**Capstone Project Submission**

| **Team Member’s Name, Email and Contribution:** |
| --- |
| Individual Team Member Name: Pankaj Ramanlal Beldar  Email id : [pankajrbell@gmail.com](mailto:pankajrbell@gmail.com)  This is the individual Capstone Project. The work done in this project is Exploratory Data Analysis, Outliers Identification and Removal, Encoding Categorical Variables, Feature Selection and Extraction using Principal Component Analysis, Anova F value and Variance Threshold, Building Different ML models, Hyper parameter Tuning , Model Explain ability with LIME and Eli5. |
| Github Link: - <https://github.com/pankaj-beldar/Capstone_2_Appliance_Energy_Prediction>  Google Drive Link:- <https://drive.google.com/drive/folders/1BCkmDXLOwWRts3pVDx4WYIMVi2iDKa4M?usp=share_link> |
| **Summary of your Capstone project: Appliances Energy Prediction** |
| 1. The household appliance energy consumption prediction models based on Linear Regression, Lasso Regression, Ridge Regression, MLP Regressor, Decision Tree Regressor Random Forest Regressor, Adaptive Boosting Regressor, Gradient Boosting Regressor, Bagging Regressor, K Neighbors Regressor and Linear SVM are explored. 2. Upon appropriate pre-processing and fitting the fourteen models, we compare and evaluate the best model with lowest error and the highest R-squared score. The variables T6 and T\_out, T9 and T7 have high correlation with each other hence we have dropped T6 and T9. When evaluating the influence of Random Variable attribute the linear models have assigned near zero weights to the random variable, negating its influence in prediction of the target variable. 3. Extra Tree Regressor was found to be the best performing model with an R-squared score of 0.78.Data from a wireless sensor network that measures humidity and temperature has been proven to increase the prediction accuracy. The data analysis showed that data from the kitchen, laundry room, living room and bathrooms had the most important contributions. Data from the other rooms also helps in the prediction. When looking at the appliances in each room, it can be seen that the laundry, kitchen and living rooms would be expected to have the highest contributions because of the equipment present. The prediction of appliances’ consumption with data from the wireless network indicates that it can help to locate where in building the main appliances’ energy consumption contributions are found. 4. This study has found curious relationships between variables. Future work could include considering weather data such as solar radiation and precipitation. Also, occupancy and occupant’s activity information could be useful to improve the prediction and find its relationship with other parameters (exterior weather for example). The wireless sensors could also measure CO2 and noise to help in the prediction and to track the occupant’s movement from room to room and time spent in each room. 5. One of the main limitations of this study is that the analysis was done for only one house. Important information could be found when analyzing several houses, and other relationships can be studied with appliances’ energy consumption in combination with: occupant’s age, number of occupants, ownership of pets, building’s geometry etc. Another research limitation is the length of continuous analyzed data. |