

1) Creation of new factor :

I tried a creating a new factor with the combination

```
s['hr'] * s['holiday'] * s['season']
```

I Observed that 'hr' had the most correlation with data. I thought I could improve models predictability by adding having hrs depend upon season and holiday to create a more comprehensive feature

Sadly, In the measurement It has come with very low correlation, suggesting the uselessness of the combination. I attribute it to the 'holiday' variable which varies closely and brings all data into its spectrum.

2) OneHotEncoder with TargetEncoder

My testing has shown that TargetEncoder works a little better than OneHotEncoder. For Stats,.

R^2 value of OneHotEncoder : 0.942

Mean Square value of OneHotEncoder : 1808.

R^2 value of targetEncoder : 0.944

Mean Square value of targetEncoder : 1762.

There is a slight Improvement in the model fitting

3) Trying LinearRegressor

My Testing has again showed that RandomForestRegressor showed is better able to classify than LinearRegressor

For Stats,

R^2 value of RandomForestRegressor: 0.944

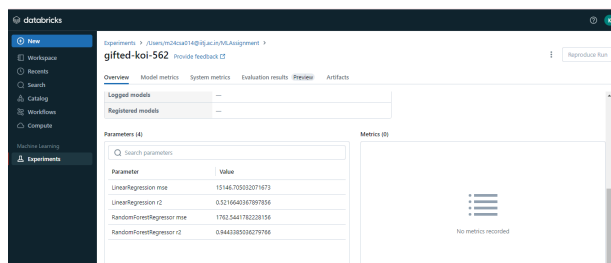
Mean Square value of RandomForestRegressor : 1762

R^2 value of LinearRegressor : 0.52

Mean Square value of LinearRegressor : 15146

4) Integrating MLFlow with the Code.

I tried using DataBricks Community for Host and used MLFlow to connect pass information to it. Here is the snippet of what is actually passed to the DataBricks



The screenshot shows the Databricks MLflow interface. The left sidebar has a navigation menu with 'New', 'Workspace', 'Recent', 'Search', 'Catalog', 'Workflow', 'Cluster', and 'Compute'. The 'Experiments' section is selected. The main area displays the details for an experiment named 'gifted-koi-562'. It includes tabs for 'Overview', 'Model metrics', 'System metrics', 'Evaluation results', 'Related', and 'Artifacts'. The 'Overview' tab is active, showing a table of parameters and a section for metrics.

Parameter	Value
LinearRegression mse	15146.70532071673
LinearRegression r2	0.5216640367697656
RandomForestRegressor mse	1762.5441782228156
RandomForestRegressor r2	0.9443385336276766

Below the parameters table, there is a section for 'Metrics (0)' which currently shows 'No metrics recorded'.