

PREDIKSI PRODUKSI TANAMAN PANGAN DI PROVINSI GORONTALO MENGGUNAKAN METODE *K-NN (K-Nearest Neighbor)*

Kartika Chandra Pelangi
Universitas Ichsan Gorontalo
apelangie@gmail.com

Pertanian sudah menjadi peranan penting bagi suatu masyarakat di Provinsi Gorontalo khususnya pada daerah lahan yang luas serta memiliki tanah yang subur. Selain itu, pertanian berperan penting dalam peningkatan suatu daerah serta perekonomian daerah itu sendiri dan sebagai sarana dalam berusaha menghidupi kebutuhan keluarga, begitu pun pangan mempunyai peranan yang sangat penting bagi kehidupan manusia karena pangan adalah bahan makanan pokok yang paling utama di butuhkan oleh tubuh manusia, Permasalahannya yaitu luas panen yang mengalami perubahan setiap tahunnya berpengaruh pada hasil produksi tanaman pangan, Karena jika luas panen tanaman pangan besar maka hasil produksi akan meningkat dan jika luas panennya kecil maka hasil produksi tanaman pangan menurun. Untuk memenuhi kebutuhan pangan di Provinsi Gorontalo yang sering mengalami perubahan pada jumlah produksi tanaman pangan yang tidak menentu tiap tahunnya, pemerintah akan terus berupaya untuk meningkatkan hasil pertanian, Maka diperlukan suatu sistem untuk melakukan prediksi terhadap hasil pertanian yang ada di Provinsi Gorontalo berdasarkan luas lahan dan produksi panen di tiap tahunnya. Tujuannya yaitu untuk mengetahui hasil produksi tanaman pangan yang ada di Dinas Pertanian Provinsi Gorontalo. Dari hasil penelitian dengan program untuk Prediksi Produksi Tanaman Pangan Di Provinsi Gorontalo menggunakan metode K-nearest neighbor di dapatkan tingkat akurasi tertinggi pada komoditi jagung dengan tingkat akurasi sebesar 92,83 %. Hasil akurasi tersebut dapat dikategorikan bahwa aplikasi yang dibuat layak untuk digunakan dalam memprediksi hasil produksi tanaman pangan.

Kata Kunci : *K-Nearest Neighbor, PHP, Prediksi Pangan*

I. PENDAHULUAN

Pertanian sudah menjadi peranan penting bagi suatu masyarakat di Provinsi Gorontalo khususnya pada daerah lahan yang luas serta memiliki tanah yang subur. Selain itu juga, pertanian mempunyai peranan yang sangat penting dalam peningkatan suatu daerah serta perekonomian daerah itu sendiri dan sebagai sarana

dalam berusaha menghidupi kebutuhan keluarga, begitu pun pangan mempunyai peran yang sangat penting bagi kehidupan manusia karena pangan merupakan bahan makanan pokok yang sangat di butuhkan oleh tubuh manusia, selain untuk dikonsumsi sehari-hari, hasil produksi dari tanaman pangan dapat dijadikan sebagai sumber usaha dalam perekonomian masyarakat yang ada di provinsi gorontalo. Jenis tanaman pangan yang ada di Provinsi Gorontalo yaitu padi, jagung, kedelai, kacang tanah, kacang hijau, ubi kayu, dan ubi jalar[1].

Pertanian merupakan salah satu sumber ekonomi bangsa. Banyak penduduk Indonesia yang berprofesi sebagai petani, produksi pertanian pada dasarnya adalah bagaimana meningkatnya hasil dengan kualitas dan kuantitas maksimal dengan menggunakan pupuk, obat minimal, untuk memanfaatkan keterbatasan sumber daya alam (air, energy dan kondisi tanah) keterbatasan pupuk dan juga obat menjadikan harga yang tinggi, kondisi cuaca juga yang tidak menentu menjadikan ketersediaan air yang cukup mengganggu pertumbuhan tanaman pangan. Dan masih banyak lagi faktor lain yang mengakibatkan penurunan hasil produksi bahkan sampai gagal panen[2].

Dinas Pertanian Provinsi Gorontalo adalah sebuah instansi yang bergerak diberbagai bidang pertanian, dalam kegiatan tahunan melakukan pencatatan hasil produksi tanaman pangan. Hasil yang didapatkan setiap tahunnya yang ada di Dinas Pertanian Provinsi Gorontalo sering mengalami perubahan. Produksi tanaman pangan juga kadang tidak konsisten perkembangannya dan tidak tergolong meningkat atau menurun disetiap tahunnya.

Permasalahannya yaitu luas panen yang mengalami perubahan setiap tahunnya berpengaruh pada hasil produksi tanaman pangan, Karena jika luas panen tanaman pangan besar maka hasil produksi akan meningkat dan jika luas panennya kecil maka hasil produksi tanaman pangan menurun. Untuk memenuhi kebutuhan pangan di Provinsi Gorontalo yang sering mengalami perubahan pada jumlah produksi tanaman pangan yang tidak menentu tiap tahunnya, pemerintah akan terus berupaya untuk meningkatkan hasil pertanian, Maka diperlukan suatu sistem untuk melakukan prediksi terhadap hasil pertanian yang ada di Provinsi Gorontalo

berdasarkan luas lahan dan produksi panen di tiap tahunnya. Tujuannya yaitu untuk mengetahui hasil produksi tanaman pangan yang ada di Dinas Pertanian Provinsi Gorontalo.

Tabel 1. 1 Sampel Data

Tahun	Jenis Tanaman Pangan	Luas Panen	Produksi
2017	Padi Sawah	74.954	342.172
	Padi Ladang	2.255	8.010
	Jagung	336.001	1.552,001
	Kedelai	3172	4998
	Kacang Tanah	1,135	1,172
	Kacang HIjau	170	129
	Ubi Kayu	683	7.838
	Ubi Jalar	59	416
2018	Padi Sawah	70.539	326.587
	Padi Ladang	15.964	27.449
	Jagung	343.241	1.554,751
	Kedelai	3.367	5.204
	Kacang Tanah	1,293	1,244
	Kacang HIjau	120	116
	Ubi Kayu	871	8.757
	Ubi Jalar	40	506
2019	Padi Sawah	21.980	110.975
	Padi Ladang	3.898	11.113
	Jagung	166.298	917.743
	Kedelai	2.541	4132
	Kacang Tanah	1,398	1,4212
	Kacang HIjau	703	831
	Ubi Kayu	654	6.675
	Ubi Jalar	13	178

(Sumber :Dinas Pertanian Provinsi Gorontalo, 2019)

II. LANDASAN TEORI

2.1 Data Mining

Data mining merupakan proses pengekstrasian informasi dari sekumpulan data yang sangat besar melalui penggunaan algoritma dan teknik penarikan dalam bidang

statistik. Data mining dan *knowledge discovery in database* (KDD) sering digunakan dalam secara bergantian untuk menjelaskan proses penggalian informasi tersembunyi yang terdapat pada suatu basis data yang besar. Data mining juga dikenal sebagai istilah *pattern recognition*. Menyebutkan bahwa KDD *Knowledge Discovery from Data*, yaitu merupakan proses terstruktur sebagai berikut:

1. *Data selection* yaitu proses untuk memilih data dari database yang sesuai dengan tujuan analisis.
2. *Data preprocessing* yaitu sebelum proses *data mining* dapat dilaksanakan, perlu dilakukan proses *cleaning* dengan tujuan untuk membuang duplikasi data, memeriksa data yang inkonsisten, dan memperbaiki kesalahan pada data, seperti kesalahan cetak (tipografi). Juga dilakukan proses *enrichment*, yaitu proses memperkaya daya yang sudah ada dengan data atau informasi lain yang relevan dan diperlukan KDD. Seperti data atau informasi eksternal.
3. *Data integration* merupakan proses untuk menggabungkan data dari beberapa sumber yang berbeda.
4. *Data transformation* merupakan proses *coding* pada data yang telah dipilih, sehingga data tersebut sesuai untuk proses *data mining*. Proses *coding* pada KDD merupakan proses kreatif dan sangat tergantung pada jenis atau polaninformasi yang akan dicari pada database.
5. *Data mining* merupakan proses penting yang menggunakan sebuah metode tertentu untuk memperoleh sebuah pola dari data.
6. *Pattern Evaluation* merupakan proses mengidentifikasi pola.
7. *Interpretation* atau *Evaluation* yaitu pola informasi yang dihasilkan dari proses data mining perlu ditampilkan dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan. Tahap ini merupakan bagian dari proses KDD yang disebut dengan *interpretation*. Tahap ini mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesa yang ada sebelumnya[6].

2.2 Metode K-Nearest Neighbor

K-Nearest Neighbor (KNN) termasuk kelompok instance-based learning. Algoritma ini juga merupakan salah satu teknik lazy learning. KNN dilakukan dengan mencari kelompok k objek dalam data training yang paling dekat (mirip) dengan objek pada data baru atau data testing. Algoritma K-Nearest Neighbor adalah sebuah metode untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut. Nearest Neighbor adalah pendekatan untuk mencari kasus dengan menghitung kedekatan antara kasus baru dan kasus lama yaitu berdasarkan pada pencocokan bobot dari sejumlah fitur yang ada.

Langkah-langkah untuk menghitung metode Algoritma *K-Nearest Neighbor*:

- a. Menentukan Parameter K (Jumlah tetangga paling dekat).

- b. Menghitung kuadrat jarak Euclid (*queri instance*) masing-masing objek terhadap data sampel yang diberikan.
- c. Kemudian mengurutkan objek-objek tersebut ke dalam kelompok yang mempunyai jarak Euclid terkecil.
- d. Mengumpulkan kategori Y (*Klasifikasi Nearest Neighbor*)
- e. Dengan menggunakan kategori *Nearest Neighbor* yang paling mayoritas maka dapat diprediksi nilai *queri instance* yang telah dihitung.

Ada banyak cara untuk mengukur jarak kedekatan antara data baru dengan data lama (data training), diantaranya euclidean distance dan manhattan distance (city block distance), yang paling sering digunakan adalah euclidean distance.

$$\sqrt{(a_1 - b_1)^2 + a_2 - b_2)^2 + \dots + a_n - b_n)^2} \dots\dots\dots (2.1)$$

Dimana :

a = a1,a2, ..., an, dan;

b = b1, b2,..., bn mewakili n nilai atribut dari dua record.

Untuk atribut dengan nilai kategori, pengukuran dengan euclidean distance tidak cocok. Sebagai penggantinya, digunakan fungsi sebagai berikut:

Different (a_i,b_i)

$$\{ 0 \text{ jika } a_i = b_i \text{ dan } 1 \text{ jika } a_i \neq b_i \dots\dots\dots (2.2)$$

dimana a_i dan b_i adalah nilai kategori.

Jika nilai variabel antara dua record yang dibandingkan sama maka nilai jaraknya 0, jika nilai variabel mempunyai kemiripan atau kedekatan maka diberi bobot 0,5, sebaliknya jika nilai variabel berbeda maka akan diberi bobot 1 artinya tidak mirip sama sekali. Misalkan variabel status dengan nilai menikah dan menikah, maka nilai kedekatannya 0, jika menikah dan janda maka nilai kedekatannya antara 0,5, sedangkan jika nilai menikah dan belum menikah maka nilai kedekatannya adalah 1.

Untuk menghitung kemiripan kasus, digunakan rumus sebagai berikut:

$$\text{similarity}(T, S) = \frac{\sum_{i=1}^n f(T_i, S_i) * w_i}{w_i} \dots\dots\dots (2.3)$$

Keterangan :

- T = Kasus baru
- S = Kasus yang ada dalam penyimpanan
- n = Jumlah atribut dalam tiap kasus
- T_i = Kasus baru dalam Atribut individu antara 1 sampai dengan n
- S_i = Kasus yang ada dalam penyimpanan Atribut individu antara 1 sampai dengan n
- f = Fungsi similarity atribut i antara kasus T dan kasus S
- w_i = Bobot yang diberikan pada atribut ke-i kedalam Atribut individu antara 1 sampai dengan n.

III.HASIL DAN PEMBAHASAN

3.1 Penerapan Metode K-Nearest Neighbor Time Series Forecasting

Contoh data Training Sebagai Berikut :

Tabel 3. 1 Data Training

BULAN	TAHUN 2018	
	TARGET(Rp) (X)	REALISASI (Rp) (y)
JULI	8.953.000	3.765.500
AGUSTUS	10.300.000	4.334.000
SEPTEMBER	9.333.000	3.615.500
OKTOBER	5.333.000	3.275.000

Selanjutnya Untuk Data Testing sebagai data yang akan di prediksi adalah sebagai berikut:

Tabel 3. 2 Data Testing

BULAN	TARGET (Rp) (X)	REALISASI (Rp) (y)
	November	8.333.000

Maka akan di prediksi beberapa realiasi untuk bulan November dengan menggunakan KNN.

Langkah pertama

Menghitung jarak antar training dengan testing

$$\text{Jarak juli} = \sqrt{(8.953.000 - 8.333.000)^2} = 620.000$$

$$\text{Jarak agustus} = \sqrt{(10.300.000 - 8.333.000)^2} = 1.967.000$$

$$\text{Jarak September} = \sqrt{(9.333.000 - 8.333.000)^2} = 1.000.000$$

$$\text{Jarak oktober} = \sqrt{(5.333.000 - 8.333.000)^2} = 3.000.000$$

Langkah kedua

Menentukan peringkat dari jarak yang telah di dapatkan jika menggunakan K=3

Tabel 3. 3 Peringkat Jarak

Peringkat	Bulan	Jarak	Target	Realisasi
1	Juli	620.000	8.953.000	3.765.500
2	September	1.000.000	9.333.000	3.615.500
3	agustus	1.967.000	10.300.000	4.334.000

Langkah ketiga

Menghitung Rata-Rata dari nilai realisasi berdasarkan nilai K sebagai hasil prediksi

$$\text{Prediksi} = (3.765.500 + 3.615.500 + 4.334.000) / 3 = 3.905.000$$

Berarti prediksi untuk realisasi bulan November adalah **Rp 3.905.000** dengan presentase (realisasi/target)*100 dengan rumus berikut

$$\text{Presentase} = \frac{3.905.000}{8.333.000} * 100$$

$$\frac{3.905.000}{8.333.000} = 46.86\%$$

Maka persentasenya adalah 46.86%

Pengujian Metode

Untuk mengetahui kehandalan dari model prediksi menggunakan Algoritma Knn maka perlu dilakukan pengujian model dengan menggunakan RMSE dan MAPE adapun langkah-langkah sebagai berikut :

1. Mengambil beberapa sampel dari dataset kemudian dilakukan prediksi kembali.

Tabel 3. 4 Prediksi Data Set

NO	Target	Realisasi	Hasil Prediksi
1	8333000	3500000	3655640
2	8333000	4375000	3655640
3	5830000	2812000	2977000
4	5830000	4212000	2977000
5	2916000	2003500	1636530
6	2916000	1523000	1636530

2. Menghitung Error Prediksi dengan cara menghitung selisih antara realisasi dengan hasil prediksi.

Tabel 3. 5 Tabel Error Prediksi

N O	Realisasi	Hasil Prediks i	Error	(Error) ²
1	3500000	3655640	-155.640	24.223.809.600
2	3500000	3655640	719.360	517.478.809.600
3	2812000	2977000	-165.000	27.225.000.000
4	4212000	2977000	1.235.000	1.525.225.000.000
5	2003500	1636530	366.970	134.666.980.900
6	1523000	1636530	-113.530	12.889.060.900

3. Menghitung *Mean Absolute Percentage Error (MAPE)*

Merupakan rata-rata dari keseluruhan presentase kesalahan (selisih) antaradata actual dengan data hasil peramalan. Ukuran akurasi dicocokkan dengan data time series, dan ditunjukkan dalam presentase.

$$\text{Mape} = \sum \frac{|y-y'|}{y} * 100\%$$

$$\text{Mape} = \sum \frac{2.755.500}{18.425.500} * 100\%$$

$$\text{Mape} = \frac{275.500.000}{18.425.500}$$

$$\text{Mape} = 15$$

Dari hasil pengujian model di dapatkan **RMSE** sebesar =86442828,2621 dan **MAPE** sebesar 15

Evaluasi Model

Untuk menghitung kesalahan (error) dalam melakukan suatu prediksi pada sitem ini, maka penulis menggunakan rumus MAPE (*Mean Absolute presentage Error*)

$$\text{Mape} = \frac{\sum \frac{|y-y'|}{y} x 100\%}{n}$$

Dimana :

y' :hasil prediksi

y :Data aktual

n :Jumlah data

IV.KESIMPULAN DAN SARAN

Kesimpulan

Berikut merupakan kesimpulan yang di dapat dari penelitian ini:

1. Sistem prediksi produksi tanaman pangan dapat digunakan untuk membantu di Dinas Pertanian Provinsi Gorontalo untuk memprediksi produksi tanaman pangan pada tahun selanjutnya.
2. Metode *K-nearest neighbor* dapat digunakan untuk memprediksi tanaman pangan secara tepat dan akurat, aplikasi yang sudah dibangun ini dapat di pakai karena memiliki hasil akurasi tertinggi pada tanaman jagung dengan tingkat akurasi sebesar 92,83%, dan memiliki tingkat error sebesar 7,169%.

Saran

Berikut merupakan saran yang di berikan oleh peneliti:

1. Peneliti selanjutnya dapat mengoptimalkan metode *K-Nearest Neighbor* dengan menambah jumlah data agar mendapatkan hasil yang lebih akurat.
2. Dapat dikembangkan dengan menambahkan variabel untuk prediksi produksi tanman pangan berdasarkan luas panen dengan menggunakan metode *K-nearest neighbor*.

REFERENSI

[1] A. A. Setiamy and E. Deliani, "Peramalan Tingkat Produksi Tanaman Pangan Dan Tanaman Perkebunan Rakyat Kabupaten Bulukumba Menggunakan Metode Exponential Smoothing (Pemulusan Eksponensial)," vol. 2, pp. 5–10, 2019.

- [2] S. Maesaroh, "Sistem Prediksi Produktifitas Pertanian Padi Menggunakan Data Mining," vol. 7, no. 2, pp. 25–30, 2017.
- [3] N. Luh and G. Pivin, "Penerapan Metode K-Nearest Neighbor Untuk Sistem Rekomendasi Pemilihan Mobil," vol. 16, no. 2, pp. 120–131, 2017.
- [4] W. Yustanti, "Algoritma K-Nearest Neighbour untuk Memprediksi Harga Jual Tanah," vol. 9, no. 1, pp. 57–68, 2012.
- [5] J. P. Informatika, L. Azmi, K. Ulfa, P. Replanting, I. Pendahuluan, and A. D. Mining, "PREDIKSI REPLANTING LAHAN KELAPA SAWIT DENGAN MENGGUNAKAN METODE K-NEAREST NEIGHBOR," vol. 18, pp. 517–522, 2019.
- [6] I. Purnamasari and M. A. Mustofa, "Penerapan Algoritma K-Nearest Neighbor (K-NN) Dengan Pencarian Optimal Untuk Prediksi Prestasi Siswa," *Optimasi Pemanfaat. Local Area Netw. Dengan Layer 7 Protoc.*, vol. 3, no. 2, pp. 1–8, 2019.
- [7] H. Fatah, A. Subekti, S. I. Komputer, and I. Komputer, "Prediksi Harga CRYPTOCURRENCY Dengan Metode K-Nearest Neighbor," vol. 14, no. 2, pp. 137–144, 2018.
- [8] P. Algoritma, C. Dan, U. Prediksi, L. Lahan, P. Tanaman, and P. Di, "Perbandingan Algoritma Cart dan K-Nearest Neighbor untuk Prediksi Luas Lahan Panen Tanaman Padi Di Kabupaten Karawang," vol. 9, pp. 74–78, 2018.
- [9] A. Rohman, "Model Algoritma K-Nearest Neighbor (K-NN) Untuk Prediksi Kelulusan Mahasiswa," 2012.
- [10] P. Bidang, K. Sains, Y. Mardi, J. Gajah, M. No, and S. Barat, "Jurnal Edik Informatika Data Mining : Klasifikasi Menggunakan Algoritma C4 . 5."
- [11] I. K. Syahputra, F. A. Bachtiar, and S. A. Wicaksono, "Implementasi Data Mining untuk Prediksi Mahasiswa Pengambil Mata Kuliah dengan Algoritme Naive Bayes," *J. Pengemb. Teknol. Inf. dan Ilmu Komput. Univ. Brawijaya*, vol. 2, no. 11, pp. 5902–5910, 2018.
- [12] F. Muttaqin and M. Al Musadieq, "Analisis dan desain sistem informasi berbasis komputer untuk persediaan barang pada toko bahan bangunan (Studi Kasus pada UD . Sumber Bumi Subur)," vol. 8, no. 1, pp. 1–7.
- [13] N. F. Bintarika, J. S. Informasi, and U. Gunadarma, "Analisis Dan Perancangan Sistem Berorientasi Objek Studi Kasus : Pembuatan SKCK Pada Polsek Cibitung," pp. 1–12, 2009.
- [14] W. N. Cholifah, S. M. Sagita, and S. Knowledge, "Pengujian Black Box Testing pada Aplikasi Action & Strategy Berbasis Android Dengan Teknologi Phonegap," vol. 3, no. 2, pp. 206–210, 2018.
- [15] M. Nuris, "White box testing pada sistem penilaian pembelajaran skripsi," 2015.