

K-Nearest Neighbors (KNN) untuk menentukan Harga Saham BBRI

K-Nearest Neighbors (KNN) to Determine BBRI Stock Price

¹Abdullah Afif Baihaqi*, ²M. Fakhriza

^{1,2}Program Studi Ilmu Komputer, Fakultas Sains dan Teknologi, Universitas Islam Negeri Sumatera Utara, Kota Medan, Sumatera Utara, Indonesia

*e-mail: afifbaii03@gmail.com, fakhriza@uinsu.ac.id

(*received*: 22 February 2025, *revised*: 26 February 2025, *accepted*: 27 February 2025)

Abstrak

Prediksi atau peramalan penjualan adalah suatu perhitungan untuk memprediksi keadaan di masa mendatang melalui pengujian keadaan di masa lalu. Metode penelitian yang digunakan dalam penelitian ini adalah metode penelitian dan pengembangan (RnD). Pemodelan dengan menggunakan algoritma Knearest Neighbor menggunakan data yang diolah dengan tahapan *Knowledge Discovery in Database* (KDD). Tujuan dari penelitian ini adalah untuk memperoleh model prediksi yang dapat melakukan praproses terhadap data produk yang terstruktur sehingga mampu mempresentasikan sebuah prediksi untuk masyarakat, guna untuk mengetahui gambaran umum terkait dalam penentuan harga saham BBRI serta memberikan kontribusi bagi para pembaca sebagai rekomendasi untuk harga saham BBRI yang sudah diklasifikasikan oleh peneliti menggunakan metode K-Nearest Neighbor. Hasil prediksi harga saham menunjukkan adanya fluktuasi nilai selama periode tersebut, di mana model mampu menangkap tren perubahan harga saham berdasarkan data historis. Sebagai contoh, pada 10 Februari 2025, harga saham diprediksi sebesar 4867,020, sedangkan pada 15 Februari 2025 naik menjadi 5101,620. Hal ini membuktikan bahwa metode k-NN mampu menganalisis pola pergerakan harga saham dengan mempertimbangkan tetangga-tetangga terdekat dari data sebelumnya. Metode k-NN terbukti efektif dalam mempelajari pola data historis dan menghasilkan prediksi yang terstruktur.

Kata kunci: metode *k-nearest neighbor*, harga saham, RMS, *rapidminer*, peramalan

Abstract

Sales prediction is a calculation aimed at forecasting future conditions by analyzing past situations. The research method used in this study is the Research and Development (RnD) method. The modeling employs the K-Nearest Neighbor algorithm, utilizing data processed through the Knowledge Discovery in Database (KDD) stages. The objective of this research is to obtain a predictive model that can preprocess structured product data, enabling it to present a forecast for the public regarding the general overview of BBRI stock price determination, as well as to provide recommendations for BBRI stock prices that have been classified by the researcher using the K-Nearest Neighbor method. The results of the stock price prediction indicate fluctuations in value during the period, where the model is capable of capturing trends in stock price changes based on historical data. For example, on February 10, 2025, the stock price is predicted to be 4867.020, while on February 15, 2025, it rises to 5101.620. This demonstrates that the k-NN method can analyze stock price movement patterns by considering the nearest neighbors from previous data. The k-NN method has proven effective in studying historical data patterns and generating structured predictions.

Keywords: *k-nearest neighbor method, stock price, RMS, rapidminer, forecasting*

1 Pendahuluan

Seiring berjalannya waktu hampir semua pekerjaan berhubungan dengan teknologi informasi, perkembangan teknologi informasi saat ini berkembang begitu pesat, mulai dari pendidikan, perbankan, pertokoan, restoran bahkan industri rumah tangga [1]. Saat ini banyak instansi dan perusahaan yang menerapkan teknologi informasi dalam segala aspek, oleh karena itu dalam dunia kerja atau bisnis,

teknologi sangat diperlukan untuk mempermudah penyelesaian masalah agar lebih efektif dan efisien, salah satunya adalah kecerdasan buatan [2].

Salah satu manfaat kecerdasan buatan adalah memprediksi harga saham. Saham adalah tanda penyertaan seseorang atau badan hukum sebagai pihak tertentu dalam suatu perseroan terbatas [3]. Saat ini, saham merupakan salah satu pemain kunci dalam perkembangan perusahaan untuk lebih mengembangkan bisnis dan efisiensi perusahaan. Meningkatnya jumlah investor dari tahun ke tahun juga mempengaruhi perkembangan portofolio investasi di Indonesia. Salah satunya adalah perusahaan BBRI [4][5].

Disaat perusahaan dengan mudah memperjual belikan sahamnya secara bebas timbul permasalahan yang terjadi pada para investor baru. Permasalahan yang terjadi kepada para investor baru adalah pergerakan harga saham yang fluktuatif atau naik turun yang tidak stabil, menjadikan perlu adanya pemodelan prediksi yang digunakan untuk menilai harga saham perusahaan yang ingin dibeli untuk dijadikan aset. Alasan lain dibuatnya pemodelan prediksi saham ini, karena investor sering melakukan investasi tanpa melihat apakah pergerakan harga saham perusahaan tersebut akan secara bertahap naik ataupun turun [6].

Pada setiap perusahaan, penjualan merupakan hal yang sangat penting agar perusahaan bisa tetap beroperasi dan mendapatkan penghasilan untuk memproduksi macam-macam produk yang akan dijual. Setiap perusahaan akan bersaing melalui mutu dan kualitas produk agar penjualan bisa tetap naik sesuai yang diharapkan. perusahaan biasanya akan menggunakan prediksi atau peramalan pada penjualan produk untuk mengetahui produk apa yang akan lebih banyak terjual pada masa yang akan datang. telah melakukan prediksi penjualan produk hemat energi yang diproduksi dengan melihat statistik data penjualan produk [7].

Untuk melakukan prediksi atau peramalan, Aplikasi ajaib mengumpulkan seluruh data penjualan dari tahun ke tahun. Dari prediksi menghasilkan suatu data yang menunjukkan suatu penjualan produk akan naik dan produk lain akan turun. Namun terdapat satu kasus dimana prediksi ini justru terbalik, dalam artian apabila salah satu produk diprediksi akan naik penjualannya, justru malah terjadi penurunan pada produk tersebut diperiode yang telah di prediksi, begitupula pada salah satu produk yang diprediksi akan mengalami penurunan penjualan justru mengalami kenaikan pada masa periode yang sudah diprediksikan. Hasil prediksi dapat dibilang kurang akurat [8].

Prediksi atau peramalan penjualan adalah suatu perhitungan untuk memprediksi keadaan di masa mendatang melalui pengujian keadaan di masa lalu. Prediksi hampir sama dengan klasifikasi dan estimasi, kecuali bahwa dalam prediksi nilai dari hasil akan ada di masa mendatang.

Meramalkan penjualan di masa mendatang artinya memperkirakan besarnya volume penjualan, bahkan menentukan potensi penjualan dan luas pasar yang dikuasai di masa mendatang. Salah satu kegunaan prediksi adalah membantu pemilik perusahaan dalam pengambilan keputusan adalah untuk menentukan fokus ke arah mana penjualan akan dilakukan sehingga diharapkan resiko kesalahan yang disebabkan oleh kesalahan perencanaan dapat ditekan seminimal mungkin [9].

Metode-metode yang digunakan untuk prediksi diantaranya metode *K- Nearest Neighbors* (KNN). Metode ini merupakan metode untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya terdekat dengan objek tersebut. Keunggulan metode ini dalam memprediksi yaitu tangguh dan efektif terhadap training data yang noisenya besar [10][11][12].

Berdasarkan latar belakang yang ada maka tujuan daripada penelitian ini adalah memperoleh model prediksi yang dapat melakukan praproses terhadap data produk yang terstruktur sehingga mampu mempresentasikan sebuah prediksi untuk masyarakat, guna untuk mengetahui gambaran umum terkait dalam penentuan harga saham BBRI. Peneliti berharap penelitian ini dapat memberikan kontribusi bagi para pembaca sebagai rekomendasi untuk harga saham BBRI yang sudah diklasifikasikan oleh peneliti menggunakan metode K-Nearest Neighbor.

2 Tinjauan Literatur

Berikut ini merupakan beberapa kajian terdahulu yang relevan dengan penelitian ini. Tujuan dari penelitian terdahulu adalah untuk mengumpulkan referensi dan data pembanding. Selain itu, jangan menarik kesejajaran antara penelitian ini dengan penelitian lainnya. Oleh karena itu, peneliti menggunakan temuan-temuan berikut dari penelitian sebelumnya dalam penelitian sebagai berikut: Hasil penelitian oleh Sihombing *et al* [13], tingkat pembelian spare part tertinggi pada tahun 2021 adalah Circlip dengan pembelian sebanyak 22 buah. Sedangkan tingkat pembelian spare part terendah

pada tahun 2021 adalah Radio SSB sebanyak 1 buah. Metode *K-Nearest Neighbors* dapat digunakan untuk menentukan tingkat prioritas pembelian spare part. Hasil dari tingkat pengelompokan spare part terdapat 3 kondisi yaitu, Prioritas Tinggi, Prioritas Sedang dan Prioritas Rendah. Hasil klasifikasi berdasarkan algoritma *K-Nearest Neighbors* menunjukkan bahwa algoritma *K-Nearest Neighbors* telah diimplementasikan dengan baik untuk permasalahan klasifikasi tingkat pembelian spare part pada divisi workshop di tahun 2021. Hasil dari klasifikasi *K-Nearest Neighbors* menunjukkan akurasi sebesar 93.33%.

Hasil penelitian Novitadewi *et al* [14], dengan menggunakan metode *K-Nearest Neighbors* (KNN) pada data penjualan produk lampu di PT. Terang Abadi Raya, penulis berhasil mengklasifikasikan 19.290 barang dengan bahasa pemrograman Python menggunakan Google Colab. Dari hasil klasifikasi tersebut, didapatkan 12.420 barang terkategori sebagai label "laris" dan 6.870 barang terkategori sebagai label "tidak laris". Oleh karena itu, dapat disimpulkan bahwa metode KNN dapat digunakan sebagai alat bantu dalam meramalkan penjualan produk di masa mendatang serta memberikan rekomendasi bagi manajemen dalam merencanakan stok produk agar toko tidak kehabisan stok atau menumpuk stok produk dengan melihat data produk yang laris dan tidak laris dari data pembelian sebelumnya.

Hasil penelitian oleh Jannah *et al* [15], hasil pengujian prediksi penjualan produk PT. Bintang Sriwijaya menggunakan metode *K-Nearest Neighbors* (KNN) maka dapat ditarik kesimpulan sebagai berikut: Faktor-faktor yang mempengaruhi penjualan produk pada PT Bintang Sriwijaya Palembang diantaranya harga produk, sales penjualan (marketing), dan persaingan pasar. Hasil prediksi penjualan produk berdasarkan nilai akurasi tertinggi yaitu $k=20$ dengan akurasi sebesar 97,3%.

Hasil penelitian oleh Yusuf *et al* [16], berdasarkan hasil percobaan dapat, nilai akurasi yang didapatkan pada penelitian ini yaitu 86.9% dengan jumlah prediksi benar 12582 dari 14478 records, tetapi nilai akurasi tidak dapat dijadikan acuan bahwa model yang dibuat dapat memprediksi dengan baik terutama saat kondisi class yang tidak balance. Hal ini dapat menyebabkan nilai akurasi sangat tinggi tetapi nilai class recall dan class precision rendah. Oleh karena itu, perlu diamati bahwa nilai class recall dan class precision memegang peranan terhadap performa model yang dibuat untuk melakukan klasifikasi terhadap data yang digunakan.

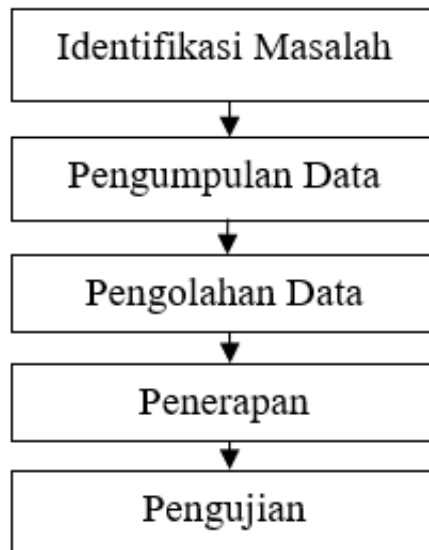
Dari beberapa kajian terdahulu yang relevan, dapat disimpulkan bahwa metode *K-Nearest Neighbors* (KNN) terbukti efektif dalam berbagai aplikasi klasifikasi dan prediksi, terutama dalam konteks pengelompokan data dan peramalan. Penelitian oleh Sihombing *et al* [13], menunjukkan bahwa KNN dapat mengklasifikasikan tingkat prioritas pembelian spare part dengan akurasi yang tinggi. Penelitian Novitadewi *et al* [14], juga menunjukkan bagaimana KNN dapat digunakan untuk mengklasifikasikan produk "laris" dan "tidak laris" pada penjualan barang, memberikan wawasan penting untuk manajemen stok. Selain itu, penelitian oleh Jannah *et al* [15], menunjukkan bahwa faktor-faktor seperti harga, sales, dan persaingan pasar dapat mempengaruhi prediksi penjualan, dengan KNN memberikan hasil prediksi yang cukup akurat. Penelitian oleh Yusuf *et al* [16], juga menekankan pentingnya memperhatikan nilai class recall dan class precision dalam evaluasi model KNN, terutama ketika data tidak seimbang.

Perbedaan penelitian ini dengan penelitian sebelumnya adalah Penelitian ini berfokus pada *K-Nearest Neighbors* (KNN) untuk menentukan harga saham BBRI, yang merupakan perbedaan utama dibandingkan dengan penelitian terdahulu. Sementara sebagian besar penelitian sebelumnya menggunakan KNN untuk aplikasi seperti klasifikasi produk atau spare part, penelitian ini akan menggunakan metode yang sama untuk prediksi harga saham, yang melibatkan data numerik dan waktu. Hal ini memberikan kontribusi baru dalam pemanfaatan algoritma KNN pada analisis pasar finansial, khususnya dalam memprediksi harga saham berdasarkan data historis dan faktor pasar lainnya. Kontribusi utama dari penelitian ini adalah penerapan KNN dalam domain pasar saham, yang relatif belum banyak diterapkan dibandingkan dengan penggunaannya dalam bidang klasifikasi produk atau spare part. Penelitian ini diharapkan dapat memberikan wawasan baru mengenai cara mengadaptasi metode KNN dalam konteks pasar modal dan prediksi harga saham yang berfluktuasi.

3 Metode Penelitian

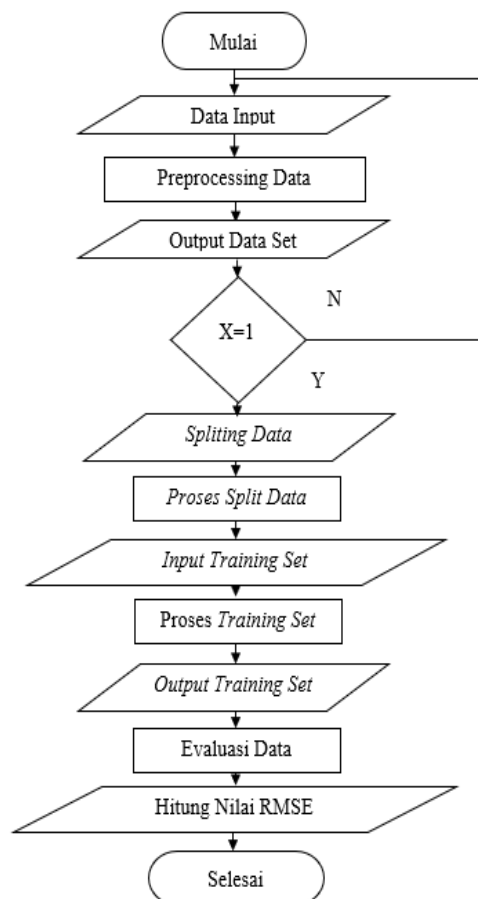
Lokasi penelitian mengambil data harga Saham BBRI di aplikasi ajaib di *Playstore* sebagai tempat penelitian karna sesuai dengan topik yang diangkat peneliti mengenai data *mining*. Metode penelitian yang digunakan dalam penelitian ini adalah metode penelitian dan pengembangan (RnD)

[17][18]. Pemodelan dengan menggunakan algoritma Knearest Neighbor menggunakan data yang diolah dengan tahapan Knowledge Discovery in Database (KDD) [19]. Berdasarkan Penentuan nilai k. Hitung jarak data training dan data uji. Pengurutan data hasil perhitungan. Langkah selanjutnya dengan menggunakan kategori K-Nearest Neighbor yang paling mayoritas. Kerangka penelitian ini dapat dilihat seperti pada Gambar 1 berikut:



Gambar 1. Kerangka Penelitian

Flowchart sistem berfungsi untuk memberi gambaran bagaimana sistem pemodelan metode K-Nearest Neighbors yang akan dibangun [20][21][22]. Adapun Flowchart sistem dapat dilihat pada Gambar 2 di bawah ini:



Gambar 2. Flowchart metode KNN

Pada Gambar 2 dijelaskan data berupa *Data Set* harga Emas yang sudah melalui acak data kemudian akan *Transformation* atau proses pembentukan data training berdasarkan data yang sudah ada, data harus di seleksi terlebih dahulu untuk menentukan atribut mana yang dapat mempengaruhi kenaikan harga Emas. dan dihitung pembobotannya, setelah pembobotan apa bila bobot tersebut bernilai $X_p > X$ maka data tersebut akan diurutkan data hasil perhitungan. Jarak yang telah didapatkan kemudian diurutkan dari yang paling dekat jaraknya sampai dengan yang paling jauh (*ascending*).

4 Hasil dan Pembahasan

Bagian ini menyajikan temuan utama penelitian serta analisisnya. Data yang diperoleh dianalisis untuk menjawab tujuan penelitian dan memberikan pemahaman yang lebih mendalam mengenai hasil yang didapatkan.

4.1 Representase Data

Algoritma KNN merupakan algoritma yang memanfaatkan nilai jarak euclidean untuk mendapatkan hasil akhir. Pada model awal, data hanya berupa data historis harga emas secara berurutan, data tersebut akan ditransformasi ke dalam bentuk yang memungkinkan untuk mendapatkan nilai jarak euclidean, oleh karena itu data awal akan ditransformasi sehingga tampilan data akan menjadi seperti Tabel 1 di bawah ini.

Tabel 1. Harga saham BBRI

Tanggal	Hari Pembukaan	Hari Penutupan
11/1/2021	3800	3800
...
12/31/2021	3900	3900
1/1/2022	4000	4000
...
12/31/2022	4900	4900
1/1/2023	4900	4900
...
12/31/2023	4900	4900
1/1/2024	4500	4500
...
10/31/2024	5500	5500

4.2 Proses Algoritma *K-Nearest Neighbour* (KNN)

Dalam penelitian ini, peneliti mengimplementasikan algoritma *K-Nearest Neighbors* (KNN) untuk memprediksi harga saham BBRI dengan memanfaatkan data historis harga saham yang telah tercatat sebelumnya. Proses prediksi harga saham dibagi menjadi dua tahap penting, yaitu tahap pelatihan dan tahap pengujian, yang bertujuan untuk memastikan bahwa model yang dibangun dapat belajar dengan baik dari data yang tersedia dan diuji dengan data yang belum pernah dilihat sebelumnya.

Pada tahap pelatihan, algoritma KNN dilatih menggunakan subset data yang telah disiapkan, di mana model mempelajari hubungan antara variabel input dalam hal ini adalah harga saham yang tercatat di periode sebelumnya dengan data yang ada, dengan fokus utama pada perhitungan jarak Euclidean untuk mengukur kedekatan antara data baru dan data historis. Selama proses ini, algoritma akan menyimpan seluruh data jarak Euclidean yang dihitung antara data historis dan data baru, yang memungkinkan model untuk "belajar" tentang pola pergerakan harga saham. Setelah model selesai dilatih, pada tahap pengujian, model KNN akan diuji menggunakan data yang sebelumnya tidak digunakan dalam pelatihan untuk mengukur sejauh mana akurasi dan kemampuan prediksi model dalam meramalkan harga saham BBRI di masa mendatang, dengan harapan dapat memberikan informasi yang akurat dan bermanfaat bagi para investor dalam pengambilan keputusan investasi mereka.

4.3 Perhitungan Jarak Euclidean

Langkah pertama dalam proses prediksi harga saham menggunakan algoritma *K-Nearest Neighbors* (K-NN) pada penelitian ini adalah melakukan perhitungan jarak Euclidean antara setiap harga awal pada data historis dengan harga terakhir yang tersedia, yaitu harga saham pada (31

Oktober 2024). Penghitungan jarak ini bertujuan untuk mencari harga-harga saham historis yang paling dekat dengan harga pada tanggal 31 Oktober 2024, yang akan digunakan sebagai dasar dalam proses prediksi harga saham BBRI. Untuk menghitung jarak tersebut, digunakan persamaan *Euclidean Distance* yang sederhana, yaitu di jelaskan dalam persamaan 1.

$$\text{Jarak Euclidean} = \sqrt{(x_2 - x_1)^2} \tag{1}$$

Dalam persamaan ini, x_1 mewakili harga saham pada setiap hari sebelumnya (harga awal yang terdapat pada data historis), dan x_2 adalah harga saham pada 31 Oktober 2024, yang diketahui bernilai 5500. Dengan menggunakan persamaan 1, maka jarak antara harga saham pada 31 Oktober 2024 dengan harga-harga pada hari-hari sebelumnya akan dihitung untuk setiap data historis yang ada. Proses ini akan memberikan gambaran mengenai kedekatan harga-harga saham historis dengan harga terakhir yang tercatat, dan harga-harga yang memiliki jarak terdekat akan dianggap lebih relevan dalam prediksi harga saham yang akan datang. Dengan demikian, tahap perhitungan jarak Euclidean ini sangat penting, karena menentukan data mana yang akan dipilih untuk digunakan dalam langkah selanjutnya, yaitu memilih tetangga terdekat untuk menghitung prediksi harga saham pada periode yang akan datang di jelaskan dalam persamaan 2.

$$\text{Jarak Euclidean} = \sqrt{(x_2 - x_1)^2} \tag{2}$$

$$\text{Harga akhir ke Harga 1 November 2021} = \sqrt{(5500 - 3800)^2} = 1700$$

Berikut adalah jarak harga akhir terhadap seluruh data awal setelah dilakukan perhitungan jarak euclidean di jelaskan dalam Tabel 2.

Tabel 2. Data euclidean distance

Tanggal	HargaPerLot	Euclidean Distance
11/1/2021	3800	1700
...
12/31/2021	3900	1600
1/1/2022	4000	1500
...
12/31/2022	4900	600
1/1/2023	4900	600
...
12/31/2023	4900	600
1/1/2024	4500	1000
...
2/29/2024	4700	800

4.4 Penentuan *Nearest Neighbour* (nilai k)

Pada algoritma K-Nearest Neighbors (K-NN), salah satu parameter penting yang harus ditentukan adalah nilai k, yaitu jumlah nearest neighbor (tetangga terdekat) yang akan digunakan dalam proses perhitungan prediksi. Variabel k ini sangat menentukan, karena nilai tersebut menunjukkan jumlah data historis terdekat yang akan dipertimbangkan dalam memprediksi nilai data baru. Semakin besar nilai k, semakin banyak data historis yang digunakan, tetapi hal ini juga dapat menyebabkan model menjadi kurang sensitif terhadap data terbaru. Sebaliknya, jika nilai k terlalu kecil, model mungkin menjadi terlalu sensitif terhadap data tertentu dan rentan terhadap noise.

Dalam penelitian ini, peneliti menetapkan nilai k = 3, yang berarti bahwa tiga tetangga terdekat akan dilibatkan dalam proses prediksi. Pemilihan nilai k ini didasarkan pada upaya untuk mencapai keseimbangan antara sensitivitas model terhadap data terdekat dan stabilitas prediksi yang dihasilkan. Dengan nilai k = 3, model akan memperhitungkan tiga harga saham awal yang memiliki jarak paling dekat dengan data terakhir, yaitu harga saham pada 31 Oktober 2024, untuk memprediksi harga saham pada 1 November 2024. Pendekatan ini diharapkan dapat menghasilkan prediksi yang lebih akurat, karena mempertimbangkan data historis yang relevan.

Setelah proses perhitungan jarak Euclidean selesai, langkah berikutnya adalah menentukan tetangga terdekat atau nearest neighbors dari data terakhir berdasarkan nilai jarak yang telah dihitung. Jarak Euclidean antara data terakhir dan semua data awal akan diurutkan dari nilai yang paling kecil (terdekat) hingga terbesar (terjauh). Dalam penelitian ini, karena nilai k = 3, maka hanya tiga data awal dengan jarak terkecil yang akan dipilih sebagai nearest neighbors. Data ini dianggap sebagai

representasi terbaik untuk membantu memprediksi harga saham yang akan datang, karena memiliki kesamaan atau kedekatan yang paling relevan dengan data terakhir.

Sebagai ilustrasi, jika jarak Euclidean telah dihitung untuk semua data historis dan hasilnya menunjukkan bahwa tiga jarak terkecil adalah sebagai berikut:

- Data historis pada 1 November 2021 dengan jarak Euclidean 1700
- Data historis pada 1 November 2022 dengan jarak Euclidean 700
- Data historis pada 1 November 2023 dengan jarak Euclidean 600

Maka, ketiga data tersebut akan dipilih sebagai nearest neighbors untuk melanjutkan ke tahap prediksi. Pemilihan ini berdasarkan asumsi bahwa data historis dengan jarak terkecil memiliki pola atau hubungan yang lebih relevan terhadap data terakhir, sehingga memberikan kontribusi yang lebih besar dalam menentukan hasil prediksi.

Langkah ini penting karena memfilter data yang tidak relevan dan hanya mempertahankan data historis yang paling mendekati pola harga saham terbaru. Dengan cara ini, algoritma K-NN dapat memberikan prediksi yang lebih terarah dan akurat berdasarkan pola yang terbentuk dari data terdekat. Selain itu, penggunaan nilai $k = 3$ juga memungkinkan model untuk menangani outlier atau anomali dalam data historis, karena rata-rata dari tiga tetangga terdekat lebih tahan terhadap efek dari satu nilai yang tidak biasa dibandingkan jika hanya menggunakan satu tetangga terdekat saja.

Proses pemilihan nearest neighbors ini menjadi salah satu langkah kunci dalam algoritma K-NN, karena memastikan bahwa prediksi didasarkan pada data historis yang paling relevan dan sesuai dengan pola yang diamati dalam data terbaru. Hal ini sekaligus menjadi dasar bagi model untuk memberikan hasil prediksi yang dapat diandalkan dan mendukung pengambilan keputusan yang lebih baik.

4.5 Perhitungan output

Setelah harga awal terdekat berhasil ditentukan berdasarkan kriteria tertentu, langkah berikutnya adalah memanfaatkan data harga lanjutan dari masing-masing harga awal tersebut sebagai acuan dalam proses prediksi harga. Dalam metode ini, nilai parameter k yang telah ditentukan, yaitu sebesar 3, menjadi dasar untuk memilih tiga data harga lanjutan yang akan digunakan sebagai referensi. Dengan demikian, proses prediksi memanfaatkan tiga data harga lanjutan terdekat yang relevan dengan harga awal sebagai input utama.

Selanjutnya, untuk menghitung hasil prediksi, digunakan pendekatan matematis sederhana berupa rata-rata atau mean. Persamaan yang digunakan untuk memperoleh nilai prediksi adalah sebagai berikut di jelaskan dalam persamaan 3.

$$\text{Hasil prediksi} = \frac{\sum \text{Harga lanjutan}}{k} \quad (3)$$

Harga lanjutan merupakan nilai-nilai harga yang dipilih sebagai acuan berdasarkan kedekatannya dengan harga awal.

k adalah jumlah data harga lanjutan yang digunakan, dalam hal ini sebanyak 3 data. Proses ini bertujuan untuk mendapatkan nilai prediksi yang merepresentasikan kondisi harga secara umum berdasarkan data harga lanjutan yang tersedia. Penggunaan metode mean ini memberikan pendekatan yang cukup sederhana namun efektif dalam menghasilkan prediksi, karena hasil rata-rata dari tiga harga lanjutan diharapkan mampu menggambarkan tren harga yang terjadi. Adapun pemilihan nilai k yang lebih besar atau lebih kecil dapat disesuaikan dengan kebutuhan dan karakteristik data yang digunakan dalam analisis, sehingga dapat menghasilkan prediksi yang lebih akurat dan sesuai konteks.

Dengan menggunakan persamaan diatas, maka hasil prediksi dapat dihitung sebagai berikut.

$$\text{Hasil prediksi 1 November 2024} = \frac{H.Lanjut 1+H.Lanjut 2+H.Lanjut 3}{k}$$

$$\text{Hasil prediksi 1 November 2024} = \frac{3800+4800+4900}{3}$$

$$\text{Hasil prediksi 1 November 2024} = \frac{13500}{3} = 4500$$

Berikut adalah Prediksi harga saham BBRI dari tanggal 1 november 2024 sampai 28 Februari 2025 di jelaskan dalam Tabel 3.

Tabel 3. Data prediksi harga saham BBRI

TanggalPrediksi	HargaPrediksi
11/1/2024	4500
...
12/31/2024	4566.666667
1/1/2025	4466.666667
...	...
2/28/2025	4560

Setelah menghitung dengan persamaan mean diperoleh hasil akhir sebesar 4500. Oleh karena itu, dengan menggunakan 362 data harga saham BBRI di sejak 01 November 2021 hingga 28 Februari 2024=5, maka dengan menggunakan algoritma K-NN dengan nilai k = 3 diperoleh hasil harga saham BBRI untuk 01 November 2024 adalah sebesar Rp. 450.000. Dimana harga saham BBRI pada bulan 01 November 2024 diprediksi mengalami penurunan dari harga pada hari sebelumnya.

4.6 Perhitungan Root Mean Square Error

Root Mean Squared Error (RMSE) adalah salah satu metrik evaluasi yang digunakan untuk mengukur seberapa jauh hasil prediksi dari nilai sebenarnya. RMSE sering digunakan karena memberikan bobot lebih besar pada kesalahan yang lebih besar, sehingga sangat sensitif terhadap error besar di jelaskan dalam persamaan 4.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - y^i)^2}{n}} \quad (4)$$

Dimana y_i : Nilai sebenarnya (observasi atau data aktual), y^i : Nilai prediksi, n : Jumlah data atau sampel, $y_i - y^i$: Selisih antara nilai sebenarnya dan nilai prediksi, disebut juga sebagai **error**.

- Tentukan Data Sebenarnya dan Prediksi: Siapkan tabel yang berisi data aktual (y_i) dan data prediksi (y^i). Seperti tabel 4 di bawah ini :

Tabel 4. Data aktual RMSE

No.	Tanggal	Y_i	y^i	Error	Error Kuadrat
1.	1 November 2021	4300	4500	-200	40000
2.	1 November 2022	4500	4500	0	0
3.	1 November 2023	4700	4500	200	40000

Langkah-langkah perhitungan *Root Mean Squared Error* seperti berikut ini :

- Hitung error di jelaskan dalam persamaan 5.

$$Error = y_i - y^i \quad (5)$$

- Data 01 November 2021 : $4300 - 4500 = -200$
- Data 01 November 2022 : $4500 - 4500 = 0$
- Data 01 November 2023 : $4700 - 4500 = 200$

- Hitung Kuadrat dari Error di jelaskan dalam persamaan 6.

$$Error \text{ Kuadrat} = (y_i - y^i)^2 \quad (6)$$

- Data 01 November 2021 = $(-200)^2 = 40000$
- Data 01 November 2022 = $0^2 = 0$
- Data 01 November 2023 = $200^2 = 40000$

- Jumlahkan semua Error Kuadrat di jelaskan dalam persamaan 7.

$$Jumlah \text{ Error Kuadrat} = \sum_{i=1}^n (y_i - y^i)^2 \quad (7)$$

$$Jumlah \text{ Error Kuadrat} = 40000 + 0 + 40000 = 80000$$

- Hitung Rata-Rata Error Kuadrat di jelaskan dalam persamaan 8.

$$Rata - rata \text{ Error Kuadrat} = \frac{Jumlah \text{ Error Kuadrat}}{n} \quad (8)$$

$$Rata - rata \text{ Error Kuadrat} = \frac{80000}{3} = 26666,67$$

- Hitung Akar Kuadrat dari Rata-rata Error Kuadrat di jelaskan dalam persamaan 9.

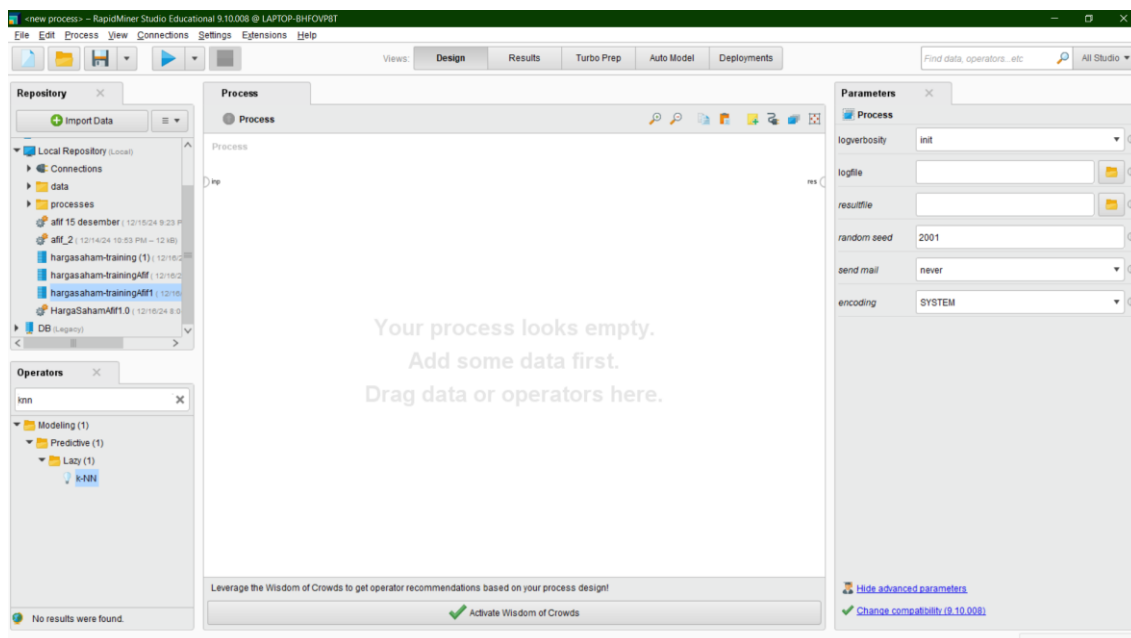
$$RMSE = \sqrt{Rata - rata \text{ Error Kuadrat}} \quad (9)$$

$$RMSE = \sqrt{26666.67} \approx 163.3$$

RMSE untuk prediksi harga saham BBRI adalah sebesar 163,3. Nilai ini menunjukkan rata-rata kesalahan prediksi terhadap harga saham aktual dalam satuan rupiah, yang berarti bahwa prediksi harga saham memiliki tingkat deviasi rata-rata sebesar 163,3 dari nilai sebenarnya. Semakin kecil nilai RMSE yang diperoleh, maka semakin baik akurasi model prediksi yang digunakan, karena menunjukkan bahwa kesalahan atau selisih antara nilai prediksi dengan nilai aktual semakin kecil. Dengan nilai RMSE yang relatif rendah, dapat dikatakan bahwa metode yang digunakan, yaitu algoritma K-Nearest Neighbors (K-NN), memiliki performa yang cukup baik dalam memprediksi harga saham BBRI berdasarkan data historis. Hasil ini memberikan kepercayaan terhadap model prediksi yang dibangun, sehingga dapat dijadikan sebagai alat bantu dalam pengambilan keputusan yang lebih akurat bagi investor atau pemangku kepentingan di pasar modal. Namun demikian, nilai RMSE ini juga dapat menjadi indikator untuk terus menyempurnakan model prediksi agar lebih adaptif terhadap pola pergerakan harga saham di masa mendatang.

4.7 Penerapan Rapid Miner

Rapid Miner merupakan salah satu perangkat lunak data mining yang digunakan untuk mengolah dataset dengan tujuan menemukan pola-pola tertentu di dalam data, sehingga dapat membantu dalam proses analisis, peramalan, atau pengambilan keputusan yang lebih akurat. Dalam pengolahan data menggunakan RapidMiner, tidak semua algoritma yang tersedia dapat langsung diterapkan pada dataset yang ada, karena setiap algoritma memiliki karakteristik, asumsi, dan cara kerja yang berbeda. Oleh karena itu, diperlukan penyesuaian pola data melalui tahap pra-pemrosesan seperti pembersihan data, normalisasi, transformasi, dan pemilihan atribut yang relevan agar algoritma yang digunakan sesuai dengan struktur data dan tujuan analisis yang ingin dicapai. Penyesuaian ini bertujuan untuk memastikan bahwa algoritma dapat mengolah data secara optimal dan memberikan hasil yang akurat serta sesuai dengan tujuan spesifik dari proyek pengolahan data tersebut, seperti klasifikasi, prediksi, clustering, atau analisis asosiasi di tampilan dalam Gambar 3.



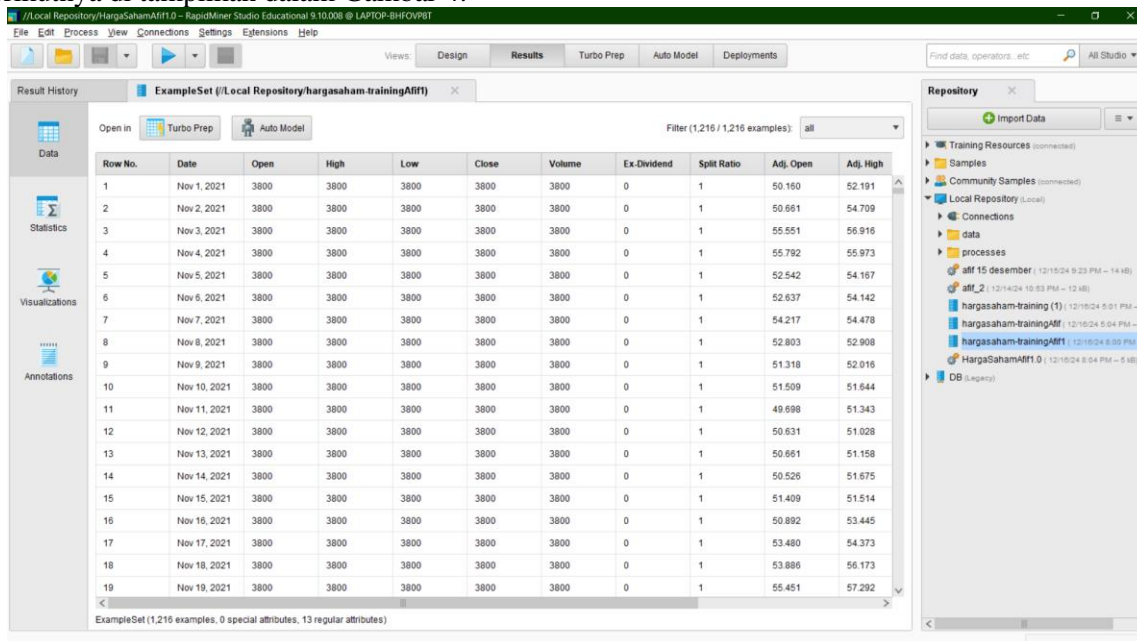
Gambar 3. Tampilan awal rapid miner studio

4.8 Susunan Operator Algoritma K-Nearest Neighbor

Setelah proses klasifikasi selesai dilakukan, langkah selanjutnya yang perlu dilakukan adalah mengimpor tabel Microsoft Excel ke dalam alur proses di RapidMiner untuk membaca data training yang tersimpan dalam format Excel. Tahap ini bertujuan untuk memastikan bahwa data yang akan digunakan dalam proses training tersedia dan dapat dibaca dengan baik oleh perangkat lunak. Proses ini dapat dilakukan dengan menggunakan metode drag and drop, di mana file yang berisi dataset, dalam hal ini File HargaSahamAfif1.0, ditarik (drag) langsung dari direktori penyimpanan komputer dan dijatuhkan (drop) ke dalam area desain proses RapidMiner. Dengan cara ini, file Excel yang berisi data harga saham akan secara otomatis terbaca dan dihubungkan ke operator Retrieve atau operator lainnya yang sesuai di dalam proses. Penting untuk memastikan bahwa struktur file Excel sudah

<http://sistemasi.ftik.unisi.ac.id>

sesuai dengan kebutuhan, seperti memiliki kolom atribut (fitur) dan label target yang jelas, agar data dapat digunakan secara optimal dalam tahap selanjutnya, seperti training model, validasi, atau evaluasi performa model klasifikasi yang telah dibuat sebelumnya. Proses pengimporan ini juga memungkinkan pengguna untuk memvisualisasikan data, mengecek kelengkapan, serta memastikan tidak ada data yang kosong atau tidak valid sebelum melanjutkan ke tahap analisis atau pemodelan berikutnya di tampilan dalam Gambar 4.



Row No.	Date	Open	High	Low	Close	Volume	Ex-Dividend	Split Ratio	Adj. Open	Adj. High
1	Nov 1, 2021	3800	3800	3800	3800	3800	0	1	50.160	52.191
2	Nov 2, 2021	3800	3800	3800	3800	3800	0	1	50.661	54.709
3	Nov 3, 2021	3800	3800	3800	3800	3800	0	1	55.551	56.916
4	Nov 4, 2021	3800	3800	3800	3800	3800	0	1	55.792	55.973
5	Nov 5, 2021	3800	3800	3800	3800	3800	0	1	52.542	54.167
6	Nov 6, 2021	3800	3800	3800	3800	3800	0	1	52.637	54.142
7	Nov 7, 2021	3800	3800	3800	3800	3800	0	1	54.217	54.478
8	Nov 8, 2021	3800	3800	3800	3800	3800	0	1	52.803	52.908
9	Nov 9, 2021	3800	3800	3800	3800	3800	0	1	51.318	52.016
10	Nov 10, 2021	3800	3800	3800	3800	3800	0	1	51.509	51.644
11	Nov 11, 2021	3800	3800	3800	3800	3800	0	1	49.698	51.343
12	Nov 12, 2021	3800	3800	3800	3800	3800	0	1	50.631	51.028
13	Nov 13, 2021	3800	3800	3800	3800	3800	0	1	50.661	51.158
14	Nov 14, 2021	3800	3800	3800	3800	3800	0	1	50.525	51.675
15	Nov 15, 2021	3800	3800	3800	3800	3800	0	1	51.409	51.514
16	Nov 16, 2021	3800	3800	3800	3800	3800	0	1	50.892	53.445
17	Nov 17, 2021	3800	3800	3800	3800	3800	0	1	53.480	54.373
18	Nov 18, 2021	3800	3800	3800	3800	3800	0	1	53.886	56.173
19	Nov 19, 2021	3800	3800	3800	3800	3800	0	1	55.451	57.292

Gambar 4. Importing data training pada operator read excel

Setelah file data berhasil diimpor ke dalam alur proses RapidMiner, langkah selanjutnya adalah menambahkan operator *Windowing* untuk mempersiapkan data dalam format yang sesuai agar dapat digunakan untuk pemodelan prediksi. Operator *Windowing* berfungsi untuk mengubah data deret waktu (time series) menjadi jendela-jendela data yang berisi beberapa observasi sebelumnya, sehingga model dapat mempelajari pola historis dari data tersebut. Untuk memulai, lakukan drag and drop operator *Windowing* ke dalam area desain proses RapidMiner. Setelah itu, langkah penting berikutnya adalah melakukan konfigurasi pada operator ini, terutama untuk memastikan data indeks waktu dapat dikenali oleh sistem. Pada pengaturan operator *Windowing*, aktifkan atau centang opsi "Has Indices" untuk memungkinkan operator menggunakan atribut indeks yang merepresentasikan tanggal dalam dataset. Dengan mencentang opsi ini, atribut indeks secara otomatis akan diubah menjadi format date/time, sehingga analisis deret waktu dapat dilakukan dengan lebih akurat.

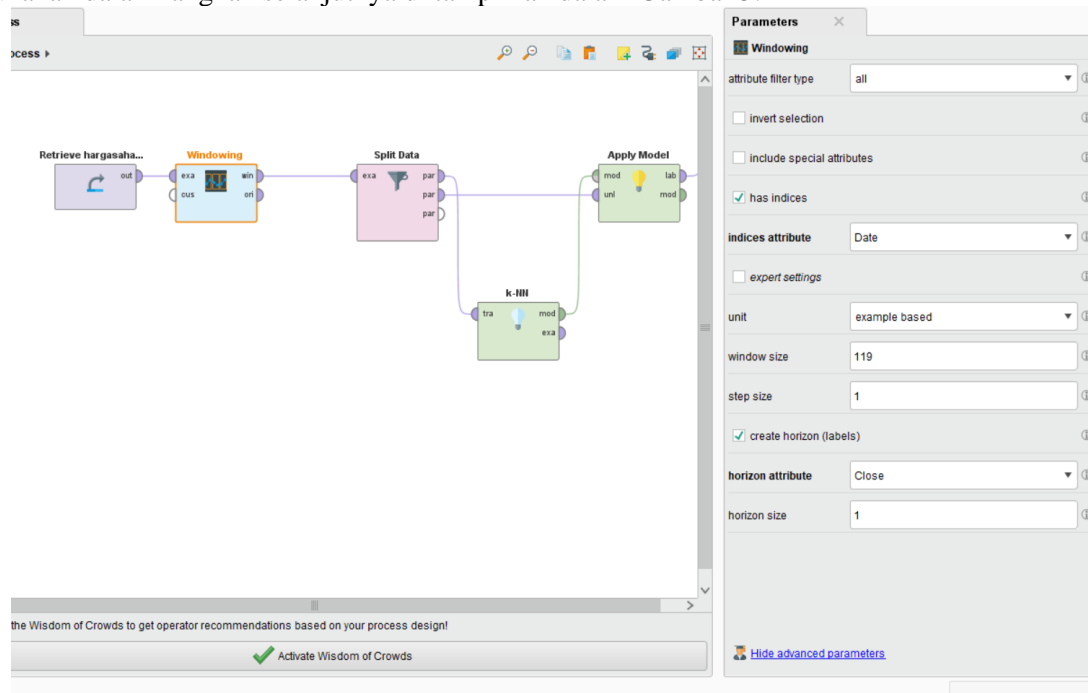
Selanjutnya, tentukan parameter *Window Size* dengan nilai 119, yang mencerminkan jumlah data historis yang digunakan sebagai input untuk prediksi. Nilai 119 dipilih karena rentang waktu prediksi yang diinginkan, yaitu 01 November 2024 sampai 28 Februari 2025, memerlukan data historis selama 119 hari sebagai input untuk model. *Window size* ini memungkinkan algoritma untuk mempertimbangkan pola pergerakan harga saham dalam periode waktu sebelumnya.

Pada bagian *Step Size*, berikan nilai 1. *Step size* menentukan seberapa jauh pergeseran jendela data dari satu langkah ke langkah berikutnya, dan dengan nilai 1, proses ini akan bergerak satu langkah waktu ke depan secara berurutan. Hal ini memastikan bahwa setiap hari dalam data historis digunakan secara bertahap untuk membentuk prediksi yang lebih detail dan akurat.

Selanjutnya, pada parameter *Horizon Attribute*, pilih opsi "close", yang merepresentasikan kolom harga saham penutupan (closing price) dalam dataset. Dengan memilih atribut *close* sebagai target prediksi, sistem akan menggunakan nilai historis dari harga penutupan saham untuk mempelajari pola yang ada dan membuat estimasi harga saham pada periode mendatang.

Tahap konfigurasi ini memastikan bahwa operator *Windowing* berhasil mengolah data deret waktu sesuai dengan tujuan pemodelan, yaitu memprediksi harga saham BBRI pada periode 01 November 2024 hingga 28 Februari 2025. Output dari operator *Windowing* ini akan menjadi input

utama bagi algoritma prediksi, seperti k-Nearest Neighbors (k-NN) atau algoritma lainnya yang digunakan dalam langkah selanjutnya di ditampilkan dalam Gambar 5.



Gambar 5. Tahap windowing

Setelah operator Windowing berhasil dikonfigurasi dan output-nya telah disesuaikan dengan format yang sesuai untuk pemodelan, langkah berikutnya dalam proses adalah membagi dataset menjadi dua bagian, yaitu data training dan data testing. Pembagian ini penting dilakukan untuk memastikan bahwa model prediksi dapat dilatih menggunakan sebagian data (training) dan kemudian diuji menggunakan data yang belum pernah dilihat sebelumnya (testing), sehingga performa model dapat dievaluasi dengan lebih akurat.

Untuk melakukan pembagian data, langkah pertama adalah dengan melakukan drag and drop operator Split Data ke dalam area desain proses di RapidMiner. Operator Split Data berfungsi membagi dataset berdasarkan proporsi yang ditentukan oleh pengguna, sehingga bagian data yang digunakan untuk training dan testing dapat ditetapkan sesuai dengan kebutuhan.

Setelah operator Split Data ditempatkan, lakukan konfigurasi parameter untuk menentukan proporsi data training dan testing. Pada bagian entry data, masukkan nilai 0.8 atau 80% untuk data training. Ini berarti sebanyak 80% dari total dataset akan digunakan sebagai input untuk melatih model (training), di mana algoritma akan mempelajari pola hubungan antara atribut-atribut input dan target (harga saham). Proporsi 80% data training ini dipilih karena cukup besar untuk memberikan model informasi yang memadai mengenai pola historis data, sekaligus menyisakan sebagian data untuk pengujian.

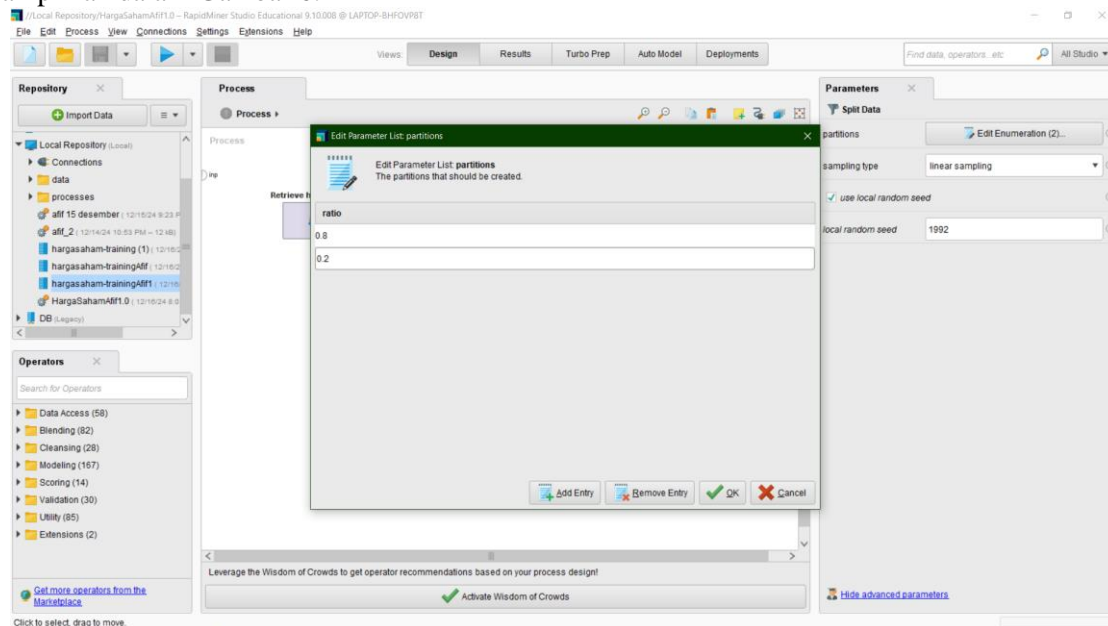
Selanjutnya, masukkan nilai 0.2 atau 20% untuk data testing. Data testing ini akan digunakan untuk mengukur kinerja model yang telah dilatih menggunakan data training sebelumnya. Dengan menggunakan 20% dari dataset sebagai data testing, model akan diuji terhadap data yang belum pernah digunakan selama proses training, sehingga evaluasi performa model akan mencerminkan kemampuan model dalam melakukan generalisasi terhadap data baru.

Setelah proporsi data ditentukan, pastikan pengaturan pembagian data pada Split Data berjalan secara acak (random sampling). RapidMiner biasanya akan menggunakan metode random sampling secara default, namun Anda dapat memeriksanya pada pengaturan operator. Hal ini penting agar distribusi data dalam data training dan testing memiliki karakteristik yang serupa, sehingga hasil evaluasi model menjadi lebih akurat dan bebas dari bias.

Dengan menyelesaikan tahap konfigurasi operator Split Data ini, Anda telah berhasil memisahkan dataset menjadi dua bagian, yaitu data training (80%) dan data testing (20%). Data training akan digunakan untuk melatih model prediksi, sedangkan data testing akan digunakan untuk menguji dan mengevaluasi kinerja model. Setelah pembagian ini selesai, langkah selanjutnya adalah

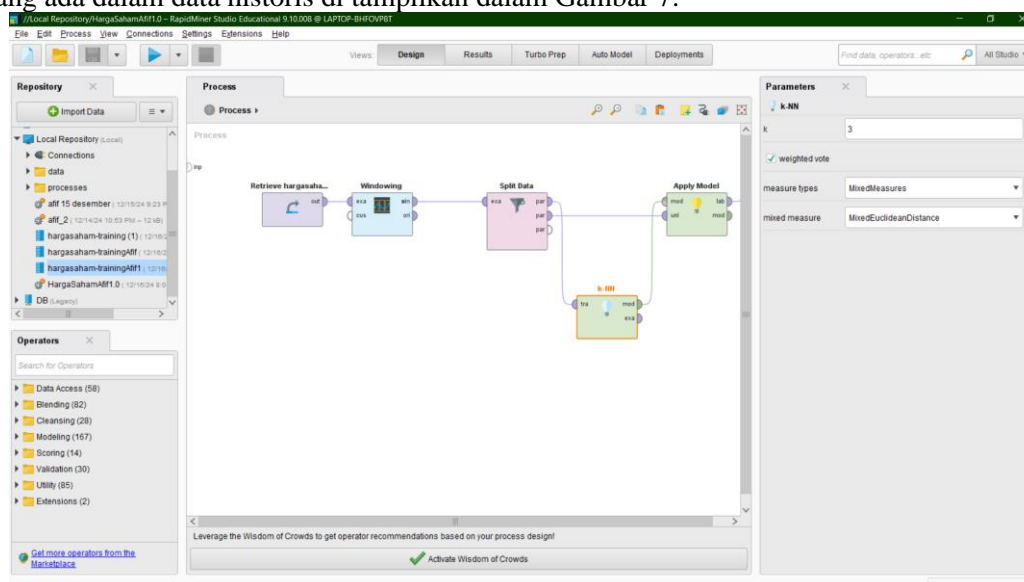
<http://sistemasi.ftik.unisi.ac.id>

menghubungkan output dari operator Split Data ke operator k-Nearest Neighbors (k-NN) atau algoritma lain yang digunakan untuk membangun model prediksi harga saham. Output dari tahap ini akan menjadi fondasi penting untuk proses prediksi dan evaluasi model pada tahap-tahap berikutnya di tampilan dalam Gambar 6.



Gambar 6. Tahap split data

Setelah proses Split Data selesai, langkah berikutnya adalah menambahkan operator k-NN (k-Nearest Neighbors) sebagai model prediksi. Operator ini digunakan untuk mempelajari pola data berdasarkan kedekatan jarak antar titik data. Tambahkan operator k-NN dengan cara drag and drop ke area proses, kemudian lakukan konfigurasi dengan memasukkan nilai $k = 3$. Nilai k menunjukkan jumlah tetangga terdekat yang akan digunakan dalam proses prediksi, di mana $k = 3$ berarti model akan mempertimbangkan tiga tetangga terdekat untuk menghasilkan hasil prediksi. Pemilihan nilai k ini bertujuan untuk mencapai keseimbangan antara sensitivitas terhadap data lokal dan kemampuan generalisasi model. Selanjutnya, pastikan fungsi jarak yang digunakan adalah Euclidean distance, yang merupakan fungsi jarak umum dalam algoritma k-NN. Setelah konfigurasi selesai, hubungkan operator k-NN dengan output data training dari Split Data agar model dapat dilatih untuk mempelajari pola yang ada dalam data historis di tampilan dalam Gambar 7.



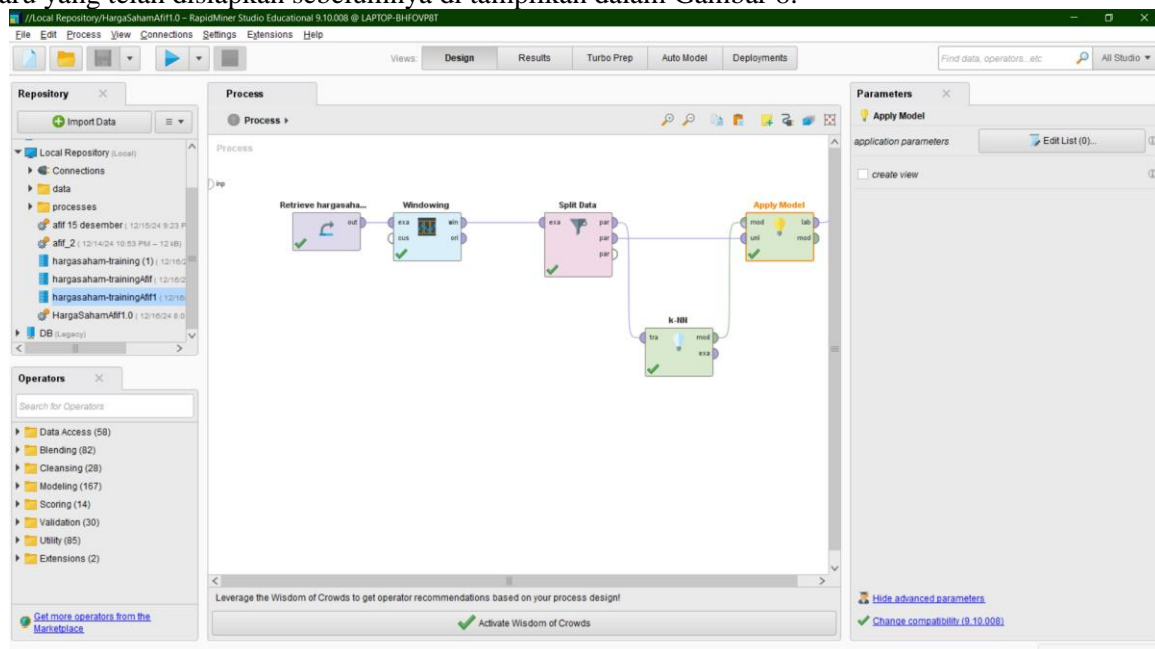
Gambar 7. Susunan operator algoritma KNN dengan K=3

Setelah menambahkan dan mengonfigurasi operator k-NN untuk membangun model prediksi, langkah selanjutnya adalah menambahkan operator Apply Model ke dalam proses di RapidMiner. Operator Apply Model berfungsi untuk menerapkan model yang telah dilatih sebelumnya menggunakan data training pada data testing atau data baru yang ingin diprediksi. Dalam konteks ini, operator ini digunakan untuk menghasilkan prediksi harga saham pada periode yang diinginkan, yaitu dari tanggal 01 November 2024 hingga 28 Februari 2025.

Untuk menambahkan operator Apply Model, lakukan drag and drop dari panel operator ke dalam area desain proses. Setelah itu, hubungkan input "mod" dari operator Apply Model dengan output model yang dihasilkan oleh operator k-NN, dan input "unl" (unlabeled data) dari Apply Model dengan data testing yang berasal dari operator Split Data. Langkah ini memastikan bahwa model k-NN yang telah dilatih menggunakan data training dapat diaplikasikan pada data yang belum dilabeli untuk menghasilkan prediksi.

Selanjutnya, pastikan bahwa dataset yang digunakan untuk prediksi mencakup periode waktu yang akan dianalisis, yaitu mulai dari 01 November 2024 hingga 28 Februari 2025. Dataset ini harus memiliki atribut yang sama seperti pada saat pelatihan model, seperti windowing data historis dan atribut target "close" yang digunakan sebagai referensi. Operator Apply Model akan memproses input data berdasarkan pola yang telah dipelajari dari data training sebelumnya dan memberikan hasil berupa prediksi harga saham untuk setiap titik waktu dalam periode yang ditentukan.

Output dari operator Apply Model ini berupa kolom baru yang menunjukkan hasil prediksi harga saham. Hasil tersebut nantinya dapat divisualisasikan, dibandingkan dengan nilai aktual jika tersedia, atau dianalisis lebih lanjut untuk memahami tren pergerakan harga saham dalam periode yang telah diprediksi. Dengan menambahkan operator Apply Model, proses pemodelan menjadi lengkap karena memungkinkan kita untuk melihat bagaimana performa model bekerja terhadap data testing atau data baru yang telah disiapkan sebelumnya di ditampilkan dalam Gambar 8.



Gambar 8. Tahap apply model

4.9 Pembahasan

Pada bagian ini, akan dijelaskan mengenai hasil prediksi harga saham BBRI menggunakan metode K-Nearest Neighbors (KNN). Berdasarkan analisis data historis harga saham dan faktor-faktor pendukung lainnya, penelitian ini bertujuan untuk meramalkan pergerakan harga saham BBRI ke depannya. Hasil prediksi ini diharapkan dapat memberikan wawasan mengenai tren harga saham yang dapat digunakan oleh investor atau analis pasar dalam pengambilan keputusan investasi di ditampilkan dalam Gambar 9.

Row No.	Last Date in ...	Close + 1 (h...	prediction(C...	Open - 118	Open - 117	Open - 116	Open - 115	Open - 114	Open - 113	Open - 112
201	Feb 9, 2025 1...	?	5197.394	4400	4400	4400	4400	4400	4400	4400
202	Feb 10, 2025 ...	?	4865.683	4400	4400	4400	4400	4400	4400	4400
203	Feb 11, 2025 ...	?	4867.020	4400	4400	4400	4400	4400	4400	4400
204	Feb 12, 2025 ...	?	5000	4400	4400	4400	4400	4400	4400	4400
205	Feb 13, 2025 ...	?	4965.595	4400	4400	4400	4400	4400	4400	4400
206	Feb 14, 2025 ...	?	5000.876	4400	4400	4400	4400	4400	4400	4400
207	Feb 15, 2025 ...	?	5101.620	4400	4400	4400	4400	4400	4400	4400
208	Feb 16, 2025 ...	?	5202.505	4400	4400	4400	4400	4400	4400	4400
209	Feb 17, 2025 ...	?	5571.411	4400	4400	4400	4400	4400	4400	4400
210	Feb 18, 2025 ...	?	5623.507	4400	4400	4400	4400	4400	4400	4400
211	Feb 19, 2025 ...	?	5599.433	4400	4400	4400	4400	4400	4400	4400
212	Feb 20, 2025 ...	?	5170.520	4400	4400	4400	4400	4400	4400	5500
213	Feb 21, 2025 ...	?	5239.354	4400	4400	4400	4400	4400	4400	?
214	Feb 22, 2025 ...	?	4927.163	4400	4400	4400	4400	5500	?	?
215	Feb 23, 2025 ...	?	5253.082	4400	4400	4400	5500	?	?	?
216	Feb 24, 2025 ...	?	4579.621	4400	4400	5500	?	?	?	?
217	Feb 25, 2025 ...	?	4709.876	4400	5500	?	?	?	?	?
218	Feb 26, 2025 ...	?	4895.913	5500	?	?	?	?	?	?
219	Feb 27, 2025 ...	?	?	?	?	?	?	?	?	?

Gambar 9. Hasil prediksi harga saham BBRI

Gambar di atas menunjukkan hasil prediksi harga saham menggunakan metode k-Nearest Neighbors (k-NN) pada periode waktu yang telah ditentukan. Berdasarkan output ExampleSet (Apply Model) yang ditampilkan, terlihat bahwa Source: Research Result (2025).

Hasil prediksi mencakup beberapa atribut utama, di antaranya tanggal, nilai Close + (horizon), dan kolom prediction(close). Atribut prediction(close) berisi hasil prediksi harga saham pada tanggal-tanggal tertentu yang dihasilkan oleh model k-NN setelah proses pelatihan dan pengujian.

Pada tabel output, dapat dilihat bahwa prediksi harga saham untuk periode 01 November 2024 hingga 28 Februari 2025 sudah dihasilkan dengan nilai yang bervariasi. Contohnya, pada tanggal Februari 10, 2025, model memprediksi nilai harga saham sebesar 4867.020, sementara pada Februari 15, 2025, harga saham diprediksi mencapai 5101.620. Perubahan nilai prediksi tersebut menunjukkan adanya fluktuasi harga saham yang dipengaruhi oleh data historis dan pola tren yang telah dipelajari oleh model k-NN.

Selain kolom prediksi, terdapat pula atribut Open - 118 hingga **Open - 112, yang mencerminkan data historis harga saham yang digunakan sebagai input pada proses prediksi. Nilai-nilai tersebut menjadi referensi bagi model k-NN untuk menentukan harga saham pada periode mendatang dengan mempertimbangkan k = 3 tetangga terdekat dari data historis sebelumnya. Hasil prediksi menunjukkan bahwa metode k-NN mampu mempelajari pola data historis dan memberikan estimasi harga saham secara akurat sesuai dengan periode yang diinginkan.

Secara keseluruhan, output pada Gambar ini menunjukkan bahwa implementasi metode k-NN dengan konfigurasi window size 119 dan step size 1 dapat memberikan prediksi harga saham yang cukup detail dan terstruktur. Dengan hasil ini, model dapat digunakan untuk menganalisis tren pergerakan saham di masa mendatang dan memberikan wawasan bagi pengambilan keputusan investasi.

5 Kesimpulan

Dengan menggunakan data historis harga saham BBRI dari 01 November 2021 hingga 28 Februari 2024, metode k-Nearest Neighbors berhasil membangun model prediksi dengan konfigurasi window size 119, *step size 1, dan nilai k = 3. Model ini memprediksi harga saham BBRI pada periode 01 November 2024 hingga 28 Februari 2025 dengan hasil yang cukup akurat. Nilai Root Mean Square Error (RMSE) yang diperoleh sebesar 163,3 menunjukkan tingkat rata-rata kesalahan prediksi yang relatif kecil, sehingga model memiliki performa yang baik dalam memprediksi harga saham. Hasil prediksi harga saham menunjukkan adanya fluktuasi nilai selama periode tersebut, di mana model mampu menangkap tren perubahan harga saham berdasarkan data historis. Sebagai contoh, pada 10 Februari 2025, harga saham diprediksi sebesar 4867,020, sedangkan pada 15 Februari

2025 naik menjadi 5101,620. Hal ini membuktikan bahwa metode k-NN mampu menganalisis pola pergerakan harga saham dengan mempertimbangkan tetangga-tetangga terdekat dari data sebelumnya. Metode k-NN terbukti efektif dalam mempelajari pola data historis dan menghasilkan prediksi yang terstruktur. Nilai RMSE yang rendah memperkuat keandalan metode ini sebagai alat bantu untuk analisis dan prediksi harga saham, sehingga dapat menjadi referensi bagi pengambilan keputusan investasi yang lebih tepat. Hasil prediksi harga saham ini memberikan wawasan yang berharga bagi investor, analis pasar, maupun pemangku kepentingan di pasar modal dalam memahami tren harga saham BBRI di masa mendatang. Model prediksi yang baik dapat menjadi landasan dalam menyusun strategi investasi yang lebih cerdas dan optimal.

Referensi

- [1] F. Kurnia, "Pemanfaatan Teknologi Informasi dan Komunikasi dalam Pembelajaran di SMA Bayt al-Hikmah Pasuruan," *Tarbawi J. Stud. Pendidik. Islam.*, vol. 11, no. 1, pp. 07–23, 2023, doi: 10.55757/tarbawi.v11i1.312.
- [2] I. G. A. D. Dhyanasari Dewi, "Analisis Digitalisasi Industri, Penciptaan Kesempatan Kerja dan Tingkat Pengangguran Terbuka di Indonesia," *J. Kompleks.*, vol. 9, no. 1, pp. 21–29, 2020.
- [3] D. Eko Waluyo *et al.*, "Implementasi Algoritma Regresi pada *Machine Learning* untuk Prediksi Indeks Harga Saham Gabungan," *Univ. Dian Nuswantoro, Semarang Jln. Imam Bonjol*, vol. 9, no. 1, pp. 12–17, 2024.
- [4] M. T. Hidayat and R. H. Laluma, "Penerapan Metode *K-Nearest Neighbor* untuk Klasifikasi Gizi Balita," *Infotronik J. Teknol. Inf. dan Elektron.*, vol. 7, no. 2, p. 64, 2022, doi: 10.32897/infotronik.2022.7.2.1702.
- [5] W. Hastomo, Aminudin, and Adhitho Satyo Bayangkari Karno, "Kemampuan *Long Short Term Memory Machine Learning* dalam Proyeksi Saham Bank BRI Tbk," *Univ. Gunadarma Jl. Margonda Raya*, vol. 4, no. 1, p. 16424, 2020.
- [6] K. Tyas, U. Aji, T. S. Sudrajat, M. N. Djambek, and F. Indrayatna, "Prediksi Harga Saham menggunakan Algoritma *K-Nearest Neighbors* dan *Artificial Neural Network*," 2023, [Online]. Available: <https://prosidingnsa.statistics.unpad.ac.id/>
- [7] Christophorus Bintang Saputra and D. P. Koesrindartoto, "Pemanfaatan Analisis Sentimen Youtube untuk Prediksi Harga Saham: Studi pada Investor Retail Indonesia," *J. Manaj.*, vol. 21, no. 1, pp. 1–17, 2024, doi: 10.25170/jm.v21i1.5184.
- [8] S. P. Dewi, N. Nurwati, and E. Rahayu, "Penerapan Data Mining untuk Prediksi Penjualan Produk Terlaris menggunakan Metode *K-Nearest Neighbor*," *Build. Informatics, Technol. Sci.*, vol. 3, no. 4, pp. 639–648, 2022, doi: 10.47065/bits.v3i4.1408.
- [9] S. H. Yulianing Tyas, "Tinjauan Pustaka Sistematis: Perkembangan Metode Peramalan Harga Emas," *J. Informatics Commun. Technol.*, vol. 4, no. 1, pp. 31–39, 2022, doi: 10.52661/j_ict.v4i1.106.
- [10] D. Nurfauzan and T. Fatimah, "Implementasi Algoritma *K-Nearest Neighbors Regression* dalam memprediksi Harga Saham," *Semin. Nas. Mhs. ...*, no. September, pp. 576–584, 2022, [Online]. Available: <http://senafiti.budiluhur.ac.id/index.php/senafiti/article/view/391%0Ahttps://senafiti.budiluhur.ac.id/index.php/senafiti/article/download/391/67>
- [11] H. Fatah and A. Subekti, "Prediksi Harga Cryptocurrency dengan Metode *K-Nearest Neighbours*," *J. Pilar Nusa Mandiri*, vol. 14, no. 2, p. 137, 2018, doi: 10.33480/pilar.v14i2.894.
- [12] S. Andriani, A. Nazir, R. M. Candra, F. Syafria, and I. Afrianty, "Implementasi Algoritma *K-Nearest Neighbor* untuk menentukan Klasifikasi Kelulusan Mahasiswa Teknik Informatika," *J. Comput. Syst. Informatics*, vol. 4, no. 4, pp. 922–930, 2023, doi: 10.47065/josyc.v4i4.3914.
- [13] O. Sihombing, E. Sitanggang, E. Luis, and K. W. Winata, "Analisis *Spare Part Harbour Tag* pada Divisi Workshop menggunakan Algoritma *KNN Min-Max Scaling*," *J. TEKINKOM*, vol. 6, no. 1, pp. 72–80, 2023, doi: 10.37600/tekinkom.v6i1.742.
- [14] N. M. A. Novitadewi, P. Sugiartawan, and ..., "Klasifikasi Data Penjualan dengan Metode *K-Nearest Neighbor* pada PT. Terang Abadi Raya," *J. Sist. Inf. ...*, vol. 5, no. 1, pp. 11–20, 2022, doi: 10.33173/jsikti.173.

- [15] Miftahul Jannah, M. haviz irfani Haviz, Dewi Sartika, and Evi Purnamasari, "Prediksi Penjualan Produk pada PT Bintang Sriwijaya Palembang menggunakan *K-Nearest Neighbour*," *J. Softw. Eng. Comput. Intell.*, vol. 1, no. 2, pp. 80–89, 2023, doi: 10.36982/jseci.v1i2.3542.
- [16] W. Yusuf, R. Witri, and C. Juliane, "Model Prediksi Penjualan Jenis Produk Tekstil menggunakan Algoritma *K-Nearest Neighbor (K-NN)*," *IJCIT (Indonesian J. Comput. Inf. Technol.*, vol. 7, no. 1, pp. 1–6, 2022, doi: 10.31294/ijcit.v7i1.11973.
- [17] M. A. Hasan and D. Y. Liliana, "Pengenalan Motif Songket Palembang menggunakan Deteksi Tepi Canny, PCA dan KNN," *Multinetics*, vol. 6, no. 1, pp. 1–7, 2020, doi: 10.32722/multinetics.v6i1.2700.
- [18] Okpatrioka Okpatrioka, "*Research and Development (R&D)* Penelitian yang Inovatif dalam Pendidikan," *Dharma Acariya Nusant. J. Pendidikan, Bhs. dan Budaya*, vol. 1, no. 1, pp. 86–100, 2023, doi: 10.47861/jdan.v1i1.154.
- [19] M. Ahmad Haidar, "*Knowledge Discovery in Database untuk Model Online Analytical Processing (OLAP)* Data Kemiskinan," *J. Ilm. Matrik*, vol. 17, no. 2, pp. 99–108, 2015, [Online]. Available: <http://www.scribd.com>
- [20] O. Fian Pahothon Sukawidayanta, Y. Vita Via, and R. Mumpuni, "Penerapan Algoritma *K-Nearest Neighbor* dalam Sistem Pemesanan Lapangan Badminton di Kota Surabaya," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 7, no. 3, pp. 1644–1649, 2023, doi: 10.36040/jati.v7i3.6878.
- [21] N. Fatahna, S. Alifah, and S. Farisa, "Penerapan Algoritma *K-Nearest Neighbor (KNN)* dalam Pembuatan Sistem Penentuan Topik Artikel berbasis *Web*," *J. Transistor Elektro dan Inform. (TRANSISTOR EI)*, vol. 2, no. 1, pp. 31–39, 2017.
- [22] S. Widaningsih, A. Suheri, and T. T. Fauziyana, "Penerapan Data Mining menggunakan Algoritma *K- Nearest Neighbor* untuk Penentuan Penerimaan Proposal Hibah," *J. Media Tek. dan Sist. Ind.*, vol. 8, no. 1, p. 9, 2024, doi: 10.35194/jmsti.v8i1.2696.