Data for Good

Challenge Statement – AI/ML

Case Description:

Your team is partnering with the United Nations to better understand Goal 4: Equitable Quality Education and Lifelong Learning for All. United Nations data is often plight by missing data. To understand how the SDG goals will be attained by year 2030, we need complete data where relevant to measure progress.

Can you identify measurement issues, understand the multi-level uniqueness and visualize the pace of growth towards 2030 target?

Tasks for Participants:

- Run an Exploratory Data Analysis
- Are there any missing data? If yes, explore ways to fill the missing data without dropping them (Focus on those with missing data from 2018).
- Split the data (years 2000 – 2018 and 2019), in order to create a forecast model for 2019. How close are the predictions to the actual values? What models were explored? How was the final model selected? Justify modelling and evaluation decisions: what metrics were used?

Data Sets Used:

United National Global SDG Database:
- Goal 4 (with 9 targets): Ensure inclusive and equitable quality education and promote lifelong learning opportunity
- Available in your team's Github repo or Jupyter Notebook Server

Deliverables:

- Digital presentation of findings/recommendation (slide deck, dashboard, poster, etc.)
  - Email by 5pm EST to data.for.good@jpmchase.com

Evaluation Metrics:

- Solutionsions will be evaluated on the data science methods used, technologies leveraged, and model's accuracy
- Presentations will be evaluated on creative, thoughtful, and meaningful insights, and effective story-telling