

# Implementasi Data Mining Untuk Estimasi Produktivitas Kacang Hijau Dengan Menggunakan Algoritma Regresi Linier Di Kabupaten Grobogan

Yulinda Kusumaningrum<sup>1</sup>, Siti Asmiatun<sup>2</sup>, Astrid Novita Putri<sup>3</sup>

<sup>1</sup>Program Studi Teknik Informatika, Fakultas Teknik Informasi dan Komunikasi, Universitas Semarang

<sup>1</sup>Jl. Soekarno Hatta, Tlogosari Kulon, Kota Semarang, Jawa Tengah 50131, e-mail: yulinda@usm.ac.id

<sup>2</sup>Jl. Soekarno Hatta, Tlogosari Kulon, Kota Semarang, Jawa Tengah 50131, e-mail: mia@usm.ac.id

<sup>3</sup>Jl. Soekarno Hatta, Tlogosari Kulon, Kota Semarang, Jawa Tengah 50131, e-mail: astrid@usm.ac.id

## ARTICLE INFO

### Article history:

Received 15 Desember 2020

Received in revised form 26 Juli 2021

Accepted 27 Juli 2021

Available online 31 Juli 2021

## ABSTRAK

The Grobogan Regency Agriculture Service is an agency that operates in the agricultural sector. One of the crop commodities in Grobogan Regency is green beans. Judging from the results obtained each year, green bean production in Grobogan Regency is inconsistent. The rise and fall of green bean productivity is influenced by several factors. Factors such as area, production, number of farmers and productivity can be estimated to determine the production of green beans in Grobogan Regency. Therefore, using the Multiple Linear Regression algorithm is expected to help to obtain results on how much green bean production is in Grobogan Regency as a reference for farmers to increase their green bean harvest each year. Based on the calculation results, it was found that the estimated productivity of green beans in Grobogan Regency reached 760.8297302 Tons/Ha, whereas previously the land was 865 Hectares (Ha)

Keywords: data mining, linear regression, estimation, productivity, green beans

## 1. Pendahuluan

Kacang hijau merupakan salah satu komoditas tanaman pangan yang penting di Indonesia. Komoditas kacang hijau merupakan salah satu komoditas dari tanaman pangan yang memiliki nilai ekspor terbesar kedua setelah jagung. Kacang hijau merupakan salah satu tanaman pangan sumber protein nabati. Kandungan protein kacang hijau sebesar 22% menempati urutan ketiga setelah kedelai dan kacang tanah. Kacang hijau berumur genjah (55-65 hari), tahan kekeringan, variasi jenis penyakit relatif sedikit, dapat ditanam pada lahan kurang subur dan harga jual relatif tinggi serta stabil [1]. Di Jawa Tengah yang sudah menjadi daerah sentra produksi kacang hijau, di antaranya Kabupaten Demak, Grobogan, Kebumen dan masih banyak lagi. Apresiasi patut diberikan kepada Kementerian Pertanian yang selalu mengoptimalkan potensi-potensi yang dimiliki oleh Indonesia terutama di sektor pertanian. Salah satunya di Kabupaten Grobogan yang menjadi salah satu Kabupaten penghasil Kacang hijau terbesar di Jawa Tengah data dari BPS Jawa Tengah [2]. Dari Badan Pusat Statistik (BPS) Jawa Tengah mengenai Luas Tanam Padi dan Palawija di Kabupaten Grobogan tahun 2018, jumlah Luas tanam Padi dan Palawija pada kacang hijau menempati urutan ke 3 setelah Padi dan Jagung. Hal ini sangat kurang untuk memenuhi produksi kacang hijau yang sangat tinggi untuk di ekspor juga [3]. Oleh sebab itu maka diperlukan estimasi produktivitas kacang hijau.

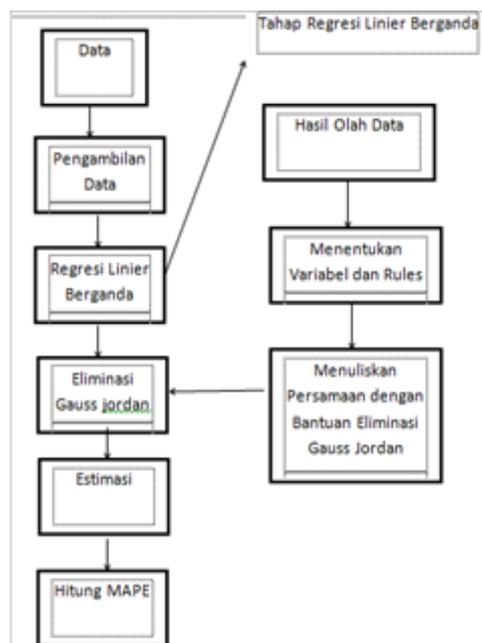
Dengan Berkembangnya Ilmu Data Mining memberikan sebuah inovasi baru dalam pendayagunaan kumpulan data yang banyak sehingga dapat bermanfaat bagi pengembangan pengetahuan, baik secara khusus pada bidang yang berkaitan dengan data tersebut maupun secara global. Banyak juga fungsi yang dapat diterapkan 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 [4]

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 [5]. Estimasi sendiri memiliki beberapa keunggulan di antara lain dapat memprediksi data time series dari beberapa tahun tertentu, yang dirangkum dan dihitung sehingga menghasilkan prediksi yang akan datang dengan hasil berupa prosentase. Beberapa prosentase yang didapat akan dihasilkan dari perhitungan estimasi dan juga dibantu dengan menggunakan Algoritma Linier Regereasion Berganda [6]

Dari latar belakang diatas dapat disimpulkan bahwa di perlukan suatu estimasi produksi berdasarkan luas lahan, produksi dan produktivitas sehingga dapat memaksimalkan hasil produksi kacang hijau di Kabupaten Grobogan.

## 2. Metode Penelitian

Penelitian ini menggunakan metode kuantitatif. Dimana metode ini digunakan untuk mengetahui pengaruh antar variabel yang akan diteliti [7]. Data primer adalah data yang diperoleh dari <http://aplikasi.pertanian.go.id/bdsp/newind.asp> dan dinas pertanian Grobogan mengenai Data Produksi, Data Produktivitas, Data Luas Panen Kacang Hijau. Pada penelitian ini terdapat beberapa tahapan penelitian seperti pada gambar di bawah ini :



### Gambar 1. Database Mirroring Architecture

Pada penelitian ini menggunakan sekunder yang diperoleh dari Badan Pusat Statistik dan Data Dinas Pertanian Republik Indonesia. Dari gambar diatas, dapat dijelaskan langkah demi langkah sebagai berikut: Dari gambar diatas, dapat dijelaskan langkah demi langkah sebagai berikut :

- Dari data yang di dapat, kemudian dikelompokkan menjadi data-data yang diperlukan yaitu hanya data kacang hijau saja
- Kemudian setelah di kelompokkan, data diolah dan dilakukan penghilangan atribut angka 0 yang tidak diperlukan guna mempermudah pada proses perhitungan.
- Setelah itu menentukan variable, perhitungan dan rules. Menuliskan persamaan Linier Regresi Berganda.
- Perhitungan dengan Eliminasi Gauss Jordan, kemudian di Estimasi hasil perhitungan tersebut. MAPE adalah untuk mengecek seberapa besar error yang didapat.

### 3. Hasil Dan Pembahasan

#### 3.1 Pengmuplan Data

Data Luas Tanam Kacang Hijau dan Dinas Pertanian Kabupaten Grobogan. Data tersebut dapat dilihat di tabel 1.

**Tabel 1 Data Sampel Produksi Kacang Hijau**

No	Kecamatan	Luas Tanam	Luas Panen	Produksi
1	Kedungjati	0	0	0
2	Karangrayung	951	915	1027
3	Penawangan	3842	3692	4562
4	Toroh	387	372	431
5	Geyer	0	0	0
6	Pulokulon	0	0	0
7	Kradenan	335	276	326
8	Gabus	556	534	662
9	Ngaringan	1230	1044	1228
10	Wirosari	1163	1187	1459
11	Tawangharjo	1860	1787	2025
12	Grobogan	1098	1055	1243
13	Purwodadi	2698	2593	3192
14	Brati	880	846	952
15	Klambu	1342	1290	1340
16	Godong	6081	5852	7250
17	Gubug	35	35	42
18	Tegowanu	1065	1023	1138
19	Tanggunharjo	865	765	882

#### 3.2 Preprocessing Data

Tahap Preprocessing Data melakukan pembersihan atau cleaning data. Berikut ditampilkan beberapa lampiran data dari 400 record data disajikan pada tabel 2 berikut:

Tabel 2. Data Sampel Preprocessing Kacang Hijau

No	Kecamatan	Luas Tanam	Luas Panen	Produksi
1	Karangrayung	951	915	1027
2	Penawangan	3842	3692	4562
3	Toroh	387	372	431
4	Kradenan	335	276	326
5	Gabus	556	534	662
6	Ngaringan	1230	1044	1228
7	Wirosari	1163	1187	1459
8	Tawangharjo	1860	1787	2025
9	Grobogan	1098	1055	1243
10	Purwodadi	2698	2593	3192
11	Brati	880	846	952
12	Klambu	1342	1290	1340
13	Godong	6081	5852	7250
14	Gubug	35	35	42
15	Tegowanu	1065	1023	1138
16	Tanggungharjo	865	765	882

### 3.3 Perhitungan Regresi

Dalam penelitian ini yang menjadi variabel dependen adalah Hasil Produksi (Y). Sedangkan variabel independen adalah Luas Tanam (X1), Luas Panen (X2). Untuk mengetahui pengaruhnya dapat digunakan rumus analisis regresi linier berganda sebagai berikut :

$$Y = a + b_1 X_1 + b_2 X_2 + e$$

Untuk mencari persamaan regresi linier berganda, terlebih dahulu menghitung koefisien-koefisien regresi dengan mencari penggandaan antara variabel yang satu dengan yang lain. Berikut hasil koefisien Regresi Linier Berganda pada tabel dibawah ini :

Tabel 3. Koefisien Regresi Linier Berganda

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>
Intercept	31,44834435	78,10503798	0,402642	0,688030415
Lpanen(ha)	1,365489249	0,224433181	6,084168	1,93191E-08
Ltanam(ha)	-0,364413745	0,217708161	-1,67386	0,097133926

Berdasarkan analisis data dengan menggunakan Excel, maka dapat diperoleh persamaan regresi

$$Y = 31,449 + 1,367_1 + -0,364_2 + e.$$

Hasil persamaan regresi diatas menunjukkan bahwa luas tanam dan luas panen sangat berpengaruh terhadap hasil produktivitas tanaman kacang hijau.

### 3.4 Estimasi produktivitas

Model regresi linear ganda tersebut kemudian diterapkan pada data testing. Hasil perhitungan estimasi dapat dilihat pada tabel 4 berikut :

**Tabel 4 Hasil prediksi produktivitas kacang hijau**

Sample	Luas panen(ha)	Luas tanam(ha)	Predicted Produksi (ton)
1	215	250	233,9250966
2	754	784	775,3268618
3	2029	2040	2058,62199
4	632	652	656,8397878
5	1397	1405	1427,035513
6	2363	2373	2393,345622
7	1763	1789	1786,8697
8	1362	1370	1391,997871
9	2201	2251	2216,594841
10	1908	1950	1926,195028
11	516	530	542,9015119
12	549	570	573,3861073
...	...	...	...
108	765	865	760,8297302

### 3.5 Menghitung Error dengan MAPE

Untuk mengetahui akurasi hasil produktivitas kacang hijau maka dilakukan perhitungan untuk mengetahui tingkat kesalahan dengan menggunakan MAPE (*Mean Absolute Percentage Error*).

**Tabel 5. Hasil Perhitungan MAPE**

Sample	Actual	Forecast	Error	Nilai Absolut Error	MAPE
			At -Ft	At -Ft	(At - Ft)/At
1	334	233,9251	100,0749	100,0749034	0,299625
2	901	775,3269	125,6731	125,6731382	0,139482
3	1932	2058,622	-126,622	126,6219903	0,065539
4	294	656,8398	-362,84	362,8397878	1,234149
5	240	1427,036	-1187,04	1187,035513	4,945981
6	961	2393,346	-1432,35	1432,345622	1,490474
7	822	1786,87	-964,87	964,8697001	1,173807
8	1381	1391,998	-10,9979	10,9978705	0,007964
9	1330	2216,595	-886,595	886,5948408	0,666613
10	1954	1926,195	27,80497	27,80497179	0,01423
11	882	542,9015	339,0985	339,0984881	0,384465
12	351	573,3861	-222,386	222,3861073	0,633579
13	981	926,9543	54,04566	54,04565927	0,055092

14	4740	4219,571	520,4289	520,4289092	0,109795
15	531	715,8111	-184,811	184,8110768	0,348043
.....	.....	.....	.....	.....	.....
Total MAPE					25,16286

Hasil pengujian prediksi hasil produksi kacang hijau dengan histori penjualan pada periode 2011-2019 menunjukkan bahwa hasil uji MAPE data yang diuji adalah presentase kesalahan kurang dari 50%. Berdasarkan tabel kriteria MAPE menunjukkan bahwa prediksi hasil produksi kacang hijau dengan metode Linear Regresi Sederhana adalah tergolong dalam kategori cukup. Hal tersebut sesuai dengan penelitian yang menyatakan bahwa produktivitas kacang hijau sangat dipengaruhi oleh luas lahan dan luas panen [8]. Hasil penelitian lain juga menyampaikana bahwa luas tanam dan luas panen itu sebanding dengan hasil produksinya [9].

#### 4. Kesimpulan dan Saran

##### 4.1 Kesimpulan

Dari hasil perhitungan diperoleh kesimpulan yaitu estimasi produktivitas kacang hijau di Kabupaten Grobogan mencapai 760,8297302 Ton/Ha yang sebelumnya adalah 865 Hektar (Ha). Hal tersebut menunjukkan produktivitas kacang hijau tahun depannya mengalami peningkatan dari tahun sebelumnya

##### 4.2 Saran

Untuk mendapatkan nilai yang lebih akurat, pada penelitian selanjutnya bisa menggunakan metode lain dalam perhitungannya.

#### Daftar Pustaka

- [1] Budhi Santoso Radjit dan Nila Prasetyaswati. 2012. Prospek Kacang Hijau Pada Musim Kemarau di Jawa Tengah. Buletin Palawija No. 24: 57-68. <http://dx.doi.org/10.21082/bulpa.v0n24.2012.p57-68>
- [2] Evi Dewi dan Rijal Adji. 2019. Estimasi Ketersediaan Lahan Pertanian Beras Menggunakan Algoritma Regresi Linear. Jurnal INFORMATION SYSTEM FOR EDUCATORS AND PROFESSIONALS Vol. 4, No. 1, Desember 2019, 75 - 84 E-ISSN: 2548-3587
- [3] Desi Putri Hastuti, Supriyono dan Sri Hartati. 2018. Pertumbuhan dan Hasil Kacang Hijau (*Vigna radiata*, L.) pada Beberapa Dosis Pupuk Organik dan Kerapatan Tanam. Journal of Sustainable Agriculture. 33(2), 89-95. Universitas Sebelas Maret.
- [4] F.O. Musalim, L.W. Santoso, dan A. Setiawan. 2012. "Pembuatan Aplikasi Analisa Sistem Pendukung Keputusan Dalam Penentuan Harga Penjualan," Universitas Kristen Petra.
- [5] Magdalena Sitindaon. 2017. Analisis Potensi Ekspor Hasil Pertanian Tanaman Pangan di Kabupaten Pati. Economics Development Analysis Journal 6 (1) (2017). Universitas Negeri Semarang
- [6] Nurmansyah, W., Sri Hartati. 2013. Prototipe Sistem Pakar Penentu Jenis Gangguan Psikologi Klinis Menggunakan *Forward Chaining* dan *Formula Bayes*. Yogyakarta. FMIPA UGM

- 
- [7] Padilah, T. N., & Adam, R. I. 2019. Analisis Regresi Linier Berganda Dalam Estimasi Produktivitas Tanaman Padi Di Kabupaten Karawang. *FIBONACCI: Jurnal Pendidikan Matematika dan Matematika*, 5(2), 117-128
- [8] Prasetyo, E., 2014. DATA MINING :Mengolah Data Menjadi Informasi Menggunakan Matlab, Yogyakarta, CV. Andi Offset
- [9] Sari, D. K., Ismullah, I. H., Sulasdi, W. N., & Harto, A. B. 2010. Estimasi Produktivitas Padi Sawah Berbasis Kalender Tanam Heterogen Menggunakan Teknologi Pengindraan Jauh. *Jurnal Itenas Rekayasa*, 14(3).
- [10] Warih, E. I. A., & Rahayu, Y. 2015. Penerapan Data Mining Untuk Menentukan Estimasi Produktivitas Tanaman Tebu Dengan Menggunakan Algoritma Linear Regresi Berganda di Kabupaten Rembang. *eprints. dinus. ac. id/16925/1/jurnal\_16115. pdf. Semarang*