**Neural Network Project**

Delivered as part of the week 6 activites

Tuwaiq Academy

**Team members:**

* Abdullah Albuqmi
* Saif Alotaibi
* Abdulelah Alowaid

**Instructors:**

* Eng. Hany El shafey
* Eng. Yazeed Alzahrani

**Table of Contents**

| **Section** | **Page** |
| --- | --- |
| **Cover Page** | 1 |
| **Table of Contents** | 2 |
| **1. Introduction** | 3 |
| **2. Dataset Overview** | 4 |
| **3. Data Preprocessing** | 5 |
| **4. Model Design & Architecture** | 6 |
| **5. Training Strategy & Monitoring** | 7 |
| **6. Results & Evaluation** | 8 |
| **7. Insights & Interpretation** | 9 |
| **8. Conclusion & Future Work** | 10 |
| **9. Acknowledgments** | 11 |

**Introduction**

Customer satisfaction plays a critical role in the airline industry, directly influencing loyalty, operational improvements, and revenue. This project aims to build a deep learning-based classifier capable of predicting whether a passenger is satisfied or dissatisfied based on demographic data, travel details, and in-flight service feedback.

The primary goal is to evaluate different neural network architectures, optimize training strategies, and identify which configurations deliver the highest predictive accuracy.

**Dataset Overview**

| **Attribute** | **Details** |
| --- | --- |
| Source | Kaggle – *Airline Passenger Satisfaction Dataset* |
| Number of Records | **89,127 rows** |
| Number of Features | **23 columns** |
| Target Variable | **Satisfaction** (0 = Not satisfied, 1 = Satisfied) |

**Key Features Used**

* **Demographic**: Gender, Age, Customer Type
* **Travel Details**: Class (Economy, Eco Plus, Business), Type of Travel (Business/Personal), Flight Distance
* **Service Ratings (1–5)**: Inflight Wi-Fi, Online Booking, Food & Drink, Inflight Entertainment, Gate Location, Leg Room, Cleanliness, etc.
* **Performance Metrics**: Departure Delay, Arrival Delay

All categorical features were encoded numerically to be suitable for deep learning.

**Data Preprocessing**

1. **Missing Values**
   * Rows with missing values in *Arrival Delay* were removed.
2. **Encoding**
   * Target Satisfaction mapped to 0/1 binary class.
   * Categorical features (Gender, Class, Travel Type) encoded via one-hot / label encoding.
3. **Normalization**
   * Numerical columns such as Flight Distance and Delay Minutes were scaled.
4. **Train-Test Split**
   * Data split into **Training (70%) / Validation (15%) / Test (15%)** sets.

**Model Design**

A **dynamic Neural Network builder** function was developed to create models with tunable:

* Optimizer (Adam, RMSprop, SGD with momentum)
* Dropout values (single or dual-layer dropout such as **0.3 / 0.5**)
* Learning Rate
* Batch Normalization

**Best Architecture (Selected Model)**

| Layer Type | Configuration |
| --- | --- |
| Dense Layers | 4 hidden layers |
| Activation | ReLU for hidden, Sigmoid output |
| Optimizer | Adam |
| Loss Function | Binary Crossentropy |
| Regularization | Dropout + Batch Normalization |

**Training & Monitoring**

A custom Keras callback ("FitMonitor") was implemented to analyze training status *during each epoch*, assessing signs of overfitting or underfitting in real time.

* RMSprop → Converged fast but plateaued early
* SGD + Momentum → Required more epochs, but produced smoother curves
* Adam → Delivered best balance between fast convergence and generalization

**Results & Evaluation**

| Model Variant | Optimizer | Dropout | Validation Accuracy |
| --- | --- | --- | --- |
| Model A | RMSprop | 0.2 | ~92% |
| Model B | SGD-M | 0.3 | ~90% |
| Model C (Best) | Adam | 0.3–0.5 | ~94–96% |

**Final Test Accuracy: ~96%**

**Fit Status: Just Right** — neither underfitted nor overfitted

Most influential features (based on model sensitivity):

* Inflight Service & Cleanliness
* Travel Class (Eco vs Business)
* Wi-Fi Rating & Leg Room Service
* Arrival Delay

**Insights & Interpretation**

* Service Quality > Demographics → Passenger perception of comfort and onboard experience drives satisfaction more than age or gender.
* Delay Tolerance Exists → Short delays were often tolerated if service ratings were high.
* Dropout Strategy Matters → Dual dropout values (0.3 / 0.5) helped significantly reduce overfitting.

**Conclusion & Future Work**

This project successfully demonstrates that deep learning can accurately predict airline passenger satisfaction using structured survey and travel data.

The Adam-based 4-layer dense neural network achieved up to 96% accuracy, and the custom FitMonitor callback improved training efficiency by dynamically evaluating fit quality.

Future Enhancements

* Deploy as an API or Web Dashboard for airlines
* Integrate Explainable AI (e.g., SHAP values)
* Predict Net Promoter Score (NPS) or Likelihood to Recommend
* Expand dataset to include airline metadata (routes, aircraft type, price class)

**Acknowledgments**

This work was developed as part of Tuwaiq Academy – Week 6 Deep Learning Activities. Special thanks to Kaggle for dataset availability and TensorFlow/Keras for enabling rapid experimentation.