Machine Learning "Hello World" Task

Hello Fellow Interns,
For your Second Week, you're being requested to go through the Dataset Namely **"_Titanic - Machine Learning from Disaster_"**. This dataset is originally from www.kaggle.com.The good thing about this dataset is that it has already been split into 2, test and train. The train set contains 12 columns, most important columns are passengerid and Survived. The Test set on the other hand has 11 columns, similar to the train set except for the survived column. Notice that the survived columns have only two values: 0 and 1, 0 means didn't survive and 1 means survived.

Dataset: https://www.kaggle.com/c/titanic/data

Your Tasks:
- Perform basic pipelining tasks such as cleaning and formatting
- Perform Regression: make a machine learning model, train it with the training data
- Test your model with the test data
- Measure Performance and Validation (Cross-Validate)

Tools & Libraries You may need:
- Python
- Matplotlib, pandas, Scikit-Learn
- Jupyter Notebook

- Download the data from the link given above.
- Now, you can start of your work by opening a Jupiter notebook in Google Colab: https://colab.research.google.com/notebooks/intro.ipynb Or Anaconda if you already have it pre installed on you PC.
- Upload and load the data on to the notebook and you're good to go.

Due Date: Monday, 14th of June 2021, 8:00 PM

For any doubts contact the following mentors:
- Vybhav Chaturvedi +91 8586094922
- Adarsh Raghav +91 8527366600

After you're done with your work show it to us and then post it on LinkedIn by Tagging all mentors and Cureya Team, Shivani Mishra.