

Prediksi Kepuasan Pelanggan Maskapai menggunakan Model Machine Learning

Predict Airline Customer Satisfaction using A Machine Learning Model

¹Yoel Dinata Suwito*, ²Yeremia Alfa Susetyo

^{1,2,3}Program Studi Teknik Informatika, Fakultas Teknologi Informasi, Satya Wacana Christian University

^{1,2,3}Jl. Dr. O. Notohamidjojo No.1 -10, Blotongan, Kec. Sidorejo, Kota Salatiga, Jawa Tengah 50715, Indonesia

*e-mail: 672021278@student.uksw.edu

(*received*: 20 November 2025, *revised*: 6 December 2025, *accepted*: 7 December 2025)

Abstrak

Kepuasan pelanggan merupakan faktor strategis bagi keberlanjutan bisnis maskapai di tengah persaingan industri penerbangan yang semakin ketat. Penelitian ini bertujuan untuk memprediksi tingkat kepuasan pelanggan maskapai menggunakan metode *Artificial Neural Network* (ANN) dengan memanfaatkan dataset publik dari Kaggle yang mencakup 22 fitur layanan penerbangan. Dua arsitektur ANN dikembangkan, dengan perbedaan utama pada jumlah *hidden layer*, jumlah neuron, serta penerapan *Batch Normalization* dan *LeakyReLU* pada model kedua. Hasil pengujian menunjukkan bahwa model ANN pertama memperoleh akurasi 92,31%, sedangkan model ANN kedua mampu mencapai performa yang jauh lebih tinggi dengan akurasi 95,75% pada data uji. Model kedua juga menampilkan keseimbangan yang kuat antara presisi dan *recall* (0,94–0,97), dengan F1-score rata-rata 0,95–0,96, serta jumlah kesalahan prediksi yang minimal. Pencapaian ini membuktikan bahwa penggunaan ANN dengan arsitektur yang lebih kompleks dapat memberikan prediksi yang sangat akurat terhadap kepuasan pelanggan. Implementasi model prediksi berbasis ANN tidak hanya meningkatkan kualitas pengalaman penumpang, tetapi juga memperkuat loyalitas pelanggan dan menjaga daya saing maskapai dalam jangka panjang.

Kata kunci: maskapai penerbangan, artificial neural network, kepuasan pelanggan, prediksi loyalitas pelanggan

Abstract

Customer satisfaction is a strategic factor for the sustainability of airline businesses amid increasingly intense competition in the aviation industry. This study aims to predict airline customer satisfaction using an Artificial Neural Network (ANN) approach by leveraging a publicly available Kaggle dataset containing 22 airline service features. Two ANN architectures were developed, differing primarily in the number of hidden layers, the number of neurons, and the application of Batch Normalization and LeakyReLU in the second model. The experimental results show that the first ANN model achieves an accuracy of 92.31%, while the second model attains significantly higher performance, with an accuracy of 95.75% on the test dataset. The second model also demonstrates a strong balance between precision and recall (0.94–0.97), with an average F1-score of 0.95–0.96 and a minimal number of misclassifications. These results confirm that employing a more complex ANN architecture can deliver highly accurate predictions of customer satisfaction. The implementation of ANN-based predictive models not only enhances passenger experience quality but also strengthens customer loyalty and helps airlines maintain long-term competitiveness.

Keywords: airline, artificial neural network (ANN), customer satisfaction, customer loyalty prediction

1 Pendahuluan

Kepuasan pelanggan menjadi faktor kunci dalam menentukan keberlanjutan dan pertumbuhan bisnis maskapai [1]. Meskipun pelanggan menuntut layanan yang lebih baik—mulai dari kenyamanan kursi, ketepatan waktu, hingga fasilitas dalam pesawat—maskapai sering kali kesulitan memprediksi

kepuasan tersebut karena sifatnya yang subjektif dan kompleks. Tantangan ini semakin mendesak di era digital, di mana kegagalan memenuhi harapan pelanggan dapat menyebar dengan cepat melalui media sosial dan berdampak negatif pada reputasi serta loyalitas pelanggan [2]. Jika masalah ini tidak segera diatasi, maskapai berisiko kehilangan daya saing di industri yang semakin kompetitif. Oleh karena itu, diperlukan metode analisis berbasis data yang mampu memprediksi tingkat kepuasan pelanggan secara akurat [3], sehingga maskapai dapat merancang strategi perbaikan layanan yang lebih tepat sasaran dan berdampak nyata [4].

Data pelanggan yang kompleks, seperti jarak penerbangan, layanan in-flight, kelas penerbangan, dan keterlambatan waktu, menimbulkan tantangan dalam memprediksi kepuasan pelanggan. Analis manual terhadap data dengan berbagai variabel ini tidak efektif, sehingga diperlukan solusi otomatisasi yang lebih canggih. Salah satu metode yang semakin relevan adalah penggunaan deep learning Artificial Neural Network (ANN), yang mampu mengolah volume data yang besar dan kompleks untuk menghasilkan prediksi yang akurat [5]. ANN ini bisa mengatasi relasi yang non-linear, memungkinkan analisis yang lebih dalam terhadap data yang kompleks dan beragam [6], sehingga memberikan solusi yang lebih efektif dalam memahami dan meningkatkan kepuasan pelanggan.

Model Artificial Neural Network (ANN) menawarkan solusi dengan kemampuannya dalam mengolah data dalam jumlah besar serta mengidentifikasi pola yang sulit ditangkap oleh metode tradisional [7]. Dengan melatih model ANN menggunakan berbagai variabel seperti kenyamanan kursi, keterlambatan, layanan dalam pesawat, dan kemudahan pemesanan, maskapai dapat memperoleh prediksi kepuasan pelanggan yang lebih akurat. Hal ini memungkinkan pengambilan keputusan berbasis data yang lebih efektif dalam meningkatkan layanan, mengurangi keluhan pelanggan, serta mempertahankan loyalitas mereka. Dengan demikian, implementasi ANN tidak hanya menjadi inovasi dalam analisis data kepuasan pelanggan, tetapi juga menjadi langkah strategis bagi maskapai dalam mempertahankan daya saing di industri penerbangan yang dinamis. Tujuan Penelitian ini adalah memprediksi kepuasan pelanggan maskapai dengan menggunakan machine learning.

2 Tinjauan Literatur

Penerapan *machine learning* untuk memprediksi kepuasan pelanggan telah menjadi topik sentral dalam penelitian terdahulu, terutama sebagai respons strategis terhadap tantangan industri pasca-pandemi COVID-19 dan persaingan yang semakin ketat. Sejalan dengan fokus studi ini, penelitian yang relevan telah membandingkan berbagai algoritma klasifikasi—seperti *K-Nearest Neighbors* (KNN), *Decision Tree* (DTC), *Logistic Regression* (LR), *Random Forest* (RF), *Naïve Bayes* (NB), dan *AdaBoost*—untuk menganalisis kepuasan pelanggan maskapai. Hasilnya menunjukkan bahwa variabel seperti *Online boarding*, *Inflight entertainment*, *Seat comfort*, dan *On-board service* memiliki korelasi kuat terhadap tingkat kepuasan. Dalam studi tersebut, *Random Forest* tercatat sebagai model terbaik dengan akurasi 89,20%, presisi 93,04%, dan F1-score 88,80% [1], yang membuktikan efektivitas *machine learning* dalam membantu maskapai meningkatkan kualitas layanan mereka.

Meskipun penelitian terdahulu berhasil mencapai akurasi yang tinggi dengan *Random Forest*, pendekatan tersebut masih memiliki keterbatasan dalam menangkap pola kompleks pada data non-linear. Oleh karena itu, penelitian ini mengusulkan penggunaan *Artificial Neural Network* (ANN) sebagai model utama. Berbeda dengan metode klasik, ANN memiliki kemampuan yang lebih baik dalam memodelkan hubungan yang abstrak dan mendalam antara fitur layanan penerbangan dan kepuasan pelanggan, sehingga diharapkan dapat menghasilkan klasifikasi yang lebih akurat.

Relevansi pendekatan ANN ini juga didukung oleh tren penelitian lintas bidang yang semakin banyak memanfaatkan kecerdasan buatan untuk meningkatkan akurasi prediksi, termasuk di sektor *e-commerce*. Sebagai contoh, studi yang menggabungkan *Structural Equation Modeling* (SEM) dan ANN terbukti efektif memprediksi niat pelanggan dalam layanan digital, seperti pada kasus *electronic word of mouth* (eWOM) di *social commerce* [8]. Pendekatan *hybrid* serupa bahkan telah diterapkan dalam sektor transportasi udara untuk menganalisis faktor layanan—seperti kenyamanan kursi dan hiburan—yang menunjukkan bahwa metode ini adaptif dan andal untuk studi kepuasan pelanggan.

Studi-studi tersebut menunjukkan bahwa ANN mampu mengidentifikasi hubungan non-linear yang tidak dapat ditangkap oleh pendekatan statistik tradisional, sehingga dapat meningkatkan akurasi prediksi model yang digunakan. Oleh karena itu, penggunaan ANN dalam penelitian ini juga relevan untuk menganalisis kepuasan pelanggan maskapai dengan memanfaatkan variabel layanan yang

beragam, sebagaimana telah diterapkan dalam penelitian e-commerce dan social commerce sebelumnya. Meskipun penelitian terdahulu telah menunjukkan bahwa ANN efektif dalam menangani data kompleks dan non-linear, studi ini berfokus pada penerapan ANN dalam prediksi kepuasan pelanggan maskapai dengan mempertimbangkan fitur-fitur spesifik yang memengaruhi pengalaman penumpang, sehingga dapat memberikan wawasan yang lebih terarah dalam meningkatkan kualitas layanan penerbangan. Dalam berbagai bidang, *Artificial Neural Network* (ANN) telah diakui sebagai metode andal untuk analisis data yang kompleks [9]. Salah satu bukti efektivitasnya terlihat dalam studi mengenai resistensi adopsi *mobile wallet* (m-wallet), yang mengombinasikan *Structural Equation Modeling* (SEM) dan ANN. Penelitian tersebut berhasil mengungkap hubungan non-linear antara berbagai faktor—seperti hambatan penggunaan, risiko, nilai, dan tradisi—terhadap keputusan konsumen [10]. Temuan ini menegaskan kemampuan ANN dalam memetakan pola perilaku yang sulit dideteksi oleh metode statistik linear biasa.

Meskipun berbeda industri, prinsip dasar penggunaan ANN untuk menggali pola kompleks data pelanggan tersebut sangat relevan dengan penelitian ini. Studi ini mengadaptasi pendekatan prediksi tersebut ke dalam konteks maskapai penerbangan, yang memiliki variabel layanan yang jauh lebih beragam dan spesifik. Dengan memanfaatkan kemampuan ANN dalam memodelkan pengalaman pelanggan yang dinamis, penelitian ini diharapkan dapat memberikan wawasan baru dan solusi yang lebih akurat untuk meningkatkan kualitas layanan di industri penerbangan, melengkapi temuan-temuan pada sektor lain sebelumnya.

3 Metode Penelitian

3.1 Sumber Data Dan Variabel Penelitian

Dalam penelitian ini, data yang digunakan diperoleh dari platform Kaggle, yang menyediakan dataset terkait kepuasan pelanggan maskapai penerbangan. Dataset ini mengandung berbagai variabel yang mencerminkan pengalaman pelanggan selama penerbangan, termasuk karakteristik individu seperti usia, jenis kelamin, dan jenis perjalanan, serta faktor-faktor layanan seperti kenyamanan kursi, kualitas makanan dan minuman, serta keterlambatan keberangkatan dan kedatangan. Dengan total 22 fitur dan satu target variabel, dataset ini memberikan wawasan komprehensif mengenai faktor-faktor yang mempengaruhi kepuasan pelanggan. Informasi ini dapat dilihat dalam Tabel 1.

Tabel 1 Target dan fitur dataset

Nomor	Atribut/Faktor	DataType	Deskripsi
1	Satisfaction	Object	Kategori kepuasan pelanggan, berisi nilai "satisfied" atau "dissatisfied".
2	Gender	Object	Jenis kelamin pelanggan, dengan nilai "Male" atau "Female".
3	Customer Type	Int64	Jenis pelanggan, terdiri dari "Loyal Customer" atau "Disloyal Customer".
4	Age	Int64	Usia pelanggan dalam angka tahun
5	Type of Travel	Object	Jenis perjalanan, terdiri dari "Business travel" atau "Personal travel".
6	Class	Int64	Kelas penerbangan, terdiri dari "Eco", "Eco Plus", atau "Business".
7	Flight Distance	Int64	Jarak penerbangan dalam satuan mil.
8	Seat comfort	Int64	Tingkat kenyamanan kursi dengan skala 0 - 5.
9	Departure/Arrival time	Int64	Kenyamanan waktu keberangkatan/kedatangan dengan skala 0 - 5.

convenient			
10	Food and drink	Int64	Penilaian terhadap makanan dan minuman yang disediakan, dalam skala 0 - 5.
11	Gate location	Int64	Penilaian lokasi gerbang keberangkatan, dalam skala 0 - 5.
12	Inflight wifi service	Int64	Kualitas layanan Wi-Fi dalam penerbangan, dalam skala 0 - 5.
13	Inflight entertainment	Int64	Penilaian terhadap hiburan dalam penerbangan, dalam skala 0 - 5.
14	Online support	Int64	Kualitas dukungan online maskapai, dalam skala 0 - 5.
15	Ease of Online booking	Int64	Kemudahan dalam melakukan pemesanan tiket secara online, dalam skala 0 - 5.
16	On-board service	Int64	Penilaian terhadap pelayanan di dalam pesawat, dalam skala 0 - 5.
17	Leg room service	Int64	Penilaian terhadap kenyamanan ruang kaki, dalam skala 0 - 5.
18	Baggage handling	Int64	Penilaian terhadap layanan penanganan bagasi, dalam skala 0 - 5.
19	Checkin service	Int64	Penilaian terhadap layanan check-in, dalam skala 0 - 5.
20	Cleanliness	Int64	Penilaian terhadap kebersihan kabin, dalam skala 0 - 5.
21	Online boarding	Int64	Penilaian terhadap kemudahan boarding secara online, dalam skala 0 - 5.
22	Departure Delay in Minutes	Int64	Waktu keterlambatan keberangkatan dalam satuan menit.
23	Arrival Delay in Minutes	Float64	Waktu keterlambatan kedatangan dalam satuan menit, memiliki nilai yang hilang (missing values).

Tabel ini menunjukkan variabel kepuasan pelanggan dikategorikan sebagai "satisfied" atau "dissatisfied," sehingga penelitian ini menjadi tugas klasifikasi biner. Beberapa fitur dalam dataset berskala ordinal (misalnya, tingkat kenyamanan kursi dari 0 hingga 5), sementara yang lain berupa data numerik kontinu, seperti jarak penerbangan dan keterlambatan dalam menit.

3.2 Model ANN 1 & 2

Penelitian akan menggunakan dua versi model arsitektur ANN. Dengan data pelatihan yang telah disiapkan, model Neural Network dibangun menggunakan library yang sesuai, seperti TensorFlow. Dalam proses ini, arsitektur jaringan ditentukan, termasuk jumlah lapisan (layers), jumlah neuron di setiap lapisan, serta pemilihan fungsi aktivasi yang tepat. Kode Snippet dapat di lihat dalam Kode Program 1 untuk model 1.0.

Dua arsitektur ANN digunakan untuk membangun model prediksi kepuasan pelanggan. Model 1 dapat dilihat dalam Tabel 2.

Tabel 2 Model ANN 1

Lapisan (layer)	Jenis Lapisan	Jumlah Neuron	Fungsi Aktivasi	Dropout Rate
Hidden Input Layer 1	Dense	32	ReLU	0.2
Hidden Input Layer 2	Dense	16	ReLU	0.2
Output Layer	Dense	2	Softmax	

Arsitektur pertama memiliki 32 neuron pada input layer, satu hidden layer (32 neuron) dengan fungsi aktivasi ReLU (Rectified Linear Unit) dan lapisan dropout untuk menghindari overfitting. Sementara itu, model kedua menggunakan arsitektur yang lebih kompleks dengan 128 input layer, 64 dan 32 neuron pada hidden layers serta menggunakan LeakyReLU untuk menangani masalah gradien yang menghilang (vanishing gradient). Pada baris kedua, terdapat lapisan pertama yang merupakan input layer dengan 32 neuron. Parameter `input_dim=X_train.shape[1]` digunakan untuk menetapkan dimensi input, yaitu jumlah fitur dalam data pelatihan. Fungsi aktivasi yang digunakan di lapisan ini adalah ReLU (Rectified Linear Unit), yang berfungsi untuk memperkenalkan non-linearitas ke dalam model. Selain itu, `kernel_initializer='he_uniform'` digunakan untuk menginisialisasi bobot dengan distribusi He Uniform, yang direkomendasikan untuk lapisan dengan fungsi aktivasi ReLU.

Pada baris keempat kode, terdapat hidden layer pertama dengan dropout 0.2, yang secara acak menonaktifkan 20% neuron selama proses pelatihan untuk mencegah overfitting. Hidden layer ini memiliki 16 neuron dengan fungsi aktivasi ReLU, dan mekanisme dropout yang sama digunakan dengan rate 20%.

Lapisan terakhir adalah output layer, yang didefinisikan dengan `Dense(2, activation='softmax')`. Output layer ini memiliki 2 neuron, sesuai dengan jumlah kelas dalam masalah klasifikasi yang sedang dihadapi. Fungsi aktivasi softmax digunakan pada output layer, yang berguna untuk menghasilkan probabilitas bagi setiap kelas, sehingga cocok untuk masalah klasifikasi multi-kelas. Secara keseluruhan, arsitektur model ini merupakan Multilayer Perceptron (MLP) dengan 2 hidden layers, dengan susunan arsitektur sebagai berikut: Input Layer (32 neuron) -> Dropout (20%) -> Hidden Layer 1 (16 neuron) -> Output Layer (2 neuron softmax).

Selanjutnya, Model 2 dapat dilihat dalam Tabel 3.

Tabel 3 Model ANN 1

Lapisan (layer)	Jumlah Neuron	Fungsi Aktivasi	Optimasi Tambahan	Dropout Rate
Hidden Input Layer 1	128	LeakyReLU alpha=0.1	Batch Normalization	0.2
Hidden Layer 2	64	LeakyReLU alpha=0.1	Batch Normalization	0.2
Hidden Layer 3	32	LeakyReLU alpha=0.1	Batch Normalization	0.2
Output Layer	2	Softmax		

Model ANN versi 2.0 memiliki arsitektur yang lebih kompleks dibandingkan dengan model sebelumnya, dengan peningkatan jumlah neuron dan penambahan mekanisme Batch Normalization di setiap hidden layer. Model ini dirancang untuk meningkatkan stabilitas dan konvergensi selama proses pelatihan. Pada lapisan pertama, terdapat 128 neuron, dengan parameter `input_dim=X_train.shape[1]` yang menetapkan jumlah fitur sebagai input model. Bobot awal diinisialisasi menggunakan He Uniform Initialization, yang cocok untuk fungsi aktivasi ReLU. Sebelum melewati fungsi aktivasi, diterapkan Batch Normalization, yang bertujuan untuk menormalisasi distribusi nilai input, mempercepat pelatihan, dan meningkatkan stabilitas model. Setelah aktivasi ReLU, diterapkan mekanisme Dropout dengan rate 0.2, yang akan secara acak menonaktifkan 20% neuron selama pelatihan untuk mencegah overfitting. Pada hidden layer kedua, terdapat 64 neuron dengan pola yang sama: bobot diinisialisasi menggunakan He Uniform, kemudian dilakukan Batch Normalization, diikuti oleh aktivasi ReLU, dan mekanisme Dropout 0.2. Pada hidden layer ketiga, terdapat 32 neuron, dengan prosedur yang serupa, termasuk Batch Normalization, aktivasi ReLU, dan Dropout 0.2.

Lapisan terakhir adalah output layer, yang terdiri dari 2 neuron dengan fungsi aktivasi softmax. Fungsi ini digunakan untuk menghitung probabilitas dari setiap kelas dalam masalah klasifikasi biner, memastikan bahwa output model berupa distribusi probabilitas yang dapat diinterpretasikan sebagai prediksi.

Secara keseluruhan, model

ini adalah Multilayer Perceptron (MLP) dengan 3 hidden layers, dengan susunan arsitektur sebagai berikut: Input Layer (128 neuron) -> Batch Normalization -> Activation (ReLU) -> Dropout (20%) -> Hidden Layer 1 (64 neuron) -> Batch Normalization -> Activation (ReLU) -> Dropout (20%) -> Hidden Layer 2 (32 neuron) -> Batch Normalization -> Activation (ReLU) -> Dropout (20%) -> Output Layer (2 neuron softmax). Dengan tambahan Batch Normalization, model ini diharapkan lebih stabil dalam pelatihan dan lebih tahan terhadap masalah vanishing/exploding gradients, serta dapat mencapai konvergensi lebih cepat dibandingkan model sebelumnya.

3.3 Alasan Pemilihan dan Penggunaan Jumlah Neuron, Batch Normalization, dan LeakyReLU

Pemilihan jumlah neuron yang terstruktur (seperti pola "mengerucut" 128 → 64 → 32 pada Model 2) bertujuan untuk melakukan ekstraksi fitur secara hierarkis. Lapisan awal dengan jumlah neuron yang lebih besar (128 neuron) bertugas menangkap fitur-fitur dasar dari data input yang kompleks. Lapisan-lapisan berikutnya dengan jumlah neuron yang semakin kecil berfungsi untuk memadatkan informasi tersebut menjadi fitur-fitur abstrak yang lebih representatif untuk klasifikasi akhir.

Strategi penambahan jumlah lapisan (*hidden layers*) pada Model 2 dibandingkan Model 1 didasarkan pada literatur yang menyatakan bahwa kedalaman jaringan (depth) berkorelasi positif dengan akurasi model dalam mempelajari pola non-linear yang kompleks. Seperti dijelaskan oleh Sharkawy [11] dan Uzair & Jamil [12], penambahan *hidden layer* memberikan kapasitas lebih bagi model untuk memetakan hubungan yang rumit antara variabel layanan penerbangan dan kepuasan pelanggan, yang tidak dapat ditangkap oleh arsitektur yang terlalu dangkal.

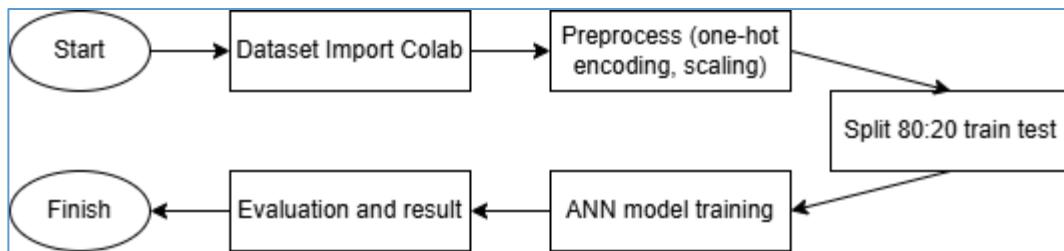
Fungsi aktivasi *LeakyReLU* (dengan alpha 0.1) dipilih untuk menggantikan ReLU standar pada model yang lebih dalam (Model 2) guna mengatasi masalah *vanishing gradient* atau "dying ReLU". Pada ReLU standar, neuron yang menerima input negatif akan menghasilkan output nol dan tidak memperbarui bobotnya, yang dapat mematikan sebagian jaringan selama pelatihan. *LeakyReLU* mengatasi ini dengan memberikan nilai gradien kecil (bukan nol) untuk input negatif, memastikan bahwa neuron tetap "hidup" dan terus belajar, sehingga menjaga aliran informasi yang efektif di seluruh lapisan jaringan.

Batch Normalization diterapkan di setiap lapisan tersembunyi pada Model 2 untuk menstabilkan dan mempercepat proses pelatihan. Teknik ini bekerja dengan menormalkan distribusi input yang masuk ke setiap lapisan (mengurangi *internal covariate shift*), sehingga model menjadi tidak terlalu sensitif terhadap inisialisasi bobot awal dan dapat menggunakan *learning rate* yang lebih tinggi. Hal ini sangat krusial untuk arsitektur yang lebih kompleks (Model 2) agar model dapat mencapai konvergensi yang optimal dengan lebih efisien dibandingkan tanpa normalisasi.

Salah satu aspek penting dalam ANN adalah fungsi aktivasi yang menentukan bagaimana sinyal dari satu neuron diteruskan ke neuron berikutnya. Dalam penelitian ini, digunakan fungsi aktivasi seperti ReLU (Rectified Linear Unit) dan Leaky ReLU untuk meningkatkan efisiensi dalam pelatihan model [13], serta mencegah permasalahan vanishing gradient yang sering terjadi pada jaringan saraf yang dalam. Selain itu, regularisasi seperti dropout digunakan untuk mengurangi kemungkinan overfitting, sehingga model dapat bekerja dengan baik pada data baru. Evaluasi model dilakukan dengan menggunakan metrik seperti akurasi, precision, recall, dan F1-score.

3.4 Alur Metodologi

Pada penelitian ini, jaringan saraf tiruan bekerja dengan cara mempelajari pola dari data historis pelanggan dan membuat prediksi berdasarkan pola tersebut. Metodologi penelitian ini menggunakan alur di diagram Gambar 1.



Gambar 1 Alur metodologi

Sebelum data digunakan dalam pelatihan model, proses pembersihan data dilakukan untuk memastikan kualitas dan konsistensinya. Data yang tidak lengkap, seperti entri yang memiliki nilai kosong atau null dalam variabel penting, harus ditangani. Jika jumlah data yang hilang signifikan, maka baris tersebut dapat dihapus, sedangkan untuk kasus kehilangan data dalam jumlah kecil, pengisian nilai berdasarkan rata-rata atau modus dapat diterapkan untuk mempertahankan integritas dataset.

Selain itu, deteksi dan penghapusan data duplikat dilakukan guna menghindari bias yang tidak perlu dalam model. Inkonistensi dalam data, seperti entri yang tidak valid atau kategori yang tumpang tindih, juga ditinjau dan dikoreksi, namun dari dataset ini tidak ada. Proses pembersihan data ini bertujuan untuk meningkatkan akurasi model dengan memastikan bahwa input yang diberikan adalah representasi terbaik dari data sebenarnya tanpa adanya noise atau informasi yang menyesatkan. Setelah proses pembersihan, transformasi data dilakukan agar sesuai dengan kebutuhan model pembelajaran mesin. Salah satu metode yang digunakan adalah one-hot encoding, yang bertujuan untuk mengkonversi variabel kategori seperti jenis pelanggan, jenis perjalanan, dan kelas penerbangan menjadi bentuk numerik tanpa mengubah hubungan hierarkis antar kategori. Hal ini penting untuk memastikan bahwa model dapat memahami perbedaan antara kategori tanpa menganggapnya sebagai nilai ordinal. Selain itu, normalisasi dilakukan pada fitur numerik seperti "Flight Distance," "Age," "Departure Delay in Minutes," dan "Arrival Delay in Minutes" menggunakan Standard Scaling. Normalisasi ini bertujuan untuk menyelaraskan skala fitur yang berbeda agar tidak ada satu fitur yang mendominasi perhitungan dalam model. Namun, variabel biner seperti "satisfaction" tidak perlu dinormalisasi karena sudah dalam skala tetap. Transformasi ini memastikan bahwa model dapat bekerja secara optimal dengan input data yang seragam.

Untuk melatih dan mengevaluasi model secara efektif, dataset dibagi menjadi dua bagian utama: data pelatihan dan data pengujian. Pembagian umum yang digunakan adalah 80% untuk pelatihan dan 20% untuk pengujian, yang merupakan rasio yang umum digunakan dalam pembelajaran mesin. Data pelatihan digunakan untuk mengajarkan model mengenali pola dalam dataset, sementara data pengujian digunakan untuk mengevaluasi seberapa baik model dapat menggeneralisasi data baru yang belum pernah dilihat sebelumnya.

Pembagian ini dilakukan secara acak untuk memastikan distribusi data yang seimbang antara kedua subset. Dalam beberapa kasus, teknik stratified sampling diterapkan untuk memastikan bahwa proporsi kelas dalam data pelatihan dan pengujian tetap konsisten, terutama karena penelitian ini merupakan klasifikasi biner dengan kategori "satisfied" dan "dissatisfied." Pemisahan data yang baik memastikan model dapat belajar secara efektif tanpa overfitting atau underfitting, sehingga hasil prediksi yang diperoleh lebih dapat diandalkan dalam skenario dunia nyata.

4 Hasil dan Pembahasan

Pengujian dilakukan dengan menggunakan data uji (X_{test} dan y_{test}) yang tidak pernah digunakan dalam proses pelatihan model. Model akan dievaluasi menggunakan metrik loss dan accuracy. Hasil yang diperoleh akan memberikan gambaran tentang seberapa baik model dapat menggeneralisasi terhadap data baru. Langkah selanjutnya adalah melakukan prediksi terhadap data uji. Prediksi ini dilakukan untuk mendapatkan kategori kepuasan pelanggan yang diperkirakan oleh model. Proses ini menggunakan fungsi `predict()` dengan hasil dikonversi menjadi kelas berdasarkan probabilitas tertinggi menggunakan `argmax()`. Selanjutnya, evaluasi performa model *Artificial Neural Network* (ANN) untuk prediksi kepuasan pelanggan dilakukan dengan menggunakan data uji (*test set*), yaitu data yang tidak digunakan selama fase pelatihan model. Landasan utama dari proses evaluasi ini adalah **matriks konfusi (confusion matrix)** yang bisa dilihat dalam Gambar 2.

		Predicted Negative	Predicted Positive
Actual Negative	Predicted Negative	True Negative (TN)	False Positive (FP)
	Predicted Positive	False Negative (FN)	True Positive (TP)

Gambar 2 Confusion matrix

Untuk memahami performa model lebih lanjut, dilakukan analisis menggunakan **classification report**. Metrik yang digunakan adalah di dalam Tabel 4.

Tabel 4 Metrik evaluasi

Metrik	Deskripsi
Precision	$Presisi = \frac{TP}{TP + FP}$
Recall	$Recall = \frac{TP}{TP + FN}$
F1-score	$F1 - Score = 2 \times \frac{Presisi \times Recall}{Presisi + Recall}$
Accuracy	$Akurasi = \frac{TP + TN}{TP + TN + FP + FN}$
Support	Jumlah sampel di setiap kelas.

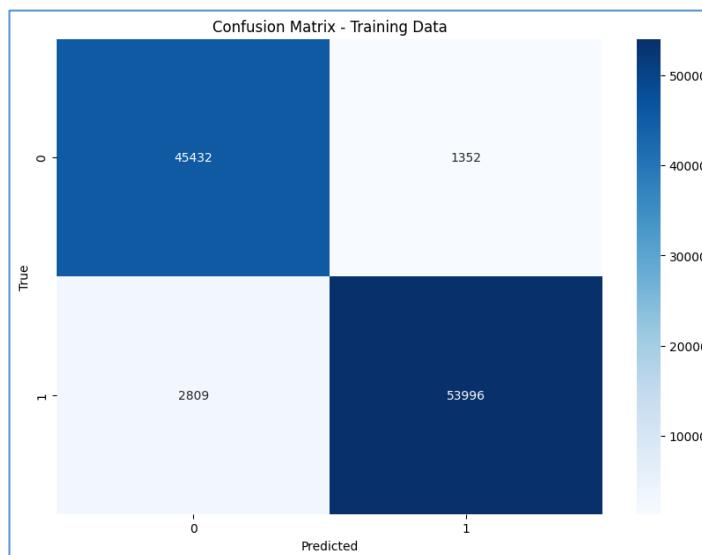
Pada penerapan Model ANN kedua, dalam hasil training mendapatkan akurasi yang sangat tinggi di angka 0.9598 atau 95.98 persen pada *training*. Untuk hasil testing mendapatkan 0.9575 atau 95.75 persen. Hasil bisa dilihat dibawah terkait akurasi, presisi, recall, f1-score dan confusion matrix untuk menentukan true positive, false positive, false negative dan false positive, serta grafik akurasi dan loss dalam Gambar 3-7.

Classification Report for Training Data:				
	precision	recall	f1-score	support
0	0.94	0.97	0.96	46784
1	0.98	0.95	0.96	56805
accuracy			0.96	103589
macro avg	0.96	0.96	0.96	103589
weighted avg	0.96	0.96	0.96	103589

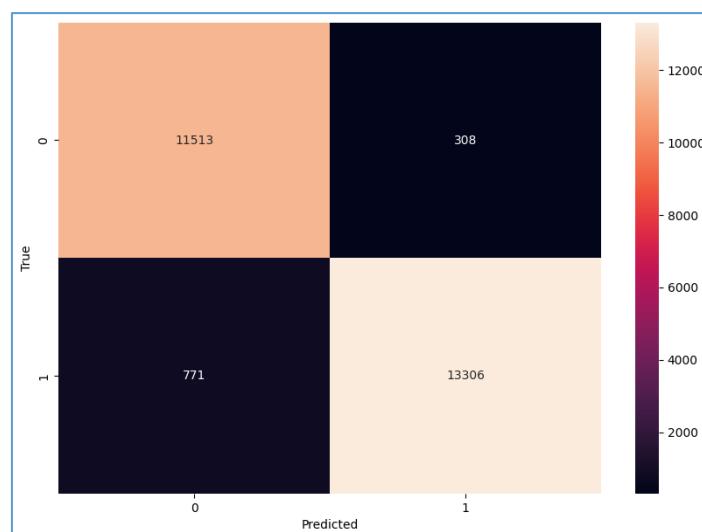
Classification Report for Testing Data:				
	precision	recall	f1-score	support
0	0.94	0.97	0.95	11821
1	0.97	0.95	0.96	14077
accuracy			0.96	25898
macro avg	0.96	0.96	0.96	25898
weighted avg	0.96	0.96	0.96	25898

Overall Training Accuracy: 0.9598
Overall Testing Accuracy: 0.9575

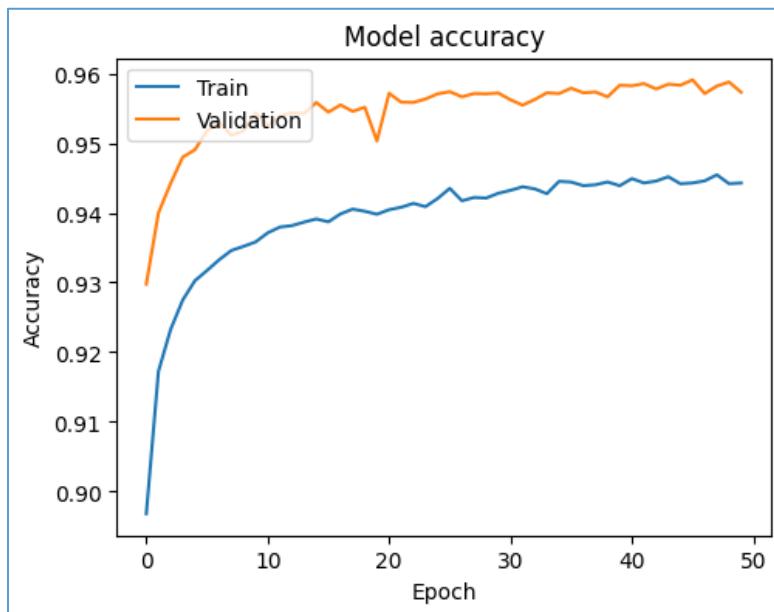
Gambar 3 Hasil classification report model 2



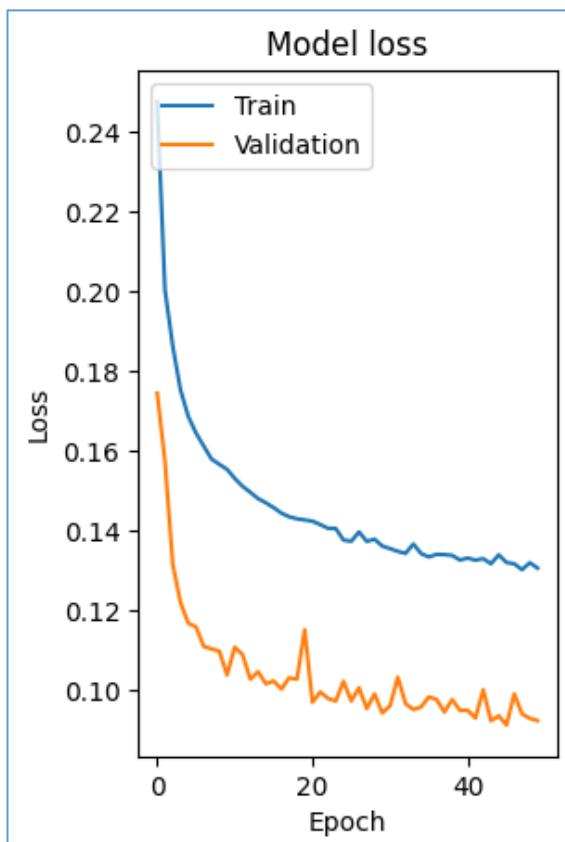
Gambar 4 Hasil confusion matrix training model 2



Gambar 5 Hasil confusion matrix testing model 2



Gambar 6 Hasil grafik akurasi model 2

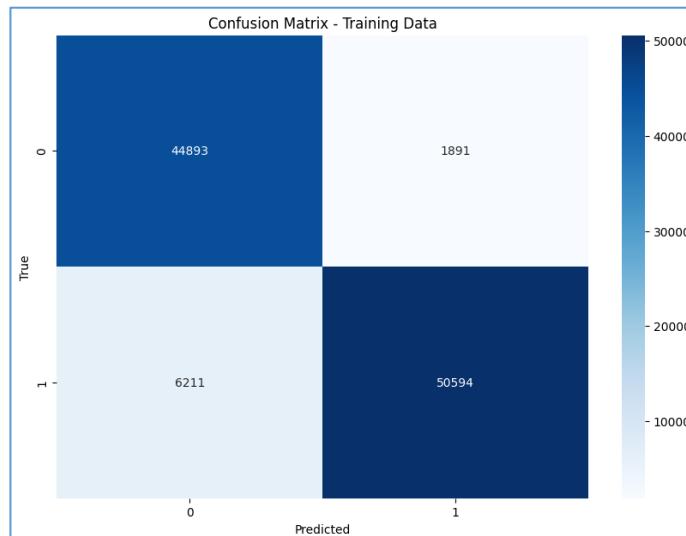


Gambar 7 Hasil Grafik Loss Model 2

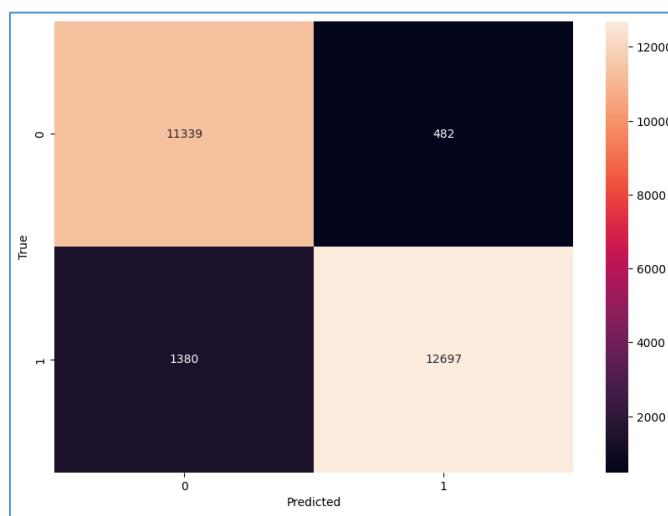
Untuk model ANN versi pertama, memiliki akurasi di 92.1 persen untuk data training, dan 0.9231 atau 92.31 persen untuk data testing, yaitu sekitar 3 persen lebih rendah dari model kedua. Hasil dapat ditemui dalam Gambar 8-12.

Classification Report for Training Data:				
	precision	recall	f1-score	support
0	0.88	0.96	0.92	46784
1	0.96	0.89	0.93	56805
accuracy			0.92	103589
macro avg		0.92	0.93	103589
weighted avg		0.93	0.92	103589
Classification Report for Testing Data:				
	precision	recall	f1-score	support
0	0.88	0.96	0.92	11821
1	0.96	0.89	0.93	14077
accuracy			0.92	25898
macro avg		0.92	0.93	25898
weighted avg		0.93	0.92	25898
Overall Training Accuracy: 0.9218				
Overall Testing Accuracy: 0.9231				

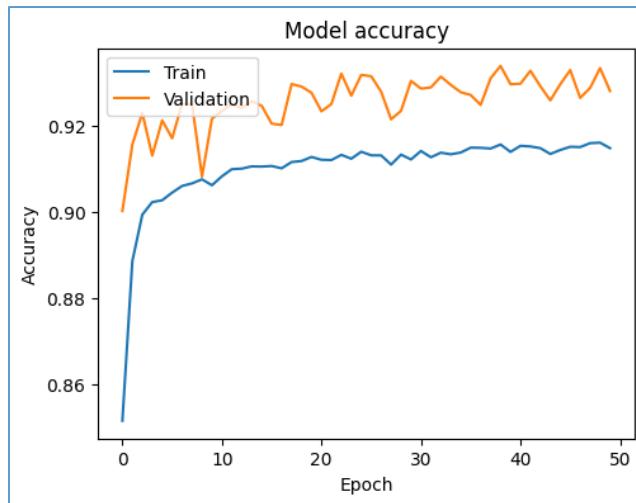
Gambar 8 Hasil Classification Report Model 1



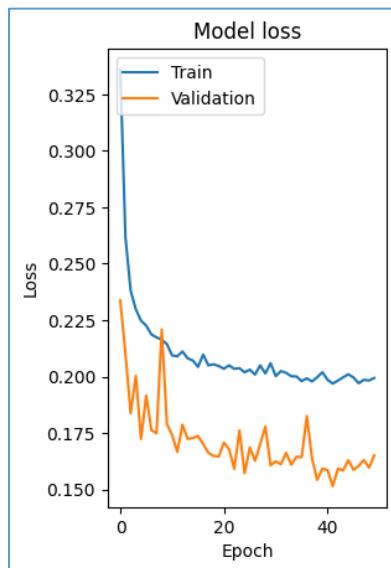
Gambar 9 Hasil confusion matrix training model 1



Gambar 10 Hasil confusion matrix testing model 1



Gambar 11 Hasil grafik akurasi model 1



Gambar 12 Hasil Grafik Loss Model 1

Model ANN kedua menunjukkan kemampuan yang seimbang dalam mengenali kedua kelas. Untuk kelas 'puas' (1), model mencapai presisi tinggi sebesar 97 persen dan recall 95 persen. Sementara itu, untuk kelas 'tidak puas' (0), model mencatatkan presisi 94 persen dan recall 97 persen. F1-score kedua kelas 0.960 untuk 'puas' dan 0.954 untuk 'tidak puas'.

Model berhasil mengklasifikasikan 13,306 pelanggan sebagai *True Positive* (benar-benar puas) dan 11,513 sebagai *True Negative* (benar-benar tidak puas). Hanya 771 kasus *False Negative* (pelanggan puas yang diprediksi tidak puas) dan 308 kasus *False Positive* (pelanggan tidak puas yang diprediksi puas).

Peningkatan jumlah *hidden layers* dari satu (pada Model 1) menjadi tiga lapisan teraktif (pada Model 2) memungkinkan model untuk mempelajari representasi fitur yang lebih hierarkis dan abstrak. Sebagaimana dikemukakan oleh Sharkawy [11], jaringan dengan lapisan yang lebih banyak (deep network) memiliki kapasitas yang lebih baik dalam memodelkan hubungan non-linear yang kompleks dibandingkan jaringan dangkal. Selain itu, penambahan jumlah neuron yang signifikan (dari 32 neuron pada Model 1 menjadi 128 neuron pada input Model 2) memberikan kapasitas komputasi yang lebih besar untuk menangkap variasi data pelanggan yang beragam. Hal ini sejalan dengan temuan Çolak [14] dan Tran et al. [15], yang menyimpulkan bahwa peningkatan jumlah neuron berkorelasi positif dengan akurasi prediksi karena model mampu menangkap pola data yang lebih halus. Konsep ini diperkuat oleh Uzair dan Jamil [12] dalam artikel mereka tahun 2020, "Effects of Hidden Layers on the Efficiency of Neural Networks". Mereka secara eksplisit menyatakan bahwa meningkatkan jumlah *hidden layer* adalah solusi efektif untuk meningkatkan akurasi neural network. Ini menunjukkan

adanya konsensus dalam literatur bahwa kompleksitas arsitektur, khususnya dalam hal jumlah layer, berkorelasi positif dengan performa.

Salah satu faktor kunci yang memungkinkan Model 2 mencapai konvergensi yang lebih baik—meskipun memiliki arsitektur yang lebih kompleks—adalah penerapan *Batch Normalization* pada setiap *hidden layer*. Pada Model 1, absennya normalisasi membuat model lebih rentan terhadap ketidakstabilan selama pelatihan. Sebaliknya, *Batch Normalization* pada Model 2 bekerja menstabilkan distribusi input di setiap lapisan, yang secara efektif mencegah masalah *overfitting* sekaligus mempercepat proses belajar. Hasilnya terlihat pada grafik *loss* yang lebih stabil dan keseimbangan yang tinggi antara presisi (0.94–0.97) dan *recall* (0.95–0.97).

Penggunaan *LeakyReLU* pada Model 2 memberikan keunggulan dibandingkan fungsi ReLU standar yang digunakan pada Model 1. Dalam jaringan yang lebih dalam, ReLU standar berisiko menyebabkan masalah *dying ReLU*, di mana neuron berhenti belajar karena output negatif dikonversi menjadi nol mutlak. Dengan menggunakan *LeakyReLU* ($\alpha=0.1$), Model 2 memastikan bahwa aliran gradien tetap terjaga bahkan untuk nilai input negatif, sehingga mencegah terjadinya *vanishing gradient*. Dampak dari perubahan ini adalah model yang lebih responsif dan mampu mendeteksi pola kepuasan pelanggan yang mungkin terlewatkan oleh model dengan aktivasi ReLU standar.

Studi oleh Sharkawy [11] menunjukkan neural network dengan enam layer menunjukkan performa yang lebih unggul dibandingkan dengan yang menggunakan satu, dua, atau tiga layer. Penelitian tersebut kemudian memperluas temuan ini dengan menambahkan lebih banyak layer neuron, dan hasilnya menunjukkan peningkatan akurasi yang signifikan. Hal ini mengindikasikan bahwa semakin banyak layer yang digunakan, semakin baik pula performa model dalam tugas-tugas tertentu. Lebih lanjut, pembahasan mengenai model convolutional neural network (CNN) oleh Josephine (2021) [16], yang berfokus pada dampak hidden layer terhadap peningkatan performa, juga mendukung argumen ini. Penelitian tersebut menyimpulkan bahwa hidden layer yang lebih banyak menghasilkan classification performance yang lebih baik. Temuan ini menggarisbawahi bahwa prinsip serupa berlaku di berbagai jenis neural network, termasuk CNN yang sering digunakan untuk tugas-tugas pengenalan gambar.

5 Kesimpulan

Berdasarkan analisis yang dilakukan terhadap dataset maskapai yang diperoleh dari platform Kaggle, model kedua ini menunjukkan performa yang luar biasa dan sangat seimbang. Dengan akurasi pengujian mencapai 95.75%, yang hampir sama persis dengan akurasi pelatihan (95.98%), model ini membuktikan kemampuannya untuk generalisasi pada data baru tanpa adanya *overfitting*. Kinerja pada setiap kelas sangatlah kuat: model ini sangat efektif dalam mengidentifikasi hampir semua pelanggan 'tidak puas' (*recall* 0.97) dan prediksinya juga sangat dapat diandalkan (presisi 0.94). Untuk kategori 'puas', model ini memiliki presisi yang sangat tinggi (0.97) dan mampu menangkap sebagian besar kasusnya (*recall* 0.95). Skor F1 yang tinggi dan hampir seragam (0.95–0.96) untuk kedua kelas menegaskan bahwa model ini tidak hanya sangat akurat tetapi juga memiliki keseimbangan yang sangat baik antara presisi dan perolehan (*recall*). Hasil dari matriks konfusi juga menegaskan efektivitas model, dengan jumlah kesalahan prediksi yang minimal (771 *False Negative* dan 308 *False Positive*) dibandingkan dengan total prediksi yang benar. Untuk pengembangan lebih lanjut, penelitian di masa depan dapat mengeksplorasi penggunaan model *ensemble learning* seperti Gradient Boosting (misalnya, XGBoost atau LightGBM), yang sering kali mampu memberikan peningkatan akurasi dengan menggabungkan beberapa model lemah menjadi satu model yang kuat. Selain itu, penerapan teknik interpretasi model seperti SHAP (*SHapley Additive exPlanations*) sangat direkomendasikan untuk mendapatkan wawasan yang lebih dalam mengenai faktor-faktor layanan yang paling signifikan dalam mempengaruhi prediksi kepuasan pelanggan, sehingga memberikan nilai bisnis yang lebih strategis bagi maskapai penerbangan.

Referensi

- [1] A. C. Y. Hong, K. W. Khaw, X. Chew, and W. C. Yeong, “*Prediction of US Airline Passenger Satisfaction using Machine Learning Algorithms*,” *Data Analytics and Applied Mathematics (DAAM)*, pp. 8–24, Apr. 2023, DOI: 10.15282/daam.v4i1.9071.

- [2] R. Shahid, A. S. Mozumder, R. Sweet, M. Hasan, M. Alam, M. Rahman, M. Prabha, M. Arif, P. Ahmed, R. Isalm, “*Predicting Customer Loyalty in the Airline Industry: A Machine Learning Approach Integrating Sentiment Analysis and User Experience.*” [Online]. Available: www.comien.org/index.php/comien
- [3] R. Pranav and H. S. Gururaja, “*Explainable Stacking Machine Learning Ensemble for Predicting Airline Customer Satisfaction,*” 2023, pp. 41–56. DOI: 10.1007/978-981-19-9225-4_4.
- [4] S. Ouf, “*An Optimized Deep Learning Approach for Improving Airline Services,*” *Computers, Materials and Continua*, Vol. 75, No. 1, pp. 1213–1233, 2023, DOI: 10.32604/cmc.2023.034399.
- [5] R. Thanga Selvi and I. Muthulakshmi, “*RETRACTED ARTICLE: An Optimal Artificial Neural Network based Big Data Application for Heart Disease Diagnosis and Classification Model,*” *J Ambient Intell Humaniz Comput*, Vol. 12, No. 6, pp. 6129–6139, Jun. 2021, DOI: 10.1007/s12652-020-02181-x.
- [6] S. Namasudra, S. Dhamodharavadhani, and R. Rathipriya, “*Nonlinear Neural Network based Forecasting Model for Predicting COVID-19 Cases,*” *Neural Process Lett*, Vol. 55, No. 1, pp. 171–191, Feb. 2023, DOI: 10.1007/s11063-021-10495-w.
- [7] J. T. Hancock and T. M. Khoshgoftaar, “*Survey on Categorical Data for Neural Networks,*” *J Big Data*, Vol. 7, No. 1, p. 28, Dec. 2020, DOI: 10.1186/s40537-020-00305-w.
- [8] A. Alnoor, V. Tiberius, A. Atiyah, K. Khaw, T. Yin, X. Chew, S. Abbas, “*How Positive and Negative Electronic Word of Mouth (eWOM) Affects Customers’ Intention to use Social Commerce? A Dual-Stage Multi Group-SEM and ANN Analysis,*” *Int J Hum Comput Interact*, Vol. 40, No. 3, pp. 808–837, Feb. 2024, DOI: 10.1080/10447318.2022.2125610.
- [9] D. Spicker, A. Nazemi, J. Hutchinson, P. Fieguth, S. Kirkpatrick, M. Wallace, K. Dodd, “*Challenges for Predictive Modeling with Neural Network Techniques using Error-Prone Dietary Intake Data,*” *Stat Med*, Vol. 44, No. 5, Feb. 2025, DOI: 10.1002/sim.70013.
- [10] L.-Y. Leong, T.-S. Hew, K.-B. Ooi, and J. Wei, “*Predicting Mobile Wallet Resistance: A Two-Staged Structural Equation Modeling-Artificial Neural Network Approach,*” *Int J Inf Manage*, Vol. 51, p. 102047, Apr. 2020, DOI: 10.1016/j.ijinfomgt.2019.102047.
- [11] A.-N. Sharkawy, “*The Effect of Increasing Hidden Layers on the Performance of the Deep Neural Network: Modelling, Investigation, and Evaluation,*” *Research on Engineering Structures and Materials*, 2024, DOI: 10.17515/resm2024.442st0909tn.
- [12] M. Uzair and N. Jamil, “*Effects of Hidden Layers on the Efficiency of Neural networks,*” in *2020 IEEE 23rd International Multitopic Conference (INMIC)*, IEEE, Nov. 2020, pp. 1–6. DOI: 10.1109/INMIC50486.2020.9318195.
- [13] X. Hu, R. Li, T. Chen, B. Zhang, L. Wang, J. Zhou, N. Gao, “*Effects of Different Activation Functions on Multilayer Perceptron Performance for Predicting Indoor Airflow Fields,*” *Build Environ*, Vol. 285, p. 113680, Nov. 2025, DOI: 10.1016/j.buildenv.2025.113680.
- [14] A. B. Çolak, “*A Novel Comparative Investigation of the Effect of the Number of Neurons on the Predictive Performance of the Artificial Neural Network: An Experimental Study on the Thermal Conductivity of <Scp> Zro₂ </Scp> Nanofluid,*” *Int J Energy Res*, Vol. 45, No. 13, pp. 18944–18956, Oct. 2021, DOI: 10.1002/er.6989.
- [15] T. T. K. Tran, S. M. Bateni, S. J. Ki, and H. Vosoughifar, “*A Review of Neural Networks for Air Temperature Forecasting,*” *Water (Basel)*, Vol. 13, No. 9, p. 1294, May 2021, DOI: 10.3390/w13091294.
- [16] V. L. Helen Josephine, A. P. Nirmala, and V. L. Alluri, “*Impact of Hidden Dense Layers in Convolutional Neural Network to Enhance Performance of Classification Model,*” *IOP Conf Ser Mater Sci Eng*, Vol. 1131, No. 1, p. 012007, Apr. 2021, DOI: 10.1088/1757-899X/1131/1/012007.