produktivitas tanaman tebu dan berkurangnya luas areal tanam tebu karena kurang menguntungkan. Bila dilihat dari statistik luas areal perkebunan tebu, maka sejak mulai tahun 1990an luas tanam tanaman tebu

terus mengalami penurunan terutama di Pulau Jawa. Hal ini terutama disebabkan oleh semakin terdesaknya lahan untuk budidaya tanaman tebu terutama di lahan sawah yang merupakan andalannya. Lahan sawah merupakan andalan karena biasanya cukup subur sehingga mampu menghasilkan produksi yang bisa dibilang lumayan banyak dari pada lahan lainnya.

Berkembangnya Ilmu Data Mining memberikan sebuah inovasi baru dalam pendayagunaan kumpulan data yang

banyak sehingga dapat bermanfaat bagi pengembangan pengetahuan, baik secara khusus pada bidang yang berkaitan dengan data tersebut maupun secara global. Banyak juga fungsi yang dapat diterapkan dalam ilmu data mining, antara lain klasterisasi, prediksi, estimasi, klasifikasi dan asosiasi. Untuk mencapai fungsi-fungsi tersebut dilakukan dengan berbagai metode (algoritma) seperti *Linier Regression* untuk estimasi, *Support Vector Machines* (SVM) untuk prediksi, *K-Means* untuk klasterisasi, *C.45* untuk klasifikasi, dan *apriori* untuk asosiasi[2].

Estimasi adalah paling tepat untuk menyelesaikan masalah ini, karena dapat diperkirakan seberapa banyak hasil produksi tanaman tebu ini. Estimasi sendiri memiliki beberapa keunggulan diantara lain dapat memprediksi data *time series* dari beberapa tahun tertentu, yang dirangkum dan dihitung sehingga menghasilkan prediksi yang akan datang dengan hasil berupa prosentase[3]. Dalam hal ini, estimasi yang akan dicari adalah dari tahun 2007 sampai tahun 2014 berdasarkan data yang telah didapat. Beberapa prosentase yang didapat akan dihasilkan dari perhitungan estimasi dan juga dibantu dengan menggunakan Algoritma Linier Regresi Berganda.

2. Landasan Teori

2.1 Data Mining

Data mining adalah proses yang mempekerjakan satu atau lebih teknik pembelajaran computer (*machine learning*) untuk menganalisis dan mengekstraksi pengetahuan (*knowledge*) secara otomatis. Data mining merupakan proses iteratif dan interaktif untuk menemukan pola atau model baru yang sempurna, bermanfaat dan dapat

dimengerti dalam suatu database yang sangat besar. Data mining berisi pencarian trend atau pola yang diinginkan dalam database besar untuk membantu pengambilan keputusan di waktu yang akan datang. Pola-pola ini dikenali oleh perangkat tertentu yang dapat memberikan suatu analisa data yang berguna dan berwawasan yang kemudian dapat dipelajari dengan lebih teliti, yang mungkin saja menggunakan perangkat pendukung keputusan yang lainnya [7].

2.2 Estimasi

Estimasi adalah suatu metode dimana kita dapat memperkirakan nilai Populasi dengan memakai nilai sampel. Estimasi biasanya diperlukan untuk mendukung keputusan yang baik, menjadwalkan pekerjaan, menentukan berapa lama proyek perlu dilakukan dan berapa biayanya, menentukan apakah proyek layak dikerjakan, mengembangkan kebutuhan arus kas, menentukan seberapa baik kemajuan proyek, menyusun anggaran time phased dan menetapkan baseline proyek [9].

2.3 Regresi Linier Berganda

Regresi linier berganda adalah analisis regresi yang menjelaskan hubungan antara peubah respon (variabel dependen) dengan faktor-faktor yang mempengaruhi lebih dari satu prediktor (variabel independen). Ketika suatu hasil/keluaran, atau kelas berupa numerik, dan semua atribut adalah numerik, regresi linear adalah teknik yang tepat untuk menyelesaikan. Ini adalah metode pokok di dalam ilmu statistik. Gunanya adalah untuk mengekspresikan kelas sebagai kombinasi linear dari atribut, dengan bobot yang telah ditentukan :

$$Y = a + a_1X_1 + a_2X_2 + \dots + a_kX_k \quad (1)$$

Di mana Y adalah kelas; X_1, X_2, \dots, X_k adalah nilai atribut; dan a, a_1, \dots, a_k adalah bobot [10].

Bobot dihitung dari data sampel. Disini notasi menjadi sedikit sulit, karena membutuhkan suatu cara untuk mengekspresikan nilai-nilai atribut untuk setiap contoh sampel. Contoh pertama semisal ada kelas, katakan Y , dan nilai atribut $X_1, X_2, \dots, X_k^{(1)}$, dimana superscript yang ditunjukkan adalah contoh pertama.

Nilai prediksi untuk kelas contoh pertama dapat ditulis sebagai :

$$an + a_1X_1 + a_2X_2 + \dots + a_kX_k = \sum_{j=0}^k a_j X_j^{(1)}$$

2.4 Eliminasi Gauss Jordan

Metode ini diberi nama Gauss-Jordan untuk menghormati Carl Friedrich Gauss dan Wilhelm Jordan. Metode ini sebenarnya adalah modifikasi dari metode Gauss, yang dijelaskan oleh Jordan di tahun 1887. Dalam eliminasi Gauss-Jordan, matriks A dieliminasi menjadi matriks identitas I . Solusinya langsung diperoleh dari vector kolom berhasil proses eliminasi[11].

$$Ax = b \rightarrow Ix = b'$$

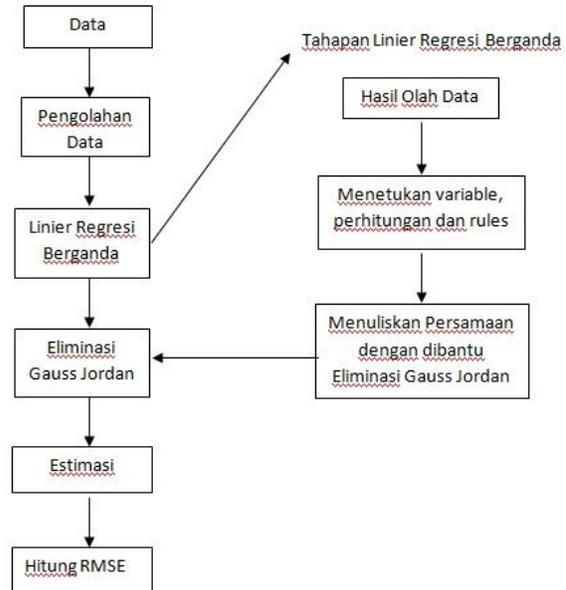
Dalam bentuk matriks, eliminasi Gauss-Jordan ditulis sebagai :

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array} \right) \rightarrow \left(\begin{array}{cccc|c} 1 & 0 & 0 & \dots & 0 & b_1' \\ 0 & 1 & 0 & \dots & 0 & b_2' \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & b_n' \end{array} \right)$$

Dalam penelitian ini, metode eliminasi Gauss-Jordan digunakan untuk membantu pencarian persamaan Linier Regresi Berganda. Untuk mempermudah

penyelesaian menggunakan algoritma linier regresi.

3. Hasil dan Pembahasan



Dari gambar diatas, dapat dijelaskan langkah demi langkah sebagai berikut:

1. Dari data yang di dapat, kemudian dikelompokkan menjadi data-data yang diperlukan yaitu hanya data tebu saja
2. Kemudian setelah di kelompokkan, data diolah dan dilakukan penghilangan atribut angka 0 yang tidak diperlukan guna mempermudah pada proses perhitungan.
3. Setelah itu menentukan variable, perhitungan dan rules. Menuliskan persamaan Linier Regresi Berganda.
4. Perhitungan dengan Eliminasi Gauss Jordan, kemudian di Estimasi hasil perhitungan tersebut. RMSE (Root Mean Squared Error), adalah untuk mengecek seberapa besar error yang didapat.

a. Menghitung Persamaan Linier

$$\sum y = \sum x_1 + \sum x_2 + \sum x_3 = \sum Y \quad (1)$$

$$\sum x_1^2 + \sum x_1 x_2 + \sum x_1 x_3 = \sum x_1 y \quad (2)$$

$$\sum x_2^2 + \sum x_1 x_2 + \sum x_2 x_3 = \sum x_2 y \quad (3)$$

$$\sum x_3^2 + \sum x_1 x_3 + \sum x_2 x_3 + \sum x_2^2 = \sum x_3 y \quad (4)$$

Sehingga menjadi

$$143 + 12839 + 147313 + 13251 = 15132$$

$$12839 + 1396533 + 14619280 + 1292524 = 1521763$$

$$147313 + 14619280 + 183233487 + 15376727 = 17441583$$

$$13251 + 1292524 + 15376727 + 1504097 = 1585881$$

Dengan menggunakan ELimasi Gauss Jordan, maka didapat hasil :

$$a = -221,141222$$

$$a_1 = -0,75616896$$

$$a_2 = -0,36199219$$

$$a_3 = 8,285397644$$

Maka diperoleh persamaan : $Y = -221,141222 + -0,75616896 X_1 + -0,36199219 X_2 + 8,285397644 X_3$

Dengan menggunakan persamaan diatas, maka dapat dihitung pula estimasi produktivitas dengan memasukkan nilai X_1 (luas), X_2 (produksi), dan X_3 (jumlah petani) untuk mendapatkan estimasi produktivitas.

b. Menghitung Estimasi Produktivitas

Untuk menghitung estimasi produktivitas tanaman tebu maka persamaan $Y = -221,141222 + -0,75616896 X_1 + -0,36199219 X_2 + 8,285397644 X_3$, Maka hasil perhitungan pada Tabel 4.4 mendapatkan nilai estimasi 15.132,00067 Kg/Ha. Jadi bila dibandingkan dengan tahun sebelumnya, produktivitas mengalami peningkatan.

c. Menghitung Root Mean Squared Error

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

Maka hasil perhitungan standart error masih tergolong besar yaitu 206,3074904895105 , karena data yang digunakan hanya 143 data, jika data yang digunakan semakin banyak maka standart error akan semakin sedikit.

4. Penutup

A. Kesimpulan

Dari hasil perhitungan, di dapat bahwa hasil estimasi produktivitas mendapatkan 15.132,00067 Kg/Ha yang sebelumnya lahanya adalah 147313 hektar (ha), berarti produktivitas tahun depannya mengalami peningkatan dari tahun sebelumnya. Maka dapat menjadi bahan evaluasi kedepannya agar hasil tanaman tebu semakin meningkat di tahun-tahun berikutnya.

B. Saran

Untuk mendapatkan nilai yang lebih akurat, penyertaan faktor-faktor lain dalam melakukan estimasi sangat diperlukan, agar mendapatkan hasil akurasi yang baik.

Daftar Pustaka

- [1] Pratiwi, R. R., Fillah Fithra Dieny. 2014. Hubungan Skor Frekuensi Diet Bebas Casein Dengan Skor Perilaku Autis. *Journal of Nutrition College*. Vol. 3., No. 1. Hal : 34-42.
- [2] Nurmansyah, W., Sri Hartati. 2013. Prototipe Sistem Pakar Penentu Jenis Gangguan Psikologi Klinis Menggunakan *Forward Chaining* dan *Formula Bayes*. Yogyakarta. FMIPA UGM.
- [3] Rohman, F. F., Ami Fauziah. 2008. Rancang Bangun Aplikasi Sistem Pakar Untuk Menentukan Jenis Gangguan Perkembangan Pada Anak. *Media Informatika*, Vol. 6, No. 1.

- [4] Sutojo, T., Edi Mulyanto, dan Vincent Suhartono. 2011. *Kecerdasan Buatan*. Yogyakarta : Penerbit Andik.
- [5] Aribowo, A.S dan Khomsah, S.. 2011. Sistem Pakar Dengan Beberapa *Knowledge Base* Menggunakan *Probabilitas Bayes* dan Mesin Inferensi *Forward Chaining*. *Semnas IF*. 51 – 58
- [6] K. Agus Wirawan, IK. Budi Susrusa, IGAA. Ambarwati, “Analisis Produktivitas Tanaman Padi di Kabupaten Badung Provinsi Bali,” Universitas Udayana, 2014.
- [7] Gorunescu, Florin, “Data Mining: Concepts, Models, and Techniques Verlag” Berlin Heidelberg: Springer, 2011.
- [8] Romeo, dkk, “Data Mining dan Knowledge Discovery in Database”, Jakarta, 2007.
- [9] Prasetyo, E., DATA MINING :Mengolah Data Menjadi Informasi Menggunakan Matlab, Yogyakarta, CV. Andi Offset, 2014.
- [10] F.O. Musalim, L.W. Santoso, dan A. Setiawan, “Pembuatan Aplikasi Analisa Sistem Pendukung Keputusan Dalam Penentuan Harga Penjualan, “ *Universitas Kristen Petra*, 2012.
- [11] Yuniarsi Rahayu, S.Si, M.Kom, ”Penerapan Metode Numerik Pada Peramalan Untuk Menghitung Koefisien-Koefisien Pada Garis Regresi Linier Berganda” Universitas Dian Nuswantara, 2011.
- [12] A.Yusuf, H. Ginardi, Isye Arieshanti, “Pengembangan Perangkat Lunak Prediktor Nilai Mahasiswa Menggunakan Metode Spectral Clustering dan Bagging Regresi Linier,” September, 2012.
- [13] Evy Thyrida Silaban, Edison Purba, Jasmani Ginting, “Pertumbuhan dan Produksi Jagung Manis Pada Berbagai Jarak Tanam dan Waktu Olah Tanah,” Universitas Sumatera Utara, 2013