CMPT 318

Group 10

Anomaly Detection of Individual Household Power Consumption

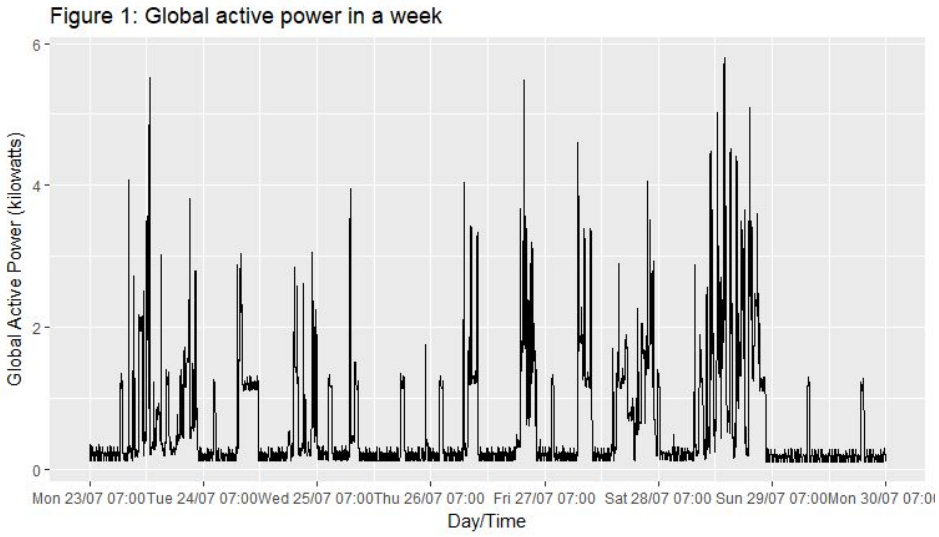
Group members:

Matheson Mawhinney

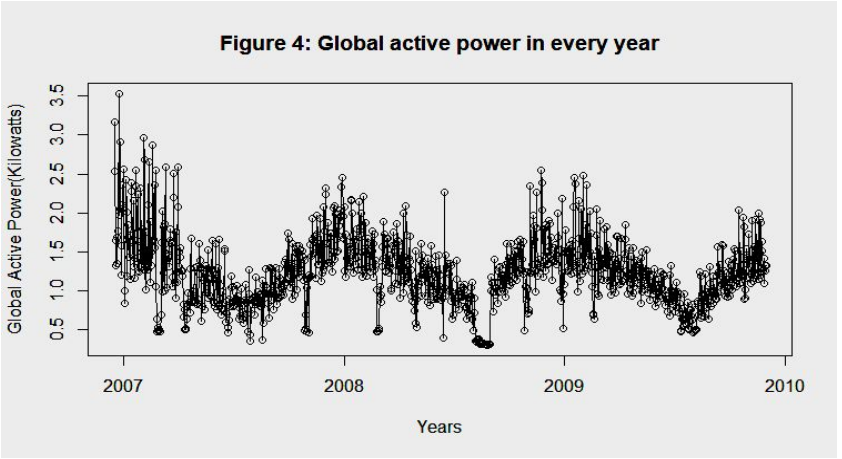
Alexis Lizardo

Truong Thinh Nguyen

Characteristic of Datasets



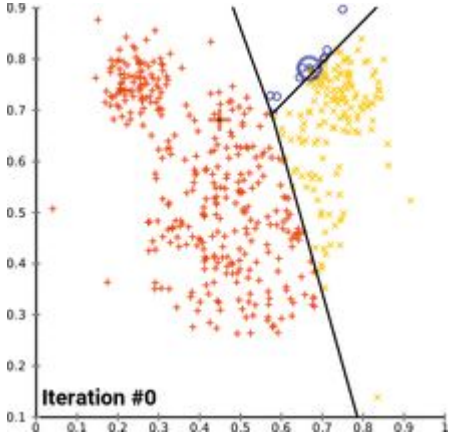
- High power consumption in morning and evening of weekdays. Higher on Friday and Saturday
- Highest throughout Sunday



- Total power consumption of each year decreases
- Peak at the start and the end of each year.
- Lowest in the middle of the year

Experiment Analysis:

- Splitting training data set: 90% for training, 10% for validating
- Set up parameters for training Markov models
- K-means clustering to find the means and standard deviations



Experiment Analysis:

- Passing means and standard deviations for each cluster into the output probability matrix (define emission distribution)
- Define Initial state distribution
- Create Univariate HMM and Multivariate HMM
- Using validation data (with noise and without noise) to test the accuracy of the models
- Test with the test set files (predict with new data).

Results:

States	Univariate HMM	MultiVariate HMM
2	-926607.9	-5292092
3	-566751.50	-4693582
4	-419092.2	-4149228

Table 1: Log Likelihood of train data

Results:

		Threshold					
	States	**0.5**	**1**	**1.25**	**1.5**	**1.75**	**2**
Univariate HMM	**2**	59064	31048	23239	17031	11585	9246
	3	43713	20852	15572	12169	9594	8213
	4	39448	19581	15632	12573	10222	8711
Multivariate HMM	**2**	57438	34935	24086	16903	14571	12960
	3	57858	29482	23066	19166	14309	10790
	4	54919	30045	23262	19491	16167	12532

Table 1: Point Anomalies (Validation noise added)

		Window Size			
	States	**5**	**10**	**15**	**20**
Univariate HMM	**2**	14339	6503	4743	3433
	3	10307	5149	4061	2958
	4	10520	5240	4028	3006
Multivariate HMM	**2**	14266	7101	4913	3578
	3	12835	6174	4046	3331
	4	13046	6267	4595	3374

Table 2: Collective Anomalies (Validation noise added)

Results:

		Threshold					
	States	**0.5**	**1**	**1.25**	**1.5**	**1.75**	**2**
Univariate HMM	**2**	279461	169086	128317	81922	43597	29992
	3	224155	109478	71283	46447	29838	22857
	4	229259	113768	81357	55863	38301	29090
Multi-variate HMM	**2**	225391	106804	71287	40329	21574	14893
	3	209873	96408	58342	37317	23151	11659
	4	214104	85758	52617	32974	19280	10564

Table 3: Point Anomalies (test1.txt)

		Window Size			
	States	**5**	**10**	**15**	**20**
Univariate HMM	**2**	29725	10609	7609	4499
	3	15646	4825	3387	1758
	4	17232	5089	3445	1746
Multivariate HMM	**2**	14764	4694	3259	1960
	3	11449	3205	2186	1211
	4	8666	2512	1784	1023

Table 4: Collective Anomalies (test1.txt)

THE END

THANK YOU FOR YOUR ATTENTION