

Prediksi Harga Beras Premium dengan Metode Algoritma *K-Nearest Neighbor*

Yuwantoro Mukhlisin¹, Mahmud Imrona², Danang Triantoro Murdiansyah³

^{1,2,3}Fakultas Informatika, Universitas Telkom, Bandung

¹yuwantoro@students.telkomuniversity.ac.id, ²imrona@telkomuniversity.ac.id,

³danangtri@telkomuniversity.ac.id

Abstrak

Pertanian merupakan salah satu sektor yang penting untuk kehidupan manusia, karena sebagian besar kebutuhan manusia dari pertanian, yaitu adalah kebutuhan pangan. Seiring berjalannya waktu, harga pangan seringkali tidak stabil, terutama harga beras, karena beras adalah makanan pokok masyarakat Indonesia. Hal ini tentu saja berpengaruh yang besar bagi masyarakat dan petani. Penyebab dari ketidakstabilan harga beras ini bisa disebabkan oleh beberapa faktor, seperti faktor lingkungan, serangan hama dan wereng, dan lahan kekeringan. Pada tugas akhir ini, akan dibahas mengenai penerapan salah satu metode *data mining* dalam proses prediksi harga beras dengan membandingkan harga beras pada tahun 2014 - 2019 dari BPS Kota Bandung. Dataset yang digunakan berasal dari BPS Kota Bandung dari tahun 2014 hingga 2019 dan BMKG Kota Bandung dengan tahun yang sama. Adapun metode yang digunakan adalah algoritma regresi *K-Nearest Neighbor* (KNN) serta untuk pengujiannya menggunakan *RMSE*. Hasil dari penelitian ini, metode *K-Nearest Neighbor* dengan model regresi dapat melakukan prediksi terhadap harga beras pada tahun 2014 - 2019 dengan nilai *RMSE* 0,125 dan parameter $K = 2$ yang sudah dinormalisasi.

Kata kunci : prediksi harga beras, *data mining*, algoritma regresi *K-Nearest Neighbor* (KNN)

Abstract

Agriculture is one of the important sectors for human life, because most of the human needs of agriculture, namely food needs. Over time, food prices are often unstable, especially the price of rice, because rice is the staple food of Indonesian people. This of course has a big effect on the community and farmers. The cause of this rice price instability can be caused by several factors, such as environmental factors, pest and plant hopper attacks, and drought land. In this observation, will be discussed regarding the application of one data mining method in the process of predicting rice prices by comparing the 2014 - 2019 rice prices from BPS Bandung. The dataset used is from the Bandung City BPS from 2014 to 2019 and BMKG Bandung City in the same year. The method used is the *K-Nearest Neighbor* (KNN) regression algorithm and for testing using *RMSE*. The results of this study, the *K-Nearest Neighbor* method with a regression model can predict rice prices in 2014 - 2019 with an *RMSE* value of 0,125 and parameter $K = 2$ which has been normalized.

Keywords: rice price prediction, data mining, *K-Nearest Neighbor* (KNN) regression algorithm

1. Pendahuluan

Latar Belakang

Tiga isu dunia saat ini sedang marak maraknya di dunia ini, seperti *Food, Water and Energy*. Salah satu dari tiga isu tersebut, *food* (pangan) adalah isu yang paling sering disorot oleh pemerintah. Seperti sektor pertanian merupakan salah satu sektor yang penting dalam kehidupan masyarakat. Indonesia merupakan negara kepulauan yang terbesar di dunia. Total luas daratan di negara ini sekitar 190 juta hektar, yang sekitar 55 juta hektarnya adalah lahan pertanian 129 juta hektarnya adalah perhutanan. Pada lahan pertanian, sekitar 24 juta hektar terdiri dari lahan yang subur [1].

Dengan ketersediaan lahan yang luas, pertanian menjadi salah satu penghasilan ekonomi negara dengan menyediakan lapangan pekerjaan dan sebagai penghasil devisa negara. Pertanian juga sebagai sumber ekonomi utama bagi masyarakat pedesaan [2]. Sebagai salah satu sumber ekonomi masyarakat, penentuan harga yang tepat sudah menjadi hal yang sangat penting. Dan pastinya petani tentu saja mempunyai keinginan untuk mendapat keuntungan dari hasil bertani.

Tetapi akhir akhir ini sering terjadi perubahan harga komoditas pertanian, seperti komoditas padi, yang dimana harga beras tersebut bisa berubah tidak menentu secara signifikan. Hal ini tentu mempengaruhi bagi petani dan masyarakat yang mengkonsumsi beras. Banyak faktor yang mempengaruhi harga beras tersebut, seperti faktor lingkungan, serangan hama dan wereng dan kekeringan di berbagai daerah. Beras merupakan makanan pokok untuk masyarakat Indonesia. Selain faktor lingkungan dan serangan hama, perubahan supply dan demand juga menyebabkan harga beras tidak stabil. Misalnya, permintaan yang tinggi pada beras dan kurangnya pasokan padi sangat berpengaruh pada harga.

Jika pasokan padi sedikit dan permintaan beras banyak, maka harga beras akan mengalami kenaikan, demikian juga untuk sebaliknya. Cuaca yang tepat untuk melakukan penanaman padi, biasanya para petani mengandalkan cuaca basah atau cuaca hujan. Karena apabila kekurangan air atau kering dapat menurunkan kualitas beras begitu juga apabila kelebihan air [3]. Maka dari itu, prediksi harga sangat penting untuk dilakukan. Dengan dilakukannya ini, bisa memudahkan petani untuk mengambil keputusan melakukan penanaman waktu yang tepat.

Dalam penelitian sebelumnya yang dilakukan oleh Mohamad Efendi Lasulika menggunakan algoritma *K-Nearest Neighbor* (KNN) tentang prediksi harga komoditi jagung diperoleh akurasi sebesar 98.7% [4]. Dalam penelitian ini algoritma yang digunakan adalah algoritma KNN berbasis *Particle Swarm Optimazation* (PSO). Implementasi algoritma *K-Nearest Neighbor* (KNN) bertujuan untuk memprediksi harga beras premium tahun 2019 dari BPS Kota Bandung dengan harga beras tahun 2019 dari proses algoritma regresi *K-Nearest Neighbor* (KNN) dan dibandingkan dengan K data tetangga terdekat untuk memperoleh nilai RMSE (*Root Mean Square Error*) yang kecil, karena nilai RMSE semakin kecil semakin akurat [5].

Topik dan Batasannya

Permasalahan yang dibahas dalam penelitian ini adalah bagaimana mengimplementasikan algoritma regresi *K-Nearest Neighbor* (KNN) dalam memprediksi harga beras premium di Kota Bandung.

Adapun batasan masalah dalam penelitian ini adalah data yang digunakan adalah data set bersumber dari BPS Kota Bandung tahun 2014 - 2019 dan untuk informasi cuaca bersumber dari BMKG Kota Bandung dengan tahun yang sama. Data beras yang diprediksi yaitu harga beras tahun 2014 - 2019 dari BPS Kota Bandung dengan harga beras tahun 2014 - 2019 dari proses algoritma regresi *K-Nearest Neighbor* (KNN) dan menghitung RMSE (*Root Mean Square Error*).

Tujuan

Adapun tujuan dari penelitian adalah mengimplementasikan algoritma regresi *K-Nearest Neighbor* (KNN) dalam memprediksi harga beras premium dan menganalisis algoritma regresi *K-Nearest Neighbor* dengan metode lain untuk memprediksi harga beras premium.

Organisasi Tulisan

Penulisan tugas akhir ini tersusun dalam beberapa bagian, yaitu sebagai berikut: bagian pertama berisi latar belakang, batasan masalah, hingga mengenai tujuan penelitian ini. Bagian kedua berisi studi terkait, yang menjelaskan hal - hal yang berkaitan dengan penelitian ini. Kemudian, bagian ketiga berisi sistem yang dibangun, akan menjelaskan rancangan sistem yang dibangun. Pada bagian keempat berisi evaluasi, mengenai hasil pengujian dan evaluasi sistem dan bagian kelima berisi kesimpulan dari penelitian.

2. Studi Terkait

2.1 Beras

Beras adalah suatu jenis tanaman yang dapat dipanen dalam jangka waktu tahunan. Tanaman ini menurutnya memiliki karakteristik setinggi 1 sampai 1,8 m dan daunnya memiliki panjang 50 sampai

100 cm [6]. Ada banyak cara untuk menggolongkan beras, yaitu berdasarkan varietas padi, seperti beras Bramo Super, Rojo Lele, beras Ramos dan lain sebagainya. Ada pula pengelompokan beras yang berdasarkan asal daerahnya, seperti beras Cianjur, beras Garut dan beras Banyuwangi.

2.2 Perum Bulog

Bulog adalah perusahaan umum milik negara yang bergerak di bidang logistik pangan. Ruang lingkup bisnis perusahaan meliputi usaha logistik atau pergudangan, survei dan pemberantasan hama, penyediaan karung plastik, usaha angkutan, perdagangan komoditi pangan dan usaha eceran. Sebagai perusahaan yang tetap mengemban tugas publik dari pemerintah, Bulog tetap melakukan kegiatan menjaga harga dasar pembelian untuk gabah, stabilisasi harga khususnya harga pokok, menyalurkan beras untuk orang miskin (Raskin) dan pengelolaan stok pangan [7].

2.3 BMKG

Badan Meteorologi Klimatologi dan Geofisika (BMKG) merupakan satu-satunya lembaga di Indonesia yang telah ditunjuk secara resmi untuk mengeluarkan informasi mengenai keadaan cuaca dan iklim. Penyebaran informasi cuaca yang cepat dari dan ke Unit Pelaksana Teknis (UPT) BMKG merupakan salah satu hal pokok yang mendukung penyelenggaraan BMKG. Salah satu metoda penyebaran data yang digunakan BMKG adalah menggunakan sistem multicast yang memberikan solusi alternatif model penyebaran informasi cuaca dan diseminasi cuaca buruk. Sistem multicast menawarkan pilihan untuk dapat mengirimkan file dari satu server ke banyak client secara bersamaan [8].

2.4 Prediksi

Prediksi adalah suatu proses memperkirakan secara sistematis tentang sesuatu yang paling mungkin terjadi di masa depan berdasarkan informasi masa lalu dan sekarang yang dimiliki, agar kesalahannya (selisih antara sesuatu yang terjadi dengan hasil perkiraan) dapat diperkecil. Prediksi tidak harus memberikan jawaban secara pasti kejadian yang akan terjadi, melainkan berusaha untuk mencari jawaban sedekat mungkin yang akan terjadi [9]. Pengertian Prediksi sama dengan ramalan atau perkiraan. Menurut kamus besar bahasa Indonesia, prediksi adalah hasil dari kegiatan memprediksi atau meramal atau memperkirakan nilai pada masa yang akan datang dengan menggunakan data masa lalu. Prediksi menunjukkan apa yang akan terjadi pada suatu keadaan tertentu dan merupakan input bagi proses perencanaan dan pengambilan keputusan.

2.5. Machine Learning

Machine learning adalah kecerdasan buatan (Artificial Intelligence) yang mempelajari bagaimana cara mengembangkan suatu sistem, membuat data dan membuat algoritma – algoritma yang bisa melakukan tugas – tugasnya dengan sendiri tanpa ada arahan dari penggunaannya dan memungkinkan programmer bisa belajar [10].

2.6. Algoritma K-Nearest Neighbor

Algoritma KNN Regressor adalah algoritma dengan mengelompokkan data berdasarkan letak ketetanggaannya, pengelompokan data tersebut bergantung pada jumlah nilai k. Nilai RMSE dihasilkan untuk mengukur tingkat error dari model yang sudah dibuat.

Untuk lebih jelasnya seperti berikut : [11]

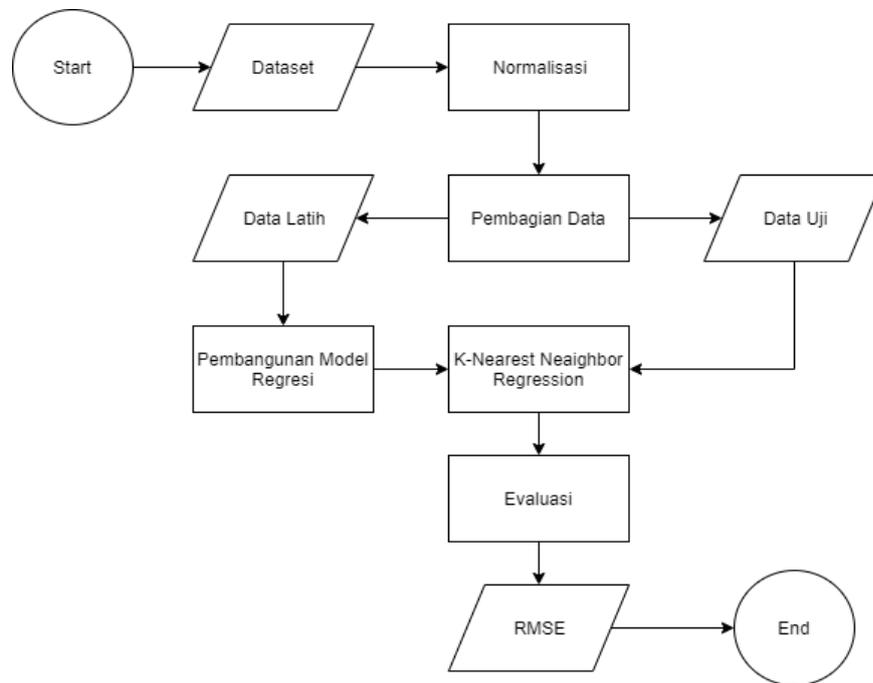
1. Menentukan K itu sebagai bilangan bulat positif berdasarkan ketersediaan data pembelajaran.
2. Memilih tetangga terdekat dari data baru sebanyak K .
3. Menghasilkan nilai prediksi dari data sampel baru.
4. Menentukan RMSE yang bergantung pada pada nilai sebenarnya dengan nilai prediksi

2.7. Data Mining

Data mining sebagai proses untuk mendapatkan informasi yang berguna dari gudang basis data yang besar. Salah satu teknik yang dibuat dalam data mining adalah bagaimana menelusuri data yang ada untuk membangun sebuah model, kemudian menggunakan model tersebut agar dapat mengenali pola data yang lain yang tidak berada dalam basis data yang tersimpan [12].

3. Sistem yang Dibangun

Proses untuk memprediksi harga beras mulai dari pengambilan data, *preprocessing*, proses, dan evaluasi terhadap hasil. Diagram alur sistem dapat dilihat pada gambar 1.



Gambar 1. Alur sistem program

3.1 Dataset

Sumber data harga beras premium pada penelitian ini diperoleh dari data set BPS (Badan Pusat Statistik) Kota Bandung dan data cuaca Kota Bandung diperoleh dari data set BMKG Kota Bandung. Data yang diperoleh berupa data harga rata - rata setiap bulannya, data harga beras premium dari tahun 2014 - 2019 dengan jumlah *record* 72. *Preprocessing* data yang dilakukan pada penelitian ini menggunakan normalisasi data. Normalisasi data ini digunakan untuk mengelompokkan data ke dalam skala atau jangkauan tertentu sehingga mempermudah dalam mengolah data. Data yang ada masih berupa data univariate, sehingga perlu diubah lagi menjadi *multivariate*, sehingga menjadi beberapa *variabel independent*. Pada tahap evaluasi ini, nilai yang dihasilkan dari pengujian berupa nilai *Root Mean Square Error* (RMSE). Nilai RMSE yang kecil menunjukkan hasil akurasi prediksi yang terbaik. Rumus RMSE dipresentasikan dengan persamaan :

3.2 Normalisasi

Pada tahap ini normalisasi data diperlukan untuk membuat data yang nilainya menyebar bisa dinormalkan dengan normalisasi dengan cara membuat rentang nilainya menjadi lebih kecil. Normalisasi yang dilakukan menggunakan min-max normalization yang membuat rentang 0-1. Persamaan min-max normalization dapat dilihat sebagai berikut :

$$X^* = \frac{X - \min(X)}{\max(x) - \min(X)} \quad (1)$$

Dimana X adalah rata-rata dari suatu data, $\min(X)$ adalah nilai minimal dari data X, dan $\max(X)$ adalah nilai maksimal dari data X.

3.3 Pembagian Data

Memprediksi suatu model yang dibuat harus terdapat dataset yang akan dipecah menjadi dua bagian yaitu data latih dan data uji. Dataset terdiri dari empat parameter yaitu bulan, suhu rata-rata, LPM sebagai variabel x dan label harga beras sebagai parameter y. Data latih digunakan untuk membuat model regresi yang diujikan terhadap data uji, data latih yang alokasikan kurun waktu tahun 2014 - 2019 dengan *record* sebanyak 12 data untuk masing-masing tahun, dan penulis mengambil data uji pada tahun 2014-2019.

3.4 Pembangunan Model Regresi KNN

Regresi KNN menggunakan jarak kedekatan antar tetangga yang direpresentasikan nilainya dengan variabel “k”. Jarak kedekatan antar tetangga dihitung menggunakan persamaan *euclidean distance* sebagai berikut.

$$dist((x_1, y_1, \dots), (x_2, y_2, \dots)) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + \dots} \quad (2)$$

Dimana, x_1 dan x_2 merupakan suatu titik yang ditinjau dari sumbu x dan y. y_1 dan y_2 merupakan titik acuan yang ditinjau dari sumbu x dan y. X_1 merupakan parameter dataset dari bulan, x_2 merupakan parameter dataset dari suhu rata-rata, y_1 merupakan parameter dataset dari jumlah curah hujan dan y_2 merupakan parameter dataset dari LPM.

Setelah titik-titik dibandingkan, pilih jarak yang terkecil, ambil sebanyak “k” sehingga mendapatkan nilai jarak terkecil sebanyak “k”, setelah itu hitung rata-rata nilai *euclidean distance*. Misalkan *euclidean distance* titik A ke titik B,C,D,E,F masing-masing berjarak 5 cm, 8 cm, 9 cm, 4 cm, dan 7 cm dengan parameter $k = 3$ maka ambil jarak yang terkecil sebanyak 3 nilai yaitu titik B,E,F dengan nilai 5 cm, 4 cm, dan 6 cm, setelah itu hitung nilai rata-rata dari ketiga titik tersebut yaitu jumlahkan ketiga titik kemudian dibagi banyaknya titik. Perhitungan tersebut menghasilkan nilai 5 dari perhitungan $(5+4+6)/3$.

3.5 Evaluasi

Evaluasi merupakan penilaian dari model regresi yang sudah dibuat. Metode regresi dikatakan baik jika memiliki *Root Mean Square Error* (RMSE) yang artinya hasil prediksi regresi yang dibuat oleh model regresi lebih mendekati nilai sebenarnya. Cara menghitung RMSE dapat dilihat pada persamaan.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \quad (3)$$

Dimana n adalah banyaknya data test, \hat{y} merupakan nilai prediksi harga beras dari model yang sudah dibuat, sedangkan y adalah nilai harga beras sebenarnya.

4. Evaluasi

Pengujian yang dilakukan penulis di penelitian ini dilakukan 2 skenario untuk menemukan RMSE paling rendah untuk menghasilkan model yang baik menggunakan KNN *regression* dan membandingkan RMSE KNN *regression* dengan RMSE metode regresi lainnya.

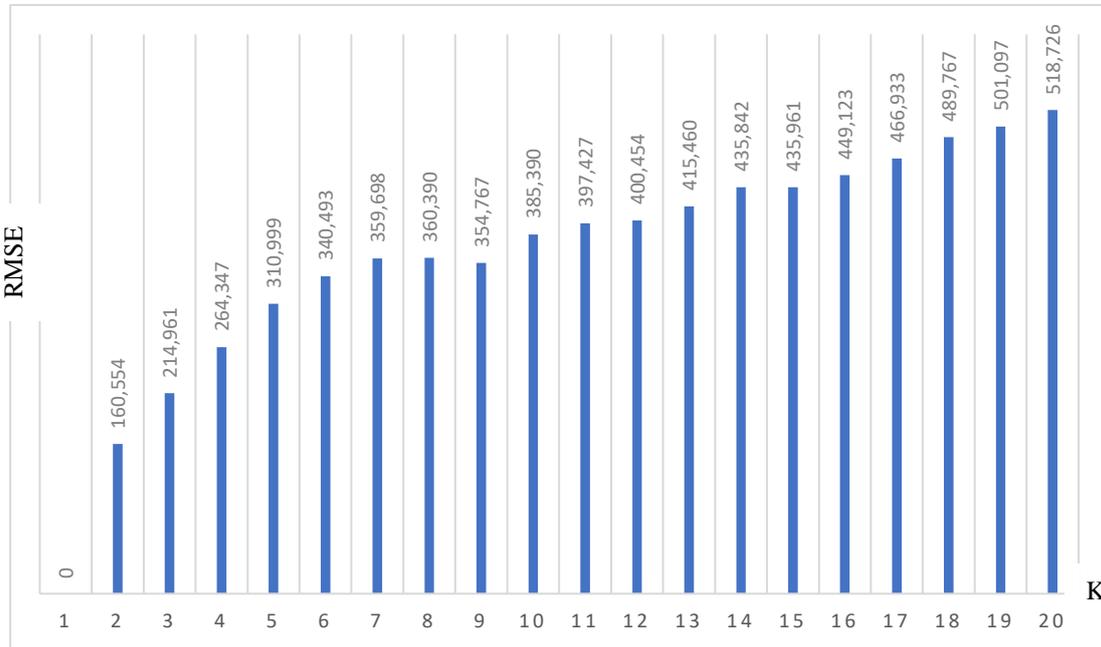
4.1 Hasil Pengujian dan Analisis Skenario Pertama

Skenario pertama bertujuan untuk menentukan nilai k yang paling optimal yang digunakan untuk memprediksi harga beras dengan RMSE yang rendah sehingga model yang dihasilkan menjadi baik. Penulis mendapatkan nilai $k = 2$ merupakan nilai terbaik untuk data yang non-normalisasi. dari rentang 1-20. Nilai $k = 2$ mendapatkan nilai RMSE 352,450 yang paling rendah daripada yang lainnya. Berikut adalah tabel nilai k non-normalisasi dan hasil RMSE.



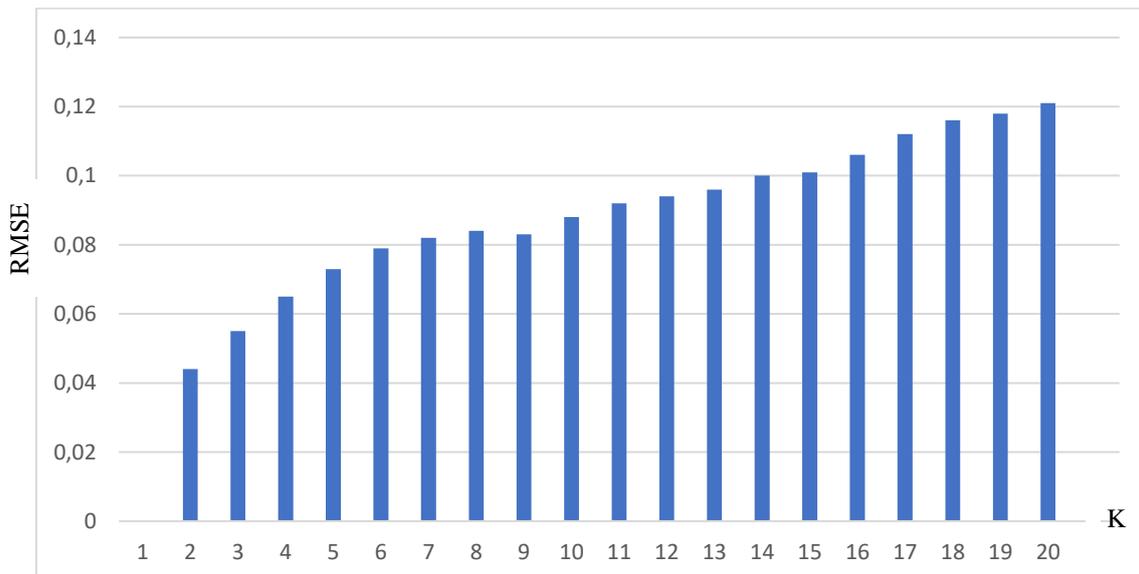
Gambar 2. RMSE setiap nilai k non-normalisasi

Penulis mendapatkan nilai k = 2 merupakan nilai terbaik untuk data denormalisasi. dari rentang 1-20. Nilai k = 2 mendapatkan nilai RMSE 160,554 yang paling rendah daripada yang lainnya. Berikut adalah tabel nilai k denormalisasi dan hasil RMSE.



Gambar 3. RMSE setiap nilai k ketika denormalisasi

Penulis mendapatkan nilai k = 2 merupakan nilai terbaik untuk data denormalisasi. dari rentang 1-20. Nilai k = 2 mendapatkan nilai RMSE 0,044 yang paling rendah daripada yang lainnya. Berikut adalah tabel nilai k denormalisasi dan hasil RMSE.

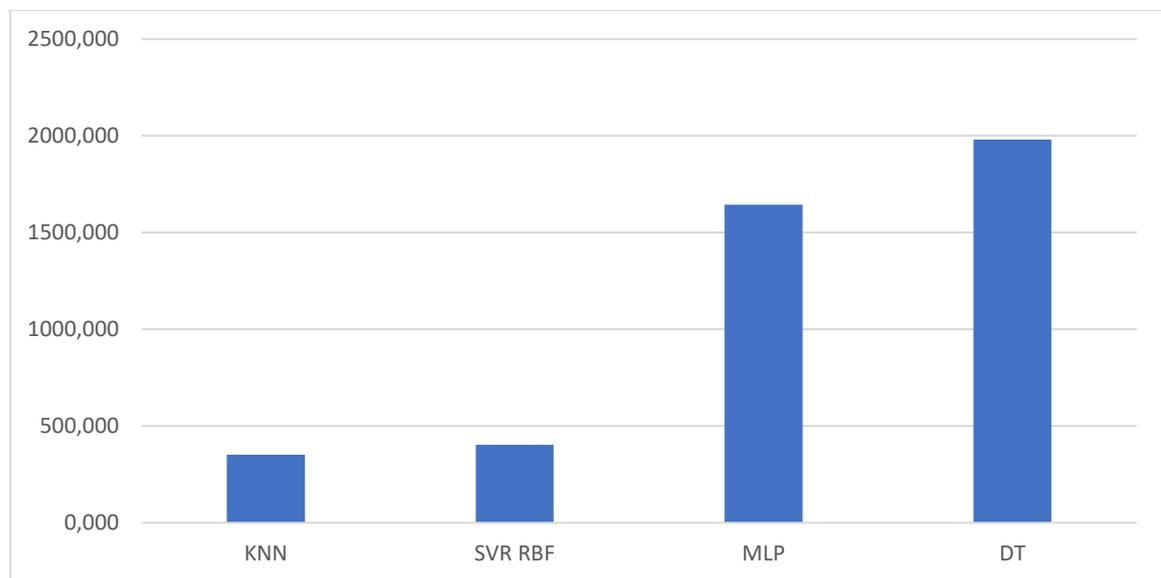


Gambar 4. RMSE setiap nilai k ketika denormalisasi

Nilai $k = 2$, artinya setiap titik mencari titik tetangga terdekatnya sebanyak dua titik terdekat, kemudian menjumlahkan hasil dari kedua tetangga tersebut yang kemudian dibagi dua, didapatlah nilai rata-rata dari $k = 2$. ini membuktikan bahwa jika data dikelompokkan menjadi 2 bagian akan menghasilkan nilai RMSE yang lebih baik. Metode yang digunakan adalah metode *brute force* untuk menelusuri setiap nilai k sehingga menghasilkan hasil RMSE yang paling baik, nilai k optimal akan berubah-ubah bergantung pada data yang ada. Hasil normalisasi data menghasilkan nilai RMSE lebih baik daripada data tanpa normalisasi, normalisasi data berpengaruh penting dalam penentuan nilai k yang optimal, karena KNN sangat berpengaruh dengan data *outlier* artinya data dengan range yang tidak wajar menjauhi pengelompokan data, dengan adanya normalisasi, data dikelompokkan dalam rentang 0-1 sehingga data *outlier* mudah dikelompokkan karena rentang nilainya tidak terlampau jauh.

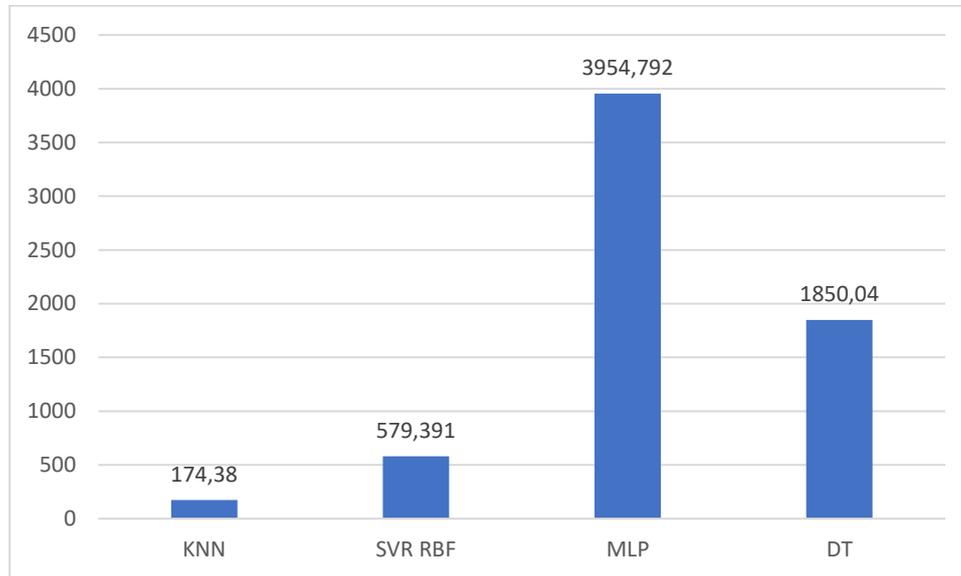
4.2 Hasil Pengujian dan Analisis Hasil Skenario Kedua

Skenario kedua bertujuan untuk membandingkan RMSE metode regresi KNN dengan metode SVR (Support Vector Regression), MLP (Multi Layer Perceptron), Decision Tree Regressor. Penulis mendapatkan hasil regresi menggunakan metode KNN Regressor menghasilkan RMSE yang paling rendah untuk data yang non-normalisasi, yaitu 352,450 dan diikuti oleh metode SVR RBF menghasilkan RMSE sebesar 402,787. Decision Tree Regressor sebesar 1980,05 dan yang paling besar RMSE terdapat pada MLP sebesar 1644,192.



Gambar 5. Perbandingan RMSE KNN Regressor non-normalisasi dengan metode lain

Selanjutnya, bertujuan untuk membandingkan RMSE metode regresi KNN dengan metode SVR (Support Vector Regression), MLP (Multi Layer Perceptron), Decision Tree Regressor. Penulis mendapatkan hasil regresi menggunakan metode KNN Regressor menghasilkan RMSE yang paling rendah untuk data yang denormalisasi, yaitu 174,38, kemudian diikuti oleh metode SVR RBF menghasilkan RMSE sebesar 579,391. Decision Tree Regressor sebesar 1850,04 dan yang paling besar RMSE terdapat pada MLP sebesar 3954,792.



Gambar 6. Perbandingan RMSE KNN Regressor denormalisasi dengan metode lain

Diagram tersebut menunjukkan bahwa metode KNN Regressor lebih unggul daripada metode lainnya karena KNN flexible akan sebaran data yang linier maupun non linier dengan membagi sebaran data berdasarkan kelompok-kelompoknya sehingga jika mendapatkan nilai k yang optimal tersebut dapat mengurangi nilai RMSE, model yang dihasilkan KNN tidak kaku artinya modelnya tidak menggunakan nilai persamaan garis yang optimal untuk menghitung nilai RMSE setiap titik terhadap persamaan garis yang sudah dimodelkan. SVR RBF nilainya hampir mendekati KNN Regressor karena sifatnya sama dengan KNN yang *flexible* dengan cara membuat garis pemisah antar kelompoknya, garis tersebut akan digunakan sebagai patokan penentuan RMSE. Akan tetapi untuk metode MLP menghasilkan nilai RMSE yang tinggi karena MLP merupakan Deep Learning yang datanya tidak cocok untuk data yang sedikit karena proses belajarnya yang lama bergantung pada *learning rate* yang dideklarasikan, jika data sedikit maka MLP susah untuk belajar karena informasi yang didapatkan kurang. Metode selanjutnya yaitu DT Regressor menghasilkan RMSE paling buruk karena DT Regressor susah untuk membuat prediksi nilai RMSE yang bentuk nilainya kontinyu berdasarkan pohon keputusan yang dibangun.

5. Kesimpulan

Berdasarkan pengujian diatas, penulis dapat menarik beberapa kesimpulan.

1. Pada skenario pertama, diperoleh kesimpulan nilai k yang paling optimal dari rentang nilai $k = 1$ sampai 20 menghasilkan nilai RMSE 352,450 pada nilai $k = 2$ yang merupakan nilai terbaik dan nilai terendah dari yang lainnya untuk data yang non-normalisasi. Untuk data yang denormalisasi, menghasilkan RMSE sebesar 160,554 pada nilai $k = 2$ dari rentang nilai $k = 1$ sampai 20. Sedangkan untuk data yang normalisasi, menghasilkan RMSE sebesar 0,044 pada nilai $k = 2$ yang merupakan nilai terbaik dan terendah.

2. Pada skenario kedua, diperoleh kesimpulan nilai K yang paling rendah, menghasilkan nilai non-normalisasi RMSE sebesar 352,040 pada metode KNN Regressor. Kemudian, diikuti metode SVR RBF menghasilkan RMSE sebesar 402,787. Decision Tree Regressor sebesar 1980,05 dan yang paling besar RMSE terdapat pada MLP sebesar 1644,192. Sedangkan untuk metode yang denormalisasi, menghasilkan nilai RMSE sebesar 174,38 pada metode KNN Regressor, kemudian diikuti metode SVR RBF menghasilkan nilai RMSE sebesar 579,391. Decision Tree Regressor sebesar 1850,04 dan yang paling besar nilai RMSE terdapat pada metode MLP sebesar 3954,792.

3. Metode KNN Regressor lebih unggul daripada metode lainnya, karena algoritma KNN *flexible* akan sebaran data yang linier maupun non linier dengan membagi sebaran data berdasarkan kelompok-kelompoknya sehingga jika mendapatkan nilai k yang optimal tersebut dapat mengurangi nilai RMSE, model yang dihasilkan KNN tidak kaku, artinya modelnya tidak menggunakan nilai persamaan garis yang optimal untuk menghitung nilai RMSE setiap titik terhadap persamaan garis yang sudah dimodelkan.

4. Ada metode yang nilai RMSE-nya mendekati dengan metode KNN Regressor, yaitu metode SVR RBF. Metode SVR RBF nilainya hampir mendekati KNN Regressor, karena sifatnya sama dengan KNN yang *flexible* dengan cara membuat garis pemisah antar kelompoknya, garis tersebut akan digunakan sebagai patokan penentuan RMSE. Akan tetapi untuk metode MLP menghasilkan nilai RMSE yang tinggi karena MLP merupakan Deep Learning yang datanya tidak cocok untuk data yang sedikit karena proses belajarnya yang lama bergantung pada *learning rate* yang dideklarasikan, jika data sedikit maka MLP susah untuk belajar karena informasi yang didapatkan kurang. Metode selanjutnya yaitu DT Regressor menghasilkan RMSE paling buruk karena DT Regressor susah untuk membuat prediksi nilai RMSE yang bentuk nilainya kontinyu berdasarkan pohon keputusan yang dibangun

Daftar Pustaka

- [1] E. Quincieu, "Asian Development Bank Papers on Indonesia," *Summary of Indonesia's Agriculture, Natural Resources, and Environment Sector Assesment*, 2015.
- [2] K. Pertanian, "Rencana Strategis Kementerian Pertanian tahun 2015 - 2019," *Kementerian Pertanian Republik Indonesia*, 2015.
- [3] P. H. S. Indonesia, "Tehnik dan Budidaya Penanaman Padi System of Rice Intensification (SRI)," Pusat Pelatihan Kewirausahaan Sampoerna , Pasuruan, 2018.
- [4] M. E. Lasulika, "Prediksi Harga Komoditi Jagung Menggunakan K-NN dan Particle Swarm Optimazation sebagai Fitur Seleksi," *ILKOM Jurnal Ilmiah*, vol. 9, no. 3, p. 233, 2017.
- [5] W. W. C. Sri, "Komparasi Beberapa Metode Estimasi Kesalahan Pengukuran," *Jurnal Penelitian dan Evaluasi Pendidikan*, vol. 13, no. 2, pp. 182-197, 2013.
- [6] I. Student, "4 Pengertian Beras Menurut Para Ahli," [Online]. Available: <https://www.indonesiastudents.com/pengertian-beras-menurut-para-ahli/>.
- [7] P. Bulog, "Sekilas Perum Bulog," [Online]. Available: <http://www.bulog.co.id/sekilas.php>. [Diakses January 2020].
- [8] Y. N. N. Bikan, "Kajian Implementasi Sistem Multicast dalam Sistem Jaringan BMKG," 2015.
- [9] Herdianto, "Prediksi Kerusakan Motor Induksi Menggunakan Metode Jaringan Saraf Tiruan Backpropagation," dalam *Tesis*, Universitas Sumatera Utara, Medan, 2013.
- [10] U. D. Arni, "Definisi dan Konsep Utama Machine Learning," Kamis Agustus 2018. [Online]. Available: <https://garudacyber.co.id/artikel/915-definisi-dan-konsep-utama-machine-learning>.
- [11] M. B. S. Imandoust, "Application of K-Nearest Neighbor (KNN) Approach for Predicting Economic Events : Theoretical Background," *Int. Journal of Engineering Research and Applications*, vol. 3, no. 5, pp. 605-610, 2013.
- [12] E. Prasetyo, *Data Mining Konsep dan Aplikasi Menggunakan Matlab*, Yogyakarta: CV. Andi Offset, 2012.