<u>Air Traffic Management:</u> Neural Network Method

Instructions:

NN & LR&ANFIS	• Description of steps taken to construct the NN, LR and ANFIS models and interpreting results, including 1. Choice of model structures, 2. Which sub-dataset(s) is(are) used for training, testing and validation, 3. Quantify and visualise the performance of the predictive model on all sub-datasets, 4. Conducting the statistical tests of the regression models, 5. Assessing transparency and generalisation capability of the model.	35%

<u>Choice of model structures:</u>

- Multilayer Perception (MLP)
- Uses 3 Hidden Layers (128, 64, 32 neurons for respectively)
    - All using activation ReLU - good for learning non-linear patterns
- Output Layer = Single Linear Neuron → Continuous Taxi-Time Prediction
- Model = Adam for learning rate 0.001
- Trained Using = Mean Squared Error (MSE)
- Forward Propagation = x_train through all layers, produces predicted taxi time
- Loss Function = difference between predicted vs actual y_train
- Back Propagation = compute gradient of loss with all weights
- Weight updates in 32 batches

-

Sub-dataset(s) to use:

- Training: 80%
- Testing: 10%
- Validating: 10%

Results:

- Mean Absolute Error : 2.107103002645689
- Root Mean Squared Error: 2.6150633207971357
- R^2: 0.6652033884022683

T-test

- t-statistic = -16.117669702233762
- p-value = 1.0945576693013282e-49

95% CI

- Mean Residual = -1.3785734906035556
- 95% CI for Mean Residual = (np.float64(-1.5465138911930207), np.float64(-1.2106330900140905))

Conducting the statistical tests:

1. Evaluation on Performance and Reliability of NN (after MAE and RMSE) - Statistical Analysis on Residuals
    a. One Sample T-testing = determines how different the mean residual is to 0
        i. Mean residual: 0, model not systematically overpredict/underpredict
        ii. Mean residual big difference to 0, model is bias
        iii. Justification: a low MAE does not guarantee the model is unbiased across all predictions, t-testing quantifies this by checking if there are deviations from zero, and how much it deviates.
    b. 95% confidence interval = estimates the uncertainty in model prediction biasness
        i. If 95% confidence interval (CI): contains 0, model has no significant systematic error
        ii. If 95% confidence interval (CI): contains no 0, model consistently over/under estimate taxi time
        iii. Justification: gives a range of estimations, reinforcing t-test.
2. Overall these were chosen as RMSE and MAE describes magnitude of error, but not:
    a. Location of error: centered around zero or not
    b. Bias or Unbias of the model
    c. Reliability of the average prediction error

Assessing Transparency and Generalisation Capability:

Comparison	• Discussing the pros and cons of the NN, LR and ANFIS approaches in taxi time prediction from the aspects of prediction accuracy, generalisation capability and model transparency. • Quantifying and visualising the comparison results in the report.	20%

Discussion on the pros and cons:

Quantifying and visualising results:

Conclusions	• Drawing conclusions of the study, as well as further crucial steps for machine learning approaches to be applied in real operational environment at airports.	10%

Conclusions:

Code:

```
import tensorflow as tf
from tensorflow.keras import layers, models
from sklearn.metrics import mean_absolute_error, mean_squared_error, r2_score
import numpy as np

tf.random.set_seed(42)

n_features = X_train.shape[1]

nn_model = models.Sequential([
   layers.Input(shape=(n_features,)),
   layers.Dense(128, activation='relu'),
   layers.Dense(64, activation='relu'),
   layers.Dense(32, activation='relu'),
   layers.Dense(1)
])

nn_model.compile(
   optimizer=tf.keras.optimizers.Adam(1e-3),
   loss='mse',
   metrics=['mae']
)
```