Energy Forecasting of Home Appliances

Project Overview

- The Project aims to predict the energy consumption of home appliances.With the advent of smart homes and rising energy for energy management, existing smart home systems can benefit from accurate prediction. If the energy usage can be predicted for every possible state of appliance, then device can be optimized for energy savings as well.

- We are going to use different machine learning techniques. Appliances usage is target variable while weather data are the features.

DataSource Link -
http://archive.ics.uci.edu/ml/datasets/Appliances+Energy+prediction

Data Exploration

NAME	DESCRIPTION	UNIT
	Features	
T1	Kitchen Temperature	°C
T2	Living Room Temperature	°C
T3	Laundry Room Temperature	°C
T4	Office Temperature	°C
T5	Bathroom Temperature	°C
T6	Temperature outside Building (North)	°C
T7	Ironing Room Temperature	°C
T8	Teenager Room Temperature	°C
T9	Parents Room Temperature	°C
T_out	Outside Temperature (Weather Station)	°C
T_dewpoint	Dewpoint Temperature (Weather Station)	°C
RH_1	Kitchen Humidity	%
RH_2	Living Room Humidity	%
RH_3	Laundry Room Humidity	%
RH_4	Office Humidity	%
RH_5	Bathroom Humidity	%
RH_6	Humidity outside Building (North)	%
RH_7	Ironing Room Humidity	%
RH_8	Teenager Room Humidity	%
RH_9	Parents Room Humidity	%
RH_out	Outside Humidity (Weather Station)	%
Pressure	Outside Pressure (Weather Station)	mm Hg
Wind speed	Outside Windspeed (Weather Station)	m/s
Visibility	Visibility (Weather Station)	km

- No.of Features affecting = 24
- No.of instances in training data = 15264
- No.of instance in testing data = 4430

Descriptive Statistics

	T1	T2	T3	T4	T5	T6	T7	T8	T9
count	14999.000000	14999.000000	14999.000000	14999.000000	14999.000000	14999.000000	14999.000000	14999.000000	14999.000000
mean	21.150587	19.658846	21.563215	20.136699	18.938280	6.177918	19.534581	21.437572	18.742925
std	1.235399	1.524600	1.525767	1.540831	1.318201	4.731859	1.603107	1.722168	1.454431
min	16.790000	16.100000	17.200000	15.100000	15.330000	-6.065000	15.390000	16.306667	14.890000
25%	20.390000	18.600000	20.463333	19.166667	18.000000	2.900000	18.463333	20.340000	17.790000
50%	21.290000	19.600000	21.600000	20.204938	18.890000	6.090000	19.500000	21.633333	18.600000
75%	22.033333	20.666667	22.600000	21.260000	20.000000	9.463333	20.890000	22.700000	20.041667
max	24.100000	24.600000	27.600000	23.760000	22.967778	21.290000	23.566667	25.200000	23.840000

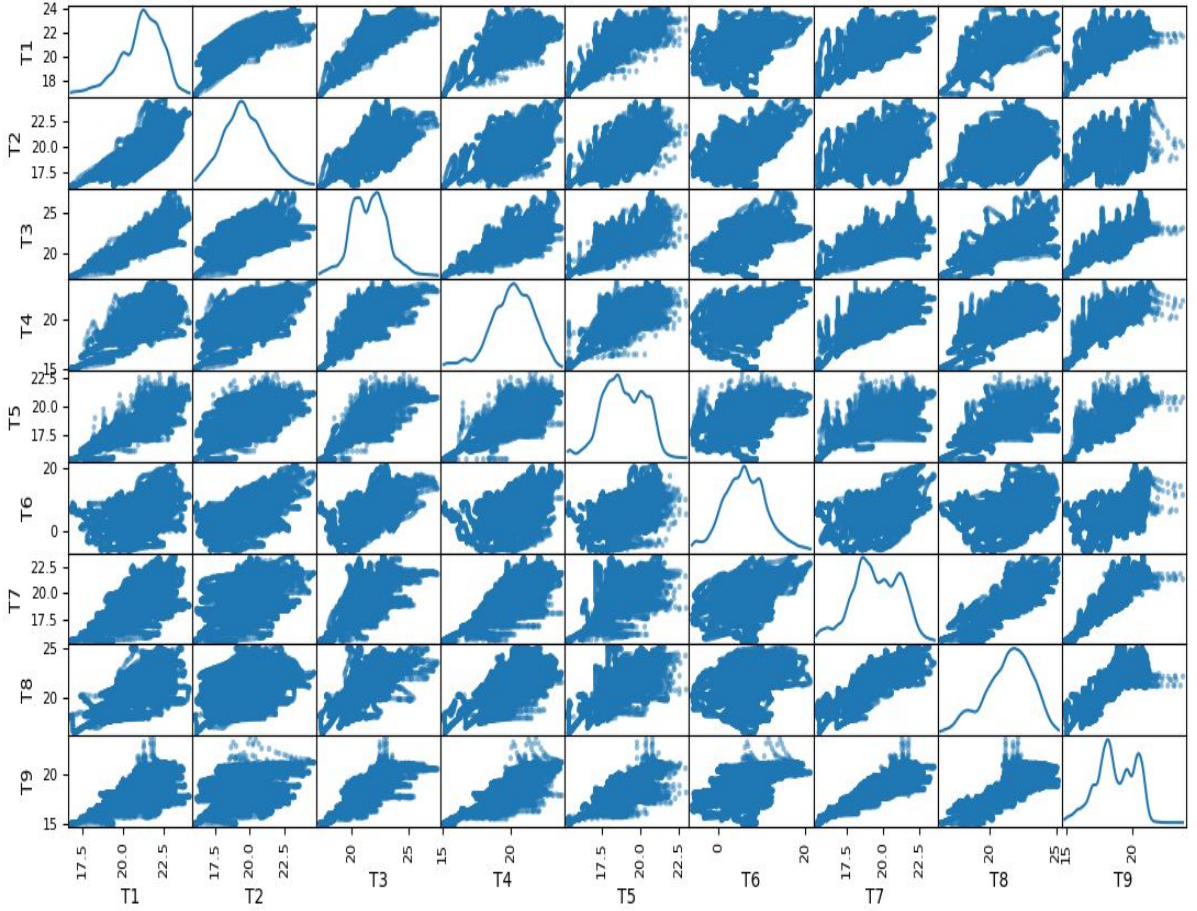
	RH_1	RH_2	RH_3	RH_4	RH_5	RH_6	RH_7	RH_8	RH_9
count	14999.000000	14999.000000	14999.000000	14999.000000	14999.000000	14999.000000	14999.000000	14999.000000	14999.000000
mean	40.223960	40.688812	39.545665	39.119283	51.539788	63.992611	35.246859	43.006928	41.670245
std	3.570467	3.387503	3.149552	4.269538	8.974617	27.818672	5.040493	5.270991	3.957531
min	27.733333	25.763333	32.626667	27.660000	35.363333	1.000000	23.200000	29.600000	31.033333
25%	37.590000	38.400000	37.230000	35.700000	45.900000	44.658333	31.500000	39.126667	38.700000
50%	39.790000	40.760000	38.833333	38.500000	49.326667	69.100000	34.754000	42.585000	40.966667
75%	42.933333	43.290000	42.000000	42.090000	54.120992	88.091667	38.745000	46.530000	44.248214
max	63.360000	56.026667	50.163333	51.090000	96.321667	99.900000	51.400000	58.780000	53.326667

- One interesting observation can be seen in Appliances column that although the max consumption is 1080Wh, 75% of values are less than 100Wh. This shows that there are fewer cases when Appliance energy consumption is very high.

	T_out	Tdewpoint	RH_out	Press_mm_hg	Windspeed	Visibility
count	14999.000000	14999.000000	14999.000000	14999.000000	14999.000000	14999.000000
mean	5.905138	2.761939	81.749428	755.512080	4.299798	38.760284
std	4.190974	3.577463	12.981362	7.929200	2.589544	12.570824
min	-5.000000	-6.600000	31.000000	729.300000	0.000000	1.000000
25%	2.980000	0.367000	73.000000	750.650000	2.000000	29.000000
50%	5.850000	2.830000	85.333333	756.166667	4.000000	40.000000
75%	8.970000	5.300000	92.000000	761.766667	6.000000	40.000000
max	19.700000	11.400000	100.000000	772.300000	14.000000	66.000000

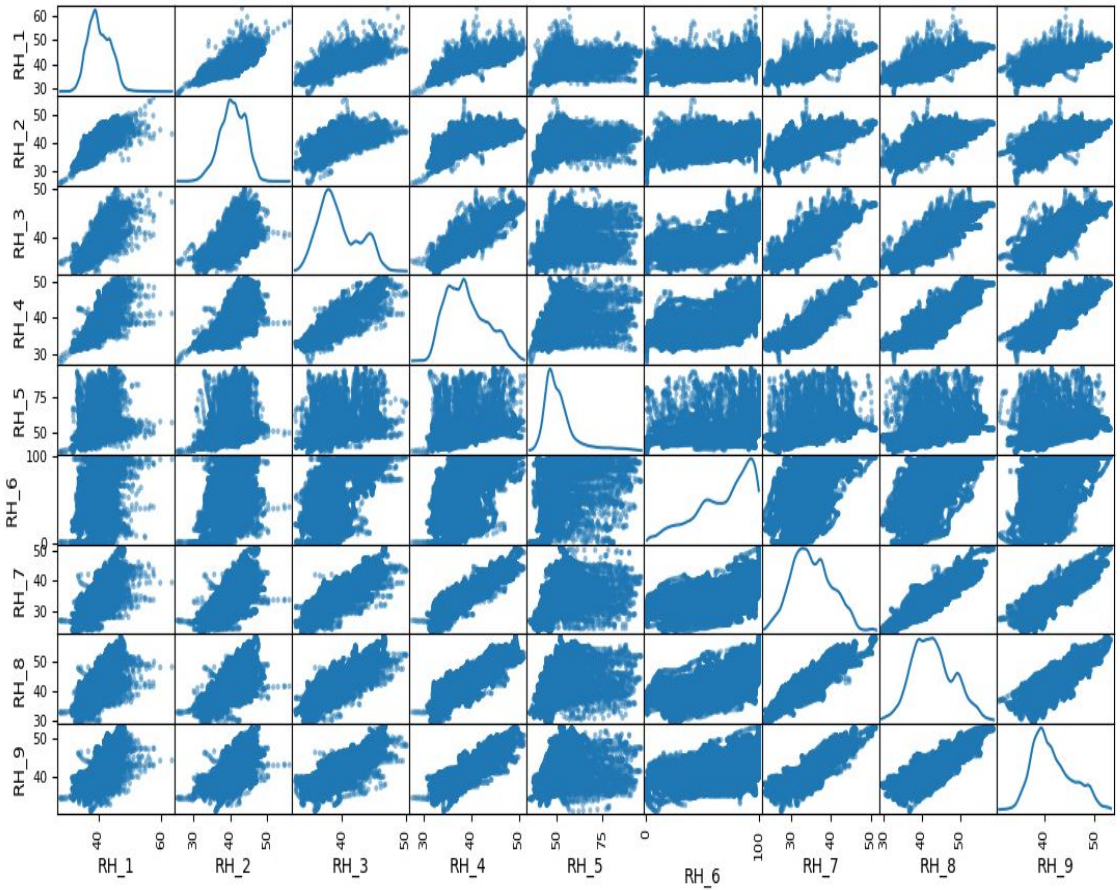
	Appliances
count	14999.000000
mean	98.751250
std	106.425424
min	10.000000
25%	50.000000
50%	60.000000
75%	100.000000
max	1080.000000

Temperature Correlation Plot

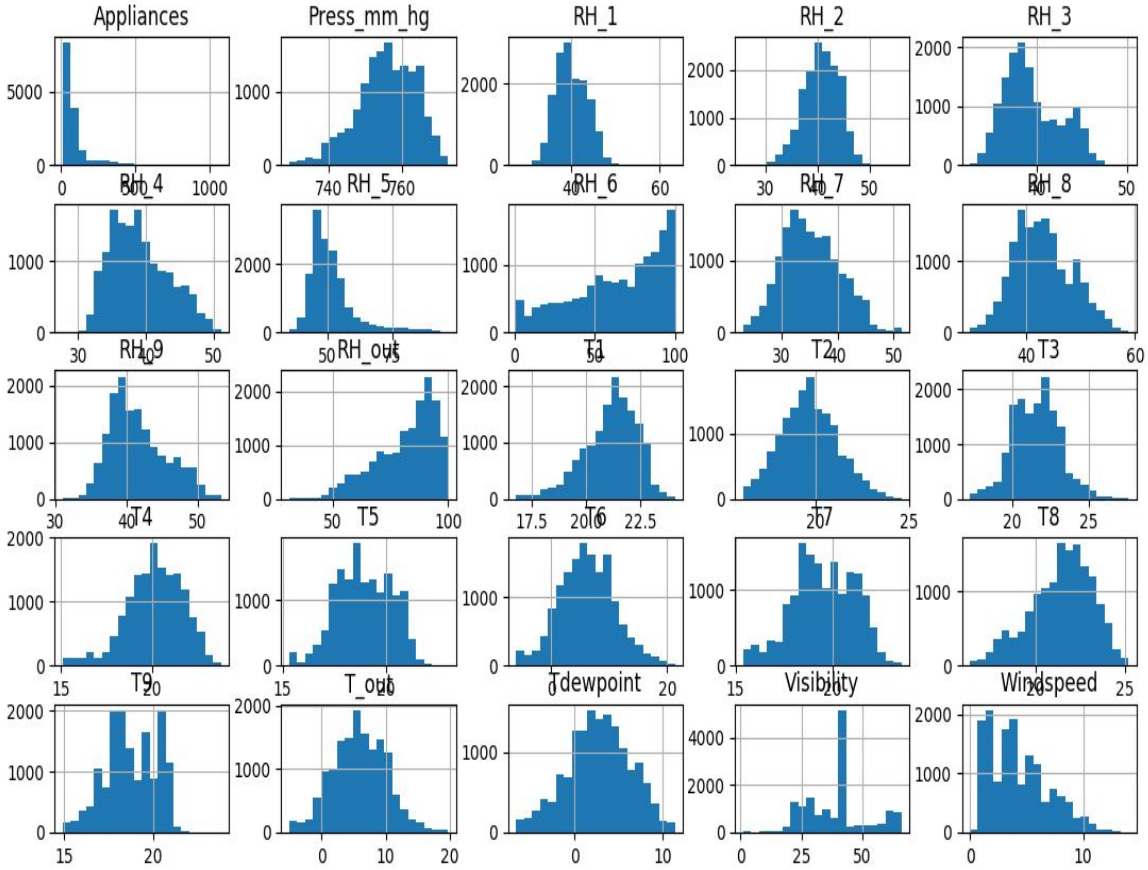


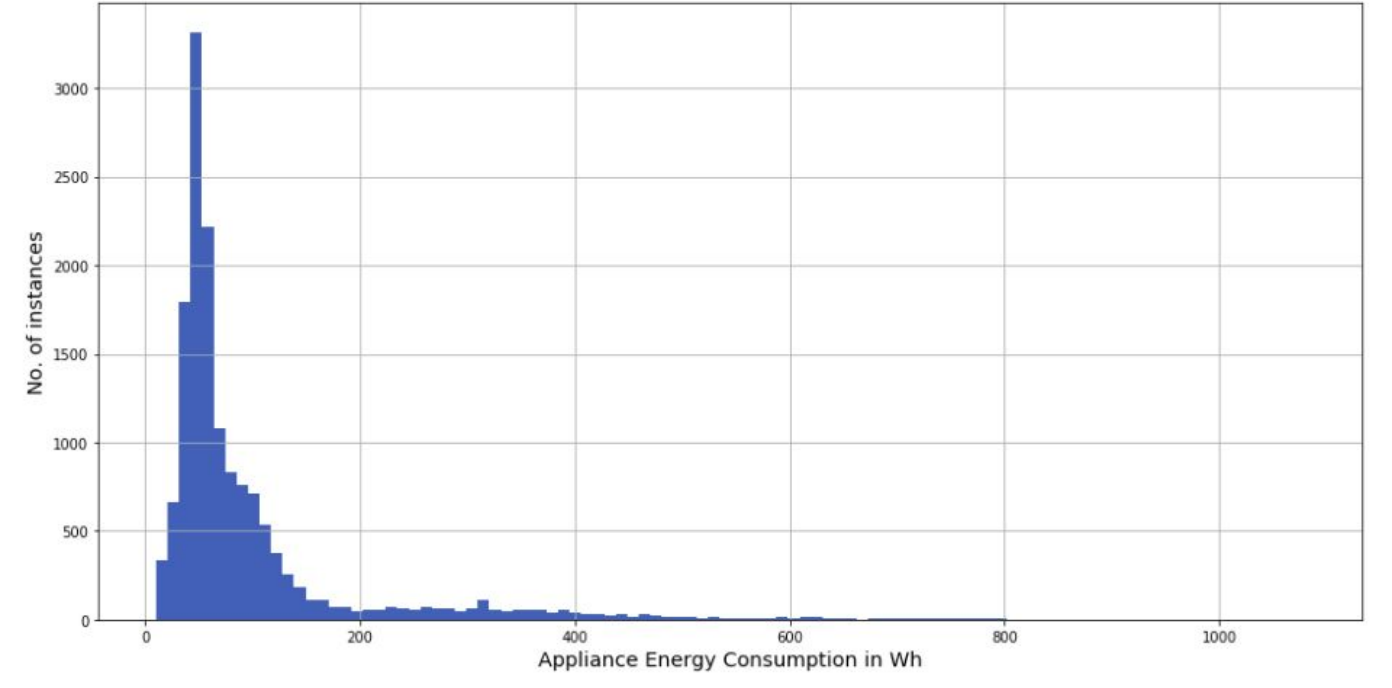
Humidity Correlation Plot

We can see that RH_3 and RH_4 are highly correlated.



Histogram of each feature





- We can see that most values are in the range of 0-200 Wh, strengthening our assumption that there are few cases of high energy consumption. The percentage of values within this range is calculated below.

- Percentage of dataset in range of 0-200 Wh = 89.713%

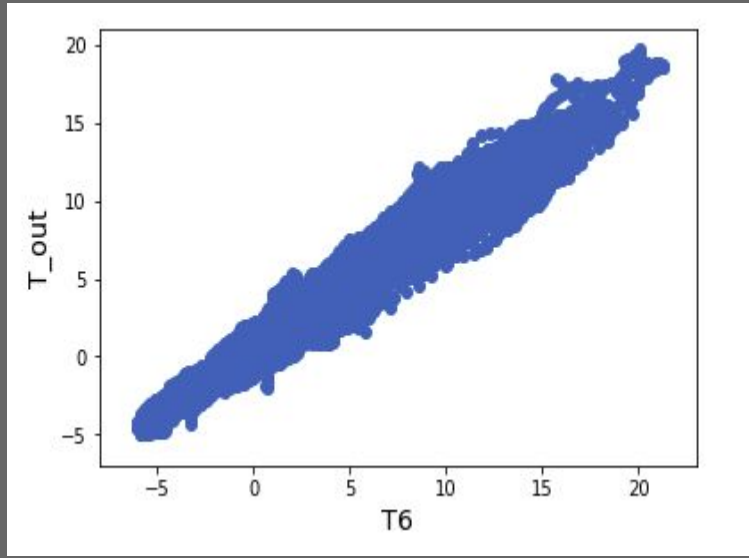
Highly Correlated Variables

- We can observe that some of the variables are highly correlatable. We can find this out using correlation coefficient and p-value.

- It was found there are two pair of variables which are highly correlated.
-
- Column pair : RH_3, RH_4
  - Correlation coefficient : 0.9027793129209211
  - p-value : 0.0

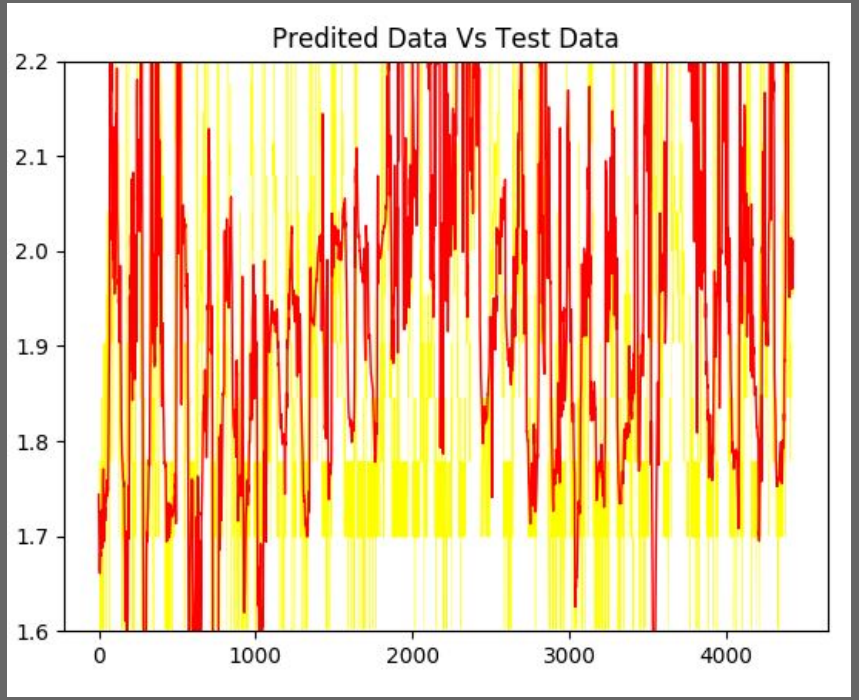
- Column pair : T6, T_out
  - Correlation coefficient : 0.97282072559015 95

- So, we will remove RH_4 and T_out from data for better prediction.



Final Prediction

- It can be observed that on an average, humidity affects power consumption more than temperature.
- Also, out of weather parameters, Humidity and Atmospheric pressure affect power consumption more significantly than others.
- Wind Speed and Visibility shouldn't affect the power consumption inside the home.
- Although natural humidity cannot be controlled, controlling humidity inside the home can lead to energy savings.



Machine Learning Technique	Mean Absolute Error
Linear Regression	7.1620058591
Ridge Regression	7.1620058591
Extra Trees Regressor	7.1491790723
Deep Neural Network	6.5456917995