JRIIN : Jurnal Riset Informatika dan Inovasi
Volume 2, No. 11 April Tahun 2025
ISSN 3025-0919 (media online)
Hal 2021-2028

Analisis Dan Prediksi Rating Game Berdasarkan Review Pengguna Steam Menggunakan Algoritma K–Nearest Neighbor

Maulana Fansyuri[1*], Abdul Hanif[2], Alfreza Routya Faizan[3], Fadhil Nata Pratama[4], Verrel Aulia Rahman[5]

[1,2,3,4,5,6] Fakultas Ilmu Komputer, Program Studi Teknik Informatika, Universitas Pamulang, Tangerang Selatan, Indonesia
Email: [1*]dosen02359@unpam.ac.id, [2]abdulhanif1803@gmail.com, [3]alfrezarf@gmail.com, [4]fadhilpratama798@gmail.com, [5]verrelaulia@gmail.com
(* : coressponding author)

Abstrak − Industri game digital terus berkembang pesat, dengan platform seperti Steam yang menawarkan ribuan game kepada jutaan pengguna. Namun banyaknya pilihan membuat pengguna kesulitan menemukan game yang sesuai dengan kebutuhannya. Penelitian ini bertujuan untuk menganalisis dan memprediksi rekomendasi pengguna untuk rating game di Steam menggunakan algoritma K-nearest neighbour (KNN). Kumpulan data yang digunakan terdiri dari 2.800 review pengguna terhadap 7 game yang diperoleh dari Kaggle.com. Melakukan proses pengolahan data secara sistematis, meliputi pembersihan data, transformasi teks, dan segmentasi data menjadi data latih dan data uji. Algoritma KNN dengan k=5 diimplementasikan menggunakan software RapidMiner yang mendukung analisis data halus dengan metode non parametrik. Hasil penelitian menunjukkan bahwa algoritma KNN dapat memprediksi statistik permainan dengan akurat. Penggunaan pemrosesan data terstruktur dan perangkat lunak pendukung menghasilkan model klasifikasi label berpemilik. Selain itu, penelitian ini menunjukkan potensi pengembangan model lebih lanjut dengan memperluas kumpulan data, memasukkan lebih banyak variabel, dan membandingkan berbagai variabel untuk penilaian kinerja. Selain memberikan solusi praktis, penelitian ini juga memberikan kerangka teoritis untuk penerapan algoritma KNN pada sistem rekomendasi big data, yang dapat meningkatkan pengalaman pengguna dalam mengeksplorasi game baru.

Kata Kunci: Sistem Rekomendasi, K-Nearest Neighbors (KNN), Steam, Prediksi Rating, Ulasan Pengguna, Personalisasi

Abstract− *The digital gaming industry continues to grow rapidly, with platforms like Steam offering thousands of games to millions of users. However, the abundance of options makes it difficult for users to find games that suit their needs. This research aims to analyse and predict user recommendations for game ratings on Steam using the K-Nearest Neighbours (KNN) algorithm. The dataset used consists of 2,800 user reviews of 7 games obtained from Kaggle.com. A systematic data processing approach was conducted, including data cleaning, text transformation, and segmentation into training and testing datasets. The KNN algorithm with k=5 was implemented using RapidMiner software, which supports fine-grained data analysis with non-parametric methods. The study results indicate that the KNN algorithm can accurately predict game statistics. The use of structured data processing and supporting software produced a proprietary label classification model. Furthermore, the study highlights the potential for further model development by expanding the dataset, incorporating more variables, and comparing various variables for performance evaluation. In addition to providing practical solutions, this research offers a theoretical framework for applying the KNN algorithm to big data recommendation systems, which can enhance user experience in exploring new games.*

Keywords*: Recommendation System, K-Nearest Neighbors (KNN), Steam, Rating Prediction, User Reviews, Personalization*

1. PENDAHULUAN

Industri permainan digital terus berkembang pesat, terutama dengan adanya platform distribusi digital seperti Steam yang menawarkan ribuan judul permainan kepada jutaan pengguna di seluruh dunia. Meskipun platform ini memudahkan akses ke berbagai permainan, banyaknya pilihan yang tersedia sering kali membuat pengguna kesulitan dalam memilih permainan yang sesuai dengan preferensi mereka. Hal ini menjadi tantangan besar, mengingat preferensi setiap pengguna sangat beragam dan dinamis. Oleh karena itu, dibutuhkan sistem yang dapat memberikan rekomendasi permainan yang relevan berdasarkan preferensi dan riwayat permainan pengguna. Sistem rekomendasi berbasis kecerdasan buatan (Ricci et al., 2011) telah menjadi solusi yang umum

JRIIN : Jurnal Riset Informatika dan Inovasi
Volume 2, No. 11 April Tahun 2025
ISSN 3025-0919 (media online)
Hal 2021-2028

digunakan untuk membantu pengguna menemukan permainan yang sesuai dengan keinginan mereka, meskipun banyak pilihan yang tersedia.

Sistem rekomendasi berbasis data besar dapat memanfaatkan data pengguna, termasuk ulasan, rating, dan perilaku bermain, untuk memprediksi permainan yang disukai oleh pengguna. Salah satu algoritma yang banyak digunakan dalam sistem rekomendasi adalah K-Nearest Neighbors (KNN), yang mengukur kesamaan antar data melalui metrik jarak, seperti Euclidean distance (Han et al., 2012). KNN dipilih karena kemampuannya dalam menangani data non-linear dan fleksibilitasnya dalam memprediksi rating atau preferensi berdasarkan data yang ada. Algoritma ini juga dapat diterapkan pada data besar, seperti yang terdapat di platform Steam, untuk menghasilkan rekomendasi yang personal dan akurat.

Penelitian ini bertujuan untuk mengembangkan dan menganalisis sistem rekomendasi permainan di platform Steam menggunakan algoritma KNN. Fokus penelitian ini adalah pada prediksi rating permainan dan rekomendasi berdasarkan ulasan pengguna di Steam. Dataset yang digunakan dalam penelitian ini terdiri dari 2.800 ulasan yang diambil dari tujuh permainan yang diperoleh melalui Kaggle.com. Dataset ini mencakup informasi tentang rating permainan, ulasan yang diberikan oleh pengguna, serta durasi waktu bermain, yang penting dalam membangun model rekomendasi yang akurat. Proses analisis data dilakukan melalui tahapan preprocessing yang mencakup pembersihan data, transformasi teks, dan pembagian data menjadi data latih dan data uji (Tan et al., 2018).

Salah satu tantangan utama dalam membangun sistem rekomendasi berbasis KNN adalah mengatasi kompleksitas data non-linear yang sering muncul dalam ulasan pengguna. Untuk itu, penelitian ini mengimplementasikan teknik text vectorization untuk mengubah teks ulasan menjadi representasi numerik yang dapat diproses lebih lanjut oleh algoritma KNN. Dengan menggunakan perangkat lunak RapidMiner, algoritma KNN diterapkan dengan nilai parameter k = 5, yang kemudian diuji untuk mengetahui tingkat akurasi prediksi rating yang dihasilkan. Hasil yang diharapkan adalah kemampuan algoritma KNN untuk memberikan prediksi yang relevan dan akurat tentang rating permainan yang dapat disarankan kepada pengguna.

Penelitian ini juga berkontribusi pada pengembangan kajian teoritik dalam penerapan algoritma KNN untuk sistem rekomendasi berbasis data besar, khususnya dalam konteks platform permainan digital. Hasil penelitian ini diharapkan dapat memberikan wawasan baru mengenai penerapan algoritma KNN dalam sistem rekomendasi dan memberikan kontribusi terhadap pengembangan teknologi personalisasi yang lebih efektif di platform digital seperti Steam. Penelitian ini juga membuka peluang untuk penelitian lebih lanjut yang dapat mengembangkan model ini dengan memperbesar variasi dataset dan mengintegrasikan algoritma lain untuk meningkatkan performa model rekomendasi (Aggarwal, 2016).

Sistem rekomendasi berbasis KNN termasuk dalam kategori sistem rekomendasi berbasis kolaboratif, yang memanfaatkan data pengguna lain untuk memberikan saran yang relevan. Sebagai metode non-parametrik, KNN tidak memerlukan asumsi mengenai distribusi data, yang menjadikannya fleksibel untuk diterapkan pada berbagai jenis data, termasuk data ulasan pengguna yang tidak terstruktur (Han et al., 2012). Penelitian sebelumnya telah menunjukkan efektivitas algoritma KNN dalam sistem rekomendasi, baik dalam konteks produk e-commerce maupun hiburan digital. Penelitian ini bertujuan untuk menguji dan mengembangkan penggunaan algoritma KNN dalam konteks platform permainan digital, dengan memanfaatkan dataset ulasan pengguna dari Steam untuk memberikan rekomendasi yang lebih akurat dan relevan.

2. METODE

Penelitian ini mengadopsi pendekatan berbasis algoritma K-Nearest Neighbors (KNN) untuk menganalisis dan memprediksi rating permainan berdasarkan ulasan pengguna pada platform distribusi permainan Steam. Tahapan yang dilakukan dalam penelitian ini terdiri dari pengumpulan data, implementasi algoritma, preprocessing data, serta evaluasi model. Berikut adalah penjelasan lebih rinci mengenai tiap tahapan yang dilakukan:

JRIIN : Jurnal Riset Informatika dan Inovasi
Volume 2, No. 11 April Tahun 2025
ISSN 3025-0919 (media online)
Hal 2021-2028

2.1 Pengumpulan Data

Pengumpulan data merupakan langkah awal yang penting dalam penelitian ini, karena data yang relevan dan berkualitas akan sangat mempengaruhi hasil analisis dan prediksi yang dihasilkan. Data yang digunakan dalam penelitian ini diperoleh dari Kaggle.com, sebuah platform sumber terbuka yang menawarkan berbagai dataset untuk penelitian data sains dan pembelajaran mesin. Kaggle memiliki reputasi yang baik dalam menyediakan data yang terstruktur dan telah diproses dengan baik, yang mempermudah peneliti dalam menganalisis masalah tertentu tanpa harus memulai dari awal.

Dalam penelitian ini, dataset yang digunakan terdiri dari ulasan pengguna yang berasal dari tujuh game yang berbeda di Steam, dengan masing-masing game memiliki 400 ulasan, sehingga total data yang digunakan mencapai 2.800 ulasan. Setiap entri dalam dataset ini mencakup beberapa atribut penting, antara lain: ID permainan, nama permainan, jumlah rekomendasi positif yang diterima permainan, jumlah total ulasan yang tersedia, rating rate, nilai rating, serta label rating. Data ini penting karena memungkinkan peneliti untuk menganalisis pola ulasan dan preferensi pengguna serta membuat prediksi yang lebih tepat tentang rating yang akan diberikan oleh pengguna kepada permainan tersebut. Data ini juga memungkinkan peneliti untuk mengidentifikasi hubungan antara faktor-faktor seperti jumlah ulasan, rating rate, dan rekomendasi positif dengan rating yang diberikan oleh pengguna.

2.2 Algoritma K-Nearest Neighbor (KNN)

Algoritma yang digunakan dalam penelitian ini adalah K-Nearest Neighbors (KNN) dengan nilai k = 5. KNN adalah metode non-parametrik yang banyak digunakan dalam bidang klasifikasi dan regresi. Keunggulan utama dari KNN adalah kesederhanaannya dalam menentukan kelas atau nilai prediksi berdasarkan kedekatannya dengan data lain yang sudah dikenal. Prinsip dasar dari KNN adalah, ketika diberikan data baru yang belum memiliki label, algoritma ini akan mencari k data terdekat (tetangga) dalam ruang fitur dan kemudian memberikan label atau nilai berdasarkan mayoritas dari k tetangga terdekat tersebut.

Dalam konteks penelitian ini, KNN digunakan untuk memprediksi rating permainan berdasarkan ulasan yang diberikan oleh pengguna. Nilai k diatur sebesar 5, yang berarti bahwa algoritma akan mempertimbangkan lima tetangga terdekat untuk memberikan prediksi rating untuk permainan yang sedang dianalisis. KNN dipilih karena kemampuannya dalam menangani data non-linear, serta fleksibilitasnya dalam menangani data yang kompleks, seperti data ulasan pengguna yang berbentuk teks. Dengan menggunakan KNN, peneliti berharap dapat memanfaatkan kesamaan antara ulasan pengguna untuk memprediksi rating yang akan diberikan kepada permainan yang belum memiliki rating.

KNN juga memiliki keunggulan dalam penerapannya pada dataset besar seperti yang digunakan dalam penelitian ini, karena tidak membutuhkan pembelajaran model yang rumit. Sebagai metode instance-based learning, KNN menyimpan semua data yang ada dan membuat prediksi secara langsung saat diperlukan, membuatnya sangat efisien untuk dataset dengan ukuran yang besar. Proses pembelajaran pada KNN terjadi saat tahap prediksi, bukan selama pelatihan, yang membuat algoritma ini cocok untuk situasi di mana data tidak sepenuhnya diketahui sebelumnya**.**

2.3 Tahapan Preprocessing Data

Preprocessing data adalah salah satu langkah paling penting dalam penelitian ini, karena kualitas data yang digunakan sangat memengaruhi hasil analisis. Proses ini bertujuan untuk menyiapkan data agar siap digunakan dalam algoritma KNN, dengan memastikan bahwa data tersebut bersih, terstruktur dengan baik, dan relevan dengan tujuan penelitian. Preprocessing dalam penelitian ini terdiri dari beberapa tahapan penting, yaitu pembersihan data, transformasi teks, dan pembagian data menjadi data latih dan data uji.

2.3.1 Pembersihan Data (Data Cleaning)

Langkah pertama dalam preprocessing adalah pembersihan data, yang bertujuan untuk menghilangkan data yang duplikat atau tidak relevan dan menangani nilai yang hilang (missing

JRIIN : Jurnal Riset Informatika dan Inovasi
Volume 2, No. 11 April Tahun 2025
ISSN 3025-0919 (media online)
Hal 2021-2028

values). Dalam dataset ini, jika ditemukan data yang tidak lengkap atau formatnya tidak sesuai, langkah-langkah perbaikan dilakukan melalui teknik imputasi atau dengan menghapus entri yang tidak memberikan kontribusi penting terhadap analisis. Selain itu, dilakukan juga normalisasi data untuk memastikan konsistensi format. Misalnya, konversi huruf kapital menjadi huruf kecil dan penghapusan spasi atau karakter yang tidak diperlukan dilakukan untuk memastikan bahwa data memiliki format yang seragam. Data yang tidak terstruktur atau memiliki inkonsistensi dapat memengaruhi hasil analisis dan prediksi, oleh karena itu langkah ini sangat krusial.

2.3.2 Transformasi Teks

Karena dataset berisi ulasan dalam bentuk teks, diperlukan langkah-langkah untuk mengubah teks menjadi format numerik yang dapat diproses oleh algoritma KNN. Proses ini dimulai dengan tokenisasi, yaitu memecah teks ulasan menjadi kata-kata atau token yang lebih kecil. Tokenisasi membantu algoritma untuk memecah teks menjadi unit-unit yang lebih mudah diproses dan dianalisis. Selanjutnya, stopwords (kata-kata yang tidak membawa informasi penting, seperti "dan", "atau", dan "yang") dihapus untuk mengurangi dimensi data yang tidak relevan. Stemming dilakukan untuk mengubah kata menjadi bentuk dasar, misalnya mengubah kata "memainkan" menjadi "main", untuk mengurangi variasi kata yang tidak diperlukan. Setelah itu, teks yang telah diproses diubah menjadi representasi numerik menggunakan metode Term Frequency-Inverse Document Frequency (TF-IDF). Metode ini memberikan bobot pada kata-kata yang sering muncul di teks tetapi jarang ditemukan di seluruh dokumen, sehingga menekankan kata-kata yang penting dan relevan.

2.3.3 Pemisahan Data (Data Splitting)

Setelah data diproses, langkah selanjutnya adalah membagi dataset menjadi dua bagian, yaitu data latih dan data uji. Data latih akan digunakan untuk melatih model, sedangkan data uji akan digunakan untuk mengevaluasi kinerja model. Dalam penelitian ini, pembagian data dilakukan dengan rasio 87.5% untuk data latih dan 12.5% untuk data uji, di mana 2.800 data digunakan untuk melatih model dan 400 data digunakan untuk menguji model. Pembagian data yang proporsional ini memastikan bahwa model memiliki cukup data untuk belajar dan dapat diuji dengan data yang belum pernah dilihat sebelumnya, sehingga evaluasi model lebih objektif.

2.3.4 Implementasi Algoritma KNN

Setelah data diproses, algoritma KNN diterapkan dengan nilai k=5. Dalam hal ini, prediksi rating permainan dilakukan dengan membandingkan ulasan baru dengan 5 ulasan terdekat yang ada dalam dataset. Proses pengukuran jarak antar data dilakukan menggunakan Euclidean distance, yang mengukur jarak antara dua titik dalam ruang fitur. Setelah jarak antar data dihitung, algoritma akan memilih 5 tetangga terdekat dan memberikan prediksi berdasarkan mayoritas rating dari tetangga tersebut.

Implementasi algoritma dilakukan menggunakan perangkat lunak RapidMiner, yang dipilih karena kemampuannya dalam melakukan analisis data secara mendalam, mengimplementasikan algoritma secara efisien, serta mendukung evaluasi model yang mudah. RapidMiner memungkinkan analisis yang cepat dan efektif tanpa perlu menulis banyak kode, sehingga peneliti dapat lebih fokus pada aspek pemodelan dan evaluasi.

2.3.5 Alat yang Digunakan

Dalam Penelitian Ini, Rapidminer Digunakan Sebagai Perangkat Lunak Utama Untuk Analisis Data Dan Penerapan Algoritma Knn. Rapidminer Dipilih Karena Kemampuannya Dalam Mengelola Dan Menganalisis Data Secara Efisien, Serta Memungkinkan Pengguna Untuk Melakukan Penerapan Algoritma Dengan Mudah Tanpa Memerlukan Keterampilan Pemrograman Yang Mendalam. Selain Itu, Rapidminer Memiliki Fitur Visualisasi Yang Membantu Peneliti Untuk Memahami Hasil Analisis Dengan Lebih Baik Dan Memberikan Panduan Dalam Mengevaluasi Performa Model Yang Dihasilkan.

JRIIN : Jurnal Riset Informatika dan Inovasi
Volume 2, No. 11 April Tahun 2025
ISSN 3025-0919 (media online)
Hal 2021-2028

3. ANALISA DAN PEMBAHASAN

3.1 Dataset

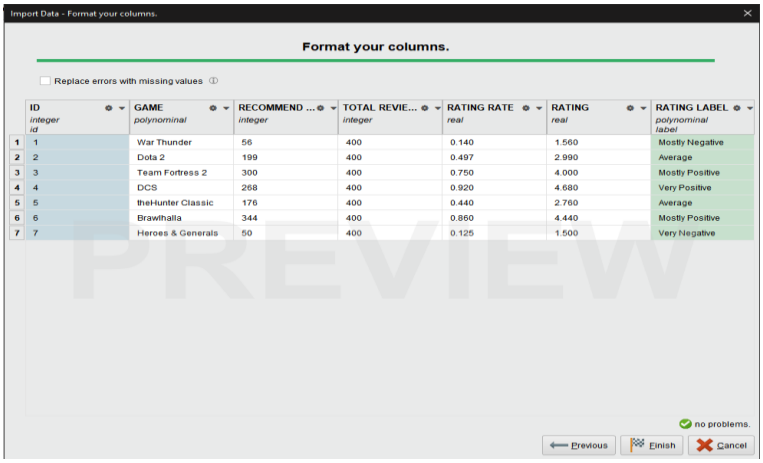
Tahapan pertama dalam analisis ini adalah memilih data yang bersumber dari situs www.kaggle.com dengan judul *"Game Rating Analysis"*. Dataset ini digunakan dalam analisis ulasan pengguna yang berasal dari tujuh game berbeda di Steam. Setiap game memiliki 400 ulasan, sehingga total data yang dianalisis mencapai 2.800 ulasan. Penelitian ini menggunakan satu dataset yang terdiri dari 7 game dengan 7 atribut. Dari 7 atribut ini, 5 atribut bertipe numerik dan 2 atribut bertipe kategorik.

Tabel 1. Jenis jenis database

Atribut	Tipe	Keterangan
ID	Numerik	ID unik untuk setiap game
GAME	Kategorik	Nama game yang dianalisis
RECOMMEND COUNT	Numerik	Jumlah ulasan yang merekomendasikan game
TOTAL REVIEW	Numerik	Total jumlah ulasan yang diterima game
RATING RATE	Numerik	Rasio antara jumlah rekomendasi dan total ulasan
RATING	Numerik	Skor rating yang dihitung berdasarkan jumlah rekomendasi
RATING LABEL	Kategorik	Kategori penilaian game (Mostly Negative, Average, Mostly Positive, Very Positive)

3.2 Transformation (Transformasi Data)

Data training menggunakan total 2.800 data game yang terbagi menjadi 7 game yang berbeda, yaitu: War Thunder, DOTA 2, Team Fortress 2, DCS, theHunter Classic, Brawlhalla, dan Heroes & General yang akan terlihat seperti berikut:

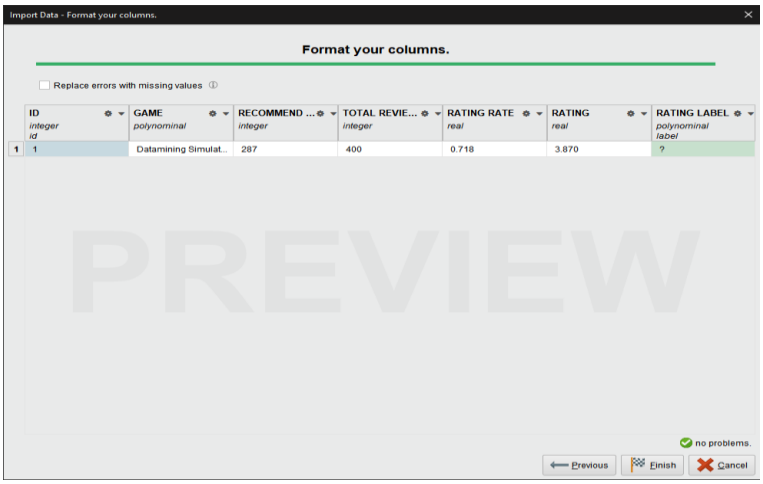


	ID integer id	GAME polynominal	RECOMMEND ... integer	TOTAL REVIE... integer	RATING RATE real	RATING real	RATING LABEL polynominal label
1	1	War Thunder	56	400	0.140	1.560	Mostly Negative
2	2	Dota 2	199	400	0.497	2.990	Average
3	3	Team Fortress 2	300	400	0.750	4.000	Mostly Positive
4	4	DCS	268	400	0.920	4.680	Very Positive
5	5	theHunter Classic	176	400	0.440	2.760	Average
6	6	Brawlhalla	344	400	0.860	4.440	Mostly Positive
7	7	Heroes & Generals	50	400	0.125	1.500	Very Negative

Gambar 1. Format Column Dataset Training

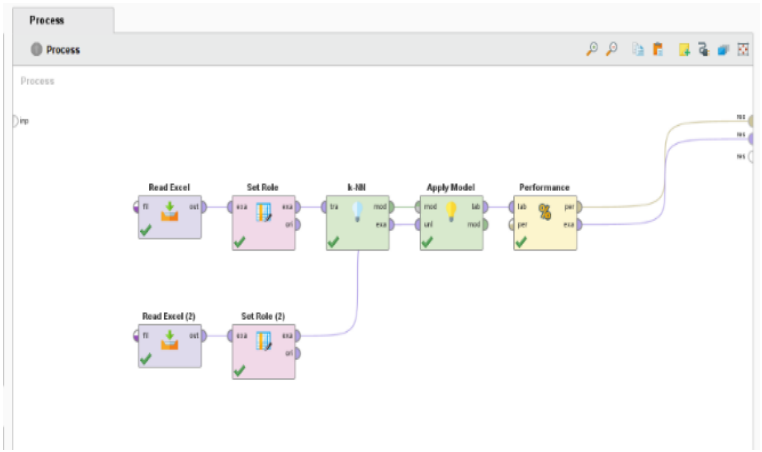
JRIIN : Jurnal Riset Informatika dan Inovasi
Volume 2, No. 11 April Tahun 2025
ISSN 3025-0919 (media online)
Hal 2021-2028

Data testing menggunakan total 400 data. Setiap data memiliki atribut: id, game (Datamining Simulator 2024) , recommend positive, total reviews, rating rate, rating, dan rating label sebagai atribut yang akan diprediksi. Sehingga tamplan data testing akan terlihat seperti berikut.



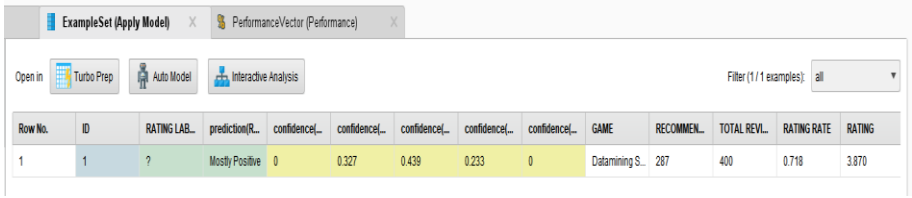
Gambar 2. Format Dataset Testing

Design yang digunakan pada RapidMiner menggunakan 2 buah read excel untuk data latih dan data uji, 2 buah set role untuk menentukan label atau keterangan, 1 buah knn untuk penerapan algoritma, 1 buah apply model, 1 buah apply model, dan 1 buah klasifikasi performance. Sehingga akan terlihat seperti berikut:



Gambar 3. Model View RapidMiner

Setelah model selesai dibuat dan seluruh data dimasukkan, langkah selanjutnya adalah mengeksekusinya untuk memperoleh hasil



Gambar 1. Hasil Apply Model

JRIIN : Jurnal Riset Informatika dan Inovasi
Volume 2, No. 11 April Tahun 2025
ISSN 3025-0919 (media online)
Hal 2021-2028

3.3 Evaluasi

Berdasarkan hasil dari gambar 4, didapatkan hasil prediksi data dengan ID 1 mendapatkan rating "Mostly Positive" dengan tingkat kepercayaan sebesar 43,9%. hasil predikisi ini didapatkan dengan melihat tingkat kepercayaan pada tabel convidence dan nilai paling tinggi ada pada tabel confidence (mostly positive) dengan nilai 0.429 atau 42.9%. sedangkan tabel confidence (avarge) mempunyai nilai 0.327 atau 32.7% dan tabel dan tabel confidence (very positive) yang hanya mempunyai nilai 0.233 atau 23.3%.Data ini didapat dari data "Datamining Simulator" dengan 287 recommend, 400 total reviews, 0,718 rating rate, dan 3.870 rating

4. KESIMPULAN

Hasil analisis menunjukkan bahwa penerapan algoritma K-Nearest Neighbor (K-NN) mampu memprediksi rating game berdasarkan ulasan pengguna di platform Steam dengan tingkat akurasi yang memadai. Model yang dikembangkan memanfaatkan preprocessing data yang sistematis serta implementasi algoritma yang terstruktur, menghasilkan kinerja yang efektif dalam klasifikasi rating. Penggunaan perangkat lunak RapidMiner juga terbukti andal dalam mendukung proses analisis dan pengolahan data.Peneliti menyarankan untuk menambah jumlah dan variasi dataset guna meningkatkan akurasi serta kemampuan generalisasi model. Eksplorasi algoritma lain, seperti Random Forest atau Neural Network, juga direkomendasikan untuk membandingkan performa dan mengidentifikasi metode yang lebih optimal. Selain itu, penerapan teknik pengolahan data yang lebih kompleks, seperti Natural Language Processing (NLP) untuk analisis sentimen, dapat digunakan untuk memperdalam pemahaman terhadap pola ulasan pengguna. Integrasi sistem berbasis kecerdasan buatan yang lebih adaptif juga diharapkan mampu meningkatkan kinerja model dalam memproses data secara real-time dan memberikan hasil prediksi yang lebih akurat dan relevan.

REFERENCES

Al-Balgaist, S. S. Y. (2023). Prediksi jumlah rilis game pada platform Steam dengan menggunakan algoritma K-Nearest Neighbor Regression. *Dissertasi*. Universitas Islam Negeri Maulana Malik Ibrahim.

Daulay, R. S. (2024). Analisis Kritis dan Pengembangan Algoritma K-Nearest Neighbor (KNN): Sebuah Tinjauan Literatur. Jurnal Pendidikan Sains dan Komputer, 4(02), 131-141.

Jalil, A., Homaidi, A., & Fatah, Z. (2024). Implementasi algoritma Support Vector Machine untuk klasifikasi status stunting pada balita. G-Tech: Jurnal Teknologi Terapan, 8(3), 2070-2079.

Mustafa, M. S., & Simpen, I. W. (2019, August). Implementasi Algoritma K-Nearest Neighbor (KNN) Untuk Memprediksi Pasien Terkena Penyakit Diabetes Pada Puskesmas Manyampa Kabupaten Bulukumba. In SISITI: Seminar Ilmiah Sistem Informasi dan Teknologi Informasi (Vol. 8, No. 1).

Nasri, E., & AW, A. S. (2020). Aplikasi Seleksi Penentuan Nasabah Untuk Penjualan Barang Secara Kredit Dengan Algoritma K-Nearest Neighbor. Jurnal Ilmiah Sains dan Teknologi, 4(1), 1-11.

Aggarwal, C. C. (2016). *Recommender Systems: The Textbook*. Springer.

Han, J., Kamber, M., & Pei, J. (2012). *Data Mining: Concepts and Techniques* (3rd ed.).Morgan Kaufmann.

Ricci, F., Rokach, L., & Shapira, B. (2011). *Introduction to Recommender Systems Handbook*. Springer.

Tan, P.-N., Steinbach, M., & Kumar, V. (2018). *Introduction to Data Mining* (2nd ed.). Pearson.

Yulisa, I. D., Testiana, G., & Putra, I. S. (2022). *Data Mining: Algoritma K-Nearest Neighbor dan Penerapannya*.