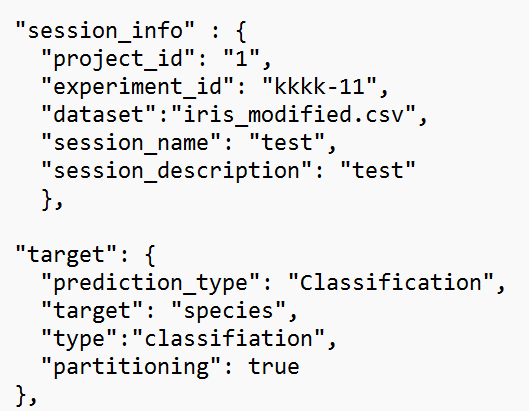
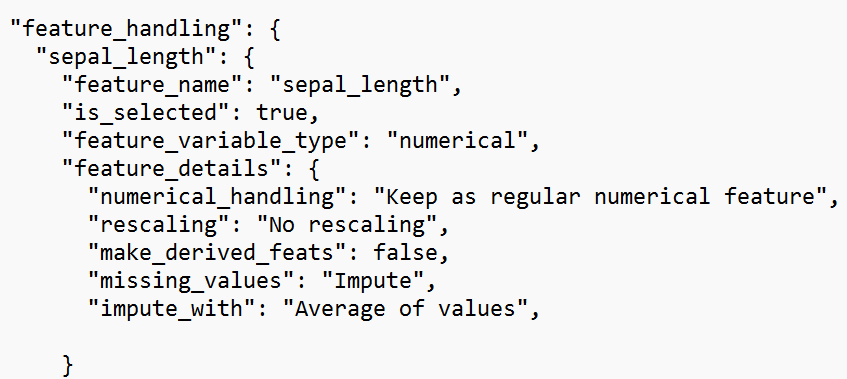
**Hackathon Stage 2 Problem**

As part of the coding assessment, you will write code to parse the JSON file provided (algoparams\_from\_ui) and kick off in sequence the following machine learning steps programmatically.

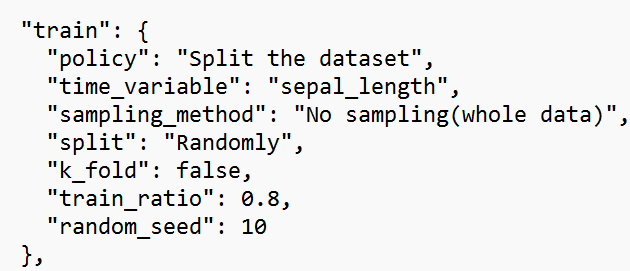
Keep in mind your final code should be able to parse any Json that follows this format. It is recommended that you write UDFs for automating these steps. It is crucial you have a generic parse that can read the various steps like feature handling, feature generation and model building using Grid search after parsing hyper params.

1) Identify the data to be used, the type of problem to be solved and the target variable of the problem. Create a dataframe based on the datafile provided.



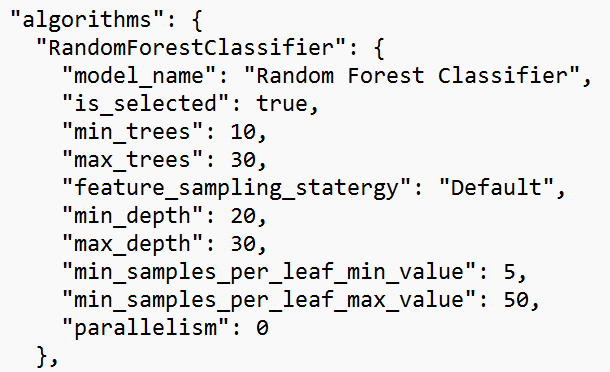
2) Read the features (which are column names in the csv) where “is\_selected”=True and create the X and the Y from the dataframe.

3) Split the data into training and validation sets based on the given split ratio and the random seed value.



4) Parse the Json and identify the model where the algorithm “is\_selected”=True.

Identify the hyperparameters to be used to build those models. Build the selected models and perform hyperparameter tuning using GridSearchCV.



5) Evaluate the models based on the type of problem. Print the results which can be used to compare and conclude upon which is a better model. Use the following evaluation metrics:

Classification metrics: confusion matrix, classification report

Regression metrics: R-squared, Adj R-squared, RMSE

**Note:**

1. **Plagiarism will lead to elimination.**
2. **Please do submit the code even if you have solved the problem partially.**
3. **Perform the step 4 on atleast 3 algorithms as a prototype, the more the better.**
4. **Please make sure you submit everything via a Github link and pls upload all assets and files.**