**Report: Predict Bike Sharing Demand with AutoGluon Solution**

Sneha Suman

**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?**

Initially, I realized that some predictions had negative values, which are not valid for bike counts. I had to clip all negative values to zero before submission. This step ensured that the predictions met the competition requirements.

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

The top-ranked model during initial training was LightGBM (GBM), which AutoGluon selected automatically based on root mean squared error (RMSE).

**Exploratory Data Analysis and Feature Creation**

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

From EDA, I observed that the hour of the day and day of the week had strong influence on bike demand. I created new features:

* hour from the datetime column
* dayofweek from the datetime column

These features helped the model understand daily and weekly usage patterns.

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

After adding the hour and dayofweek features, the Kaggle score improved significantly from 1.80001 to 0.62599. This improvement is because the model could now capture temporal patterns in demand more effectively.

**Hyperparameter Tuning**

**How much better did your model perform after trying different hyperparameters?**

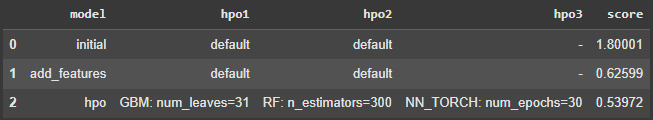
After tuning hyperparameters for GBM, RF, and XGBoost, the Kaggle score improved further to 0.53972, indicating that tuning helped optimize model complexity and generalization.

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

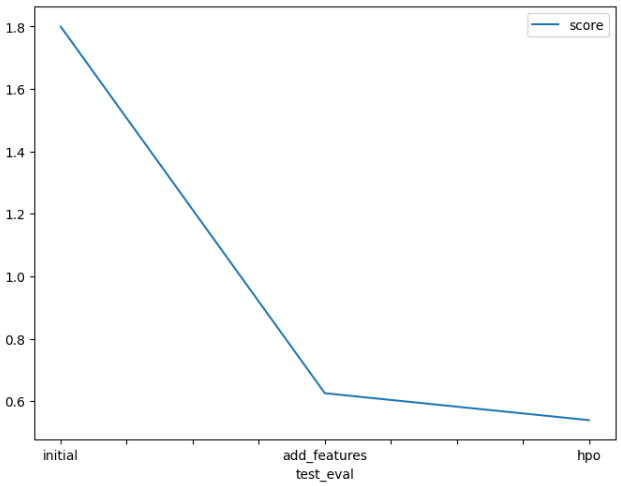
With more time, I would:

* Add more time-based features like is\_weekend, is\_rush\_hour, or season
* Engineer weather interaction terms
* Perform cross-validation and use ensemble blending
* Analyze error distribution to guide additional features

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

****

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

****