Defined Architecture of the ANN Model

Project Name: Customer Churn Prediction

1. Input Layer

Number of Input Features:

The input layer receives 8 features after preprocessing and encoding:

Gender

SeniorCitizen

Dependents

Tenure

PhoneService

MultipleLines

InternetService

Contract

MonthlyCharges

Preprocessing:

Categorical features have been encoded using LabelEncoder.

Numerical features such as MonthlyCharges and tenure have been scaled using StandardScaler.

2. Hidden Layers

Architecture

Two hidden layers of 64 neurons were selected to sufficiently learn the non-linear associations between observations.

Activation Function:

ReLU was chosen as the activation function since it's efficient to handle non-linearity.

Dropout:

No explicit dropouts were made, but for reducing overfitting, extra regularization may be done.

Weight Initialization:

Default initialization in MLPClassifier using adam solver.

3. Output Layer

Number of Neurons:

1 neuron - binary classification for "Churn" prediction encoded into 0 and 1.

Activation Function:

Softmax activation is handled internally by the MLPClassifier for classification problems.

Loss Function:

The adam solver makes use of Cross-Entropy Loss for binary classification by default.

4. Optimization

Optimizer:

adam optimizer was chosen because it is efficient, and it does adaptive learning.

Learning Rate:

By default, the learning rate was taken as 0.001.

Max Iteration:

Number of iterations set at 500 to converge.

Batch Size:

MLPClassifier does internal batching.