

# Project 2

Rodrigo Petricoli

12/3/2021

## Overview

The proposed model used for this Project is an Artificial Neural Network. For this project the feature engineering was already done so all that was left was to separate the data into training and testing sets (by converting them first to NumPy arrays). I used 70% of the data set to train the data and the other 30% to test it.

## Model

The model selected was made up off an input layer with 78 nodes (one for each feature of the data set), and two hidden layers, of 64 and 128 nodes each. After some trial and error, I found that even by adding many more layers with any number of nodes the model didn't perform better, so I chose to use 'only' two hidden layers. For the activation function, I selected the *tanh* function. I tried the ReLU, the sigmoid, and the Leaky ReLU, and found that the model performed slightly better using the *tanh*. I also added Dropouts for the model at 0.5 for each layer and added EarlyStopping with patience at 5 to try avoiding over fitting and to use computing power more efficiently.

## Results

After training the model, I tested it and computed a classification matrix. The results were as follows:

	precision	recall	f1-score	support
0	0.82	0.79	0.80	14637
1	0.79	0.82	0.81	14340
accuracy			0.80	28977
macro avg	0.80	0.80	0.80	28977
weighted avg	0.80	0.80	0.80	28977