

## ***Analysis of the Effectiveness of Data Warehousing in Management Information Systems Using the Neural Networks Method***

Analisis Efektivitas Data Warehousing pada Sistem Informasi Manajemen menggunakan Metode Neural Networks

**Muliani Br. Ginting, Alam Rahmatulloh**

Informatika, Fakultas Teknik, Universitas Siliwangi, Indonesia

217006079@student.unsil.ac.id, \*alam@unsil.ac.id

\*: Penulis korespondensi (corresponding author)

### ***Informasi Artikel***

Received: August 2025

Revised: November 2025

Accepted: October 2025

Published: October 2025

### ***Abstract***

*Purpose: The purpose of this research is to investigate the effectiveness of data warehousing and the application of Neural Networks methods in analyzing bicycle travel app user data, with a focus on enhancing the annual membership of app users in North America.*

*Design/methodology/approach: This study utilizes a dataset that includes membership and usage data from relevant bicycle travel apps. It involves comparing the performance of different Neural Networks architectures, such as Feedforward Neural Networks, Convolutional Neural Networks (CNN), and other suitable models, to evaluate their effectiveness in predicting user membership.*

*Findings/result: The analysis results demonstrate that the implementation of Neural Networks can improve prediction accuracy, with the most effective model achieving 76.03% accuracy. The research also highlights the importance of preprocessing steps, such as data normalization and transformation, in contributing significantly to model performance. However, challenges such as overfitting were identified, suggesting the need for further testing with model and parameter variations.*

*Originality/value/state of the art: This research provides valuable insights for application developers and policy makers, helping them create data-driven strategies to improve the bicycle travel management information system. It also supports efforts to sustainably grow user membership. The study contributes to the field by exploring the practical application of Neural Networks for data analysis in the context of bicycle travel management, filling a gap in current research on effective predictive models for user membership growth.*

## Abstrak

Keywords:

*Data Warehousing; Data Preprocessing; Management Information System; Neural Network; Overfitting*

Kata kunci: *Data Warehousing; Neural Networks; Overfitting; Preprocessing Data; Sistem Informasi Manajemen.*

Tujuan: Penelitian ini bertujuan untuk menyelidiki efektivitas data warehousing dan penerapan metode *Neural Networks* dalam analisis data pengguna aplikasi perjalanan sepeda, dengan fokus pada peningkatan keanggotaan tahunan pengguna aplikasi di Amerika Utara.

Perancangan/metode/pendekatan: Penelitian ini menggunakan dataset yang mencakup data keanggotaan dan penggunaan dari aplikasi perjalanan sepeda yang relevan. Metode yang digunakan melibatkan perbandingan kinerja berbagai arsitektur *Neural Networks*, seperti *Feedforward Neural Networks*, *Convolutional Neural Networks (CNN)*, dan model lain yang sesuai, untuk mengevaluasi kemampuan masing-masing model dalam memprediksi keanggotaan pengguna.

Hasil: Hasil analisis menunjukkan bahwa penerapan *Neural Networks* dapat meningkatkan akurasi prediksi, dengan model terbaik mencapai akurasi 76,03%. Proses *preprocessing data*, seperti normalisasi dan transformasi data, berkontribusi secara signifikan terhadap peningkatan kinerja model. Namun, tantangan seperti *overfitting* juga ditemukan, yang menunjukkan perlunya pengujian tambahan dengan variasi model dan parameter yang lebih beragam.

Keaslian/state of the art: Penelitian ini memberikan wawasan baru yang berharga bagi pengembang aplikasi dan pembuat kebijakan dalam merumuskan strategi berbasis data untuk meningkatkan sistem informasi manajemen perjalanan sepeda. Studi ini juga berkontribusi pada perkembangan penelitian di bidang penerapan *Neural Networks* untuk analisis data dalam konteks manajemen perjalanan sepeda, mengisi kekosongan dalam penelitian terkait model prediktif yang efektif untuk pertumbuhan keanggotaan pengguna.

## 1. Pendahuluan

Perkembangan teknologi informasi telah memberi pengaruh yang signifikan terhadap beberapa sektor, termasuk dalam pengelolaan data di industri sepeda perjalanan. *Data warehousing*, sebagai sistem yang menyimpan dan mengelola data dalam jumlah besar, memainkan peranan penting dalam memberikan informasi yang akurat dan cepat. Dalam konteks manajemen perjalanan sepeda, kemampuan untuk mengakses dan menganalisis data secara efisien menjadi krusial untuk meningkatkan pengalaman pengguna dan operasional perusahaan.

*Data warehouse* membantu organisasi mengelola dan menyimpan data secara efisien, memfasilitasi pengambilan keputusan berbasis data yang lebih baik [1]. Kutipan ini menyoroti bahwa *data warehouse* bukan hanya sebuah tempat penyimpanan data, melainkan juga

merupakan pendekatan terpadu yang memanfaatkan teknologi untuk mendukung pengelolaan data historis. Data historis tersebut diperoleh dari berbagai sistem operasional yang ada di dalam organisasi, seperti aplikasi yang digunakan untuk operasional sehari-hari. Dengan adanya *data warehouse*, organisasi dapat mempertahankan data historis ini secara sistematis, memungkinkan mereka untuk melakukan analisis jangka panjang, mengidentifikasi pola, dan mendukung pengambilan keputusan strategis berdasarkan informasi yang lebih lengkap dan terpercaya. Sistem Informasi Manajemen (SIM) dirancang secara terstruktur yang akurat dan tepat waktu, baik untuk keperluan manajemen eksternal maupun untuk mendukung kegiatan operasional internal suatu organisasi. Sistem ini bertujuan untuk memperlancar proses manajemen, meningkatkan efektivitas perencanaan dan pengendalian, serta mendukung manajemen dalam proses pengambilan keputusan yang lebih informatif dan efisien [2]. Dengan menyediakan akses cepat terhadap informasi, SIM bertujuan untuk mempermudah proses manajemen, mendukung perencanaan, dan memperkuat pengendalian operasional. Selain itu, SIM juga berfungsi untuk memberikan informasi yang dibutuhkan dalam proses pengambilan keputusan, membantu manajemen dalam merespons kebutuhan bisnis dan menghadapi tantangan dengan lebih efektif dan efisien. Sistem informasi manajemen perjalanan sepeda di Amerika Utara menghadapi berbagai tantangan, seperti pengelolaan data yang terfragmentasi dan kesulitan dalam mendapatkan wawasan yang mendalam dari data yang tersedia. Ketidaktepatan dalam pengelolaan data, seperti yang disebutkan dalam penelitian, dapat mengakibatkan keputusan yang kurang akurat dan berpotensi menurunkan tingkat kepuasan pelanggan [3]. Menunjukkan bahwa kualitas data yang kurang akurat bisa memengaruhi seluruh proses pengambilan keputusan dalam organisasi, terutama yang berkaitan dengan pelayanan pelanggan. Ketika data yang digunakan tidak dikelola dengan baik, risiko kesalahan dalam keputusan strategis meningkat, yang pada akhirnya berpotensi menurunkan kepercayaan dan kepuasan pelanggan terhadap layanan yang diberikan. *Data warehouse* menyediakan sarana untuk menyimpan yang dihimpun dari berbagai sumber internal maupun eksternal [4]. Data warehouse berfungsi sebagai platform untuk mengumpulkan dan menyimpan data dari berbagai sumber, baik yang berasal dari dalam organisasi maupun dari luar. Dengan pengelolaan data yang terorganisir, data warehouse memberikan akses ke informasi yang lebih lengkap dan relevan, sehingga membantu organisasi dalam membuat keputusan yang lebih tepat dan strategis.

Data warehouse merancang sebagai sistem komputer untuk menganalisis data yang dimiliki oleh suatu organisasi [5]. Pengembangan sistem data warehousing yang efisien menjadi kunci utama untuk meningkatkan kinerja sistem informasi manajemen, karena data warehouse berperan sebagai pusat penyimpanan data terstruktur yang mengintegrasikan berbagai informasi dari seluruh bagian organisasi. Dengan sifatnya yang terintegrasi dan terpusat, *data warehouse* memastikan bahwa data yang tersedia lebih mudah diakses dan lebih akurat, sehingga sangat mendukung proses pengambilan keputusan. Ini menyediakan data yang konsisten dan relevan untuk berbagai analisis dan evaluasi yang dibutuhkan oleh manajemen, memungkinkan organisasi untuk membuat keputusan yang lebih informatif dan tepat.

Untuk mengatasi permasalahan dalam penelitian, peneliti memanfaatkan metode *Neural Networks* untuk menganalisis efektivitas *data warehousing*. Model Neural Network dapat mendukung identifikasi [6]. *Neural Networks*, yang dikenal karena kemampuannya dalam mengenali pola dan membuat prediksi dari data besar, dapat memberikan solusi yang inovatif dalam pengelolaan informasi perjalanan sepeda. Penggunaan *Neural Networks* dalam analisis data dapat meningkatkan akurasi prediksi serta efisiensi dalam proses pengambilan keputusan [3]. Metode *Neural Networks*, yang dikenal dengan kemampuannya dalam menangani data yang

kompleks, dapat menghasilkan prediksi yang lebih akurat. Selain itu, penggunaan teknologi ini juga mempercepat proses analisis data, memungkinkan pengambilan keputusan dilakukan dengan lebih efisien dan berdasarkan informasi yang lebih akurat. *Neural Networks* memungkinkan organisasi untuk mengenali pola-pola tersembunyi dalam data, memberikan hasil analisis yang dapat diandalkan dan mendukung keputusan strategis yang lebih baik. Beberapa penelitian terkait penerapan Data Warehousing dan *Neural Networks* dalam analisis sistem informasi manajemen perjalanan sepeda telah dilakukan dengan berbagai pendekatan dan algoritma yang berbeda. Sebagai contoh, penelitian oleh Azzahra dkk. [7] menggunakan metode ensemble seperti Bagging dan Adaboost untuk analisis data perjalanan sepeda dan prediksi keanggotaan. Metode tersebut menghasilkan akurasi yang cukup tinggi, yaitu 94.63% untuk Adaboost dan 95.02% untuk Bagging. Meskipun demikian, kedua studi ini menghadapi tantangan pada aspek sensitivitas dan nilai f1-score yang menunjukkan tingkat false negative yang tinggi. Hal ini menunjukkan bahwa meskipun akurasi tinggi, model ini mungkin kurang efektif dalam mendeteksi kategori tertentu dalam data.

Sementara itu, penelitian lain yang menggunakan Support Vector Machine (SVM), seperti yang dilakukan oleh Yulianto [8], mendapatkan akurasi 100% dalam prediksi dataset akademik, tetapi dataset ini hanya terdiri dari jumlah data yang terbatas. Keterbatasan dataset ini menjadi masalah utama karena hasil prediksi cenderung terbatas pada pola yang ada dalam dataset tanpa dapat menggeneralisasi ke data baru. Penelitian ini menunjukkan kelebihan dalam hal akurasi, tetapi memerlukan evaluasi lebih lanjut dengan dataset yang lebih beragam.

Di sisi lain, penelitian Hilman dan Djamarudin [9] mengembangkan metode Weighted Vote (WV) sebagai teknik ensemble untuk mengoptimalkan proses ETL dalam mendukung pengambilan keputusan berbasis data warehouse. Meskipun WV memberikan akurasi sebesar 85.21%, kelemahannya terletak pada kurangnya pengujian dengan dataset yang lebih variatif, yang dapat memengaruhi generalisasi model terhadap data dengan karakteristik berbeda. Hal ini juga menjadi tantangan dalam penelitian yang lebih luas, termasuk dalam sistem manajemen perjalanan sepeda, di mana karakteristik pengguna dan pola perjalanan dapat sangat beragam. Penelitian-penelitian tersebut memberikan gambaran bahwa meskipun berbagai metode seperti *Bagging & Adaboost*, SVM, dan WV dapat memberikan hasil yang baik, tantangan seperti overfitting, ukuran dataset terbatas, dan keterbatasan dalam menggeneralisasi ke data baru tetap menjadi isu yang perlu diatasi. Dalam konteks penelitian ini, penggunaan *Neural Networks* dalam sistem manajemen perjalanan sepeda memiliki potensi untuk memberikan model yang lebih kuat dan efisien. Dengan pendekatan *data warehousing*, penelitian ini juga berkontribusi pada peningkatan pengelolaan data dan pemrosesan informasi dalam jumlah besar, yang dapat meningkatkan pengambilan keputusan berbasis data.

Penelitian ini bertujuan untuk mengeksplorasi bagaimana penerapan *data warehousing* dan *Neural Networks* dapat meningkatkan efektivitas sistem informasi manajemen perjalanan sepeda. Dengan menggunakan dataset yang mencakup informasi penggunaan sepeda, pola perjalanan, dan preferensi pengguna, penelitian ini berusaha untuk mengevaluasi kinerja sistem dalam memberikan informasi yang relevan dan akurat.

Hasil dari penelitian ini diharapkan mampu memberikan dasar yang berdampak besar pada pengembangan sistem informasi manajemen perjalanan sepeda yang lebih baik, serta memberikan wawasan mengenai manfaat dan tantangan yang dihadapi dalam penerapan teknologi *data warehousing* dan *Neural Networks*. Melalui penelitian ini, diharapkan dapat ditemukan solusi yang mampu meningkatkan pengalaman pengguna dan efisiensi operasional di industri perjalanan sepeda.

## 2. Metode/Perancangan

Metode yang diterapkan dalam penelitian ini untuk menganalisis efektivitas *data warehousing* pada sistem informasi manajemen perjalanan sepeda di Amerika Utara. Metode yang digunakan terdiri dari tahapan studi literatur, pengumpulan data, pengolahan data, dan implementasi algoritma. Pada **Gambar 1.** merupakan tahapan untuk memberikan pemahaman yang jelas mengenai proses yang dilakukan dalam penelitian ini.



**Gambar 1.** Tahapan Penelitian

### 2.1. Studi Literatur

Langkah pertama dalam penelitian ini melibatkan studi literatur, yang merupakan langkah penting dalam merancang dan melaksanakan analisis yang efektif. Pada tahap ini, peneliti mencari dan mengkaji literatur yang relevan, meliputi jurnal ilmiah, buku, dan sumber online yang membahas topik-topik penting, seperti *data warehousing*, *Neural Networks*, dan Manajemen Perjalanan Sepeda. Peneliti melakukan penelusuran terhadap teori dan konsep yang relevan mengenai *data warehousing* serta penerapan *Neural Networks* dalam analisis data. Dalam kajian ini, peneliti mempelajari konsep, implementasi, dan penerapan *Data warehousing* dalam sistem informasi, serta mendalami algoritma *Neural Networks*, cara kerja, dan aplikasinya dalam analisis big data.

Selain itu, peneliti juga meneliti bagaimana data dari penggunaan sepeda diintegrasikan dan dianalisis dalam sistem manajemen perjalanan. Studi literatur ini memberikan landasan dasar teori yang kuat bagi penelitian ini, dengan tujuan memahami teori dasar dan metode yang telah digunakan dalam penelitian sebelumnya. Pengetahuan yang diperoleh dari studi literatur ini akan membantu peneliti dalam mendesain model analisis yang tepat dan mengembangkan solusi yang lebih baik, serta menjadi landasan dalam menentukan metodologi yang akan diterapkan dalam penelitian ini.

### 2.2. Pengumpulan Data

Setelah menyelesaikan studi literatur, peneliti melanjutkan ke tahap pengumpulan data. Pada penelitian ini, peneliti menggunakan dataset dari sistem berbagi sepeda *Divvy Bicycle Sharing System* di kota Chicago, Amerika Utara, yang diambil dari Kaggle, sebuah platform berbagi dataset yang menyediakan berbagai informasi tentang perjalanan sepeda. Dataset ini mencakup berbagai atribut penting, seperti 'Rider ID', 'Trip Duration', 'Start Time dan End Time', dan 'Start Station dan End Station'. Selain itu, dataset ini juga mencakup data demografi pengguna, termasuk 'Membership Type', 'gender', dan 'Rider Gender and Age'.

Data ini mencakup jutaan catatan perjalanan yang dikumpulkan selama beberapa tahun, menyediakan informasi yang sangat kaya untuk dianalisis menggunakan sistem *Data warehousing*. Secara umum, semakin banyak data yang diintegrasikan ke dalam *data warehouse*, semakin beragam dan komprehensif informasi yang dapat dihasilkan dari sistem tersebut [10]. Dengan menampung lebih banyak data dari berbagai sumber, *data warehouse* mampu memberikan pandangan yang lebih menyeluruh dan detail. Hal ini memungkinkan organisasi untuk mendapatkan informasi yang lebih lengkap, yang dapat digunakan dalam

analisis mendalam dan mendukung keputusan yang lebih komprehensif. Data yang terakumulasi dari berbagai aspek operasional dan strategis organisasi meningkatkan kualitas dan kedalaman informasi yang tersedia. Peneliti memastikan bahwa dataset yang diunduh memiliki kualitas yang memadai untuk analisis lebih lanjut. Informasi yang terkumpul akan digunakan untuk menganalisis pola penggunaan sepeda, efektivitas keanggotaan tahunan, serta mendukung sistem manajemen informasi terkait pengambilan keputusan strategis. *Neural Networks* akan digunakan untuk mengidentifikasi tren dan pola, seperti rute perjalanan yang paling sering digunakan, waktu puncak penggunaan, serta hubungan antara jenis keanggotaan dengan durasi dan frekuensi penggunaan sepeda.

### 2.3. Pengolahan Data

Sebelum melanjutkan dengan proses ETL, penting untuk memahami konsep integrasi data. Integrasi data merujuk pada proses menggabungkan data dari berbagai sumber yang berbeda menjadi satu kesatuan untuk menghasilkan satu set data yang terkoordinasi dan terstruktur. Tujuan dari integrasi data adalah untuk menyatukan informasi yang terpisah-pisah agar dapat dianalisis secara komprehensif, memberikan wawasan yang lebih lengkap, dan mendukung pengambilan keputusan yang lebih baik. Proses ETL (Extract, Transform, Load) digunakan untuk mengekstrak, mengolah, dan mengintegrasikan data dari berbagai sumber internal dan eksternal sebelum dimasukkan ke dalam data warehouse untuk diproses lebih lanjut [11]. Proses ini sangat penting karena memungkinkan untuk mengintegrasikan data terdistribusi dari sistem yang berbeda, mengubahnya menjadi format yang konsisten dan berguna, dan menyimpannya di gudang data untuk analisis lebih lanjut. Pendekatan ini memungkinkan peneliti memanfaatkan kualitas dan keragaman data secara optimal untuk membantu pengambilan keputusan yang lebih efektif.

Setelah pengumpulan data, tahap berikutnya yaitu pengolahan data. Data yang diperoleh dari Kaggle diolah menggunakan Pentaho, yang merupakan platform Extract, Transform, Load (ETL). Proses-proses tersebut dilakukan secara bertahap, dimulai dari setiap sumber data dalam basis data yang terdistribusi, kemudian digabungkan dalam sistem xtrex [8]. Proses ETL mencakup pengambilan data dari database transaksi, yang meliputi dimensi seperti waktu, supplier, toko, kategori, dan barang, kemudian menyimpannya dalam *data warehouse* [7]. Proses ETL dilakukan dengan langkah-langkah sebagai berikut:

Pertama, file input digunakan untuk mengimpor data yang diambil dari Kaggle ke dalam Pentaho. Dalam tahap ini, peneliti memilih file dataset yang relevan untuk analisis. Selanjutnya, setelah data diimpor, langkah berikutnya adalah menggunakan komponen *Select Value*. Di sini, peneliti melakukan pemilihan atribut yang akan digunakan dalam analisis, menghapus atribut yang tidak relevan atau duplikat untuk memastikan bahwa hanya informasi penting yang diproses. Penggabungan data dari berbagai sumber yang berbeda menghasilkan data terintegrasi [12]. Proses ini membantu peneliti dalam memfokuskan analisis pada data yang relevan dan krusial. Setelah pemilihan nilai selesai, data yang telah diproses disimpan dalam format Excel menggunakan komponen *Excel Output*. File Excel ini berfungsi sebagai hasil akhir dari proses ETL, yang siap digunakan untuk analisis lebih lanjut dengan *Neural Networks*. Dalam membangun *data warehouse* untuk analisis mendukung pengambilan keputusan yang cepat dan akurat sangat bergantung pada efektivitas proses ETL yang menransfer data dari database OLTP ke *data warehouse* [9]. Proses *Extract, Transform, Load* (ETL) dalam memastikan keberhasilan pembangunan *data warehouse*. Proses ETL bertugas untuk mengambil data dari sistem transaksi operasional (OLTP), menransformasikannya agar sesuai dengan format yang

dibutuhkan, dan memuatnya ke dalam *data warehouse*. Data warehouse menyediakan informasi yang sesuai dengan dimensi dan fakta yang dibutuhkan [13]. Keberhasilan dalam tahap ini sangat mempengaruhi kualitas data yang tersedia dalam *data warehouse*, yang pada gilirannya mendukung analisis yang akurat dan pengambilan keputusan yang cepat dan tepat. Pentaho Data Integration digunakan untuk Extract, Transform, Load (ETL), yang mengelola dan mengalirkan data dari sumber ke *data warehouse*. Setelah data tersedia, Power Business Intelligence digunakan untuk memvisualisasikan data dalam bentuk grafik yang mudah dipahami. Ini mempermudah pemangku kepentingan dalam menganalisis data dan membuat keputusan dengan lebih cepat dan efisien, tanpa harus memproses data secara manual. Dalam membangun data warehouse, perhatian utama diberikan pada kualitas data, karena kualitas data berpengaruh langsung terhadap proses ETL [9]. Penelitian ini juga melakukan proses ETL dengan menggunakan *tools* dari pentaho data integration. Dengan menyimpan data dalam format Excel, peneliti memastikan bahwa data dapat dengan mudah diakses dan dianalisis dalam langkah-langkah berikutnya.

Tahap pengolahan data diawali dengan proses pembersihan (*data cleaning*) terhadap dataset yang telah dikumpulkan dari *Divvy Bicycle Sharing System*. Pada tahap ini, peneliti menangani data yang tidak lengkap atau tidak valid, seperti catatan perjalanan yang memiliki nilai yang tidak logis, misalnya durasi perjalanan yang negatif atau nol. Data yang tidak relevan atau cacat akan dihapus, sedangkan data yang hilang (*missing values*) diimputasi atau ditangani menggunakan metode tertentu sesuai dengan karakteristik dataset. Selanjutnya, peneliti melakukan proses transformasi data, yang mencakup normalisasi atau standarisasi variabel numerik, seperti durasi perjalanan, untuk memastikan bahwa data berada dalam skala yang seragam dan dapat dianalisis secara komprehensif. Selain itu, data kategorikal seperti jenis keanggotaan, jenis kelamin, dan status stasiun dikonversi menjadi representasi numerik menggunakan metode *encoding*, sehingga dapat diproses oleh algoritma *Neural Networks*.

Setelah transformasi, peneliti melakukan agregasi dan pengelompokan (*clustering*) terhadap data. Proses ini bertujuan untuk mengidentifikasi pola-pola tertentu, misalnya dengan mengelompokkan data berdasarkan lokasi stasiun, waktu perjalanan, atau jenis keanggotaan. Tahap ini membantu peneliti dalam memahami pola penggunaan sepeda secara lebih mendetail. Data yang telah melalui tahap pembersihan, transformasi, dan pengelompokan kemudian diintegrasikan ke dalam *Data warehouse* untuk penyimpanan yang lebih terstruktur dan efisien. Integrasi ini sangat penting karena memungkinkan akses cepat dan analisis data yang lebih mendalam dalam konteks Sistem Informasi Manajemen (SIM).

*Data warehouse* berfungsi untuk mempercepat pengumpulan data yang diperlukan dengan cara yang efisien, menghasilkan informasi yang ringkas, akurat, dan relevan, sehingga dapat membantu dalam membuat keputusan dengan cepat dan tepat [14]. Dengan data yang rapi dan saling terhubung, data warehouse membantu organisasi mendapatkan informasi yang jelas dan akurat dengan cepat. Informasi yang diperoleh dari *data warehouse* memudahkan pemangku kepentingan dalam mengambil keputusan, karena mereka memiliki akses langsung ke data yang dibutuhkan untuk analisis strategis. *Data warehouse* berfungsi sebagai penyimpanan data historis yang mendukung analisis dan pengambilan keputusan. Informasi yang dihasilkan bersifat *read-only*, sehingga hanya dapat diakses untuk dibaca dan dianalisis [15]. Data yang terstruktur dalam *Data warehouse* akan digunakan dalam tahap pelatihan model *Neural Networks* untuk analisis lebih lanjut dalam rangka mengevaluasi pola perjalanan sepeda serta efektivitas keanggotaan tahunan. Proses pengolahan data ini memastikan bahwa dataset yang

digunakan telah siap untuk dianalisis secara optimal, mendukung keputusan strategis dalam pengelolaan sistem informasi dan pengambilan keputusan berbasis data.

#### 2.4. Implementasi Algoritma

Tahap berikutnya setelah pengolahan data adalah penerapan algoritma. Dalam tahap ini, peneliti melatih model *Neural Networks* untuk menganalisis efektivitas *data warehousing*. *Neural Networks* berguna untuk memetakan hubungan kompleks antara input dan output guna mengidentifikasi pola dalam data [9]. Implementasi algoritma *Neural Networks* dimulai dengan mengimpor library yang diperlukan, seperti Pandas, numpy, dan Keras, untuk analisis data dan pengembangan model.

Setelah itu, dataset dibaca dari file CSV dan dieksplorasi untuk memahami struktur serta karakteristik datanya. Proses ini mencakup penghapusan nilai yang hilang dan konversi kolom tanggal ke format datetime, diikuti dengan penambahan kolom durasi perjalanan yang dihitung dalam menit. Kategori pada kolom member\_casual kemudian diubah menjadi format numerik untuk memudahkan pelatihan model.

Selanjutnya, fitur yang relevan dipilih, dengan memisahkan variabel fitur dari variabel target. Dataset dibagi menjadi data pelatihan dan pengujian menggunakan metode `train_test_split`, dan fitur distandarisasi menggunakan `standardscaler` agar model dapat belajar secara optimal. Model *Neural Networks* dibangun menggunakan Keras dengan arsitektur yang terdiri dari dua lapisan tersembunyi dan satu lapisan output yang memanfaatkan fungsi aktivasi sigmoid untuk melakukan klasifikasi biner.

Model ini dapat berupa beberapa jenis, termasuk *multilayer perceptron* (MLP) dan *Convolutional Neural Networks* (CNN), yang akan diimplementasikan untuk mempelajari pola dalam data perjalanan sepeda. Model dikompilasi menggunakan optimizer Adam dan fungsi loss *binary crossentropy* untuk memastikan proses pelatihan berjalan optimal dalam klasifikasi biner. Untuk mencegah overfitting, `callback early stopping` diterapkan, yang menghentikan pelatihan jika tidak ada peningkatan dalam validasi. Model dilatih menggunakan data pelatihan selama beberapa epoch, dengan hasil pelatihan dan validasi yang diukur pada setiap epoch. Evaluasi model juga dilakukan menggunakan set validasi untuk memastikan kinerja yang optimal.

Setelah pelatihan, model dievaluasi pada data pengujian untuk mengukur akurasinya, yang kemudian divisualisasikan dalam grafik untuk membandingkan akurasi pelatihan dan validasi. Terakhir, model yang telah dilatih disimpan untuk digunakan di masa depan. Proses ini memberikan wawasan yang lebih mendalam mengenai perilaku pengguna dalam konteks aplikasi data yang sedang dianalisis, dengan memprediksi status keanggotaan berdasarkan berbagai fitur yang relevan. Implementasi algoritma ini bertujuan untuk memberikan insight yang mendalam terkait pola penggunaan sepeda, serta membantu dalam pengambilan keputusan strategis dalam manajemen sistem informasi. Sistem informasi dapat mempercepat respon terhadap perubahan lingkungan dan merampingkan proses, sehingga membantu perusahaan mencapai keunggulan kompetitif [15]. Dengan menyediakan data yang tepat waktu dan relevan, sistem informasi memungkinkan perusahaan untuk merespon perubahan pasar dan lingkungan dengan lebih cepat dan efisien. Ini mendukung perusahaan dalam tetap adaptif serta mengambil keputusan yang lebih efektif, yang pada akhirnya dapat memperkuat daya saing mereka di pasar.

### 3. Hasil dan Pembahasan

#### 3.1. Hasil Penelitian

**Tabel 1.** Hasil akurasi pada beberapa epoch utama

EPOCH	AKURASI PELATIHAN	AKURASI VALIDASI	LOSS PELATIHAN	LOSS VALIDASI
1	49.96%	76.92%	0.6860	0.5660
5	79.57%	76.92%	0.4989	0.5399
10	74.66% (uji)	-	0.5768	-

Pada penelitian ini menggunakan data perjalanan sepeda dari sistem manajemen informasi di Amerika Utara. Data ini diproses melalui tahapan pembersihan data, transformasi, dan normalisasi sebelum diimplementasikan ke dalam model *Neural Networks*. Pada **Tabel 1**, terdapat proses pelatihan model *Neural Networks* yang dilakukan menggunakan 50 epoch, dan hasil akurasi model dievaluasi baik pada data pelatihan maupun validasi. Berdasarkan hasil penelitian, diperoleh informasi sebagai berikut:

Epoch 1: Akurasi pelatihan mencapai 49.96%, sementara akurasi validasi berada pada 76.92%.

Pada epoch pertama, nilai loss sebesar 0.6860 untuk pelatihan dan 0.5660 untuk validasi.

Epoch 5: Model mencapai akurasi pelatihan 79.57%, dengan akurasi validasi tetap konsisten di 76.92%. Loss untuk pelatihan sebesar 0.4986 dan loss untuk validasi 0.5399.

Epoch 10: Akurasi pengujian akhir menggunakan data uji mencapai 74.66%, dengan loss sebesar 0.5768.

#### 3.2. Pembahasan

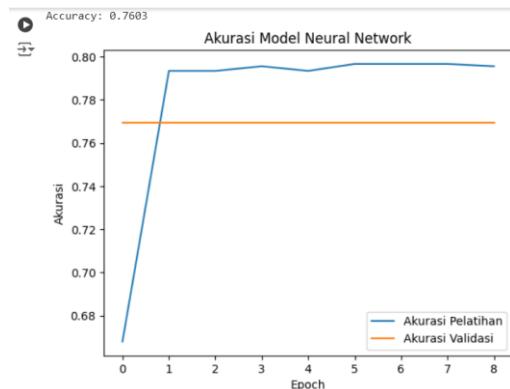
##### 3.2.1. Analisis *Data warehousing*

Tujuan dari penelitian ini adalah untuk menganalisis seberapa efektif penerapan sistem data warehousing dalam mendukung sistem informasi manajemen perjalanan sepeda. Penggunaan *Data warehousing* memungkinkan penyimpanan data yang terstruktur, sehingga dapat digunakan untuk pengambilan keputusan yang lebih baik, terutama dalam memahami pola perjalanan pengguna sepeda di Amerika Utara.

Dalam konteks ini, *data warehousing* sangat berperan dalam pengolahan volume besar data perjalanan sepeda, yang meliputi informasi stasiun, waktu mulai dan berakhir perjalanan, serta tipe keanggotaan (member atau casual). Dengan pengolahan data yang efisien menggunakan teknik *Data warehousing*, data tersebut siap untuk dianalisis menggunakan algoritma *machine learning*, termasuk *Neural Networks*.

##### 3.2.2. Implementasi *Neural Networks*

Pada **Gambar 2**, menunjukkan kurva akurasi model *Neural Networks* selama proses pelatihan dan validasi. Model *Neural Networks* digunakan untuk mengklasifikasikan pengguna sepeda ke dalam dua kategori: member dan casual. Hasil pelatihan menunjukkan bahwa model mampu mengenali pola penggunaan sepeda dengan akurasi yang cukup tinggi. Pada akhir pelatihan, model mencapai akurasi pengujian sebesar 74.66%, menunjukkan bahwa model mampu memprediksi dengan cukup akurat, meskipun terdapat sedikit overfitting pada data pelatihan. Performa terbaik terjadi pada epoch ketujuh, di mana akurasi pelatihan mencapai 80.98%. Namun, akurasi validasi tetap konsisten pada 76.92%, yang mengindikasikan bahwa model memiliki generalisasi yang baik terhadap data yang belum pernah dilihat.



**Gambar 2.** Akurasi Model *Neural Networks*

### 3.2.3. Efektivitas Sistem Informasi Manajemen

Berdasarkan hasil analisis, dapat disimpulkan bahwa penerapan sistem data warehousing dalam sistem informasi manajemen perjalanan sepeda di Amerika Utara terbukti efektif dalam menyimpan serta mengelola data dalam jumlah besar dan kompleks. Selain itu, hasil eksperimen juga menunjukkan bahwa sistem ini dapat digunakan untuk menghasilkan insight penting terkait perilaku pengguna, seperti perbedaan pola perjalanan antara pengguna member dan casual. Penggunaan model *Neural Networks* dalam analisis ini memungkinkan sistem untuk melakukan klasifikasi otomatis, yang berpotensi meningkatkan efektivitas dalam pengambilan keputusan, terutama dalam menyusun strategi peningkatan jumlah keanggotaan sepeda.

### 3.2.4. Tantangan dan Solusi

Selama proses pelatihan, salah satu tantangan yang dihadapi adalah overfitting pada data pelatihan. Hal ini terlihat dari perbedaan akurasi antara data pelatihan dan validasi pada beberapa epoch. Untuk mengatasi tantangan ini, peneliti menggunakan teknik regularisasi dan dropout pada layer model untuk mengurangi *overfitting*.

Selain itu, dataset yang digunakan mengandung beberapa nilai yang kosong atau tidak lengkap pada fitur, sehingga diperlukan langkah *preprocessing* yang ekstensif untuk memastikan kualitas data sebelum pelatihan model.

## 4. Kesimpulan dan Saran

Hasil penelitian menunjukkan bahwa penerapan data warehousing dan penggunaan metode Neural Networks memiliki dampak yang signifikan terhadap peningkatan efektivitas sistem informasi manajemen perjalanan sepeda di Amerika Utara. Penggunaan *Neural Networks* dalam analisis data pengguna berhasil menghasilkan model yang memiliki akurasi tinggi dalam memprediksi keanggotaan tahunan pengguna aplikasi. Model *Neural Networks* yang diterapkan menunjukkan performa yang baik, dengan nilai akurasi mencapai 76.03% pada pengujian data. Analisis menunjukkan bahwa *preprocessing* data, termasuk normalisasi dan transformasi, berkontribusi pada peningkatan kualitas data yang digunakan untuk pelatihan model. Meskipun demikian, terdapat tantangan dalam hal overfitting, yang dialami oleh beberapa model *Neural Networks*. Untuk mengatasi masalah ini, perlu dilakukan pengujian lebih lanjut dengan berbagai arsitektur dan parameter untuk meningkatkan kinerja model. Penelitian ini juga menunjukkan pentingnya pengelolaan data dalam *data warehousing*, yang tidak hanya meningkatkan efisiensi

penyimpanan tetapi juga memperbaiki aksesibilitas dan analisis data. Dengan menggunakan pendekatan *data warehousing* yang tepat, organisasi dapat lebih mudah mengolah data dan mengambil keputusan berbasis data yang lebih baik.

Sebagai saran, penelitian ini merekomendasikan eksplorasi lebih lanjut terhadap model *Neural Networks* yang berbeda dan pengujian pada dataset yang lebih besar dan beragam. Selain itu, penelitian di masa depan dapat mencakup eksplorasi integrasi faktor-faktor tambahan yang mungkin mempengaruhi efektivitas sistem, serta penerapan metode evaluasi yang lebih menyeluruh, seperti cross-validation, untuk memastikan bahwa model yang digunakan dapat digeneralisasi dengan baik. Diharapkan hasil penelitian ini dapat memberikan wawasan yang berharga bagi pengembang aplikasi dan pemangku kebijakan untuk meningkatkan sistem informasi manajemen perjalanan sepeda, serta mendorong pengembangan strategi berbasis data yang dapat memperkuat keanggotaan pengguna.

## Daftar Pustaka

- [1] F. Fahrianto, “Data Warehouse dan Data Mining,” *GAES-PACE Book Publisher*, p. 193, 2016, [Online]. Available: [https://books.google.com/books?hl=en%5C&lr=%5C&id=3bGpEAAAQBAJ%5C&oi=fnd%5C&pg=PA1%5C&dq=penggunaan+teknologi+kecerdasan+buatan+dalam+pengembangan+sistem+informasi+manajemen+pergudangan%5C&ots=\\_zK9UNdFoq%5C&sig=rO00pAHE3rVs77a1pf9aZTn2\\_yE](https://books.google.com/books?hl=en%5C&lr=%5C&id=3bGpEAAAQBAJ%5C&oi=fnd%5C&pg=PA1%5C&dq=penggunaan+teknologi+kecerdasan+buatan+dalam+pengembangan+sistem+informasi+manajemen+pergudangan%5C&ots=_zK9UNdFoq%5C&sig=rO00pAHE3rVs77a1pf9aZTn2_yE)
- [2] T. Pipit Mulyiah, Dyah Aminatun, Sukma Septian Nasution, Tommy Hastomo, Setiana Sri Wahyuni Sitepu, *濟無No Title No Title No Title*, vol. 7, no. 2. 2020.
- [3] D. Alexander and I. N. Pujawan, “Perancangan Sistem Warehouse Berbasis Teknologi Ocr Untuk Meningkatkan Efektivitas Dan Efisiensi,” *Jurnal Ilmiah Teknik Industri*, vol. 12, no. 1, pp. 1–11, 2024, doi: 10.24912/jitiuntar.v12i1.28100.
- [4] M. Made Hanindia Prami Swari, Fawwaz Ali Akbar, “Data Warehouse Implementation Techniques In Data Processing (Case Study Data Sales at PT Spirit Sejahtera Bersama),” *Journal Mantik*, vol. 3, no. 3, pp. 135–142, 2019, [Online]. Available: <https://iocscience.org/ejournal/index.php/mantik/article/view/400>
- [5] T. F. Efendi and M. Krisanty, “Warehouse Data System Analysis PT. Kanaan Global Indonesia,” *International Journal of Computer and Information System (IJCIS)*, vol. 1, no. 3, pp. 70–73, 2020, doi: 10.29040/ijcis.v1i2.26.
- [6] R. Pratama, R. Herdiana, R. Hamonangan, and S. Anwar, “Analisis Prediksi Kelulusan Mahasiswa Menggunakan Metode Artificial Neural Network,” *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 8, no. 1, pp. 687–693, 2024, doi: 10.36040/jati.v8i1.8762.
- [7] S. P. Azzahra, Y. A. Apriyanto, and A. Wijaya, “Analisis Dan Perancangan Data Warehouse Untuk Pengelolaan Stok Barang Pada Cv Aneka Artha Niaga,” *Journal Of Informatics And Busisnes*, vol. 01, no. 03, pp. 103–112, 2023.
- [8] A. A. Yulianto, “Extract Transform Load (ETL) Process in Distributed Database Academic Data Warehouse,” *APTIKOM Journal on Computer Science and Information Technologies*, vol. 4, no. 2, pp. 61–68, 2019, doi: 10.11591/aptikom.j.csit.36.

- [9] M. Hilman and D. Djamarudin, "Analisis Faktor Optimalisasi Proses Etl Pada Data Warehouse Sebagai Pendukung Pengambilan Keputusan Management Dengan Business Intelligence," *Faktor Exacta*, vol. 11, no. 1, p. 24, 2018, doi: 10.30998/faktorexacta.v11i1.2325.
- [10] N. T. S. Saptadi and H. C. Marwi, "Rancang Bangun Model Layanan Fungsi Menggunakan Data Warehouse Dalam Penyusunan Blue Print Rumah Sakit," *Researchgate.Net*, no. November 2014, 2020, [Online]. Available: [https://www.researchgate.net/profile/Norbertus-Saptadi/publication/332079934\\_Rancang\\_Bangun\\_Model\\_Layanan\\_Fungsi\\_Mengguna kan\\_Data\\_Warehouse\\_Dalam\\_Penyusunan\\_Blueprint\\_Rumah\\_Sakit\\_November\\_2014/links/5c9e354d299bf1116950040a/Rancang-Bangun-Model-Layanan-Fu](https://www.researchgate.net/profile/Norbertus-Saptadi/publication/332079934_Rancang_Bangun_Model_Layanan_Fungsi_Mengguna kan_Data_Warehouse_Dalam_Penyusunan_Blueprint_Rumah_Sakit_November_2014/links/5c9e354d299bf1116950040a/Rancang-Bangun-Model-Layanan-Fu)
- [11] S. Sucipto, S. Sucipto, and A. Nugroho, "Analisis Data Warehouse Pada Perpustakaan Man X Untuk Efisiensi Manajemen," *Fountain of Informatics Journal*, vol. 5, no. 3, p. 17, 2020, doi: 10.21111/fij.v5i3.4988.
- [12] m t hidayatullah and s r asroni, "Penerapan Algoritma Neural Network Untuk Memprediksi Kelayakan Calon Asisten Dosen," *Repository.Umy.Ac.Id*, no. XXX, 2015, [Online]. Available: [http://repository.ums.ac.id/bitstream/handle/123456789/30148/g\\_bab\\_3.pdf?sequence=7&isallowed=y](http://repository.ums.ac.id/bitstream/handle/123456789/30148/g_bab_3.pdf?sequence=7&isallowed=y)
- [13] M. Hendrawaty, "Analisis dan Perancangan Data Warehouse pada PT. Polaris Sapta Manggala," *Jurnal Ilmiah Binary STMIK Bina Nusantara Jaya Lubuklinggau*, vol. 6, no. 1, pp. 8–16, 2024, doi: 10.52303/jb.v6i1.137.
- [14] S. M. Qibtiyah, A. Nugroho, and R. Firliana, "Sistem Informasi Pengolahan Data Peminjaman Buku di Taman Baca Dengan Menggunakan Data Warehouse," *Jurnal Ilmu Komputer dan Sistem Informasi (JIKOMSI)*, vol. 5, no. 2, pp. 68–75, 2022, doi: 10.55338/jikomsi.v5i2.310.
- [15] S. Wahono and H. Ali, "Peranan Data Warehouse, Software Dan Brainware Terhadap Pengambilan Keputusan (Literature Review Executive Support Sistem for Business)," *Jurnal Ekonomi Manajemen Sistem Informasi*, vol. 3, no. 2, pp. 225–239, 2021, doi: 10.31933/jemsi.v3i2.781.