

Prediksi Harga Jual Beras Eceran Menggunakan Metode *K-Nearest Neighbor* Di Kabupaten Manokwari

Feny Rosalina Pare^{1*}, Lion Ferdinand Marini², Hermansa³
 Teknik informatika, Universitas Papua, Papua Barat, Indonesia.
 *e-mail *Corresponding Author*: haryanto@dinamika.ac.id

Abstrack

The retail price of rice sold in markets in Manokwari district fluctuates every year. This can make it difficult for traders to determine the ideal retail selling price for rice to the people of Manokwari, because some people in Manokwari have a mediocre income. This research aims to determine the prediction of retail rice selling prices based on year. The method used is the classification method with the K-Nearest Neighbor algorithm. Testing was carried out using the Rapid Miner Studio tools. The output of the Root Mean Square Error (RMSE) test results is 981,735, while for testing the Mean Absolute Error (MAE) is used with an accuracy value of 627,074. Even though the MAE accuracy value is smaller than RMSE, the error value varies. Based on the results of research that has been carried out, a Prediction Average value of 13438.362 is obtained, which shows that the average rice price predicted by the model is slightly higher than the actual one, and a Squared Correlation (R-squared) value of 0.393 which explains that there is around 39% variability Rice prices can be explained by the model used in the research.

Keywords: Prediction; Rice Prices; K-Nearest Neighbor Algorithm; Rapid Miner Studio.

Keywords: Algorithm, Rice, Price, K-NN, Prediction

Abstrak

Harga beras eceran yang dijual di pasar-pasar yang ada di kabupaten Manokwari mengalami fluktuasi setiap tahunnya. Hal ini dapat membuat para pedagang sulit untuk menentukan harga jual beras eceran yang ideal kepada para masyarakat manokwari, karena bagi sebagian masyarakat manokwari memiliki pendapatan yang pas-pasan. Penelitian ini bertujuan untuk memprediksi harga jual beras eceran berdasarkan tahun. Metode yang digunakan adalah metode klasifikasi dengan algoritma *K-Nearest Neighbor*. Pengujian dilakukan dengan menggunakan *tools Rapid Miner Studio*. Output hasil pengujian *Root Mean Square Error* (RMSE) sebesar 981.735, sedangkan untuk pengujian menggunakan *Mean Absolute Error* (MAE) dengan nilai akurasi sebesar 627.074. Meskipun nilai akurasi MAE lebih kecil dari RMSE, namun nilai kesalahan bervariasi. Berdasarkan hasil penelitian yang telah dilakukan, diperoleh nilai *Prediction Average* sebesar 13438.362 yang menunjukkan bahwa harga beras rata-rata yang diprediksi oleh model sedikit lebih tinggi dari yang sebenarnya, dan nilai *Squared Correlation (R-squared)* sebesar 0.393 yang menjelaskan bahwa sekitar 39 % variabilitas harga beras dapat dijelaskan oleh model yang digunakan pada penelitian.

Kata kunci: *Prediksi; Harga Beras; Algoritma K-Nearest Neighbor; Rapid Miner Studio*

1. Pendahuluan

Kabupaten Manokwari merupakan salah satu kabupaten yang terletak di Provinsi Papua Barat, Indonesia. Di Kabupaten Manokwari sendiri terdapat dua pasar sentral yaitu pasar sanggeng dan pasar wosi. Pasar sanggeng sendiri adalah pasar tertua yang ada di Kabupaten Manokwari yang berada di Kelurahan Sanggeng, Distrik Manokwari Barat, Kabupaten Manokwari, yang beralamat di Jl. Yos Sudarso dan pada tahun 2018 pasar ini sempat terbakar. Selain pasar sanggeng ada juga pasar wosi yang beralamat di Jl. Pasir, Kelurahan Wosi, Kecamatan Manokwari Barat, Kabupaten Manokwari, Provinsi Papua Barat. Namun luasan pasar wosi ini masih lebih luas pasar sanggeng. Kedua pasar ini juga menjual berbagai macam kebutuhan masyarakat Manokwari seperti pakaian, alat elektronik, perhiasan, alat makan serta menjual berbagai macam bahan makanan pokok, salah satu bahan makanan pokoknya yaitu beras.

Beras merupakan salah satu bahan makanan pokok yang sangat penting bagi sebagian masyarakat Indonesia, khususnya masyarakat Manokwari, karena beras sendiri merupakan salah satu kebutuhan primer yang hampir setiap hari dibutuhkan dalam kehidupan sehari-hari untuk memperoleh asupan gizi yang baik, walaupun harga beras naik ataupun tidak naik, tetap masyarakat akan membelinya dan juga beras sendiri memiliki nilai yang strategis baik dari aspek ekonomi, lingkungan hidup, sosial dan politik [1].

Beras yang dijual pada pasar sanggeng dan pasar wosi yang ada di Kabupaten Manokwari biasa dijual eceran atau dalam per kilo (Kg) yang memiliki harga yang berbeda-beda setiap tahunnya tergantung dari merek dan kualitas berasnya. Karena harga jual beras sendiri juga bisa menjadi hal penting dalam meningkatkan produksi beras [2]. Dan juga jika harga beras yang naik setiap tahunnya akan mempengaruhi kehidupan masyarakat khususnya Manokwari, terutama bagi mereka yang memiliki penghasilan yang terbatas dan akan berdampak bagi tubuh mereka karena kurang asupan gizi.

Untuk itu diperlukannya cara yang dapat membantu dalam memprediksi perkembangan harga jual beras eceran per kilo (Kg) menggunakan metode KNN di Kabupaten Manokwari yang akan menjadi salah satu dasar bagi pemerintah dalam mengambil keputusan yang baik dalam mencegah terjadinya inflasi yang akan berdampak pada stabilitas ekonomi di Manokwari untuk masa yang akan datang dan juga buat para pedagang harus mempunyai perhitungan yang baik dalam menentukan harga jual beras eceran per kilo (Kg) kepada para konsumen dalam memperoleh keuntungan dimasa yang akan datang [3].

KNN merupakan metode yang mengklasifikasikan suatu objek berdasarkan atribut dari contoh data yang akan di uji [4]. Oleh karena itu, penelitian ini dilakukan agar para pedagang dan juga pemerintah dapat mengambil keputusan dalam penentuan harga jual beras eceran kepada masyarakat dimasa mendatang.

2. Tinjauan Pustaka

Pada bagian ini merupakan kumpulan hasil penelitian dari peneliti terdahulu yang melakukan penelitian berkaitan dengan teknik data mining menggunakan metode *K-Nearest Neighbor (KNN)* dalam memprediksi sebuah data. Berikut penelitian dengan menggunakan metode *KNN* yang sudah pernah dilakukan sebelumnya.

Hasil penelitian yang dilakukan oleh (Mukhlisin et al., 2019) dengan judul "Prediksi Harga Beras Premium Dengan Metode *K-Nearest Neighbor*" Pada penelitian ini menggunakan data dari Badan Pusat Statistik (BPS) Bandung dari tahun 2014 sampai 2019 dan juga data BMKG dengan tahun yang sama. Penelitian ini melakukan pengujian menggunakan RMSE dan hasil dari penelitian memprediksi harga beras premium dari tahun 2014 sampai 2019 menggunakan metode KNN mendapatkan nilai RMSE 0,125 dan nilai parameter $K=2$ yang sudah di normalisasikan yang artinya metode KNN dengan pengujian RMSE menghasilkan nilai yang baik dalam memprediksi harga beras premium.

Lalu ada juga hasil penelitian dari (Rismala et al., 2023) dengan judul "Penerapan Metode *K-Nearest Neighbor* Untuk Prediksi Harga Jagung Dengan Pengujian RMSE" Tujuan dari penelitian ini yaitu untuk memprediksi harga jagung menggunakan metode KNN di kabupaten tulungagung yang menghasilkan harga prediksi pada harga jagung bulan mei dengan pengujian menggunakan metode RMSE untuk mengetahui tingkat akurasi, dimana hasil pengujian menggunakan metode ini menghasilkan tingkat *error* 185,1497 dan persentase sebesar 95%.

Selanjutnya hasil penelitian dari (Ilmananda & Alfianty, 2022), dengan judul "Prediksi Rata-Rata Harga Beras Nasional Dengan Kualitas Premium di Tingkat Penggilingan Hingga Akhir Tahun 2022" Data yang digunakan pada penelitian ini yaitu data dari tahun 2018 sampai tahun 2022 yang bersumber dari Badan Pusat Statistik (BPS) Indonesia. Penelitian ini dilakukan dengan pendekatan kuantitatif dengan metode *double exponential smoothing* dan ukuran kesalahan dalam penelitian tersebut menggunakan *mean absolute percentage error (mape)* dengan di peroleh hasil peramalan terbaik dengan nilai mape sebesar 1,00097% dan di peroleh hasil prediksi dengan rata-rata harga beras pada bulan maret hingga desember tahun 2022 berkisar antara Rp. 9.836 sampai dengan Rp. 9.919. Peralaman ini dilakukan untuk menunjukkan bahwa rata-rata harga beras pada tahun 2022 cenderung stabil dan hanya mengalami kenaikan secara bertahap. Pada ketiga penelitian sebelumnya memprediksi harga beras menggunakan metode KNN, memiliki beberapa perbedaan dengan penelitian saat ini. Pada penelitian yang dilakukan oleh Mukhlisin et. al 2019 [5] memiliki perbedaan penggunaan

metode KNN yang berbasis *Particle Swarm Optimization* (PSO), sedangkan pada penelitian ini menggunakan metode KNN dengan melakukan perhitungan jarak.

Pada penelitian [6] memiliki perbedaan yaitu objek yang digunakan pada penelitian tersebut adalah harga jagung, sedangkan pada penelitian ini menggunakan harga beras medium dan premium. Lalu pada penelitian [7] memiliki perbedaan pada metode untuk menguji tingkat kesalahannya menggunakan *Mean Absolute Percentage Error* (MAPE), sedangkan pada penelitian ini menggunakan pengujian RMSE.

3. Metodologi

Penelitian ini dilakukan di manokwari. Alat dan bahan pada penelitian ini menggunakan laptop acer, Ms. Excel dan Rapidminer Studio. Pada penelitian ini juga menggunakan teknik pengumpulan data yaitu studi dokumen, yang dimana datanya diambil pada website SP2KP. Metode Analisis yang dilakukan pada penelitian ini yaitu menggunakan tahapan *Knowledge Discovery in Database* (KDD) dalam memprediksi harga beras. Adapun tahapan analisis datanya, sebagai berikut:



Gambar 1 Alur Analisa Data

Studi Pustaka Pada tahap ini dimana mencari dan mempelajari berbagai sumber referensi baik dari jurnal, buku digital maupun sumber lainnya yang dianggap dapat membantu dalam penelitian ini yang berkaitan dengan *data mining*, prediksi dalam menggunakan algoritma KNN. Pengumpulan Data Mengumpulkan data pada penelitian ini yaitu data harga beras eceran yang didapat dari *website* Sistem Pemantauan dan Kebutuhan Pokok Kementerian Perdagangan (SP2KP) dari tahun 2018 sampai 2023. Seleksi Data Pada proses ini data yang dikumpulkan dari *website* SP2KP dari tahun 2018 sampai 2023 akan diseleksi sebelum masuk ke dalam proses data mining. Pembersihan Data Selanjutnya proses *data cleaning*, dimana data yang diperoleh dari data yang telah di seleksi akan di bersihkan apabila terdapat duplikat data dan akan di lakukan perbaikan data apabila terdapat kesalahan pada data. Transformasi Data Pada proses ini dimana akan mengubah bentuk data yang sesuai dalam proses data mining dengan cara melakukan normalisasi atau memilih atribut-atribut dari data sebelumnya. *Data Mining* Proses data mining yang dilakukan dalam memprediksi menggunakan algoritma KNN untuk memberikan informasi mengenai harga beras eceran yang ada di Kabupaten Manokwari. Proses *data mining* juga dapat di jelaskan secara aplikatif dengan *software* RapidMiner. *Interpretation/ Evaluation* Pada proses akan diperolehnya pola atau informasi yang dihasilkan dari proses data mining. Informasi yang akan dihasilkan oleh *software* RapidMiner akan menampilkan dan memberikan informasi mengenai harga beras eceran yang ada di kabupaten

manokwari. Kesimpulan Hasil Prediksi Dari hasil prediksi harga beras ini pemerintah kabupaten manokwari nantinya dapat mengambil kebijakan yang baik kepada pedagang dan masyarakat dalam menetapkan harga beras eceran untuk masa yang akan datang. Menurut W. Yustanti 2012 dalam *K-Nearest Neighbor* (KNN) merupakan salah satu algoritma klasifikasi yang bertujuan untuk mengklasifikasikan objek baru berdasarkan atribut atau ciri dari sampel data yang sudah ada pada system. Ada beberapa langkah-langkah perhitungan dalam algoritma KNN sebagai berikut [4].

- 1) Menentukan nilai k atau tetangga terdekat pada data yang akan di uji.
- 2) Menghitung jarak *Eucliden Distance* pada masing-masing objek terhadap data yang ada.
- 3) Kemudian urutkan objek-objek tersebut dari yang terkecil sampai terbesar.
- 4) Mengelompokkan data sejumlah K yang telah di tentukan sebelumnya.
- 5) Lalu memilih kategori yang paling mayoritas atau paling banyak muncul.

4. Hasil dan Pembahasan

Tahapan dalam penelitian ini menggunakan tahapan *Knowledge Discovery In Database* (KDD) seperti yang telah dijelaskan pada bab sebelumnya. Semua tahapan yang dilakukan pada penelitian ini menggunakan *Software Ms. Excel* dan *RapidMiner Studio*.

4.1. Data Selection

Data yang didapatkan dari website Sistem Pemantauan dan Kebutuhan Pasar Kementerian Perdagangan (SP2KP) di Kabupaten Manokwari adalah data harga beras eceran per kilo (Kg) yang didapatkan per harinya dari tahun 2018 sampai 2023. Berikut data awalnya dapat dilihat pada Gambar 2.

Gambar 2. Data Awal

Data awal yang diperoleh pada Gambar 2 merupakan dataset yang diambil langsung dari website (SP2KP) Kabupaten Manokwari yang disusun kedalam excel dengan atribut tahun, bulan, tanggal, satuan, dan jenis beras, dimana jenis beras ini yaitu beras mendium dan premium.

Proses pertama dilakukan pada pemilihan data yaitu memasukkan operator "*read excel*" yang berfungsi untuk membaca data dari file excel peneliti yang telah dikumpulkan dalam bentuk excel yaitu data harga beras mendium dan premium [8]. Selanjutnya maka akan dilakukan pengimporan data dengan tahapan sebagai berikut. Adapun proses *select atribut* pada tahapan ini ditunjukkan pada bagian "*Read Excel*" pada gambar di bawah ini.

4.2. Cleaning Data

Data yang diperoleh dari website Sistem Pemantauan dan Kebutuhan Pasar Kementerian Perdagangan (SP2KP) Kabupaten Manokwari, tidak terdapat duplikasi atau data yang kosong, sehingga tidak perlu dilakukannya pembersihan data. Adapun proses *cleaning* data ditunjukkan pada keterangan di bawah ini [9].

```
1. 1. # Mengimpor library yang diperlukan untuk manipulasi data
2. 2. import pandas as pd
3. 3.
4. 4. # Memuat file Excel ke dalam DataFrame dari jalur file yang ditentukan
5. 5. file_path = 'path_to_your_excel_file.xlsx' # Ganti dengan jalur file aktual Anda
6. 6.
7. 7. # Membaca data dari lembar khusus di file Excel
8. 8. # Ini mengasumsikan nama lembar adalah 'Sheet1' dan berisi kolom 'Tgl', 'Medium',
'Premium'
9. 9. df = pd.read_excel(file_path, sheet_name='Sheet1')
10. 10.
11. 11. # Mengasumsikan kolom 'Tgl' berulang setiap hari dan direset menjadi 1 di awal
setiap bulan,
12. 12. # kita akan menggunakan pola ini untuk mengidentifikasi awal bulan baru.
13. 13.
14. 14. # Menginisialisasi sebuah daftar kosong untuk menyimpan harga rata-rata per bulan
15. 15. monthly_averages = []
16. 16.
17. 17. # Menginisialisasi variabel untuk melacak data bulan saat ini
18. 18. current_month_data = []
19. 19. current_month = 1
20. 20. current_year = 2018
21. 21.
22. 22. # Melakukan iterasi pada setiap baris dalam DataFrame untuk menghitung rata-rata
bulanan
23. 23. for index, row in df.iterrows():
24. 24.     # Jika kita menemui '1' di kolom 'Tgl' dan bukan baris pertama,
25. 25.     # itu menandakan awal bulan baru
26. 26.     if row['Tgl'] == 1 and index != 0:
27. 27.         # Menghitung harga rata-rata untuk bulan saat ini
28. 28.         month_data_df = pd.DataFrame(current_month_data)
29. 29.         average_medium = month_data_df['Medium'].mean()
30. 30.         average_premium = month_data_df['Premium'].mean()
31. 31.
32. 32.         # Menambahkan rata-rata ke daftar bersama dengan tahun dan bulan yang
sesuai
33. 33.         monthly_averages.append({
34. 34.             'Tahun': current_year,
35. 35.             'Bulan': current_month,
36. 36.             'Rata-rata Medium': average_medium,
37. 37.             'Rata-rata Premium': average_premium
38. 38.         })
39. 39.
40. 40.         # Mereset data bulan saat ini dan memperbarui bulan dan tahun jika
diperlukan
41. 41.         current_month_data = []
42. 42.         current_month += 1
43. 43.         if current_month > 12:
44. 44.             current_month = 1
45. 45.             current_year += 1
46. 46.
47. 47.         # Menambahkan data baris saat ini ke data bulan saat ini
48. 48.         current_month_data.append(row)
49. 49.
50. 50. # Jangan lupa untuk menghitung untuk bulan terakhir setelah loop berakhir
51. 51. # Ini diperlukan karena loop mungkin tidak menangkap data bulan terakhir
52. 52. if current_month_data:
53. 53.     month_data_df = pd.DataFrame(current_month_data)
54. 54.     average_medium = month_data_df['Medium'].mean()
55. 55.     average_premium = month_data_df['Premium'].mean()
56. 56.     monthly_averages.append({
57. 57.         'Tahun': current_year,
58. 58.         'Bulan': current_month,
59. 59.         'Rata-rata Medium': average_medium,
60. 60.         'Rata-rata Premium': average_premium
61. 61.     })
62. 62.
```

```

63. 63. # Mengonversi daftar rata-rata bulanan ke DataFrame
64. 64. monthly_averages_df = pd.DataFrame(monthly_averages)
65. 65.

66. 66. # Menyimpan DataFrame rata-rata bulanan ke file Excel baru
67. 67. output_file_path = 'path_to_your_output_excel_file.xlsx' # Ganti dengan jalur
file output yang diinginkan
68. 68. monthly_averages_df.to_excel(output_file_path, index=False)
69. 69.
70. 70. # Jalur ke file output disimpan dalam variabel 'output_file_path'
71. 71.
72. 72.

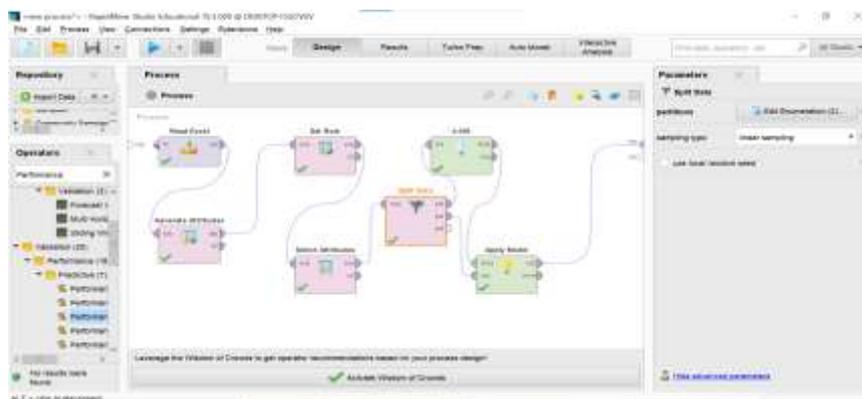
```

Gambar 3. Proses *Cleaning Data*

4.3 Transformation

Pada tahapan ini akan mengubah data dari atribut-atribut hasil pembersihan menjadi data yang siap diminingkan dengan melakukan perhitungan jarak terhadap data harga beras tersebut menggunakan rumus *Ecludien Distance*. Pada penelitian ini dilakukan dua proses transformasi data yaitu sebagai berikut.

- 1) Proses pertama adalah membagi data (*split data*) Pada klasifikasi *split data* operator digunakan untuk membagi data menjadi data *training* dan data *testing*, yang kemudian dibandingkan.



Gambar 4. Proses Split Data

- 2) Dalam proses data mining klasifikasi, diperlukan sebuah atribut spesial yang dijadikan sebagai label. Atribut label pada klasifikasi berfungsi sebagai atribut penentu, nilai dari atribut label akan menjadi hasil akhir dari klasifikasi [10]. Operator set role digunakan untuk merubah atribut regular menjadi atribut spesial, dalam hal ini yaitu label.

4.4 Data Mining

Langkah-langkah klasifikasi dengan menerapkan algoritma *K-Nearest Neighbor* ini dirancang dengan menggunakan tools Rapidminer. Pada proyek tugas akhir ini, digunakan beberapa operator yang ada pada Rapidminer, yaitu sebagai berikut [11].

- 1) *K-Nearest Neighbor* (KNN)

Algoritma yang dipakai untuk proyek tugas akhir ini yaitu algoritma *K-Nearest Neighbor* (KNN). Setelah dataset sudah diolah pada tahap sebelumnya, kemudian dataset akan diproses dengan menggunakan operator KNN pada Rapidminer [12]. Adapun variable yang digunakan dalam menghitung jaraknya, yaitu harga beras, jenis beras dan tahunnya.

Menghitung jarak menggunakan rumus *Euclidean Distance*. Berikut rumusnya:

$$d = \sqrt{\sum_{i=1}^p (x_{2i} - x_{1i})^2} \dots \dots \dots (1)$$

keterangan:

d = jarak euclidien

x_2^i = nilai pada data uji ke = i

x_1^i = nilai pada data latih ke = i

p = jumlah atribut yang digunakan

a. Bulan Januari 2018

$$\begin{aligned} & \sqrt{(10.000 - 10.000)^2} + \sqrt{(10.000 - 10.000)^2} + \sqrt{(10.000 - 10.000)^2} + \sqrt{(10.000 - 10.000)^2} + \\ & \sqrt{(10.000 - 10.000)^2} + \sqrt{(10.000 - 10.000)^2} + \sqrt{(10.000 - 10.000)^2} + \sqrt{(10.000 - 10.000)^2} + \\ & \sqrt{(10.000 - 10.000)^2} + \sqrt{(10.000 - 10.000)^2} + \sqrt{(10.000 - 10.000)^2} + \sqrt{(10.000 - 10.000)^2} + \\ & \sqrt{(10.000 - 10.000)^2} + \sqrt{(10.000 - 10.000)^2} + \sqrt{(10.000 - 10.000)^2} + \sqrt{(10.000 - 10.000)^2} + \\ & \sqrt{(10.000 - 10.000)^2} + \sqrt{(10.000 - 10.000)^2} + \sqrt{(10.000 - 10.000)^2} + \sqrt{(10.000 - 10.000)^2} + \\ & \sqrt{(10.000 - 10.000)^2} + \sqrt{(10.000 - 10.000)^2} + \sqrt{(10.000 - 10.000)^2} + \sqrt{(10.000 - 10.000)^2} + \\ & \sqrt{(10.000 - 10.000)^2} + \sqrt{(10.000 - 10.000)^2} + \sqrt{(10.000 - 10.000)^2} + \sqrt{(10.000 - 10.000)^2} = 0 \end{aligned}$$

b. Bulan Februari 2018

$$\begin{aligned} & \sqrt{(10.000 - 10.000)^2} + \sqrt{(10.000 - 10.000)^2} + \sqrt{(10.000 - 10.000)^2} + \sqrt{(10.000 - 10.000)^2} + \\ & \sqrt{(10.000 - 10.000)^2} + \sqrt{(10.000 - 10.000)^2} + \sqrt{(10.000 - 10.000)^2} + \sqrt{(10.000 - 10.000)^2} + \\ & \sqrt{(10.000 - 10.000)^2} + \sqrt{(10.000 - 10.000)^2} + \sqrt{(10.000 - 10.000)^2} + \sqrt{(10.000 - 10.000)^2} + \\ & \sqrt{(10.000 - 10.000)^2} + \sqrt{(10.000 - 10.000)^2} + \sqrt{(10.000 - 10.000)^2} + \sqrt{(10.000 - 10.000)^2} + \\ & \sqrt{(10.000 - 10.000)^2} + \sqrt{(10.000 - 10.000)^2} + \sqrt{(10.000 - 10.000)^2} + \sqrt{(10.000 - 10.000)^2} + \\ & \sqrt{(10.000 - 10.000)^2} + \sqrt{(10.000 - 10.000)^2} + \sqrt{(10.000 - 10.000)^2} + \sqrt{(10.000 - 10.000)^2} = 0 \end{aligned}$$

c. Bulan Maret 2018

$$\begin{aligned} & \sqrt{(14.000 - 13.142)^2} + \sqrt{(14.000 - 13.142)^2} + \sqrt{(14.000 - 13.142)^2} + \sqrt{(14.000 - 13.142)^2} + \\ & \sqrt{(14.000 - 13.142)^2} + \sqrt{(14.000 - 13.142)^2} + \sqrt{(14.000 - 13.142)^2} + \sqrt{(14.000 - 13.142)^2} + \\ & \sqrt{(14.000 - 13.142)^2} + \sqrt{(14.000 - 13.142)^2} + \sqrt{(14.000 - 13.142)^2} + \sqrt{(14.000 - 13.142)^2} + \\ & \sqrt{(14.000 - 13.142)^2} + \sqrt{(14.000 - 13.142)^2} + \sqrt{(14.000 - 13.142)^2} + \sqrt{(14.000 - 13.142)^2} + \\ & \sqrt{(14.000 - 13.142)^2} + \sqrt{(14.000 - 13.142)^2} + \sqrt{(14.000 - 13.142)^2} + \sqrt{(14.000 - 13.142)^2} + \\ & \sqrt{(14.000 - 13.142)^2} + \sqrt{(14.000 - 13.142)^2} + \sqrt{(14.000 - 13.142)^2} + \sqrt{(14.000 - 13.142)^2} + \\ & \sqrt{(14.000 - 13.142)^2} + \sqrt{(14.000 - 13.142)^2} = 4777.142 \end{aligned}$$

d. Bulan April 2018

$$\begin{aligned} & \sqrt{(15.000 - 12.333)^2} + \sqrt{(14.300 - 13.142)^2} + \sqrt{(14.300 - 13.285)^2} + \sqrt{(14.300 - 13.285)^2} + \\ & \sqrt{(14.300 - 13.214)^2} + \sqrt{(14.300 - 13.285)^2} + \sqrt{(14.958 - 11.833)^2} + \sqrt{(15.000 - 12.499)^2} + \\ & \sqrt{(14.250 - 13.285)^2} + \sqrt{(14.250 - 13.285)^2} + \sqrt{(14.300 - 13.285)^2} + \sqrt{(14.300 - 13.285)^2} + \\ & \sqrt{(14.300 - 13.285)^2} + \sqrt{(14.333 - 10.500)^2} + \sqrt{(15.000 - 11.666)^2} + \sqrt{(14.300 - 13.285)^2} + \\ & \sqrt{(14.300 - 13.285)^2} + \sqrt{(14.300 - 13.285)^2} + \sqrt{(14.300 - 13.285)^2} + \sqrt{(14.300 - 13.285)^2} + \\ & \sqrt{(14.650 - 12.071)^2} + \sqrt{(14.552 - 13.226)^2} + \sqrt{(14.300 - 13.285)^2} + \sqrt{(14.300 - 13.285)^2} + \\ & \sqrt{(14.300 - 13.285)^2} + \sqrt{(14.300 - 13.285)^2} + \sqrt{(14.300 - 13.285)^2} + \sqrt{(14.500 - 13.142)^2} + \\ & \sqrt{(14.450 - 13.142)^2} + \sqrt{(14.300 - 13.258)^2} = 9099.568 \end{aligned}$$

e. Bulan Mei 2018

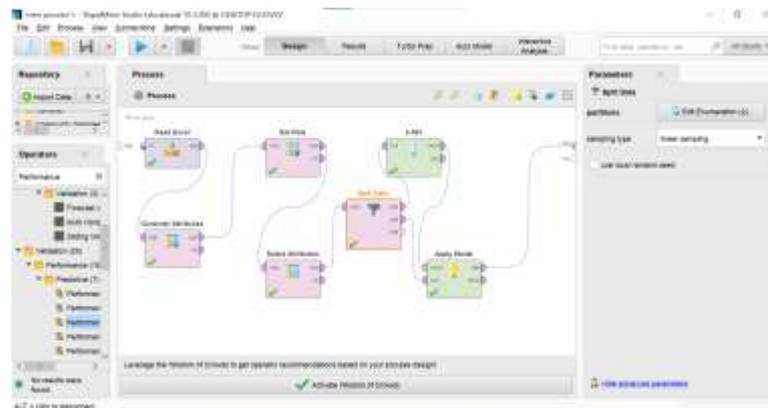
$$\begin{aligned} & \sqrt{(14.550 - 13.142)^2} + \sqrt{(14.550 - 13.142)^2} + \sqrt{(14.550 - 13.142)^2} + \sqrt{(14.550 - 13.142)^2} + \\ & \sqrt{(15.000 - 11.500)^2} + \sqrt{(14.833 - 11.500)^2} + \sqrt{(14.833 - 11.500)^2} + \sqrt{(14.833 - 11.500)^2} + \\ & \sqrt{(14.833 - 11.500)^2} + \sqrt{(14.833 - 11.500)^2} + \sqrt{(14.833 - 11.500)^2} + \sqrt{(14.833 - 11.500)^2} + \\ & \sqrt{(14.833 - 11.500)^2} + \sqrt{(14.833 - 11.500)^2} + \sqrt{(14.833 - 11.500)^2} + \sqrt{(14.833 - 11.500)^2} + \\ & \sqrt{(14.833 - 11.500)^2} + \sqrt{(14.833 - 11.500)^2} + \sqrt{(14.833 - 11.500)^2} + \sqrt{(14.833 - 11.500)^2} + \\ & \sqrt{(14.833 - 11.500)^2} + \sqrt{(14.833 - 11.500)^2} + \sqrt{(14.833 - 11.500)^2} + \sqrt{(14.833 - 11.500)^2} + \\ & \sqrt{(14.450 - 13.285)^2} + \sqrt{(14.300 - 13.285)^2} + \sqrt{(14.300 - 13.285)^2} = 16706.35 \end{aligned}$$

f. Bulan Juni 2018

$$\begin{aligned} & \sqrt{(14.550 - 13.285)^2} + \sqrt{(14.350 - 13.285)^2} + \sqrt{(14.550 - 3.285)^2} + \sqrt{(14.300 - 13.285)^2} + \\ & \sqrt{(14.300 - 13.285)^2} + \sqrt{(14.300 - 13.285)^2} + \sqrt{(14.300 - 13.285)^2} + \sqrt{(14.300 - 13.285)^2} + \\ & \sqrt{(14.300 - 13.285)^2} + \sqrt{(14.300 - 13.285)^2} + \sqrt{(14.300 - 13.285)^2} + \sqrt{(14.300 - 13.285)^2} + \\ & \sqrt{(14.300 - 13.285)^2} + \sqrt{(14.300 - 13.285)^2} + \sqrt{(14.300 - 13.285)^2} + \sqrt{(14.300 - 13.285)^2} + \\ & \sqrt{(14.300 - 13.285)^2} + \sqrt{(14.300 - 13.285)^2} + \sqrt{(14.300 - 13.214)^2} + \sqrt{(14.589 - 13.183)^2} + \\ & \sqrt{(14.650 - 12.880)^2} + \sqrt{(14.300 - 13.285)^2} + \sqrt{(14.300 - 11.550)^2} + \sqrt{(14.300 - 11.550)^2} + \\ & \sqrt{(14.300 - 11.550)^2} + \sqrt{(14.300 - 11.550)^2} = 7911.277 \end{aligned}$$

2) Apply Model

Setelah tahapan tersebut dilakukan maka akan dilakukan langkah selanjutnya yaitu penerapan model operator k-NN ke operator Apply Model, lalu dihubungkan ke set pelatihan. Adapun proses penentuan nilai k-NN pada tahapan ini ditunjukkan pada bagian "Apply Model" pada gambar di bawah ini [13]. Operator Apply Model adalah operator lanjutan dari K-Nearest Neighbor yang digunakan untuk mengaplikasikan model yang sebelumnya telah dilatih menggunakan data training pada data testing. Operator ini menghasilkan prediksi label pada data testing [14].



Gambar 5. Apply Model

4.5 Desain Akhir Penelitian

Berdasarkan beberapa tahapan-tahapan yang telah dilakukan di atas, maka dibawah ini dapat disajikan desain akhir pada penelitian ini yaitu sebagai berikut.



Gambar 6 Desain Akhir Penelitian

Pada penelitian ini, jenis sampling pada pengolahan data digunakan tipe sampling yaitu Linear Sampling dengan data *training* sebesar 80 % dari data yang telah dibersihkan, dan sebesar 20 % data testing untuk dilakukan pengujian. Adapun output hasil prediksi dengan menggunakan algoritma k-NN pada *RapidMiner* disajikan pada gambar di bawah ini [15].

Row No.	Average Medium	prediction/Average Medium
1	12380.323	12379.820
2	12365.987	12377.382
3	12581.363	12583.423
4	12479.987	12489.829
5	12587.419	12595.820
6	12581.387	12583.989
7	12583	12584.148
8	12688.581	12689.178
9	13000	12900.193
10	13000	12889.430
11	13134.089	12900.791
12	13368.383	13325.820
13	13819.682	13812.819

Gambar 7. Output Hasil Prediksi Dengan Algoritma KNN

Berdasarkan hasil pengolahan data yang dilakukan maka diperoleh hasil sebagai berikut.

- 1) *Root Mean Squared Error* (RMSE): 981.735 menunjukkan bahwa prediksi model cenderung menyimpang dari nilai sebenarnya dengan rata-rata sekitar 981 unit harga. Ini memberikan gambaran umum tentang besarnya kesalahan prediksi.
- 2) *Absolute Error* (MAE): 627.074 menunjukkan rata-rata kesalahan absolut dalam prediksi model. Meskipun angkanya lebih kecil dari RMSE, variabilitasnya tinggi (lihat nilai plus-minus), menandakan konsistensi kesalahan yang bervariasi.
- 3) *Relative Error*: Sekitar 4.35% menandakan bahwa kesalahan rata-rata model adalah 4.35% dari nilai yang benar. Ini memberikan perspektif kesalahan sebagai persentase dari harga sebenarnya.
- 4) *Normalized Absolute Error* (NAE): 0.887 menunjukkan bahwa MAE mendekati nilai rata-rata harga beras, yang berarti kesalahan cukup signifikan.
- 5) *Root Relative Squared Error*: 1.146 lebih dari 1, yang menunjukkan bahwa kesalahan model lebih besar dari variabilitas harga beras itu sendiri.
- 6) *Squared Error*: 963803.882 adalah total dari kesalahan kuadrat yang menunjukkan variasi kesalahan yang besar.
- 7) *Correlation*: 0.627 menunjukkan bahwa ada hubungan moderat antara prediksi dan nilai sebenarnya. Ini bukan hubungan yang sempurna, tetapi menunjukkan adanya pola yang model dapat tangkap.

- 8) *Squared Correlation (R-squared)*: 0.393 menjelaskan bahwa sekitar 39% variabilitas harga beras dapat dijelaskan oleh model. Meski demikian, masih ada ruang yang besar untuk perbaikan.
- 9) *Prediction Average*: 13438.362 menunjukkan bahwa harga beras rata-rata yang diprediksi oleh model sedikit lebih tinggi dari yang sebenarnya.
- 10) *Spearman's rho*: 0.824 adalah ukuran korelasi non-parametrik yang menunjukkan adanya hubungan kuat antara rangking harga beras yang sebenarnya dan yang diprediksi.
- 11) *Kendall's tau*: 0.668 juga adalah ukuran korelasi non-parametrik yang menunjukkan hubungan yang kuat tetapi sedikit lebih lemah dibandingkan *Spearman's rho*.

Nilai *Prediction Average* sebesar 13438.362 yang menunjukkan bahwa harga beras rata-rata yang diprediksi oleh model sedikit lebih tinggi dari yang sebenarnya dan nilai *Squared Correlation (R-squared)* sebesar 0.393 yang menjelaskan bahwa sekitar 39 % variabilitas harga beras dapat dijelaskan oleh model yang digunakan pada penelitian ini. Angka-angka tersebut diperoleh dari menghitung jarak dan menguji data pada *Rapidminer studio*.

5. Simpulan

Berdasarkan penelitian yang telah dilakukan, maka dapat ditarik kesimpulan bahwa algoritma menggunakan K-NN berhasil menangkap beberapa pola dalam data, seperti yang ditunjukkan oleh korelasi yang moderat hingga kuat. Namun, masih terdapat kesalahan yang signifikan dalam prediksi yang dibuat. Pada jenis data penelitian ini, algoritma K-NN bukan algoritma yang paling cocok untuk data ini tanpa penyetelan lebih lanjut atau tanpa mempertimbangkan variabel lain yang mungkin mempengaruhi harga beras. Kesalahan yang relatif tinggi dan variabilitas prediksi menunjukkan bahwa model lain perlu dilakukan peningkatan dengan menambah fitur yang relevan, penyetelan parameter, atau menggunakan metode prediksi yang berbeda.

Daftar Referensi

- [1] T. Siburian, M. Safii, and I. Parlina, "Penerapan Algoritma K-Means Clustering untuk Pengelompokan Harga Eceran Beras di Pasar Tradisional Berdasarkan Wilayah Kota," *Pros. Semin. Nas. Ris. Inf. Sci.*, vol. 1, no. September, pp. 927-936, 2019, doi: 10.30645/senaris.v1i0.101.
- [2] N. Hermanto and N. Saptana, "Kebijakan Harga Beras Ditinjau dari Dimensi Penentu Harga," *Forum Penelit. Agro Ekon.*, vol. 35, no. 1, pp. 31-43, 2018, doi: 10.21082/fae.v35n1.2017.31-43.
- [3] M.A. Sitinjak, "Indeks Harga Komsumen (IHK) di Lampung Menggunakan Autoregressive Integrated Moving Average (ARIMA)". *Indonesian Journal of Applied Mathematics*, vol. 3, no.1, pp. 15-20, 2023.
- [4] M. H. Iku, I. Surya Kumala, and Y. A. Mustofa, "Metode K-Nearest Neighbor untuk Memprediksi Harga Eceran Beras di Pasar Tradisional Gorontalo," *J. Nas. cosPhi*, vol. 3, no. 2, pp. 50-53, 2019.
- [5] Mukhlisin, M. Imrona, and D. T. Murdiansyah, "Prediksi Harga Beras Premium dengan Metode Algoritma K-Nearest Neighbor," *e-Proceeding Eng.*, vol. 7, no. 1, pp. 5-2714, 2019.
- [6] Rismala, I. Ali, and A. Rizki Rinaldi, "Penerapan Metode K-Nearest Neighbor Untuk Prediksi," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 7, no. 1, pp. 65-85, 2023.
- [7] A. S. Ilmananda and D. A. Alfianty, "Prediksi Rata-Rata Harga Beras Nasional dengan Kualitas Premium di Tingkat Penggilingan hingga Akhir Tahun 2022," *Semin. Nas. Sist. Inf.*, no. December, pp. 3385-3393, 2022.
- [8] F. D. Astuti and M. Guntara, "Analisis Performa Algoritma K-NN Dan C4.5 Pada Klasifikasi Data Penduduk Miskin," *J. Rekayasa Teknol. Inf.*, vol. 2, no. 2, pp. 135-144, 2018, doi: 10.30872/jurti.v2i2.1865.
- [9] P. Partono, H. N. Wardhani, N. I. Setyowati, A. Tsalitsa, and S. N. Putri, "Strategi Meningkatkan Kompetensi 4C (Critical Thinking, Creativity, Communication, & Collaborative)," *J. Penelit. Ilmu Pendidik.*, vol. 14, no. 1, pp. 41-52, 2021, doi: 10.21831/jpipip.v14i1.35810.
- [10] P. M. Shakeel, M. A. Burhanuddin, and M. I. Desa, "Lung cancer detection from CT image using improved profuse clustering and deep learning instantaneously trained neural networks," *Meas. J. Int. Meas. Confed.*, vol. 145, pp. 702-712, 2019, doi:

- 10.1016/j.measurement.2019.05.027.
- [11] V. N. Budiyasari, P. Studi, T. Informatika, F. Teknik, U. Nusantara, and P. Kediri, "Implementasi Data Mining Pada Penjualan kacamata Dengan Menggunakan Algoritma Apriori," *Indones. J. Comput. Inf. Technol.*, vol. 2, no. 2, pp. 31-41, 2017.
- [12] A. P. Windarto, "Penerapan Datamining Pada Ekspor Buah-Buahan Menurut Negara Tujuan Menggunakan K-Means Clustering Method," *Techno.Com*, vol. 16, no. 4, pp. 348-359, 2017, doi: 10.33633/tc.v16i4.1447.
- [13] Rahmawati and E. M. I. Putri, "Learning From Home dalam Perspektif Persepsi Mahasiswa Era Pandemi Covid-19," *Pros. Semin. Nas. Hardiknas*, vol. 1, pp. 17-24, 2020, [Online]. Available: <https://proceedings.ideaspublishing.co.id/index.php/hardiknas/article/view/3>
- [14] A. Haditsah, "Klasifikasi Masyarakat Miskin menggunakan Metode Naïve Bayes," *Ilk. J. Ilm.*, vol. 10, no. 2, pp. 160-165, 2018.
- [15] S. M. Azhari, T. H. Pudjiantoro, and I. Santikarama, "Klasterisasi Outlet Berdasarkan Data Penjualan Dengan Menggunakan Algoritma K-Medoids," *JUMANJI (Jurnal Masy. Inform. Unjani)*, vol. 5, no. 2, p. 69-79, 2021, doi: 10.26874/jumanji.v5i2.93.