IDA
Impact Design Academy
**Introduction to Machine Learning**
Due Date: Aug 18, 2023

Timing: 1 Hour                    Intro to ML                    Max mark: 50

**Attempt all five questions**

1. Write a Python function that takes in a dataset and splits it into training and testing sets, using an 80-20 split ratio. [10]

2. Implement a Python class for a decision tree node, capable of handling both categorical and numerical features. Include methods for splitting nodes based on different criteria, such as Gini impurity or entropy. [10]

3. Develop a Python function that calculates the accuracy of a classification model by comparing its predictions with the true labels of a test dataset. [10]

4. Create a Python function that performs k-fold cross-validation on a given dataset, using a specified number of folds. Print the average accuracy across all folds. [10]

5. Write Python code to implement a random forest classifier using a specified number of decision trees. Train the model on a given dataset and evaluate its accuracy on a separate test dataset. [10]

**Best wishes**
*for reference* https://www.youtube.com/watch?v=L7R4HUQ-eQ0