Home Electric Consumption forecasting

Ali Ahammad

Objective

- Building a prediction model to predict power consumption in every 15 minutes
- Algorithms used MLP and SVM models and hyper tuned parameters
- Datasets used from load Algarve forecasting using HEM where the data are already compiled, and parameters are selected

Datasets

Processing

- Separated into two dataframe depending on the suffix 1 and 2 and drop the nan rows from suffix 1 dataframe
- Concatenated two dataframes

Datasets

Daycode

- DayCode has column elements: 0.05, 0.15, 0.25, 0.3, 0.35, 0.5, 0.8, 0.7, 1.0

- Renamed the column name: "DayCode","Occupation", "Power"
- Then normality test
- Then supervised data for next 48 timestamps in 12 hours which would show every 15 minutes interval
- Then the dataset is splited to train and test

Dataset normality test

For power column

- Skewness for Power is 1.046 and kurtosis is 1.701

Performing Shapiro–Wilk test

- Shapiro Test Statistic 0.93622
- p-value 4.295927216204073e-34

Performing D'Agostino's K2 test

- D'Agostino's K2 Test Statistic 500.32654
- p-value 2.2671072101201093e-109

Performing Anderson-Darling test

- Anderson-Darling Test Statistic 37.10178617295105
- At 15.0% significance level, critical value is 0.575
- At 10.0% significance level, critical value is 0.655
- At 5.0% significance level, critical value is 0.786
- At 2.5% significance level, critical value is 0.917
- At 1.0% significance level, critical value is 1.091

Dataset normality test result

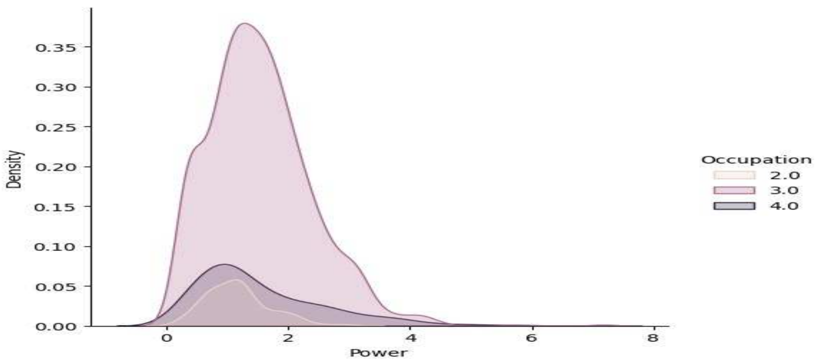
For power column

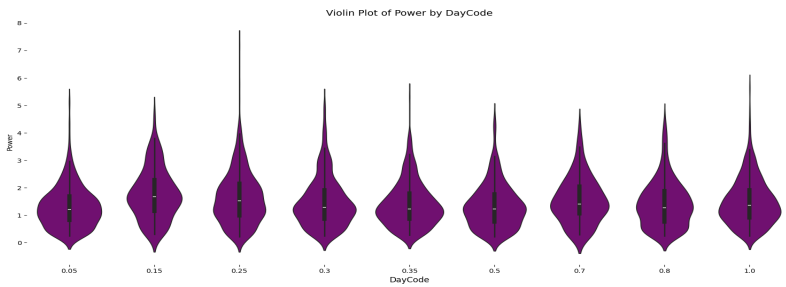
- We can reject the null hypothesis (D'Agostino's K^2), so our series is not normally distributed.
- We can reject the null hypothesis (Anderson-Darling), so our series is not normally distributed.

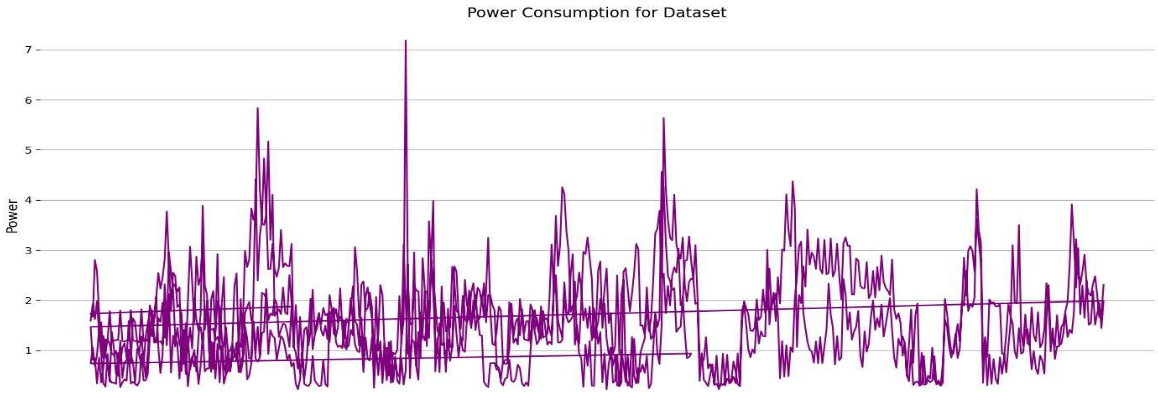
Overall normality result for 'Power': Not Normal

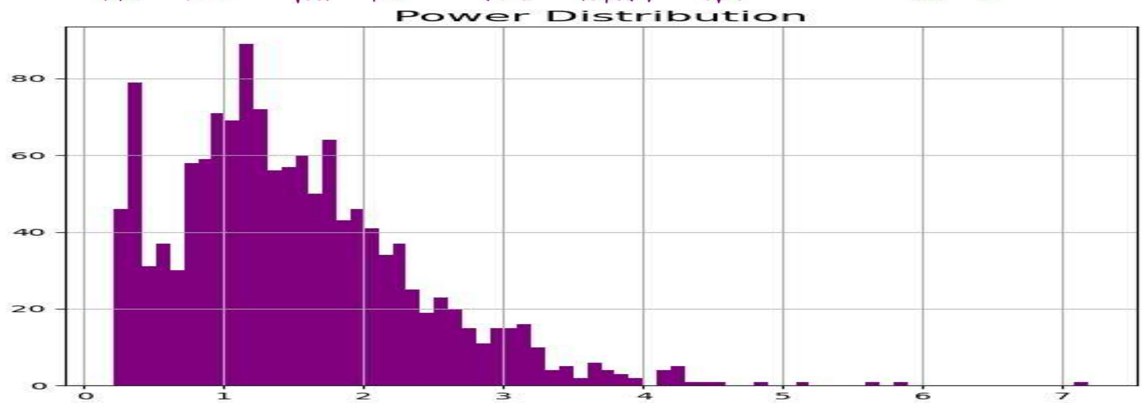
- Kurtosis of normal distribution: 1.6963600512675807
- Skewness of normal distribution: 1.04577179073994

Statistics plots







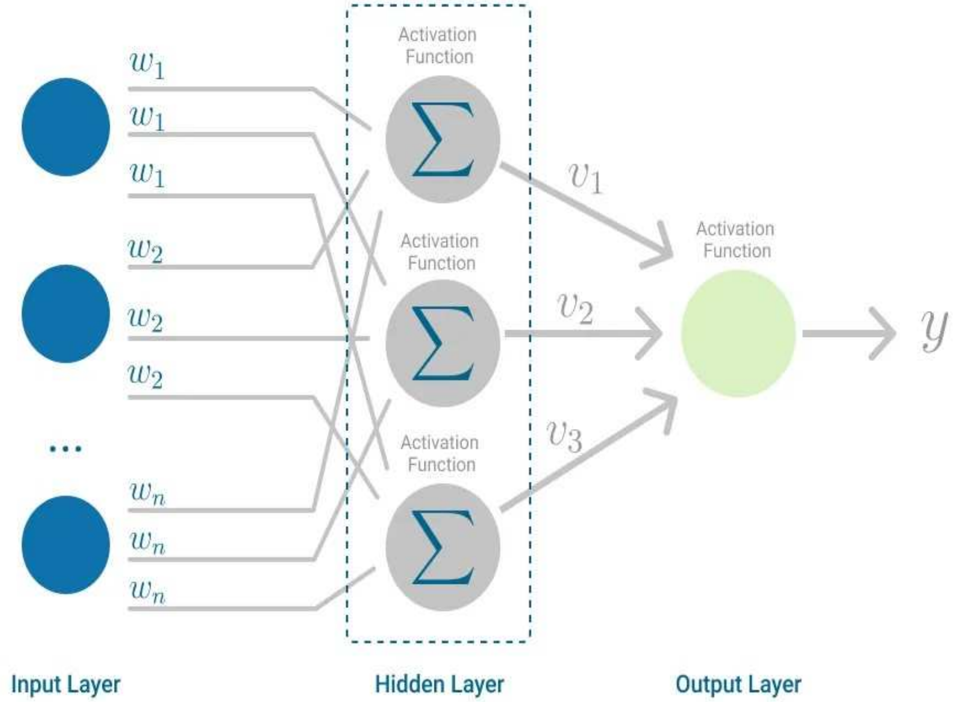


Scaling and sampling

- Used robustScaler() because it is robust to outliers in the sense that adding or removing outliers in the training set will yield approximately the same transformation
- Then we have sampled the dataset with lookback=380 and horizon 48
- For lookback, 24 hours * 4 (15-minute intervals per hour) the dataset is sampled every 15 minutes, so there are 4 samples per hour.For a 24-hour period, we need 24 hours * 4 samples/hour = 96 samples. By setting lookback = 380, we are considering a period of 380 / 4 ≈ 95 hours, which is approximately 4 days' worth of historical data.
- For horizon, The dataset is sampled every 15 minutes, so there are 4 samples per hour. To predict for the next 12 hours, we need 12 hours * 4 samples/hour = 48 samples. By setting horizon = 48, we are defining that the model should predict the next 48 samples, which corresponds to 12 hours ahead.

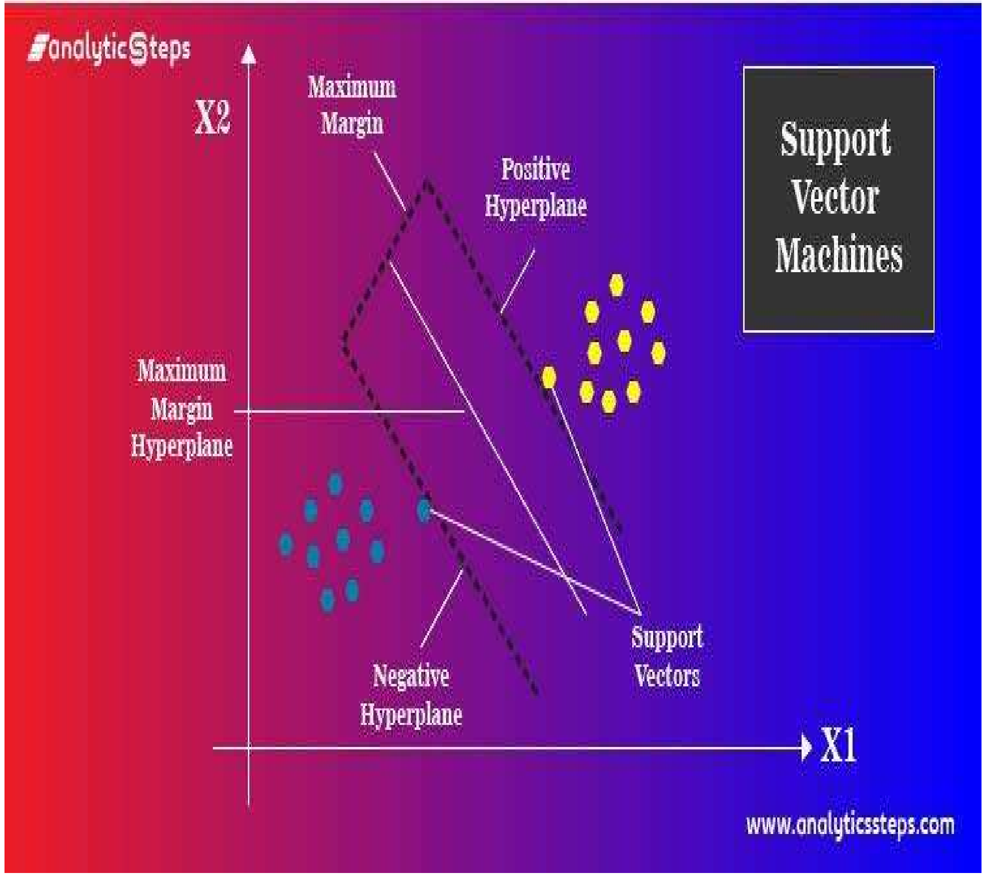
Model 1: Multi layered perceptron

- Hidden layer sizes= 100,50 (the number of neurons)
- Maximum iteration= 500
- Random state=42
- The activation function for hidden layer is by default relu
- Solver for weight optimization is adam
- Early_stopping is false



Model 2: Support Vector Machine Algorithm

- ‘rbf’ is used for kernel type.
- Other parameters are by default chosen such as gamma would be scale and it uses 1 / (n_features * X.var()) as value of gamma
- Regularization parameter C is 1.0
- Max_iter is hard limit on iteration within solver or -1 for no limit



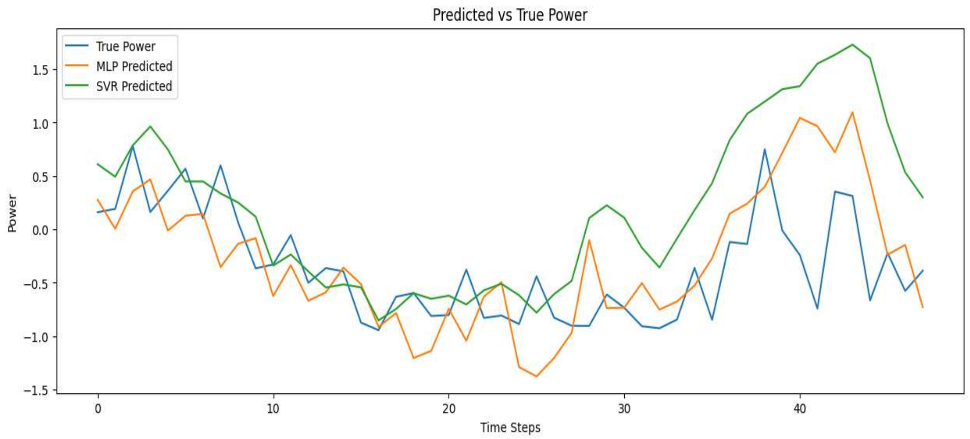
Output comparison between MLP and SVM

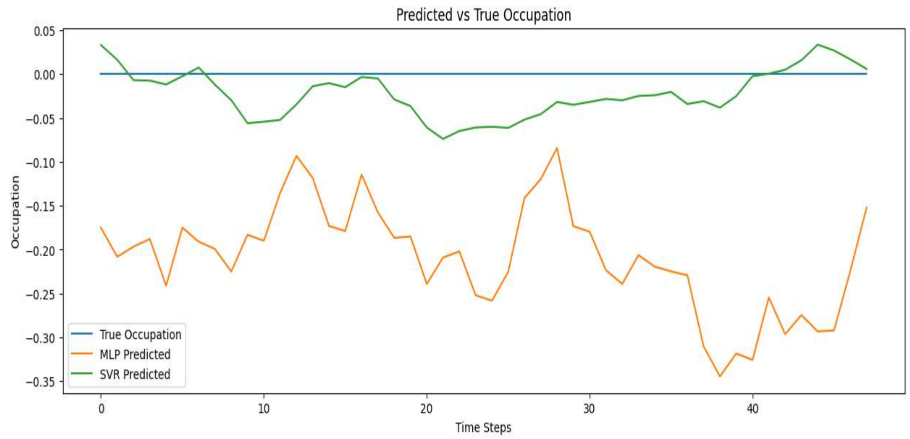
	MLP Power	SVR Power	MLP Occupation	SVR Occupation
Mean Squared Error	1.02	0.73	0.09	0.03
Mean Absolute Error	0.84	0.72	0.23	0.14
R-squared	-2.00	-1.15	0.00	0.00
Explained Variance	-0.56	-0.13	0.00	0.00

- In short, SVR consistently performs better than MLP across all metrics, especially for power prediction.
- However, both models struggle with occupation prediction, as indicated by low R^2 and explained variance.

- In both power and occupation models, SVR outperforms MLP, as it has lower MSE values.
- SVR performs better in both power and occupation models due to its lower MAE.
- For R^2 measures , both models perform poorly for power prediction. However, SVR is better for occupation prediction (though still not ideal).
- For explained variance, SVR slightly outperforms MLP for power prediction, but neither model explains much variance. For occupation, both models fail to capture any variance.

Graph comparison between MLP and SVM





Hyper-parameter tuning

SVM

- kernel: rbf, linear,
- C: 0.1, 1, 10,
- gamma: scale, auto

MLP classifier

- Hidden layer sizes: (50,), (100,), (100, 50), (50, 100)
- Max_iter: 200,500
- alpha: 0.0001, 0.001, 0.01,
- learning_rate: 'constant', 'adaptive'

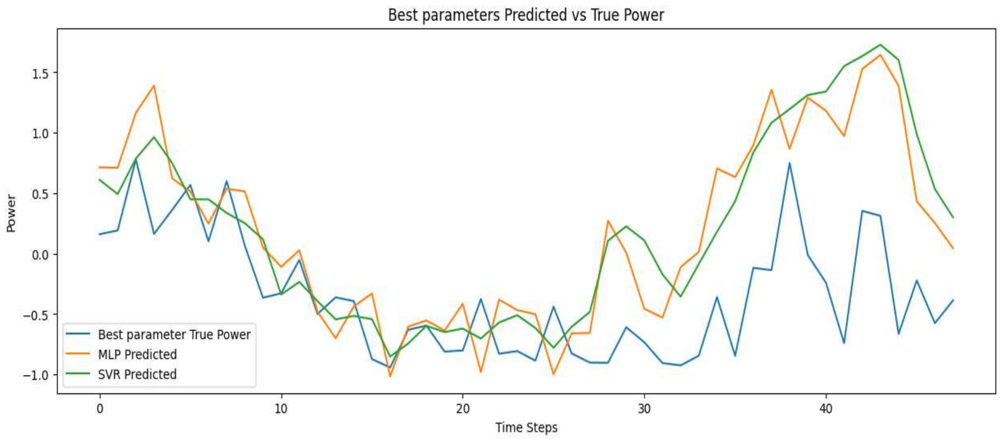
Output comparison between MLP and SVM after tuning

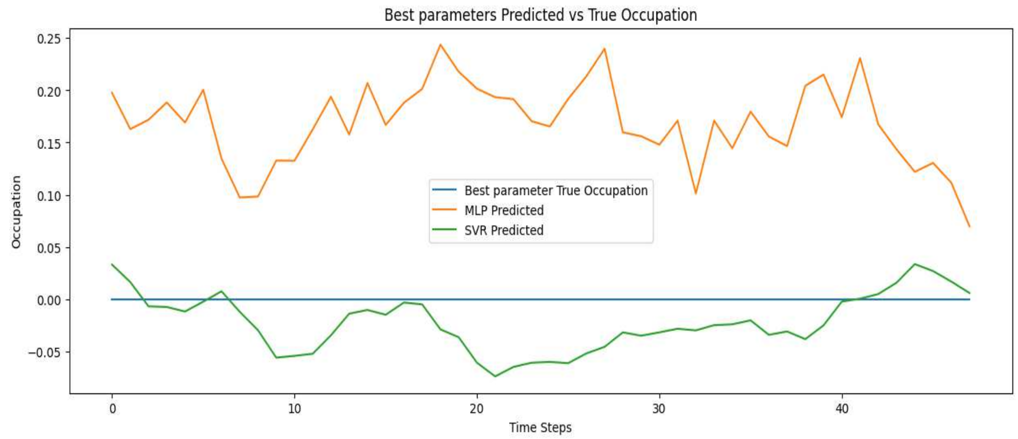
	MLP Power	SVR Power	MLP Occupation	SVR Occupation
Mean Squared Error	1.07	0.73	0.05	0.03
Mean Absolute Error	0.86	0.72	0.17	0.14
R-squared	-2.14	-1.15	0.00	0.00
Explained Variance	-0.61	-0.13	0.00	0.00

- **SVR** remains the better model overall, especially for power prediction.

Mean Squared Error (MSE): Both models show improvements in power prediction, with SVR maintaining its lower MSE. For occupation prediction, both models have lower MSE values, but SVR still performs better.
Mean Absolute Error (MAE): SVR continues to outperform MLP in both power and occupation prediction. The MAE values are lower for both models in the occupation task.
R-squared (R^2): The R^2 values remain negative for both power and occupation prediction, indicating poor model fit. SVR is still the better performer, but neither model explains much variance.
Explained Variance: SVR maintains its advantage in explained variance for power prediction. Unfortunately, both models fail to capture any variance in occupation prediction.

Graph comparison between MLP and SVM after tuning





Thank you

Ali Ahammad

Any Questions???