**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, Decision Tree Regression Random Forest Regression, Adaptive Boosting Regression, Gradient Boosting Regression, Bagging Regression, K Neighbors Regression and Linear SVM are explored. 2. Random Forest Regression was found to be the best performing model with an R-squared score of 0.64.After optimizing the hyper parameters of the Random Forest Regression, doing principle Component Analysis, its R-squared score increased from 0.62 to 0.64. 3. 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 a building the main appliances’ energy consumption contributions are found. 4. When using all the predictors the light consumption was ranked highly. However, when studying different predictor subsets, removing the light consumption appeared not to have a significant impact. This may be an indication that other features are correlated well with the light energy consumption. 5. The possible explanation for why the pressure has a strong prediction power may be related to its influence on the wind speed and higher rainfall probability which could potentially increase the occupancy of the house. 6. 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. |