

**IHC – Artificial Intelligence for Real World Application**

**Industry Assignment No. 1**

**Problem Statement**

An agency has decided to study the different factors that affect the country’s GDP. For this purpose, they have collected the data of 227 countries which includes – name of the country, region, population, area, population density, coastline area, net migration, infant mortality, GDP, literacy, phones per 1000, arable, crops, climate, birth-rate, deathrate, agriculture, industry, service, others.

**Data Set**

CSV dataset – gdpWorld.csv

**What is expected**

Identify at least six factors that affect the GDP of a country. Apply any linear and non-linear model on the dataset with these six factors and compare the two models to find which gives a more accurate prediction.

**Rubrics**

|  |  |  |
| --- | --- | --- |
| 1 | Data Collection - Completeness and correctness of data collection | 10 |
| 2 | Data preparation - cleansing and preparing data for the next steps of model development | 10 |
| 3 | Algorithm Understanding - Knowledge of the algorithm used for model development | 15 |
| 4 | Model Development - Building the model with the right selection of the parameters | 20 |
| 5 | Model Tuning - Improving the model performance by tuning the hyper parameters | 15 |
| 6 | Model Evaluation - Testing the model on sufficient sample for correctness | 10 |
| 7 | i) Final Presentation and QA ii) Demonstration of the Working Demo of the system. | 20 |

Students are also given a logical flow as a skeleton to the overall project, but they are advised to add further steps where they feel necessary based on the rubrics defined above for maximum marks.

Developing a model to find the factors that affect the GDP of a country