**Continuous Assessment Project Report**

**Nachiket Mehendale (x23272473)**

**Data Analytics for Artificial Intelligence – H9DAI**

**MSCAI\_SEP24**

**School of Computing**

**National College of Ireland**

**Title**

The title of the report is Predicting Visa Applications with Regression Algorithms. In this study, we are considering the visa applications dataset published by the Department of Justice of Ireland [1]. The report demonstrates how we addressed 5 learning outcomes by analysing the visa applications dataset - (i) Retrieve, extract, manipulate, synthesize, explore, and visualize data in preparation for data analysis and machine learning. (ii) Concepts and methods associated with the analysis of data using numerical and statistical techniques to assist on decision-making. (iii) Fundamental machine learning concepts and techniques to build and evaluate machine learning models on various problem domains. (iv) Evaluate and employ graphical tools for building comprehensive analytics processes and dashboards. (v) Critically analyze, compare, summarize, and present results to support decision making and address requirements in real-world problems.

**Instructions to run the Jupyter File:**

**Python Environment**: Python 3.12.2

**IDE**: Jupyter Notebook

**Required Libraries**:

import pandas as pd

import seaborn as sns

import numpy as np

import matplotlib.pyplot as plt

from sklearn.model\_selection import train\_test\_split, cross\_val\_predict, KFold

from sklearn.preprocessing import StandardScaler

from sklearn.metrics import mean\_squared\_error, mean\_absolute\_error, r2\_score

from sklearn.linear\_model import LinearRegression

from sklearn.ensemble import RandomForestRegressor, GradientBoostingRegressor

from sklearn.svm import SVR

**Dataset :** The same dataset file which was given during the CA Test:

“D3\_Paper and online short term visa applications received by year and nationality.csv”

NOTE: Make sure there are NO ENCDONG Issues in the dataset file. If yes then encode it in UTF and then use it for analysis.

**Jupyter File Name:** CA\_Project\_x23272473.ipynb

**How to Execute the file:**

1) First place the dataset file in a particular folder which has read/write access, for example in our case - C Drive

2)Make sure the dataset file is also NOT Locked and has READ/WRITE access.

3) Keep the jupyter file CA\_Project\_x23272473.ipynb in the Python Executable Scripts Folder.

4) Run the Jupyter Notebook command from the terminal and Open the notebook mentioned in point no 3.

5) Ensure the dataset file path in the Jupyter file and if not, provide the correct path so that it can be read by panda.

6)Run all the cells

**Note** : Make sure available libraries/packages are installed properly in your python environment.