Approach_AV_job_a_thon_November_2022

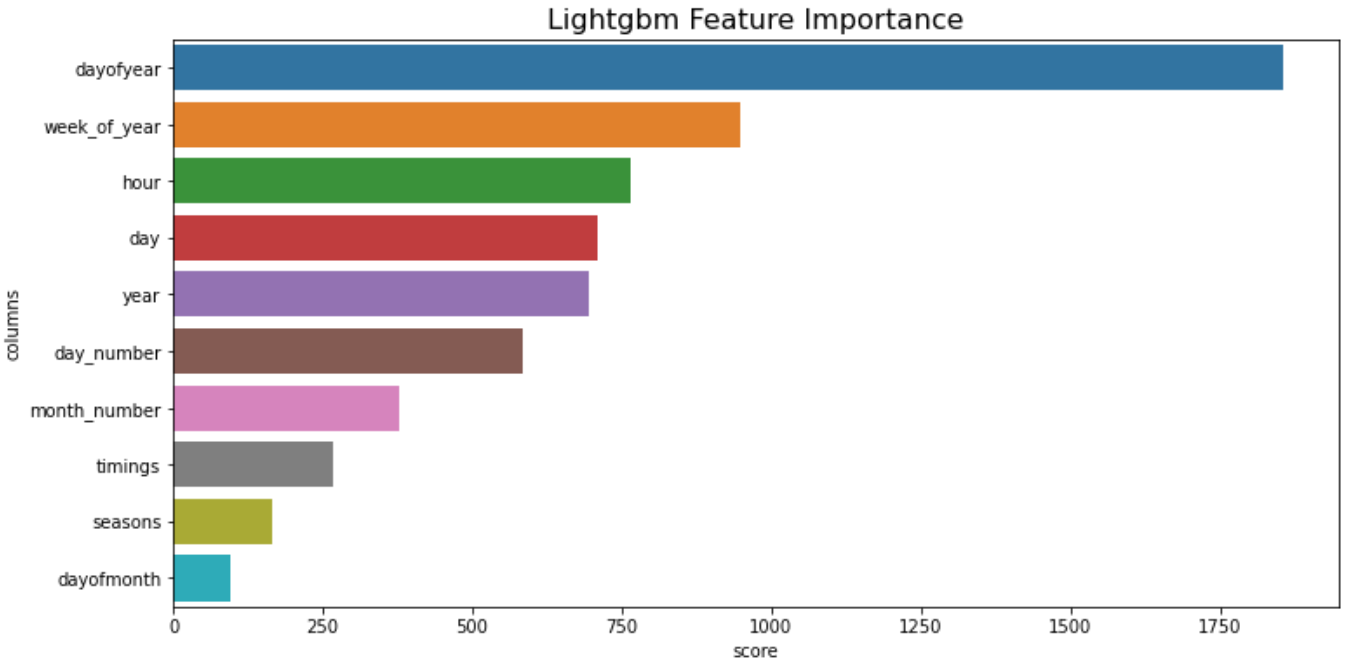
Build a machine learning/deep learning approach to forecast the total energy demand on an hourly basis for the next 3 years based on past trends.

- Basic exploratory data analysis using pandas, matplotlib, seaborn packages.
- Data pre-processing
    - Feature Engineering
        - Convert to date-time format
        - Extract date from the date_time
        - Extract day from the date
        - Extract the day name from the date
        - Extract the day number from the date
        - Extract month number from the date
        - Extract the month name from the date
        - Extract the quarter of the year
        - Extract week of the year from date
        - Extract year
        - Extract the day of the month
        - Extract day of the year
        - Create weekday column
        - Create weekend column
        - Create month start
        - Create month end

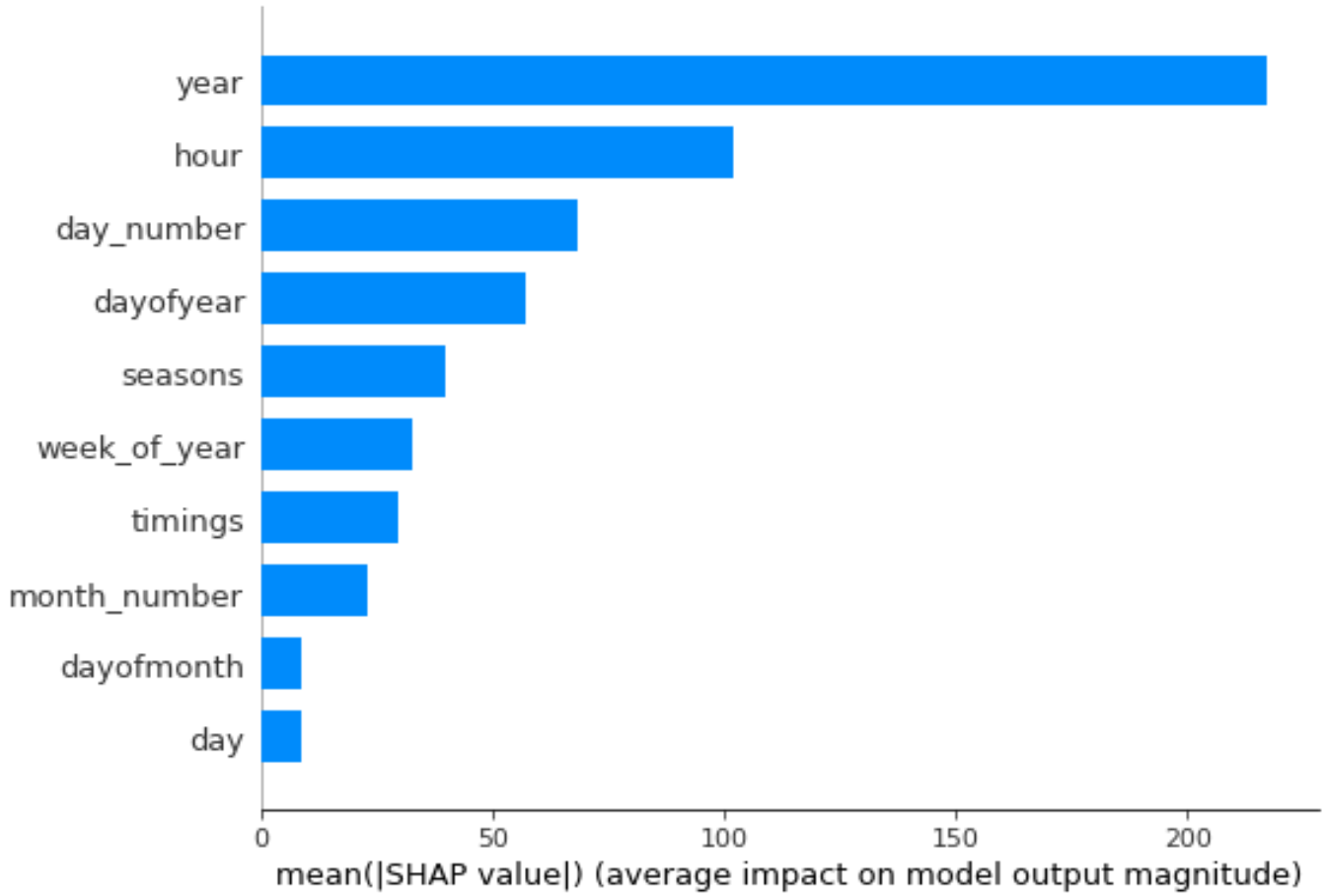
    - Create quarter start
    - Create quarter end
    - Create year start
    - Create year end
    - Convert month to seasons
    - Create hour
- The final features for the model
    0_day
    1_day_number
    2_month_number
    3_year_quarter
    4_week_of_year
    5_year
    6_dayofmonth
    7_dayofyear
    8_weekend
    9_month_start
    10_month_end
    11_quarter_start
    12_quarter_end
    13_year_start
    14_year_end
    15_hour
    16_seasons
    17_timings

- Created lightgbm regressor model with the following paremeters ,
    - bagging_fraction=0.8,
    - bagging_freq=2,
    - device_type='cpu',
    - feature_fraction=0.5,
    - lambda_l1=974,
    - lambda_l2=4,
    - learning_rate=0.07717248152232783,
    - max_depth=0,
    - min_data_in_leaf=15,
    - min_gain_to_split=6.420135636985865,
    - n_estimators=152,
    - num_leaves=45,
    - random_state=48

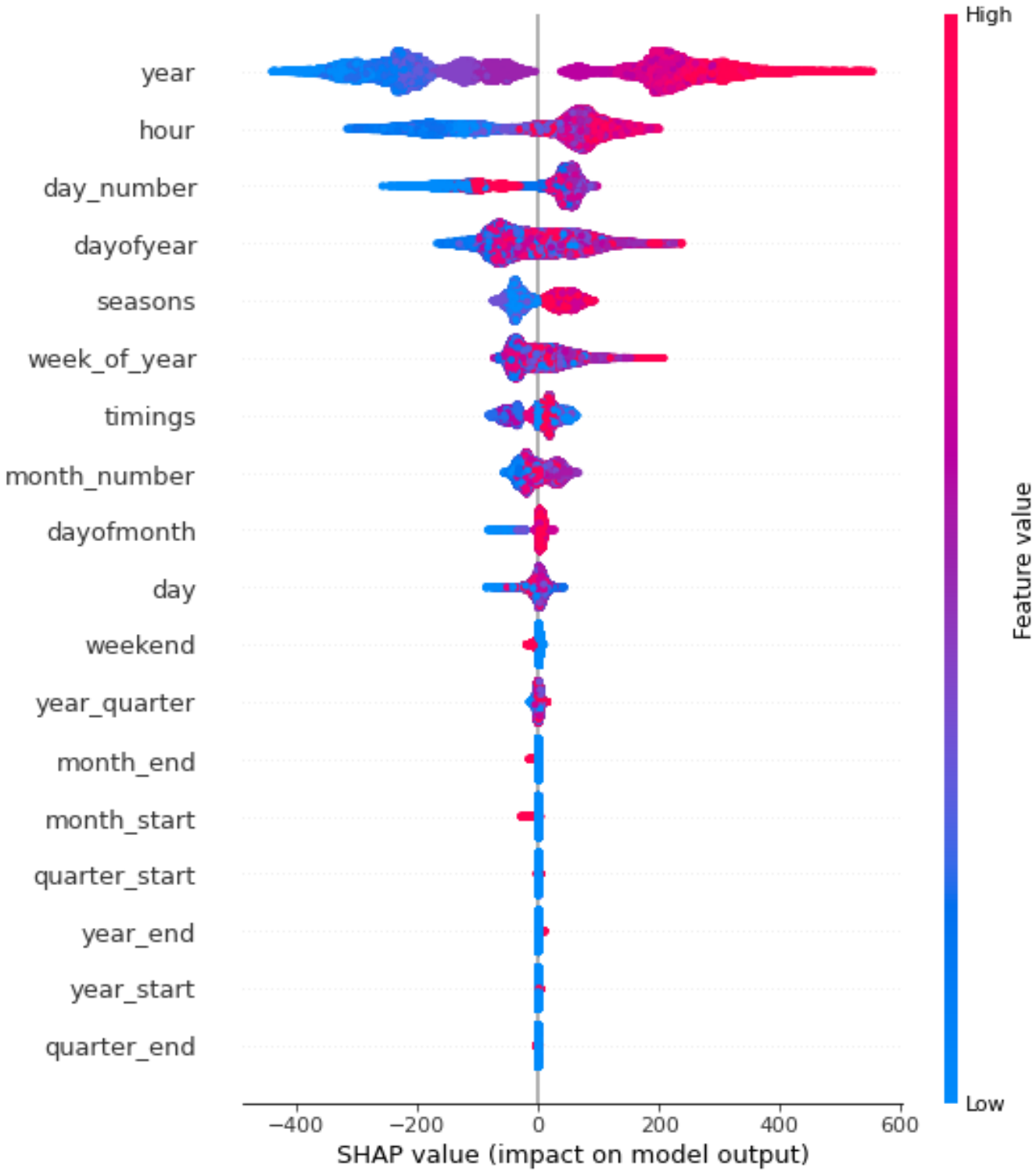
- Lightgbm Regressor - Top 10 Feature Importances



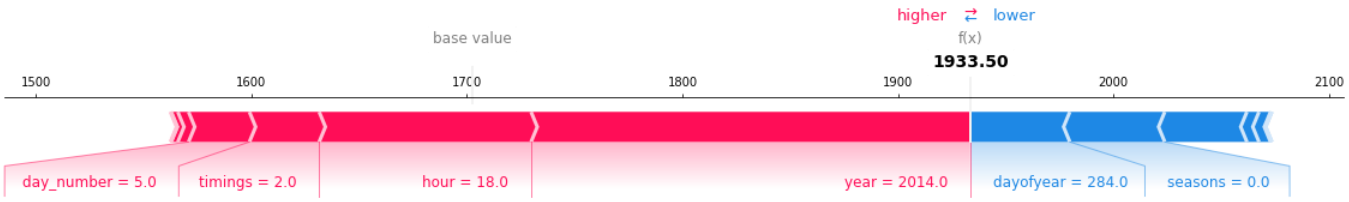
- SHAP - Lightgbm Regressor - Top 10 Feature Importances



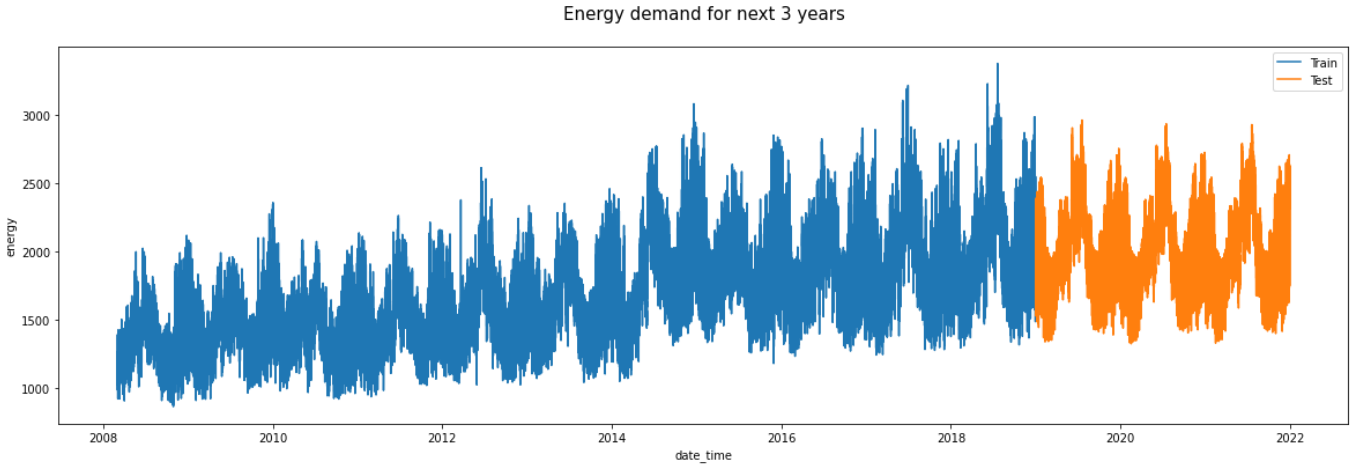
- SHAP - Top feature impact the model



- SHAP - Feature impact the single obsevartion



- Energy demand for next 3 years



- Final score
  - Public LB:
    - Score: 307.273811909926
    - Rank: 118
  - Private LB:
    - Score: 583.858113004428
    - Rank: 207