# Report: Predict Bike Sharing Demand with AutoGluon Solution

#### Ahmed Zidane

## Initial Training

### **What did you realize when you tried to submit your predictions? What changes were needed to the output of the predictor to submit your results?**

TODO: You must drop all negative values and set them to zero before submission

### What was the top ranked model that performed?

TODO: Model name : Initial model (Tablur Predictor) | score: 1.80832

## Exploratory data analysis and feature creation

### **What did the exploratory analysis find and how did you add additional features?**

TODO: EDA found the negative values, then we changed it. and get some descriptive information and we changed the datetime format , features are [ year, month, day, hour]

### How much better did your model preform after adding additional features and why do you think that is?

TODO: The score changed from 1.80832 to 0.68068

## Hyper parameter tuning

### How much better did your model preform after trying different hyper parameters?

TODO: Not much it decreases from 1.80832 to 0.68068

### If you were given more time with this dataset, where do you think you would spend more time?

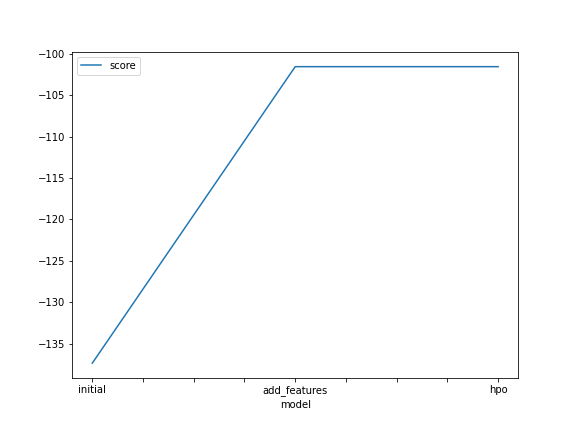
TODO: Yes, as adding additional features and changing hyperparameters led to a direct improvement in the kaggle score.

### Create a table with the models you ran, the hyperparameters modified, and the kaggle score.

| Model | hpo1 | hpo2 | hpo3 | score |
| --- | --- | --- | --- | --- |
| initial | time\_limit=600 | presets=’best\_qulity’ | Default values | 1.80832 |
| add\_features | time\_limit=600 | presets=’best\_qulity’ | problem\_type=’regression’ | 0.68874 |
| hpo | 'num\_epochs': 10, 'learning\_rate': ag.space.Real(1e-4, 1e-2, default=5e-4, log=True), 'activation': ag.space.Categorical('relu', 'softrelu', 'tanh'), 'layers': ag.space.Categorical([100], [1000], [200, 100], [300, 200, 100]), 'dropout\_prob': ag.space.Real(0.0, 0.5, default=0.1) | 'GBM': gbm\_options,  'NN': nn\_options | Tabular Autogluon | 0.68868 |

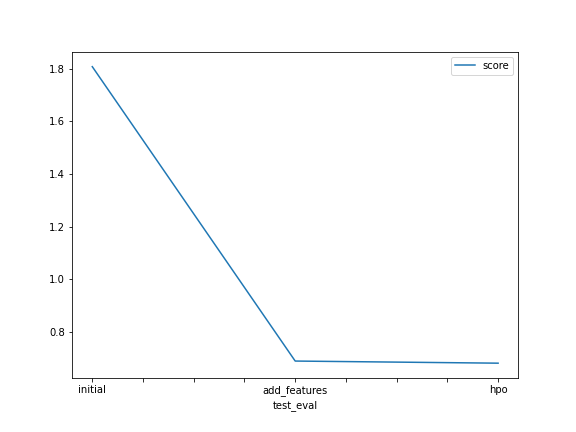
### Create a line plot showing the top model score for the three (or more) training runs during the project.

TODO: Replace the image below with your own.



### Create a line plot showing the top kaggle score for the three (or more) prediction submissions during the project.

TODO: Replace the image below with your own.



## Summary

TODO: **Add your explanation**

**We make our data in DataFrame format using Pandas, then splitting our data to train and test**

**We tried to solve the problem with three different methods**

1. **Without hyperparameters and new feature**
2. **With new features**
3. **With new features and hyperparameters**

**We used cloud based IDE , SageMaker studio built on top of Jupyter-lab with suitable python environment we used AutoML which it allows us to try many algorithms and parameters till reach maximum accuracy**