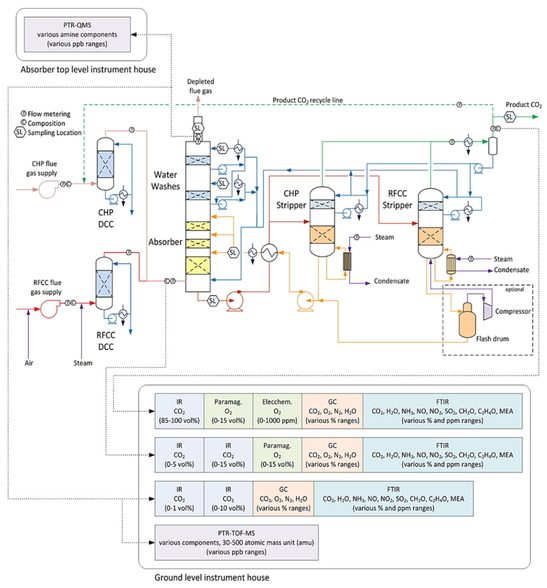
University of South-Eastern Norway

Development of machine learning model for CO_2 capture plants to predict solvent degradation

Sam Narimani
MSc. Process Technology

Introduction and background

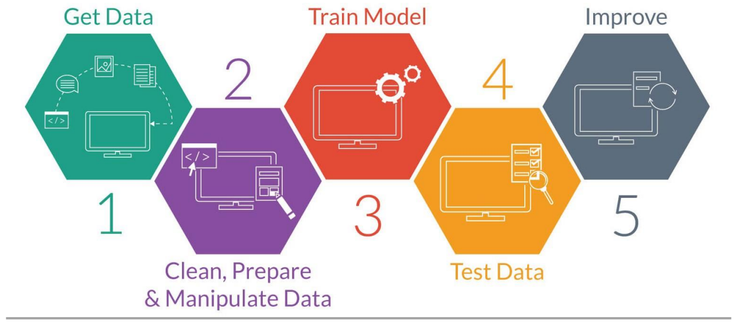
The purpose of this study is developement of machine learning methods to predict solvent degradation phenomena in a carbon capture plant. There are two main groups of solvent degradation in a carbon capture plant namely oxidative and thermal degradation. Information regarding 10 types of solvent degradation in the plant were extracted. Then, three methods of Support Vector Regression (SVR), Random Forest (RF) and Artificial Neural Netwrok (ANN) were used to find the proper model for all types of solvent degradation. To reach the best result for each method, hyperparameters were tuned by grid search and randomized search cross validation.



Plant overview [1]

Machine learning Process

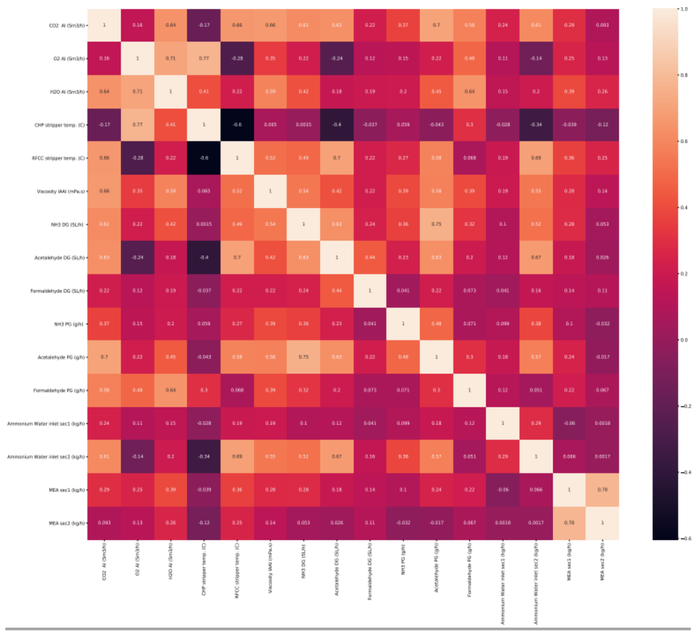
Relevant data were extracted and pre-processed before machine leaning implementations. Feature selection methods were applied for all data to find the best features for each type of solvent degradation. Then, machine learning methods were implemented to train the models. Finally, hyperparameters were tuned by grid search and randomized search cross validation. The models were evaluated and tested and improved in case of overfitting or undefitting.



Machine learning process [2]

Feature Selection

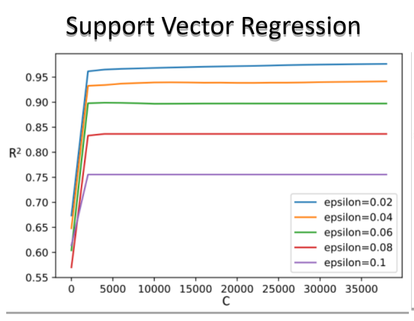
Two Feature selection methods were utilized to enhance the performance of the models. Spearman's and Pearson's methods were used as two main techniques in correlation coefficient methods and finally, Spearman's coefficients were used for further implementations since it is a non-linear rank coefficient.



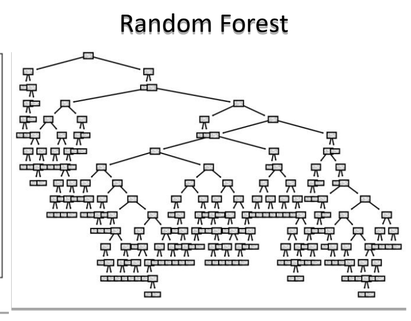
Results

Results showed that Random Forest and Artificial Neural Network were the best model for prediction of solvent degradation phenomena. Support Vector Regresion also presented acceptable results and in a few cases better results rather than the other two methods.

Support Vector Regression



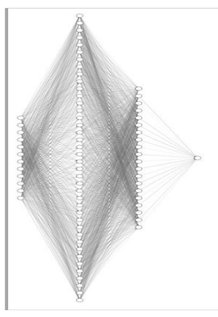
Random Forest



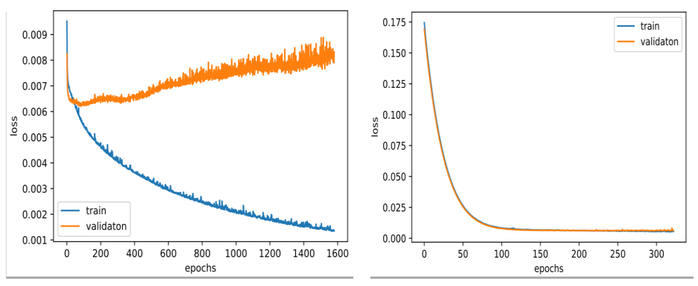
Splitting type	Train set (80%), Test set (20%)		Train set (70%), Test set (30%)		Regularization parameter	Epsilon
Outputs	Train R^2	Test R^2	Train R^2	Test R^2		
NH3 DG	0.919	0.902	0.919	0.906	47000	0.02
Acetaldehyde DG	0.554	0.502	0.566	0.485	45000	0.04
Formaldehyde DG	0.449	0.443	0.482	0.362	55000	0.05
NH3 PG	0.453	0.446	0.458	0.436	52000	0.04
Acetaldehyde PG	0.970	0.973	0.968	0.976	56000	0.02
Formaldehyde PG	0.873	0.856	0.876	0.852	50000	0.02
Ammonium Water inlet sec1	0.305	0.149	0.292	0.175	12000	0.02
Ammonium Water inlet sec2	0.947	0.887	0.947	0.915	60000	0.04
MEA sec1	0.967 0	0.9861	0.963	0.983	30000	0.02
MEA sec2	0.980	0.984	0.975	0.988	15000	0.02

Splitting type	Train set (80%), Test set (20%)		Train set (70%), Test set (30%)	
Type of degradation	Train R^2	Test R^2	Train R^2	Test R^2
NH3 DG	0.9951	0.9099	0.9999	0.8917
Acetaldehyde DG	0.9878	0.8334	0.9999	0.7684
Formaldehyde DG	0.9218	0.4841	0.5848	0.2866
NH3 PG	0.9250	0.5059	0.9414	0.5434
Acetaldehyde PG	0.9999	0.9368	0.9982	0.9146
Formaldehyde PG	0.9999	0.7695	0.9919	0.7257
Ammonium Water inlet sec1	0.9489	0.1384	0.9411	0.1594
Ammonium Water inlet sec2	0.9999	0.8467	0.9999	0.8143
MEA sec1	0.9911	0.8830	0.9999	0.8944
MEA sec2	0.9756	0.9841	0.9602	0.9585

Artificial Neural Network



# Neuron in first hidden layer	# Neuron in second hidden layer	Train R^2	Test R^2	Learning rate
120	100	0.9981	0.9231	0.0015
	60	0.9833	0.9081	0.0019
	20	0.9916	0.9002	0.0035
70	100	0.9783	0.9161	0.0035
	60	0.9821	0.9196	0.0021
	20	0.9776	0.9224	0.0023
40	100	0.9800	0.9170	0.0021
	60	0.9707	0.9154	0.0021
	20	0.9819	0.9107	0.0031



Conclusion

Various machine learning methods were used to represent the pattern between the selected features and different types of solvent degradation. Three models of ANN, RF and SVR were implemented, and the corresponding results were demonstrated. Hyperparameters in all methods were tuned to introduce the best possible models for all types of solvent degradation. To optimize the hyperparameter, grid and randomized search optimization methods were used for all models. Results demonstrated that all models forecasted very well except for some cases for instance, NH3 PG and Ammonium water inlet sec1. ANN and RF displayed the favorable results in most cases whereas SVR also presented acceptable models in a few cases. For example, R2 results for NH3 DG, MEA sec 1, MEA sec2, Acetaldehyde DG, Acetaldehyde PG and Ammonium water inlet sec2 were more than 0.90 in RF and ANN models which showed quite accurate results. SVR also appropriately predicted MEA sec1 and MEA sec2 with high train and test R2 showing more than 0.97.

References

1. Morken, A.K., et al., *Emission results of amine plant operations from MEA testing at the CO2 Technology Centre Mongstad.* Energy Procedia, 2014. **63**: p. 6023-6038.

2. Fareed, H. *Machine Learning for Dummies*. 07.04.2018 17.05.2022]; Available from: https://towardsdatascience.com/wtf-is-machine-learning-a-quick-guide-39457e49c65b.

Supervisors:
Leila Ben Saad , Ru Yan

TECHNOLOGY CENTRE MONGSTAD