FH | JOANNEUM
University of Applied Sciences

ENERGY USAGE OF AN OFFICE BUILDING

ENERGY AND TRANSPORT MANAGEMENT (MSc)

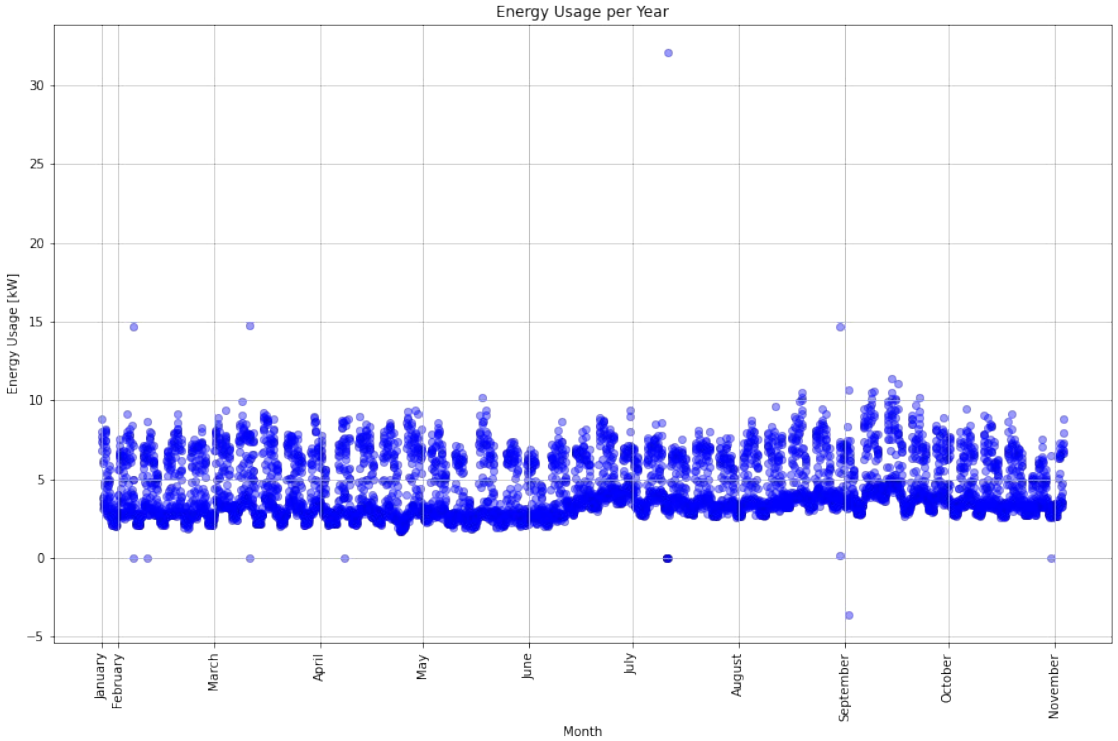
FH | JOANNEUM
University of Applied Sciences

Digital Modelling and Big Data Simulation

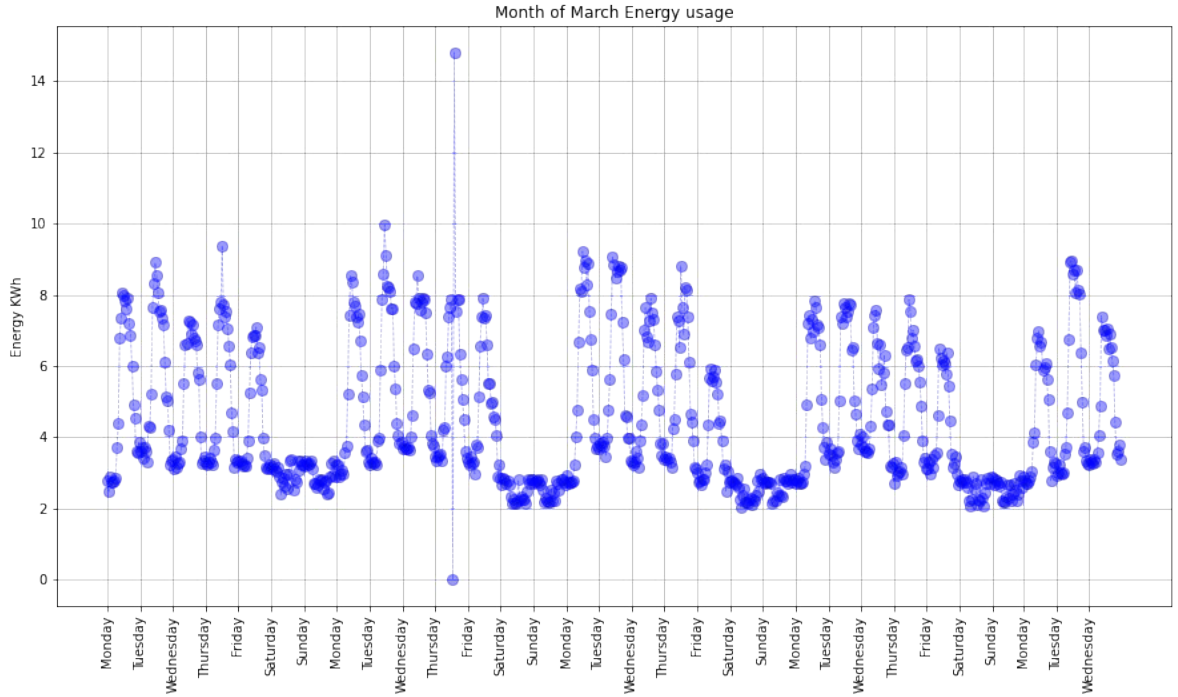
Final Project – Predict the Energy Consumption of an Office Building in Graz

Group:
Maria Angeles Doblas Florido
Matias Dogliani
Nathan Thomas Nord
Ian Calixto

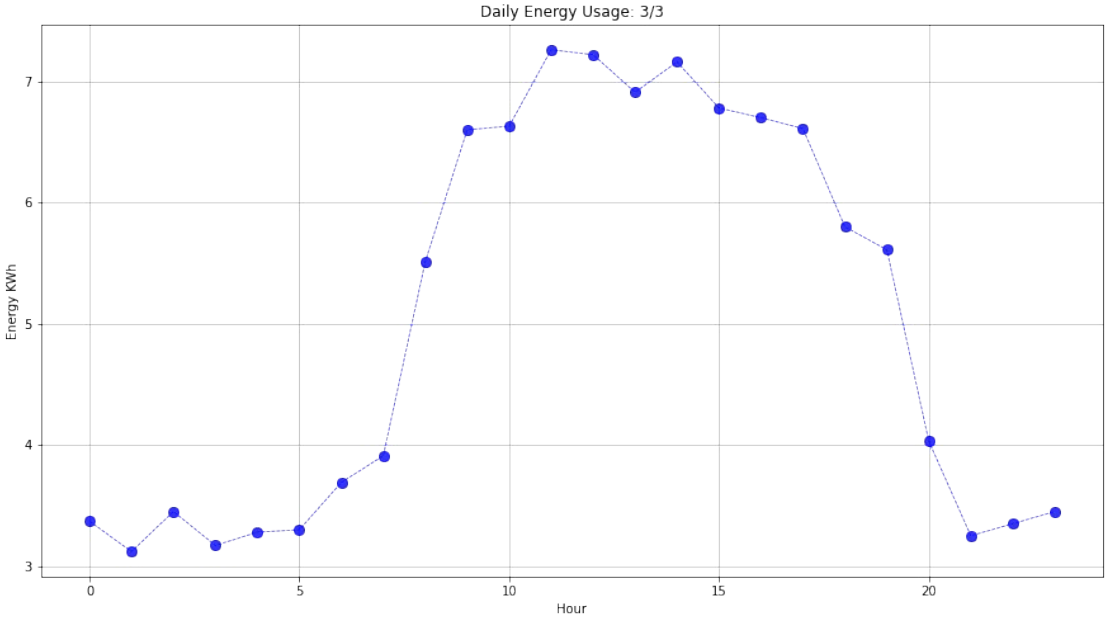
Energy Usage Plot - Year



Energy Usage Plot - Month



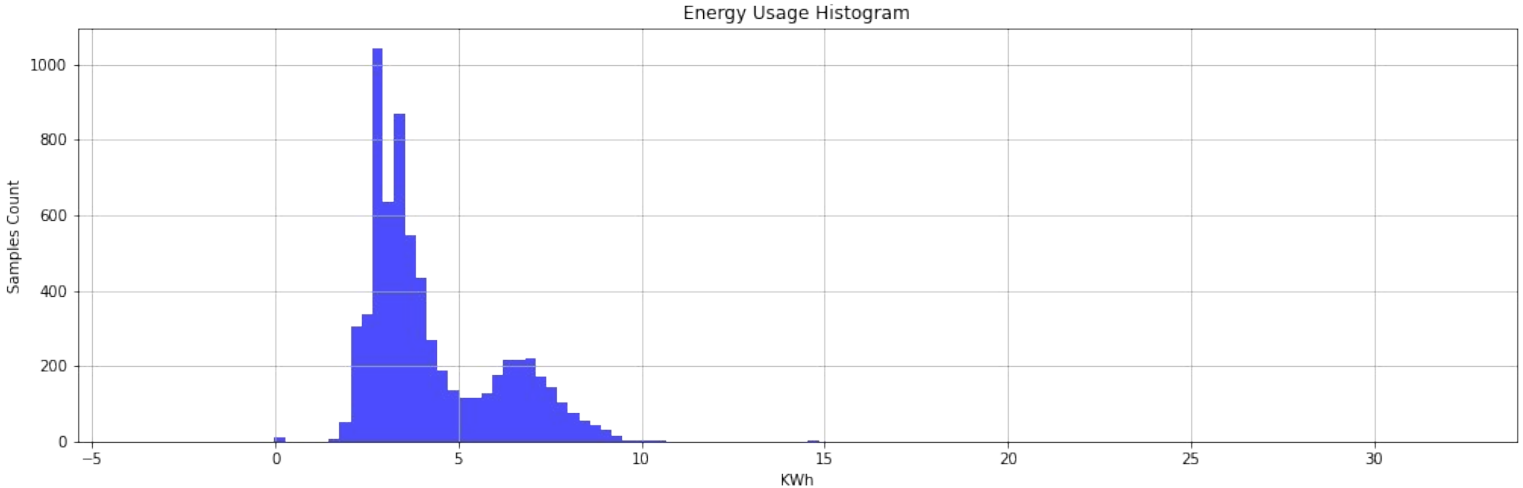
Energy Usage Plot - Day



Detecting Outliers/Anomalies

Determine Distribution of Data

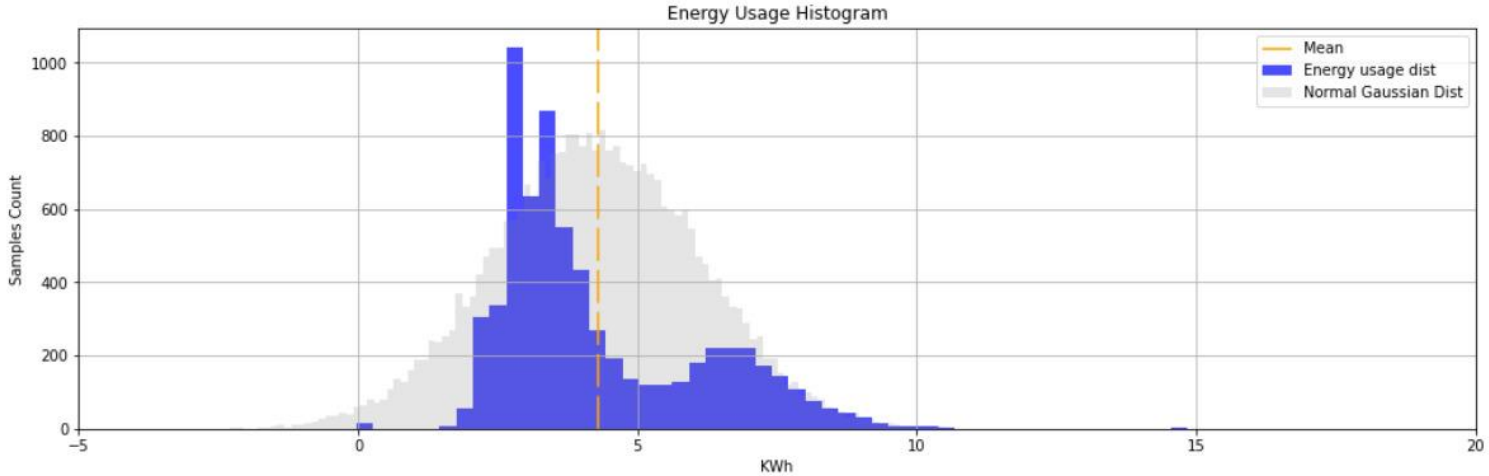
- Is data distribution normal/Gaussian?



Detecting Outliers/Anomalies

Our data vs Gaussian Distribution

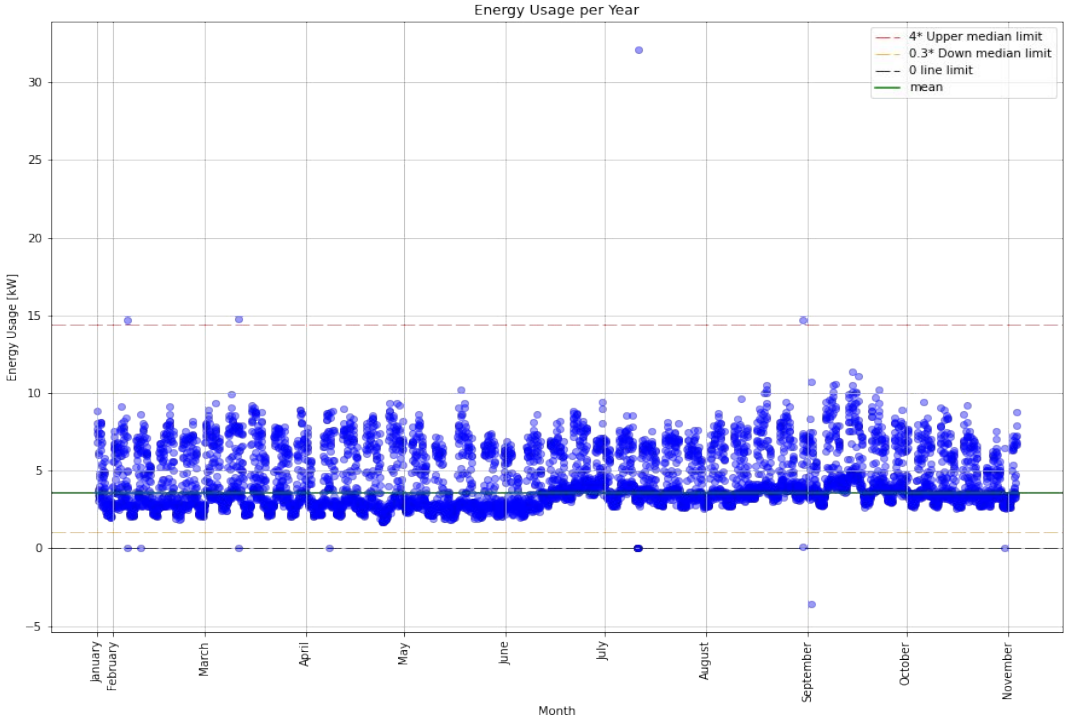
- Outlier Detection with Gaussian Distribution - Standard Deviation Method
  - Define outliers as any value outside of 4 standard deviations from the mean



Detecting Outliers/Anomalies

Rule Based Method

- Determine outliers of a non-normal distribution by defining rules based on knowledge of the data
  - Negative Values
  - Zero Values
  - Illogical Values (too high or too low)



Detecting Outliers/Anomalies

Standard Deviation Method vs Rule Based Method

- Standard Deviation Method [14.7, 14.8, 32.1, 14.7, -3.6,]
- Outliers detected by Rule Based Method: [0.0, 14.7, 0.0, 0.0, 14.8, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 0.0, 32.1, 14.7, -3.6, 0.0]

FH | JOANNEUM
University of Applied Sciences

Correcting Outliers/Anomalies

```
3945    3.26
3946    3.58
3947    3.80
3948    3.88
3949     NaN
3950     NaN
3951     NaN
3952     NaN
3953     NaN
3954     NaN
3955     NaN
3956     NaN
3957     NaN
3958    3.19
3959    3.20
Name: Energy Usage, dtype: fl
oat64
```

→

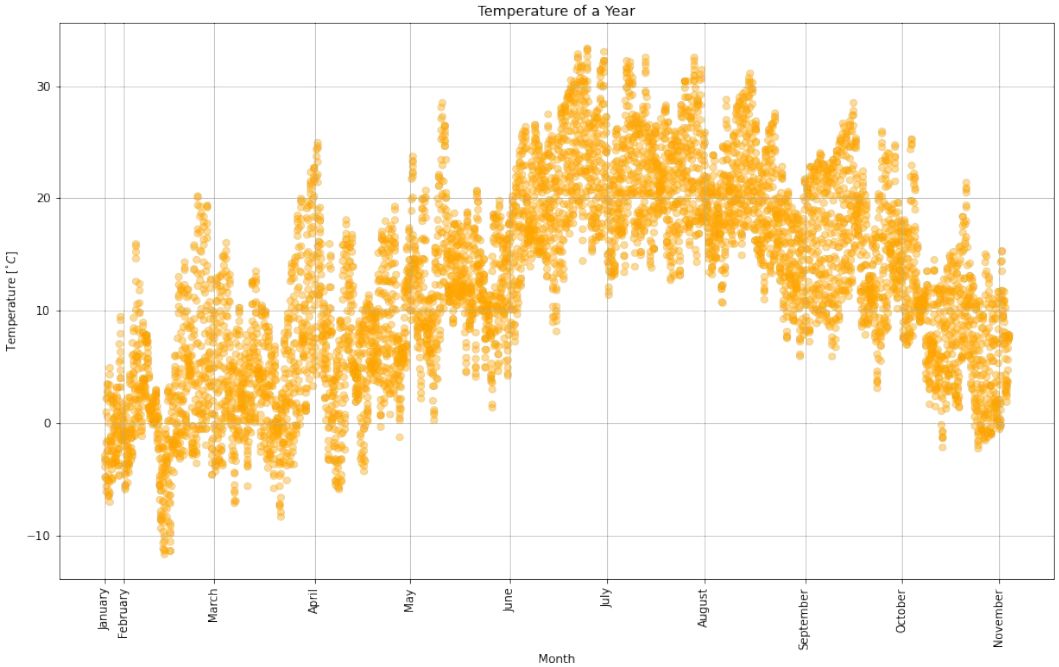
```
3945    3.260
3946    3.580
3947    3.800
3948    3.880
3949    3.811
3950    3.742
3951    3.673
3952    3.604
3953    3.535
3954    3.466
3955    3.397
3956    3.328
3957    3.259
3958    3.190
3959    3.200
Name: Energy Usage, dtype: fl
oat64
```

Replace outliers with null values

Interpolate

Preprocessing Weather Data

- Missing Data
  - Fill in missing date with null values and interpolate
- Outliers
  - by visual inspection, outliers were not detected

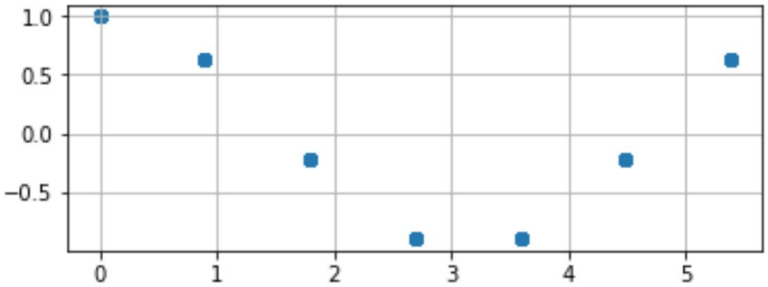


Identifying the Holidays

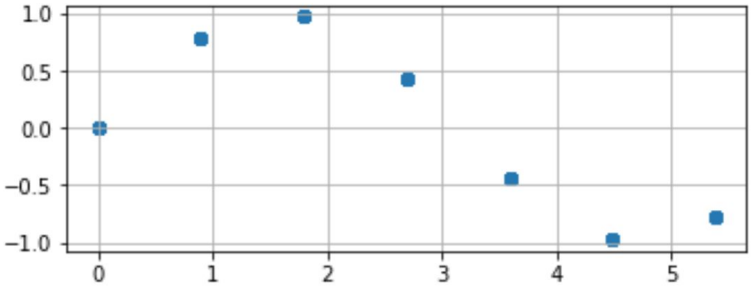
- Install holidays package
- Designate holidays as 1 and non-holidays as 0

Decoding the week days

- Normalize week days (0-6) between 0 and 2 **π**
- Take the sin and cos of normalized weekday values
- Each day of the week corresponds to a specific combination of sin and cos



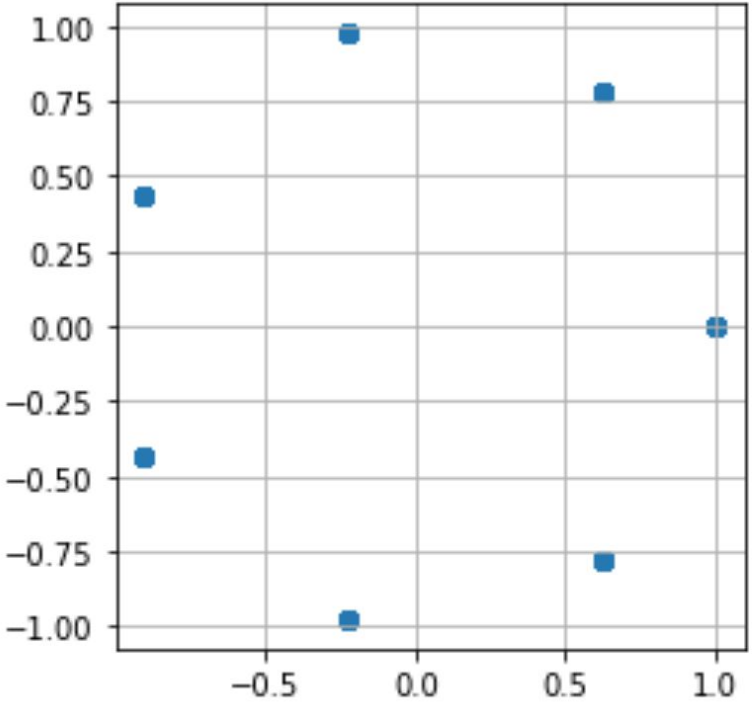
Cos vs Norm Week day



Sin vs Norm Week day

Decoding the week days

- Plotting sin vs cos demonstrates the circular nature of the weekdays



Merging Data into one Dataframe

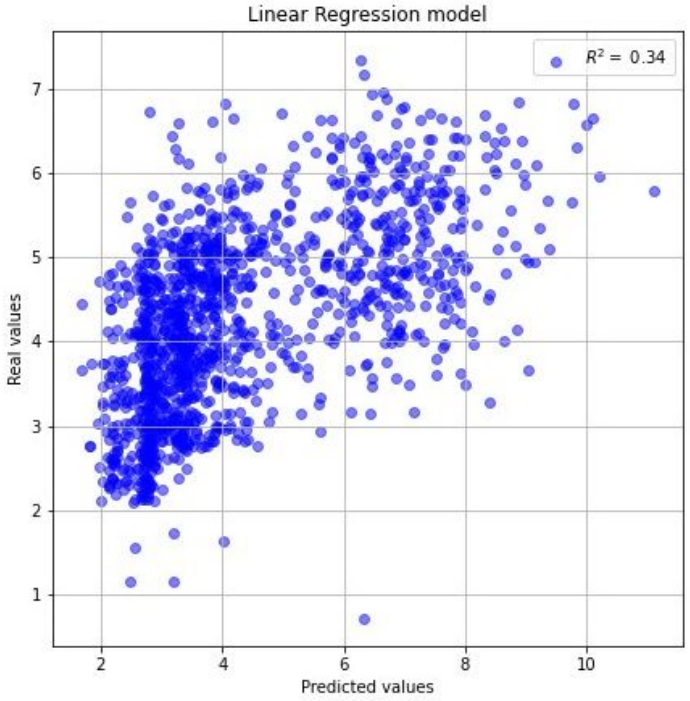
- Weather data, weekdays, and holidays were merged with the master Dataframe

Inputs

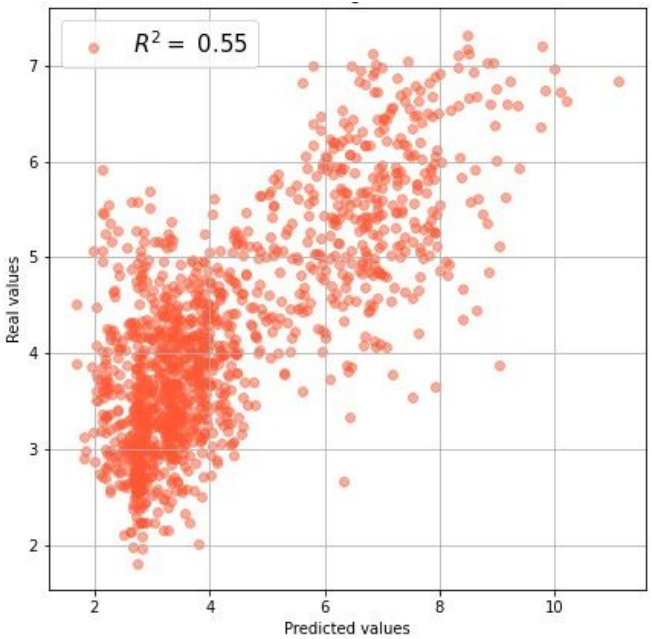
Target Variable

	year	month	day	hour	Sin	Cos	Holiday	t	Energy Usage
0	2021	1	27	11	0.974928	-0.222521	0.0	0.0	8.03
1	2021	1	27	12	0.974928	-0.222521	0.0	1.9	6.85
2	2021	1	27	13	0.974928	-0.222521	0.0	3.5	8.86
3	2021	1	27	14	0.974928	-0.222521	0.0	4.4	7.37
4	2021	1	27	15	0.974928	-0.222521	0.0	5.0	7.67
...	...	...	...	...	...	...	...	...	...
6721	2021	11	3	12	0.974928	-0.222521	0.0	5.7	6.68
6722	2021	11	3	13	0.974928	-0.222521	0.0	6.7	7.26
6723	2021	11	3	14	0.974928	-0.222521	0.0	7.3	6.81
6724	2021	11	3	15	0.974928	-0.222521	0.0	7.4	8.81
6725	2021	11	3	16	0.974928	-0.222521	0.0	7.7	7.92

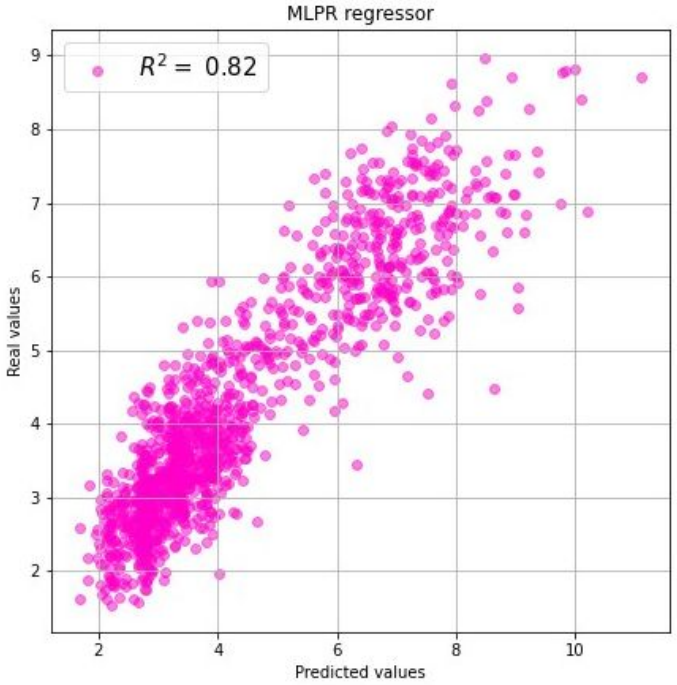
LINEAR REGRESSION



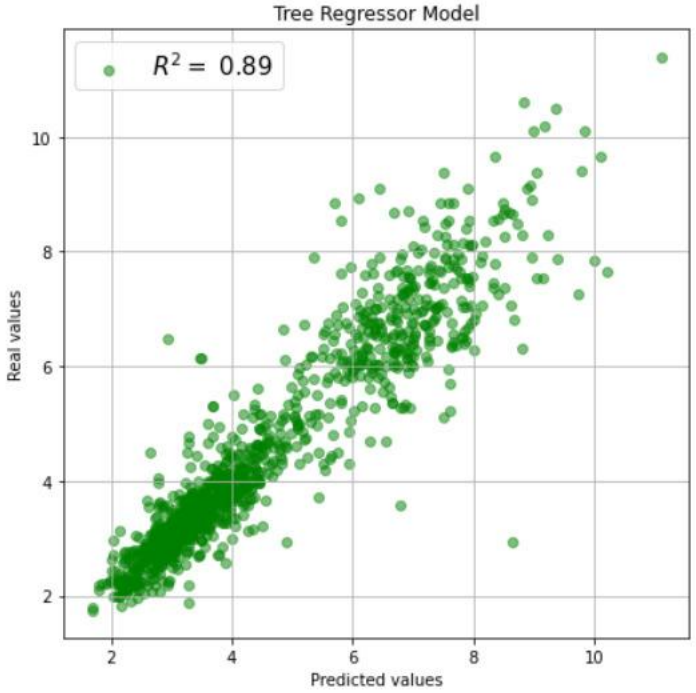
SUPPORT VECTOR REGRESSOR



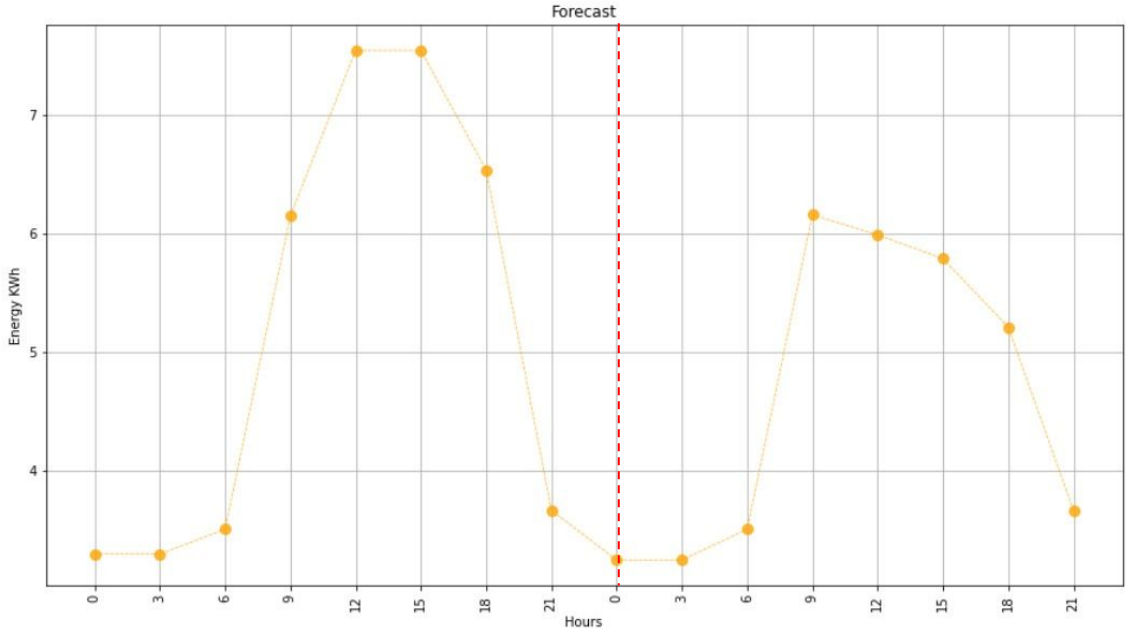
MLP REGRESSOR



DECISION TREE REGRESSOR

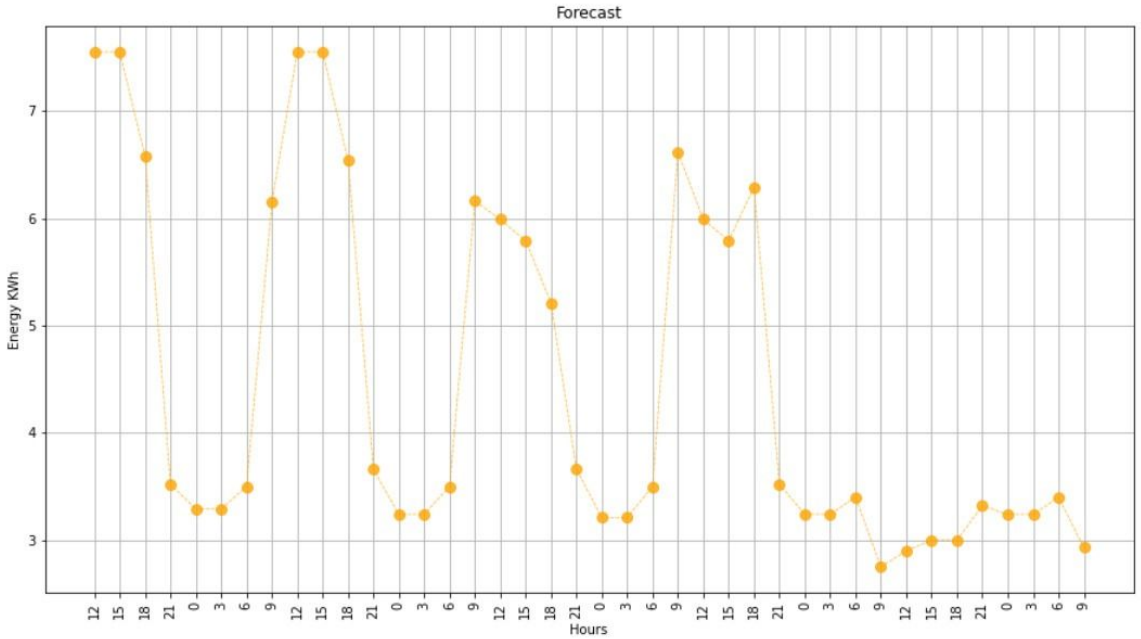


48 hour forecast using Decision Tree Regressor model



Tuesday and Wednesday forecast

5 day forecast using Decision Tree Regressor model



5 day forecast (Monday 12:00 - Saturday 9:00)

Conclusion

- Based on R-squared values, the best model out of the tested methods is the Decision Tree Regressor
- Next steps:
  - Further analysis on the different models
  - Update model: adjust inputs and hyperparameters if necessary

FH | JOANNEUM
University of Applied Sciences

Thank you!