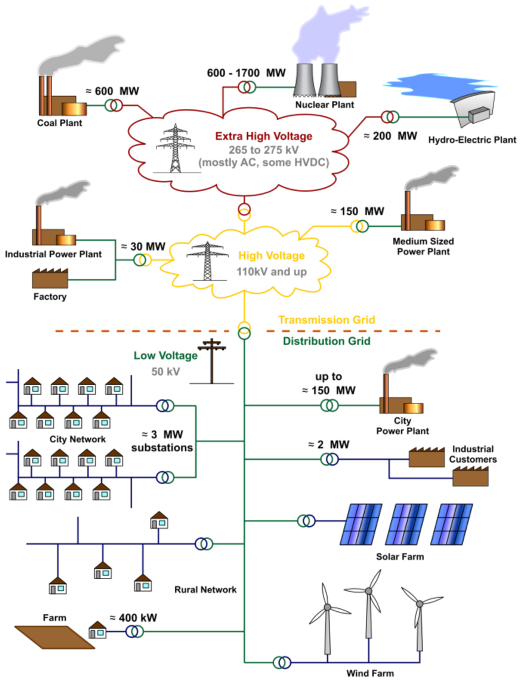
Hydro-Québec: Predicting the Hourly Ontario Energy Price in the Medium and Long Term

AUGUST 27, 2020

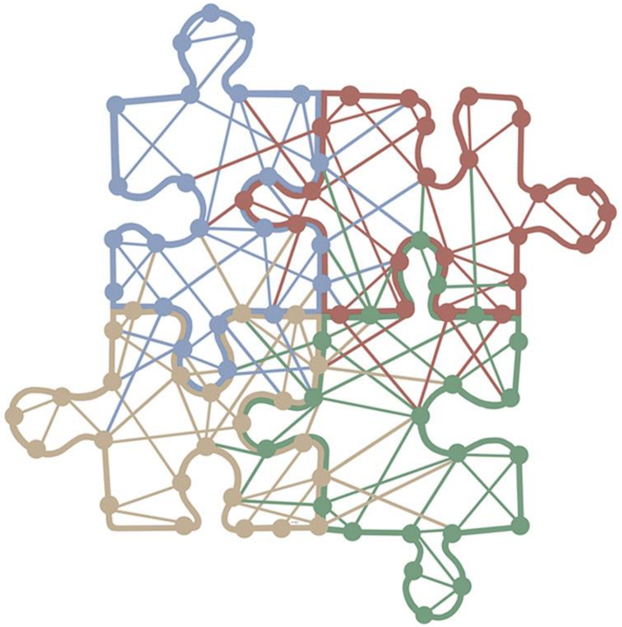
Outline

I. DATASET & GOAL



https://en.wikipedia.org/wiki/Electrical_grid

II. MACHINE LEARNING APPROACH



https://behavioralscientist.org/scaling-nudges-machine-learning/

I. Dataset and Goal

I. Dataset and Goal

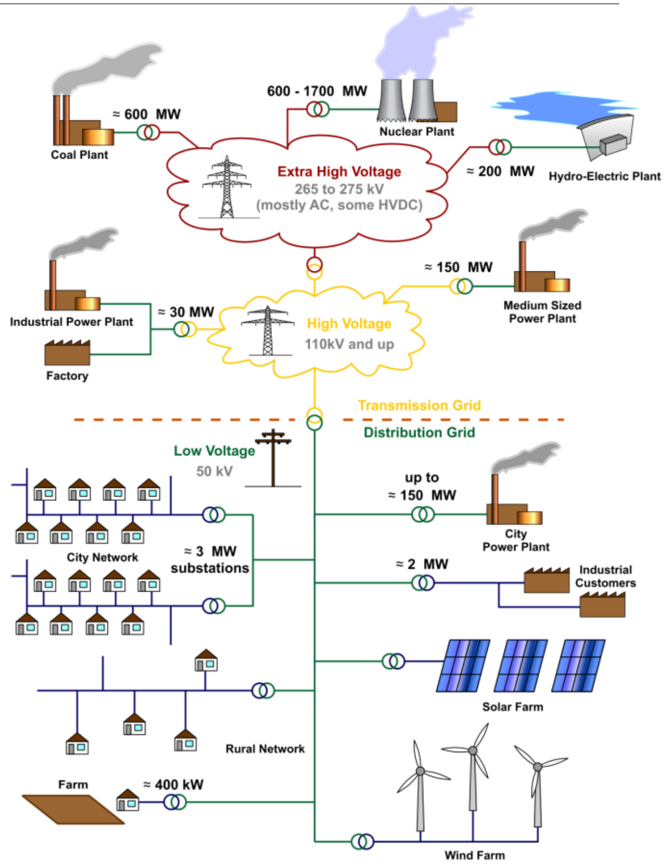
Goal

➢ Prediction of the **Ontario Energy Price**

- Medium and long-term Periods (18 months)
- For sales planning

➢ Ontario Market:

- A Difficult market to predict:
  - Many fixed price supply contracts
  - 12 % coming from wind-based resources → intermittent
  - A lot of uncertainty in demand



https://en.wikipedia.org/wiki/Electrical_grid

I. Dataset and Goal

Dataset

Available Data sets:

- Predicted weekly data (18 month predictions): 2015 – 2020
- Historical hourly data: 2017 – 2020

	HOEP	Bruce PD	East PD	Essa PD	Niagara PD	NorthEast PD	NorthWest PD	Ottawa PD	SouthWest PD	Toronto PD	...	Expected Hydro Output
count	1473.000000	1746.000000	1746.000000	1746.000000	1746.000000	1746.000000	1746.000000	1746.000000	1746.000000	1746.000000	...	1737.00000
mean	16.592956	86.197850	1302.368587	1197.743535	587.803402	1275.523675	471.597209	1228.498074	4012.024750	7388.431608	...	2409.75806
std	9.340326	22.155787	160.801969	167.620619	82.616598	177.474115	74.234597	180.238151	347.982958	843.344809	...	371.01938
min	-1.624762	47.730869	570.300757	762.932920	385.170000	856.255972	279.708361	872.306650	3132.000000	5652.000000	...	0.00000

Goal: Prediction of price from predicted parameters in weekly data

- Test:
  - Three 18-months prediction files

II. Machine Learning Approach

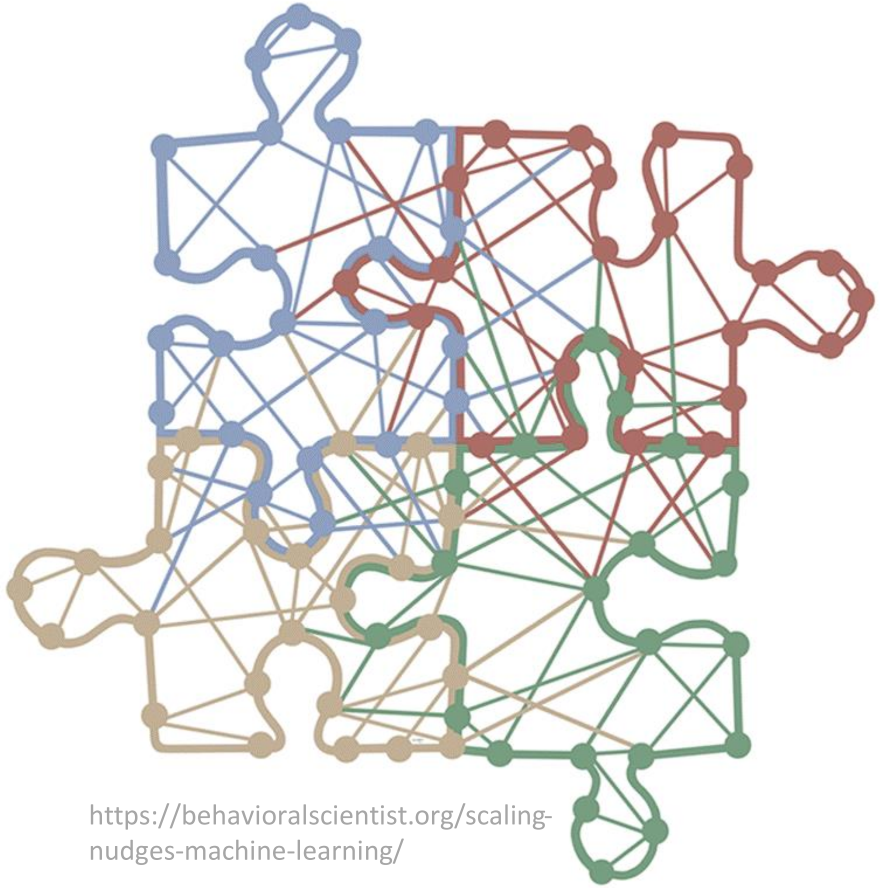
- Classical Machine Learning

II. Machine Learning Approach

General overview

- Methods:

1. **Classical machine learning (CML)**

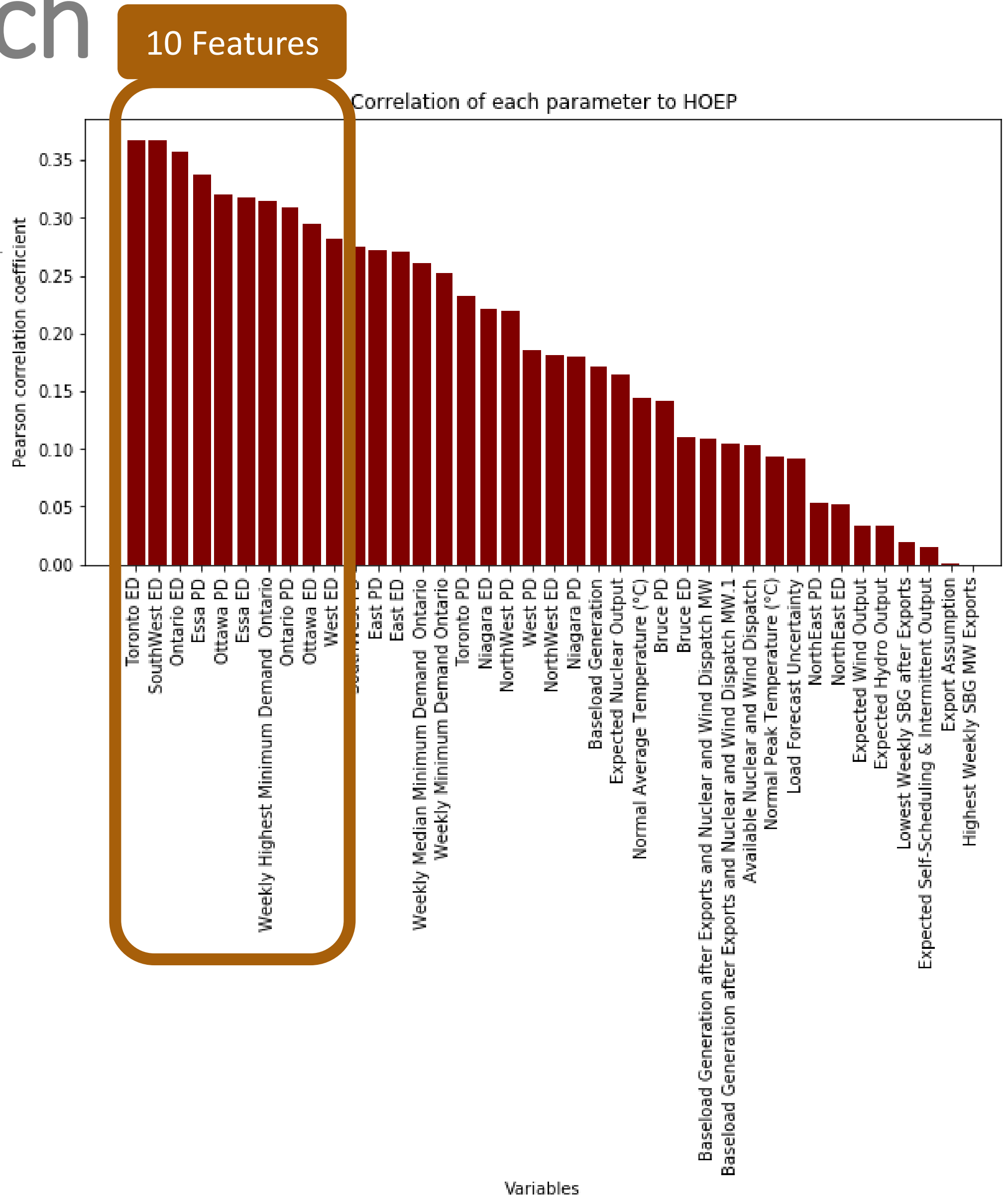


https://behavioralscientist.org/scaling-nudges-machine-learning/

II. Machine Learning Approach

CML Method: Features

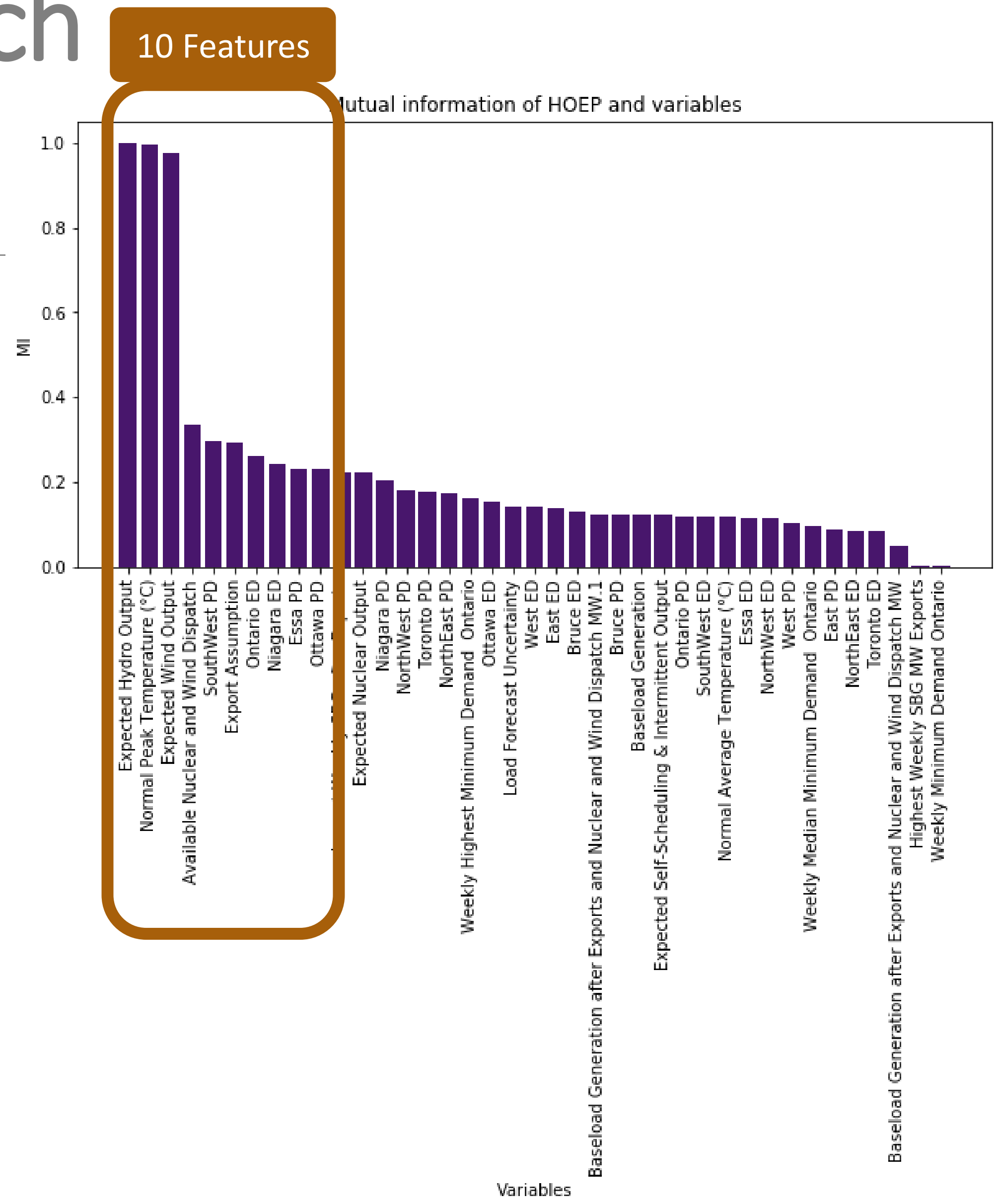
- Features to be used
    1. Pearson Correlation



II. Machine Learning Approach

CML Method: Features

- Features to be used
  1. Pearson Correlation
  2. Mutual information



II. Machine Learning Approach

CML Method: Features

- Features to be used

1. Pearson Correlation
2. Mutual information
3. Non-zero coefficients from Lasso Regression (all survived feature with alpha = 0.4)

'Bruce PD' 'NorthEast PD' 'NorthWest PD' 'Load Forecast Uncertainty' 'Essa ED' 'NorthEast ED' 'SouthWest ED' 'Toronto ED' 'Baseload Generation after Exports and Nuclear and Wind Dispatch MW.1' 'Lowest Weekly SBG after Exports'

II. Machine Learning Approach

CML Method: Features

- Features to be used

  1. Pearson Correlation
  2. Mutual information
  3. Non-zero coefficients from Lasso Regression

  - All features grouped together:
    - 24 unique features:

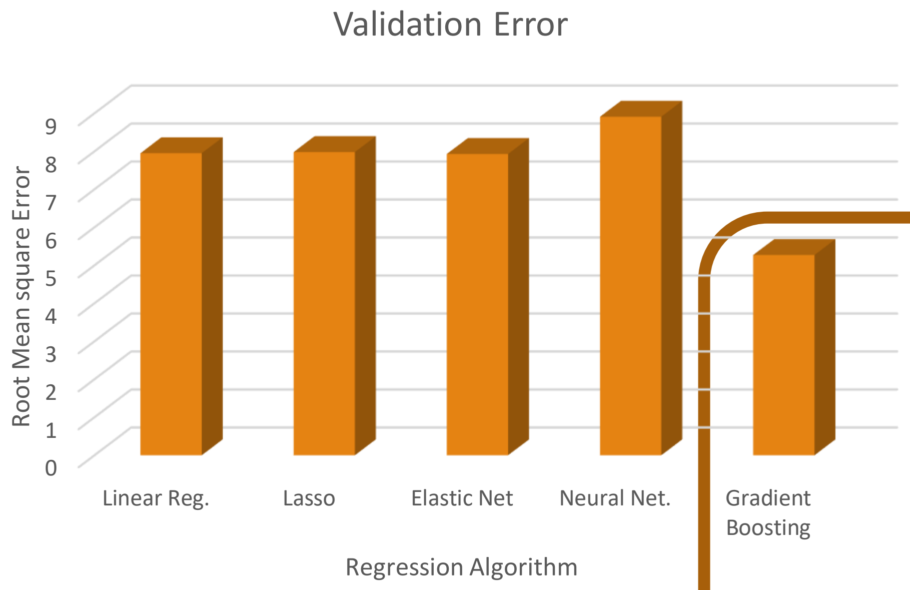
      'Toronto ED', 'SouthWest ED', 'Ontario ED', 'Essa PD', 'Ottawa PD', 'Essa ED', 'Weekly Highest Minimum Demand Ontario', 'Ontario PD', 'Ottawa ED', 'West ED', 'Expected Hydro Output', 'Normal Peak Temperature (°C)', 'Expected Wind Output', 'Available Nuclear and Wind Dispatch', 'SouthWest PD', 'Export Assumption', 'Niagara ED', 'Bruce PD', 'NorthEast PD', 'NorthWest PD', 'Load Forecast Uncertainty', 'NorthEast ED', 'Baseload Generation after Exports and Nuclear and Wind Dispatch MW.1', 'Lowest Weekly SBG after Exports'

II. Machine Learning Approach

CML Method: Regression

- Algorithms:

  1. Linear Regression
  2. Lasso
  3. Elastic Net
  4. Neural Network (MLP with two hidden layers)
  5. Gradient Boosting



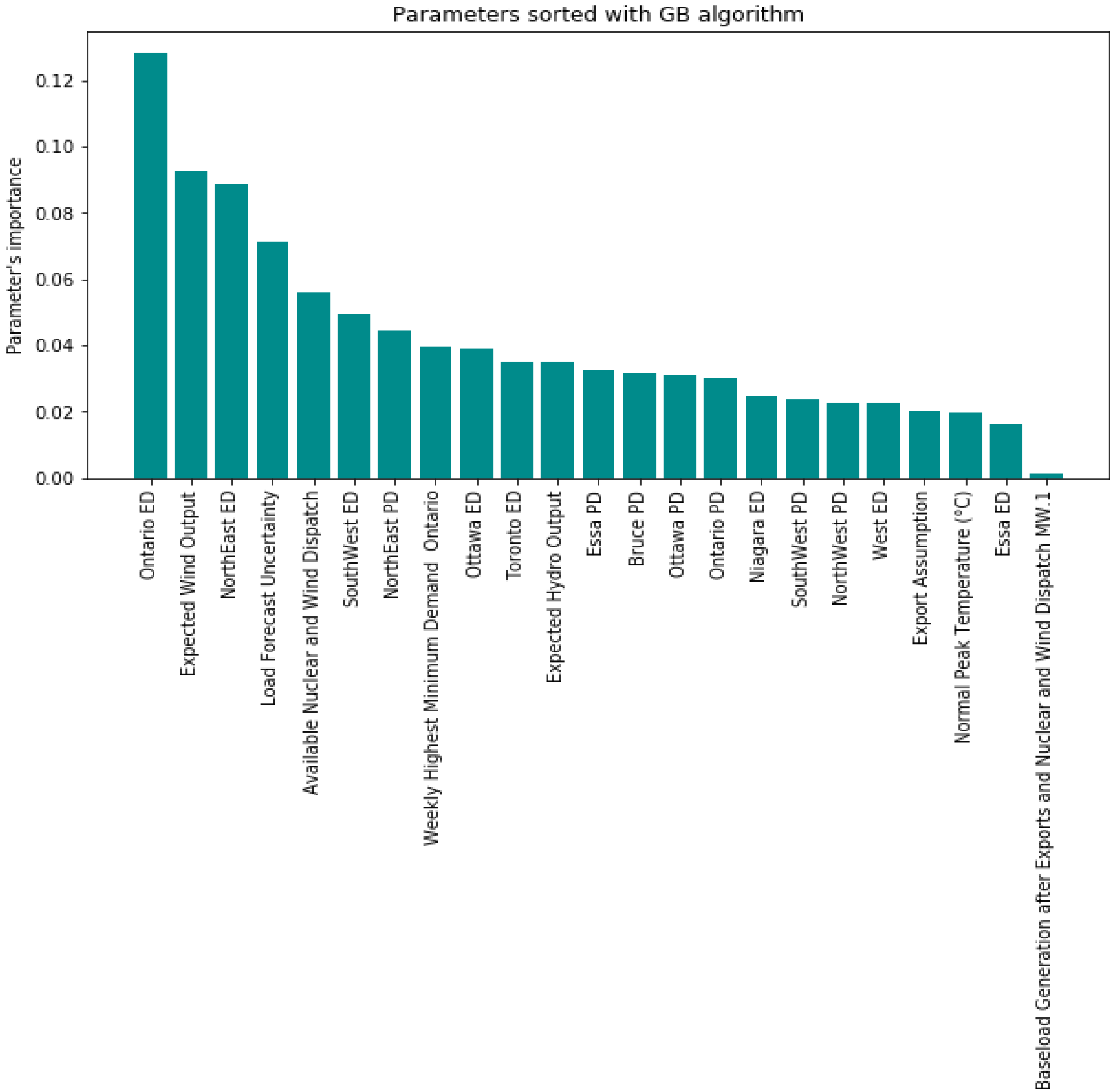
- Validation:
  - 25% of Training data

	Linear Reg.	Lasso	Elastic Net	Neural Net.	Gradient Boosting
rMSE (price)	7.95	7.98	7.93	8.91	5.15

II. Machine Learning Approach

CML Method: Regression

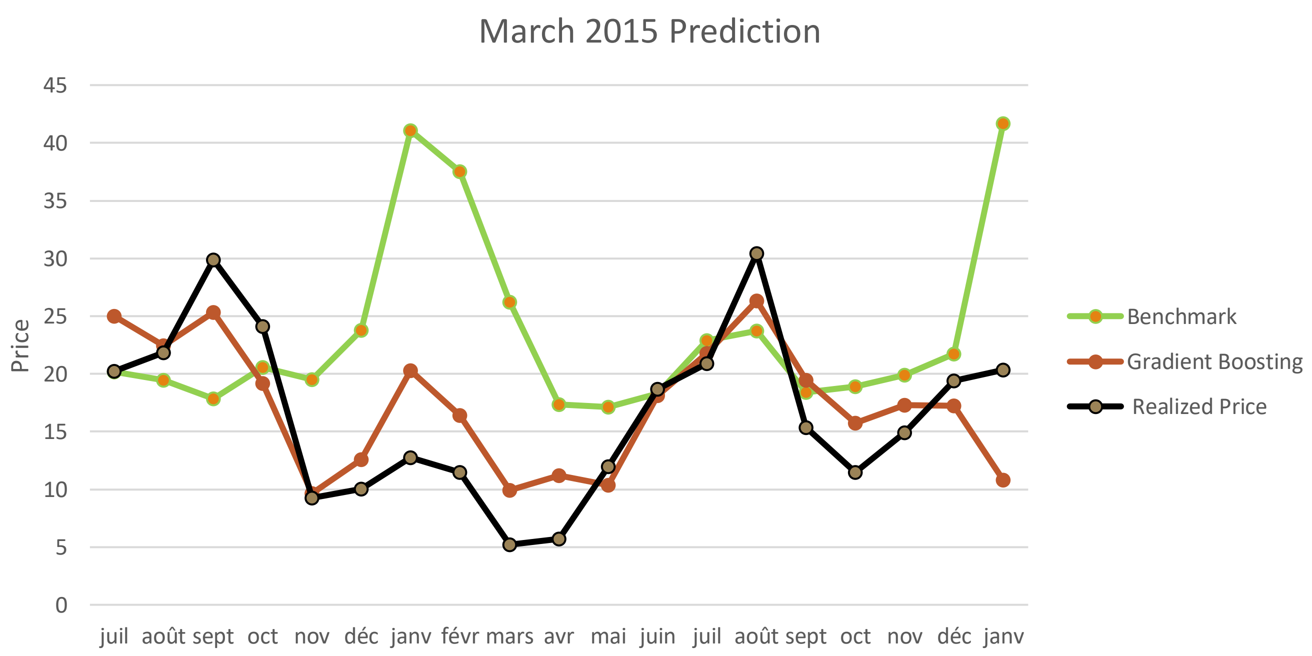
- Algorithms:
  - ✓ **Gradient Boosting**
  - Selected features **sorted** with **Gradient Boosting** algorithm:



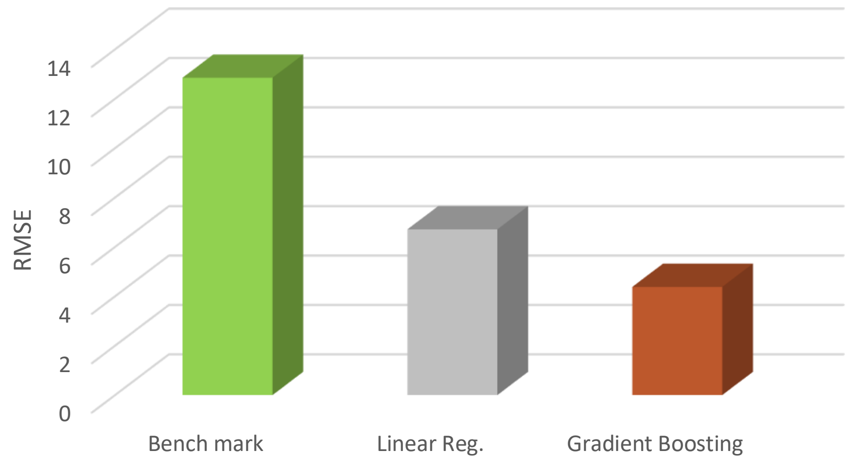
II. Machine Learning Approach

CML Results

- Goal: predicting the **real price** better than the benchmark (FWD HOEP)
  - Best approach based on validation data: Gradient Boosting



Error in predicting real price (18 month period)

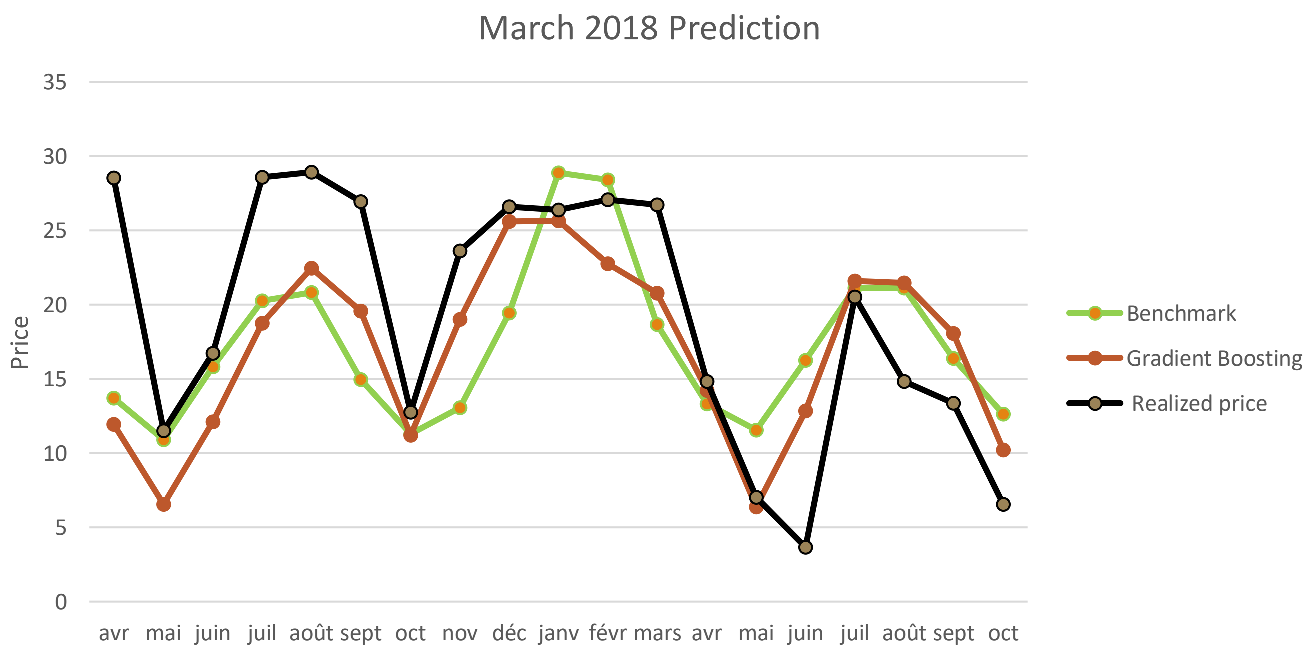


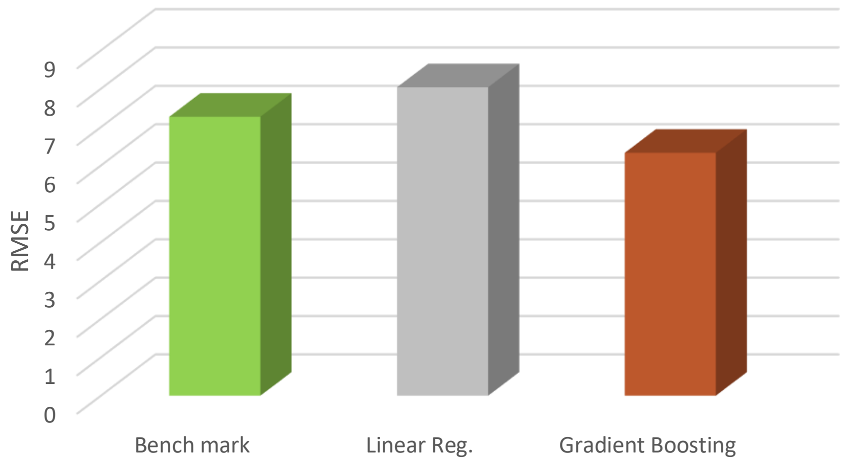
	Benchmark	Linear Reg.	Gradient Boosting
rMSE	12.85	6.71	**4.38**

II. Machine Learning Approach

CML Results

- Goal: predicting the **real price** better than the benchmark (FWD HOEP)
  - Best approach based on validation data: Gradient Boosting



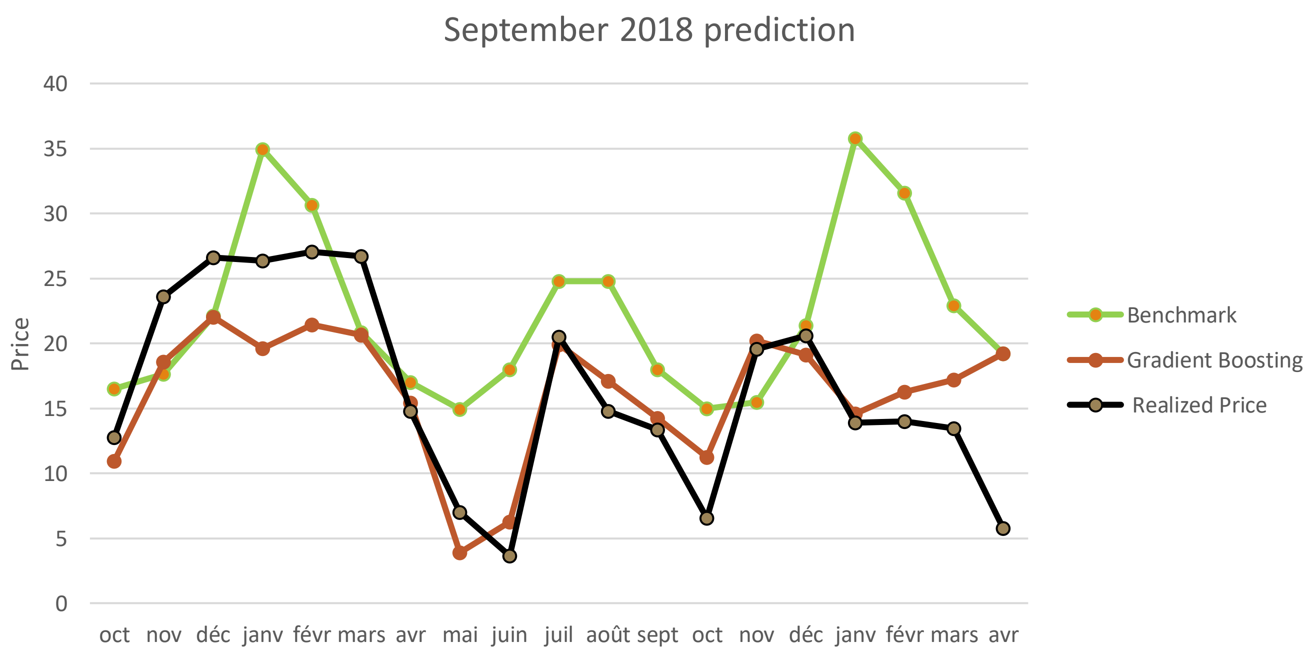


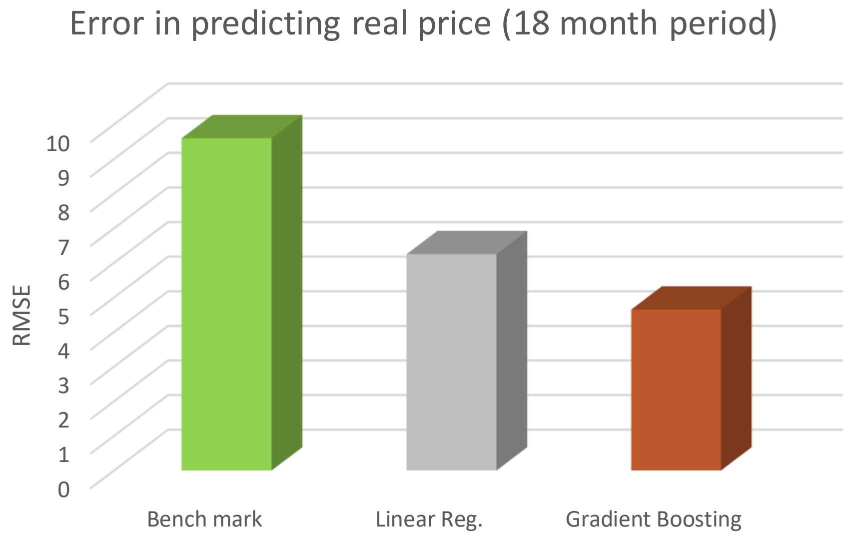
	Benchmark	Linear Reg.	**Gradient Boosting**
rMSE	7.27	8.04	**6.33**

II. Machine Learning Approach

CML Results

- Goal: predicting the **real price** better than the benchmark (FWD HOEP)
  - Best approach based on validation data: Gradient Boosting



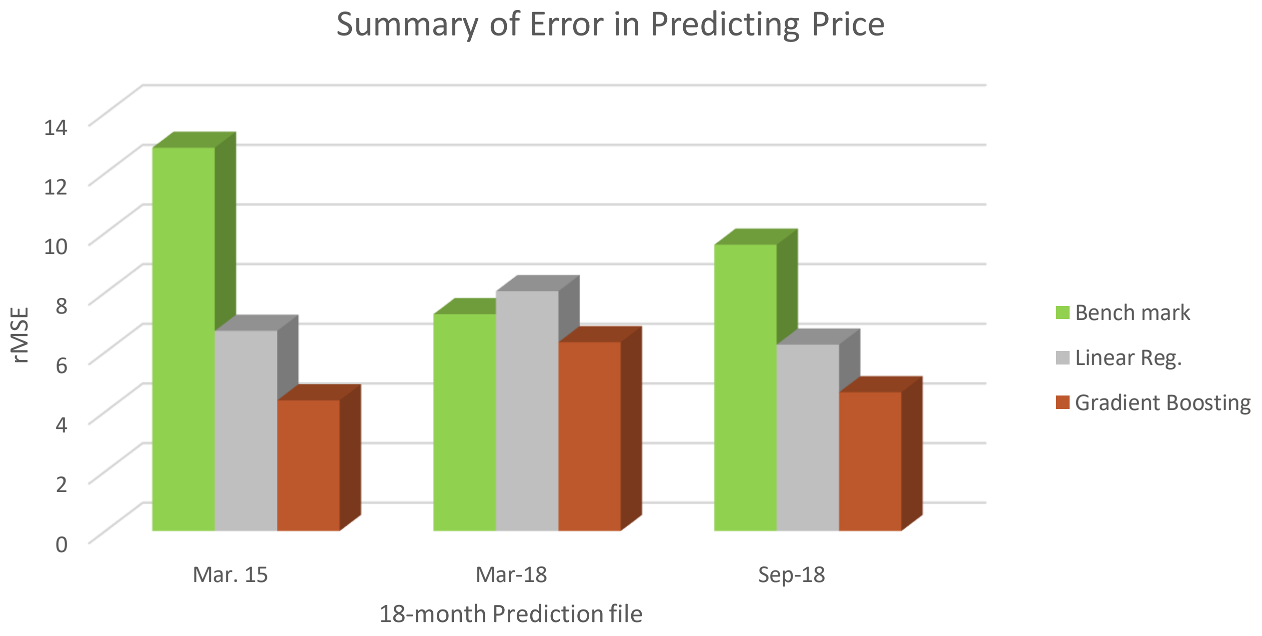


	Benchmark	Linear Reg.	**Gradient Boosting**
rMSE	9.60	6.25	**4.65**

II. Machine Learning Approach

CML Results

➢ Goal: predicting the **real price** better than the benchmark (FWD HOEP)



	Benchmark	Linear Reg.	Gradient Boosting
Mar. 15	12.85	6.71	**4.38**
Mar. 18	7.27	8.04	**6.33**
Sep. 18	9.6	6.25	**4.65**