Impact of Covid-19 Pandemic on the Global Economy Machine Learning

Data Science Immersive

MAIN TOPICS

POINTS TO TALK ABOUT

The Objective

Data

Target Column

Data pre-processing

Feature importance

Models

Results

THE OBJECTIVE

Predicting the impact of Covid-19 on the global economy especially in Gross Domestic Product Per Capita on citizens.

DATA

CODE

COUNTRY

DATE

HDI

TC

TD

STI

POP

GDPCAP

GDP is one of the most important economic indicators; it shows the short-term development of an economy.

DATA PRE-PROCESSING

FILL MISSIG VALUES

Fill missing values in HDI column by median.

DROP SOME COLUMNS

Before dummy, I dropped CODE column it was duplicated column.
After dummy, I dropped all categorical columns.

GET DUMMIES

Get dummy on one column that I had it was COUNTRY column.

1,144.411

CHINA MOST COUNTRY IN THE WILLING TO FOLLOW THE RULES OF A PANDEMIC

1,096.837

SINGAPORE COMES SECOND AFTER CHINA

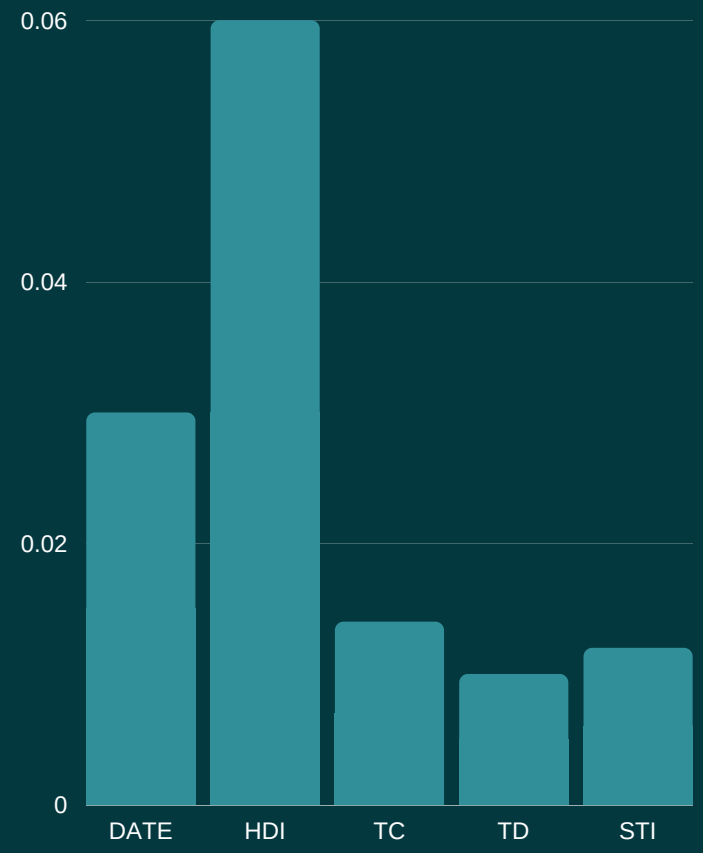
0.953

HDI IN AFGHANISTAN

Citizens of Afghanistan have more chances of surviving a COVID-19 pandemic

FEATURE IMPORTANCE

Feature importance is calculated as the decrease in node impurity weighted by the probability of reaching that node. The higher the value the more important the feature.



Models

LINEAR RGRESSION

Model is to find a relationship between features and a continuous target variable

DECISION TREE REGRESSION

The model breaks down a dataset into smaller and smaller subsets

XGBOOST REGRESSION

The model a powerful approach for building supervised regression models

RANDOM FOREST REGRESSOR

It is a supervised learning algorithm that uses ensemble learning method for regression

Results

LINEAR RGRESSION

Mean squared error is: 0.00

DECISION TREE REGRESSION

Mean squared error is: 0.00

XGBOOST REGRESSION

Mean squared error is: 0.47

RANDOMFOREST REGRESSOR

Mean squared error is: 0.04

THANK YOU FOR YOUR ATTENTION

Questions?