ENSY 5800: Applications of A.I. in Energy Systems

Fall 2021 Project Presentation

Fault diagnosis of a building air-handling unit

Shwe Myat Myo Oo

December 08, 2021

Outline

- Executive Summary
- Problem Statement & Definition
- Exploratory Data Analysis
- AI/ML Modeling & Results
- Conclusions & Recommendations

Executive Summary

Problem

- Fault detection is an integral part of any modern building system. If faults are not detected and resolved in a timely manner, health and safety issues may arise in addition to wasted energy from ineffective usage.

Objective

- Accurately detect whether a fault has occurred in the given AHU system with the available data.

Approach

- Pre-process and prepare the data, conduct EDA on the dataset, apply AI techniques (unsupervised learning, supervised learning, and ANN), check goodness measures and find the best model for fault diagnosis.

Impact

- Currently, detecting a fault in AHUs are automated via conditional programming in BMS systems (which may not be reliable) and the final determination is done by an experienced engineer (especially for false negatives)
- Accurate and prompt automated fault detection minimizes AHU downtime, resulting in improved comfort, energy savings, air quality improvement, increased equipment lifetime, and improved service scheduling

Problem Statement & Definition

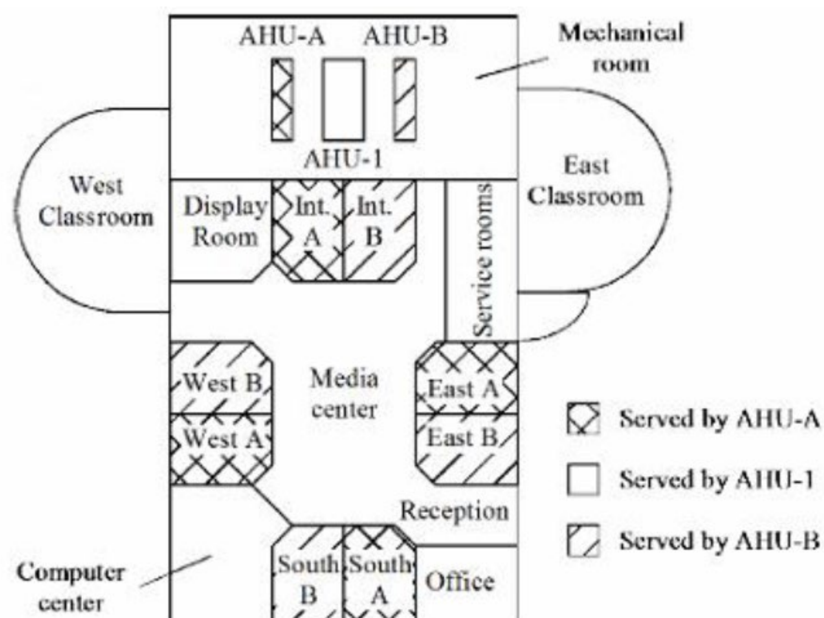
- An air handling unit (AHU) connects primary heating and cooling plants with building zones, controls building ventilation air intake, and greatly affects the energy consumed for heating, cooling, and ventilating, as well as supply air temperature and humidity levels
- In effect, AHUs manage heat/energy exchange and ventilation
- Therefore, an AHU's operation significantly impacts building energy use, health, and comfort aspects, e.g. faulty AHU ventilation → higher risk for recirculation of particulates such as COVID-19

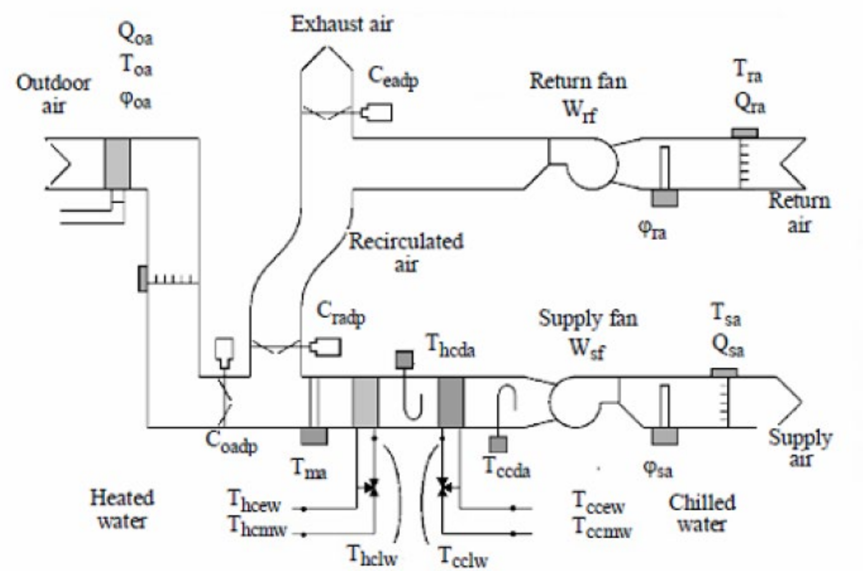
Difficulties for fault prediction in AHU systems

- Every AHU system is slightly different, as each is a custom system (especially for variable-air-volume systems)
- Multiple operational modes possible
- Frequent transient operation
- There is currently no generalized analytical model for complete analysis of an arbitrary AHU.

Problem Statement & Definition

- Fault detection for AHU-A in a small size commercial building in Iowa during summer, winter, and transition season. The faults were manually imposed into the control system. Each fault was tested for one day. The experimental dataset was provided in "MZVAV-2-1.csv" and the simulated dataset was provided in "MZVAV-2-2.csv"
- **Main variables:** temperature (outdoor air, supply air, mixed air, return air, status signals (supply air fan, return air fan), control signals (supply air fan speed, return air fan speed, exhaust air damper, outdoor air damper, return air damper, cooling coil valve, heating coil valve), occupancy mode indicator
- **Determine:** fault detection ground truth





Problem Statement & Definition

Exploratory Data Analysis

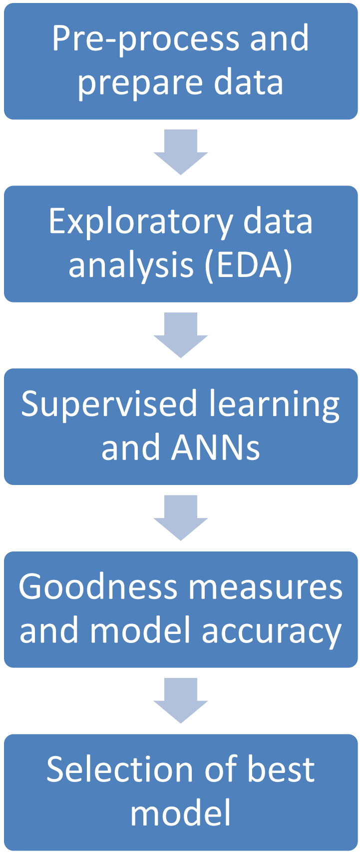
- For exploratory data analysis, I used boxplots, histograms, cross plots, correlation plots, principal components analysis, and elliptic covariance outlier detection to investigate the data.

Unsupervised learning

- For clustering, I utilized k-means clustering, hierarchical clustering, Gaussian Mixture Modeling, and self-organizing maps (SOMs)

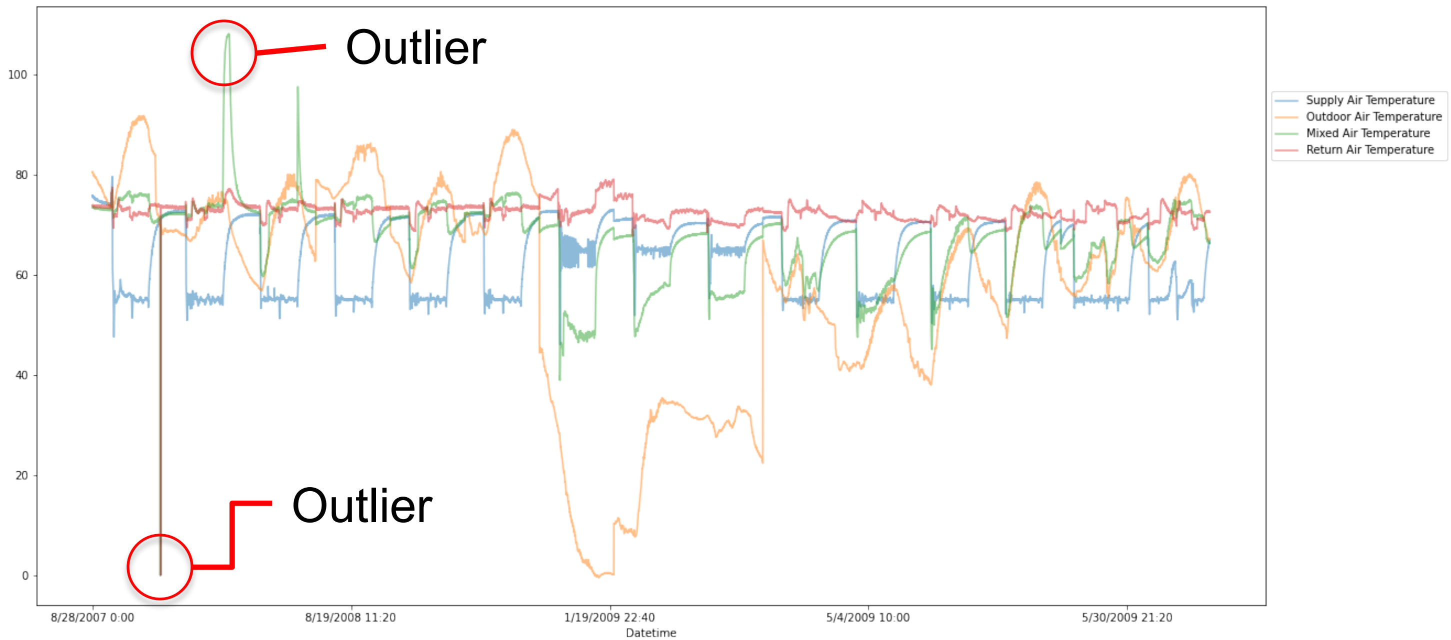
Supervised learning

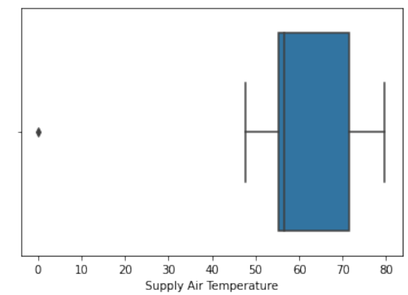
- For classification, I utilized decision trees classification, k-nearest neighbors (kNN), and multi-layer perceptrons (MLPs).
- For regression, I used decision trees regression, linear regression, and long short-term neural networks (LSTMs).

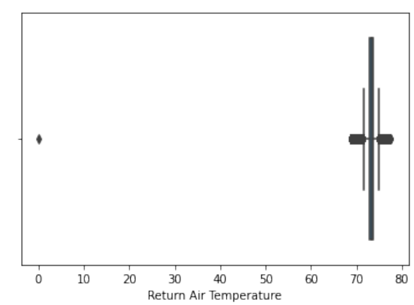


Exploratory Data Analysis

- Pre-processed raw data:
- No missing values but has clear outliers



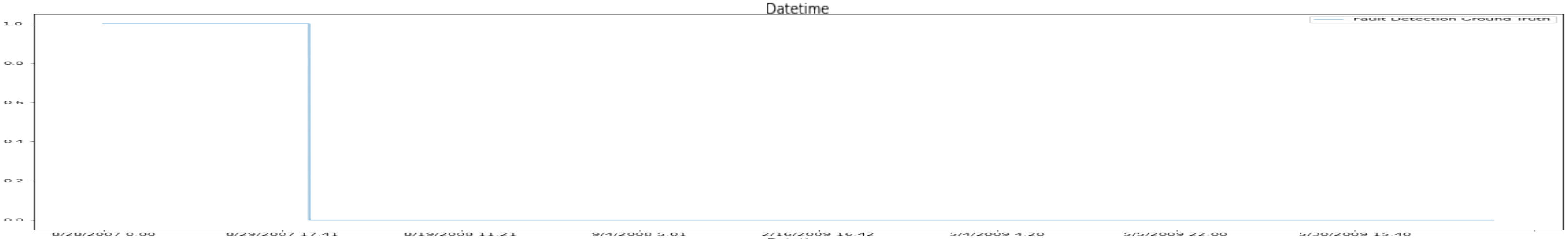




Exploratory Data Analysis

- After outlier removal using elliptic covariance method:

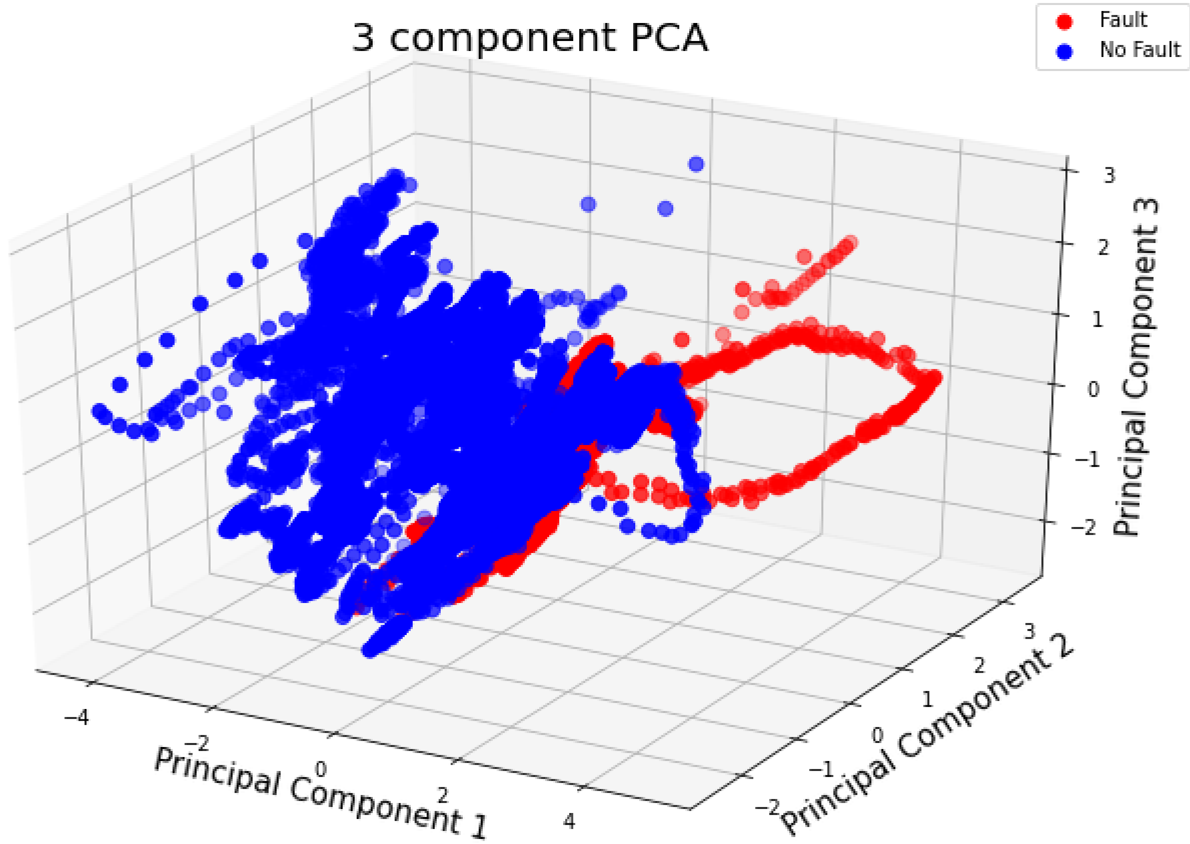
Exploratory Data Analysis



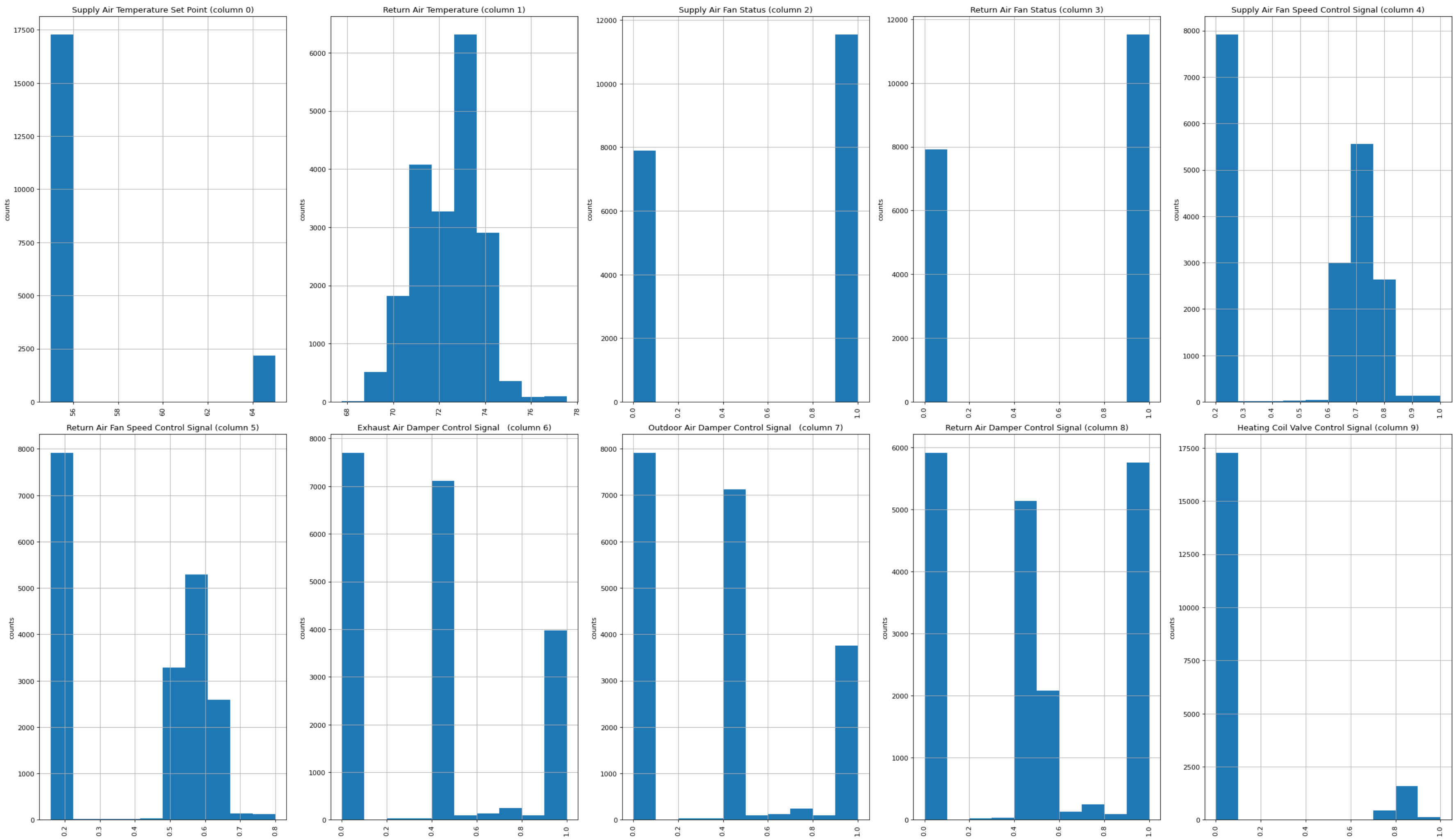
Response data
Fault = 1
No fault = 0

Exploratory Data Analysis

- Implemented 3-component PCA focusing on temperature variables
- Visually, did not find any straightforward covariance relations between principal components

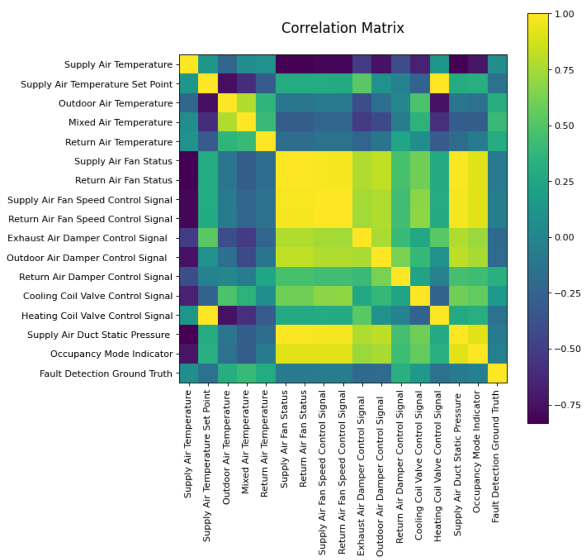


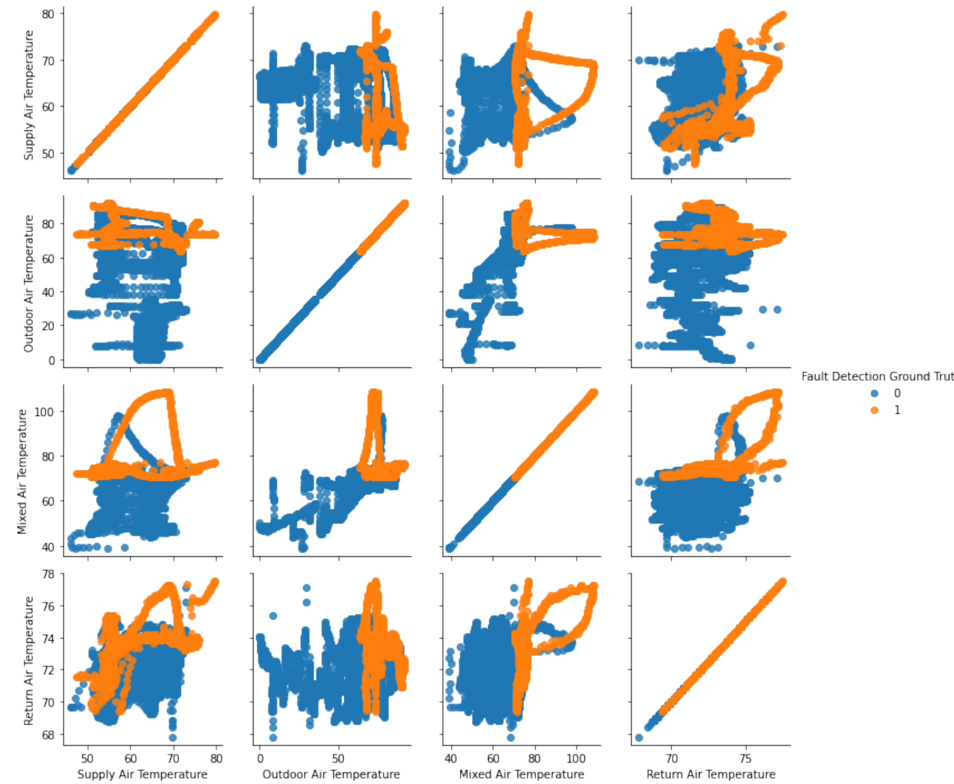
Exploratory Data Analysis



Exploratory Data Analysis

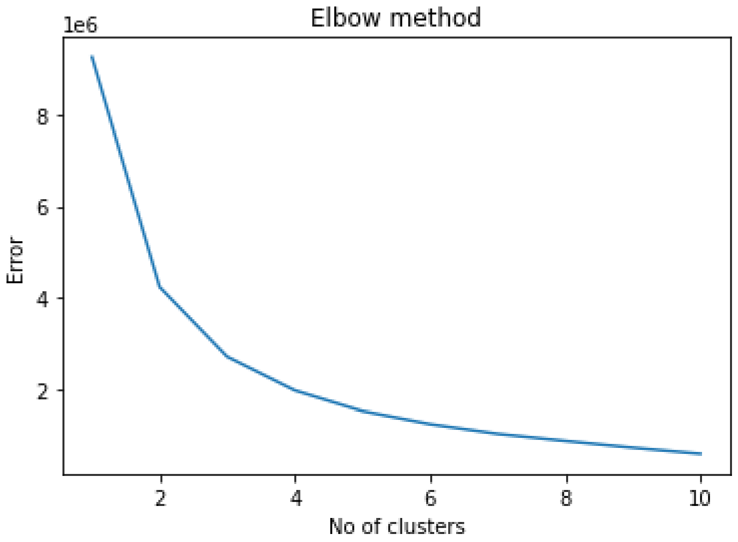
- Cross plots of temperature variables were performed to see relationships and whether there were natural fault clusters
- There were some strong correlations that were intuitive; for example, between mixed air temperature and outdoor air temperature (C = 0.783)

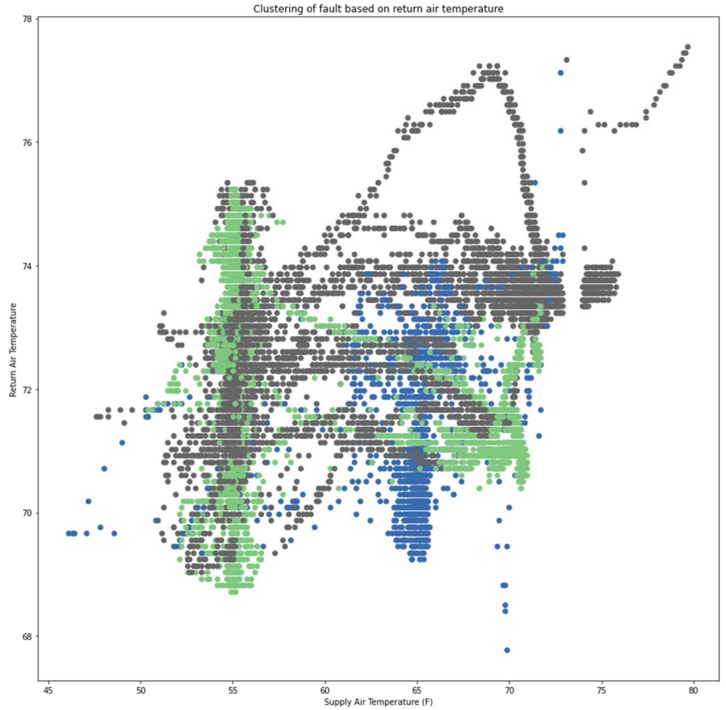


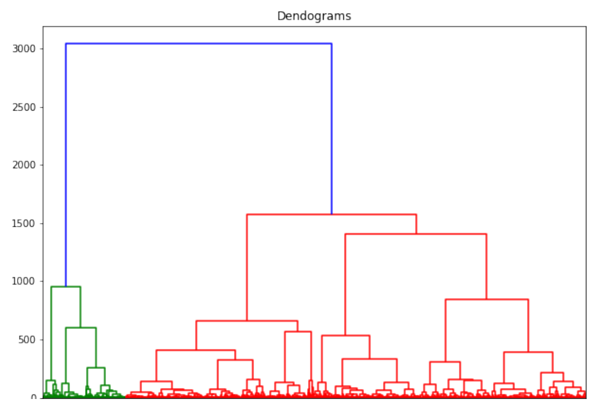


Exploratory Data Analysis

- Elbow method suggested 3 clusters, while dendrograms suggest 4 clusters
- However, clusters do not seem to have any physical significance

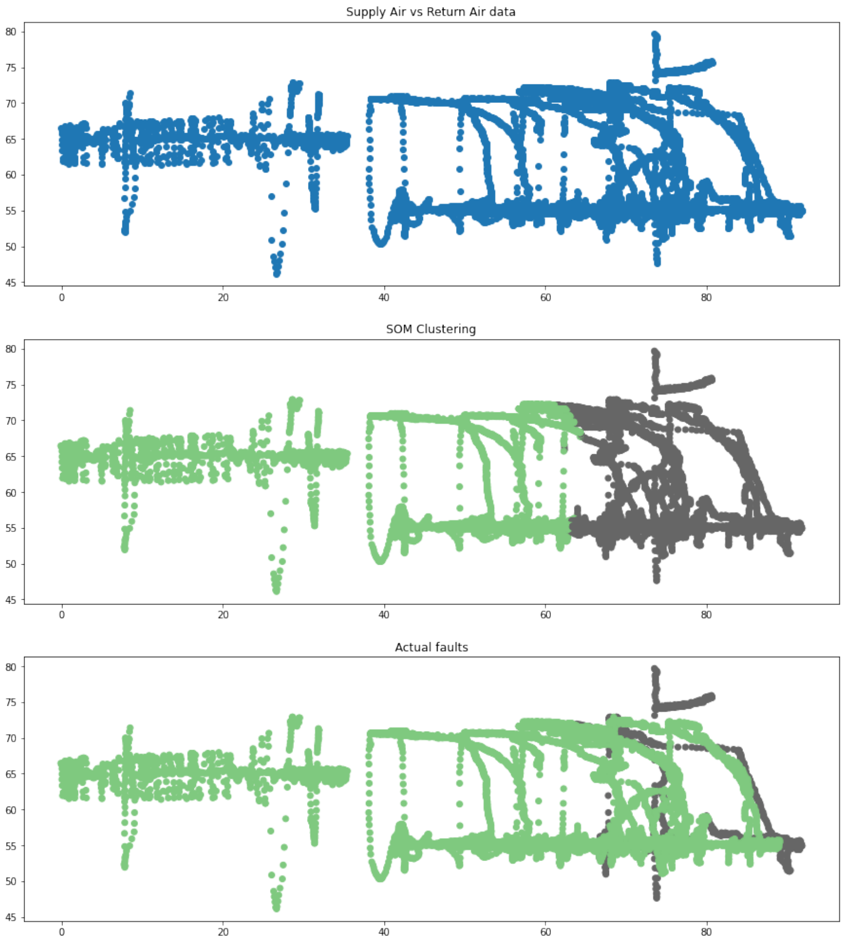




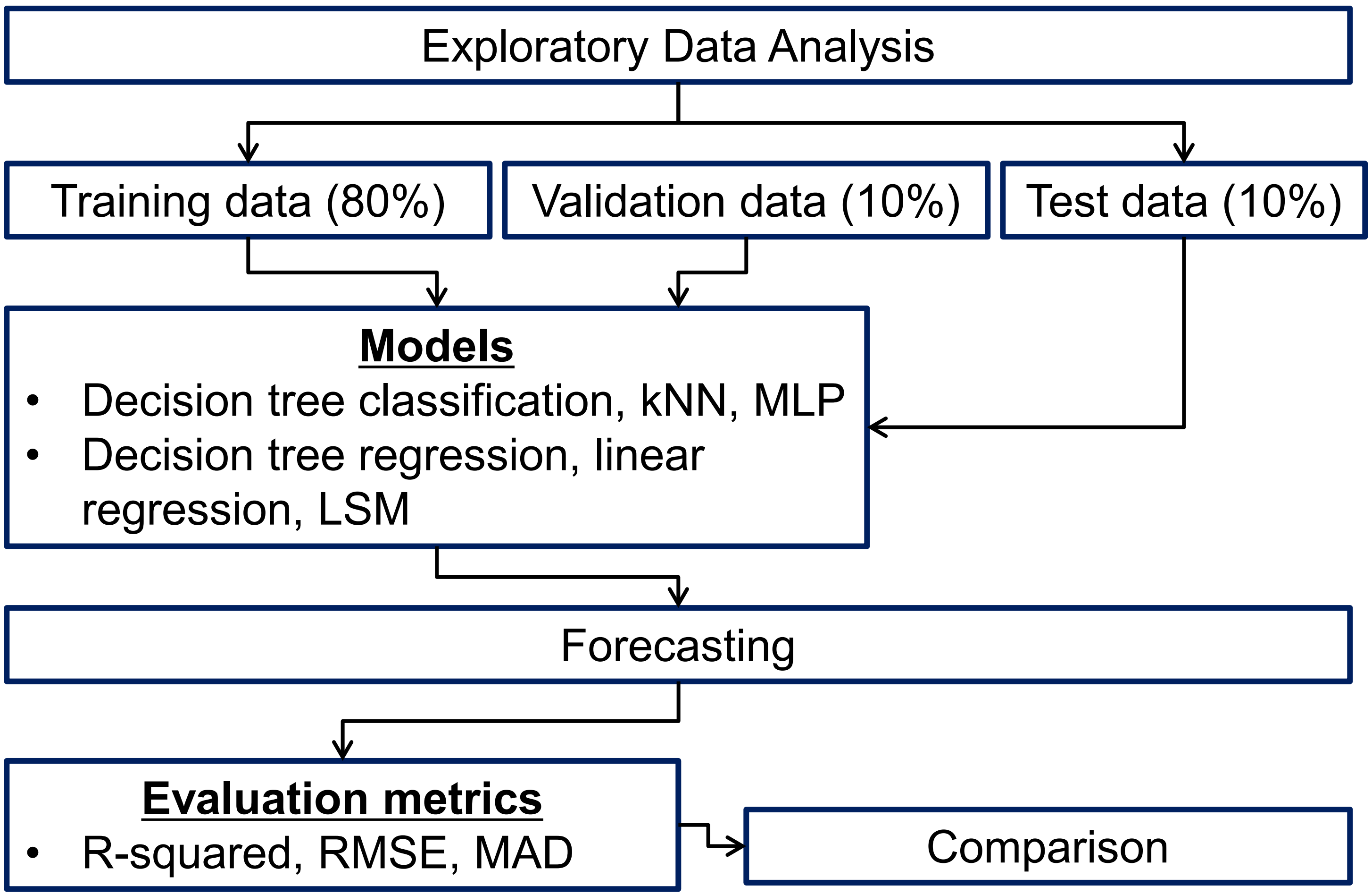


Exploratory Data Analysis

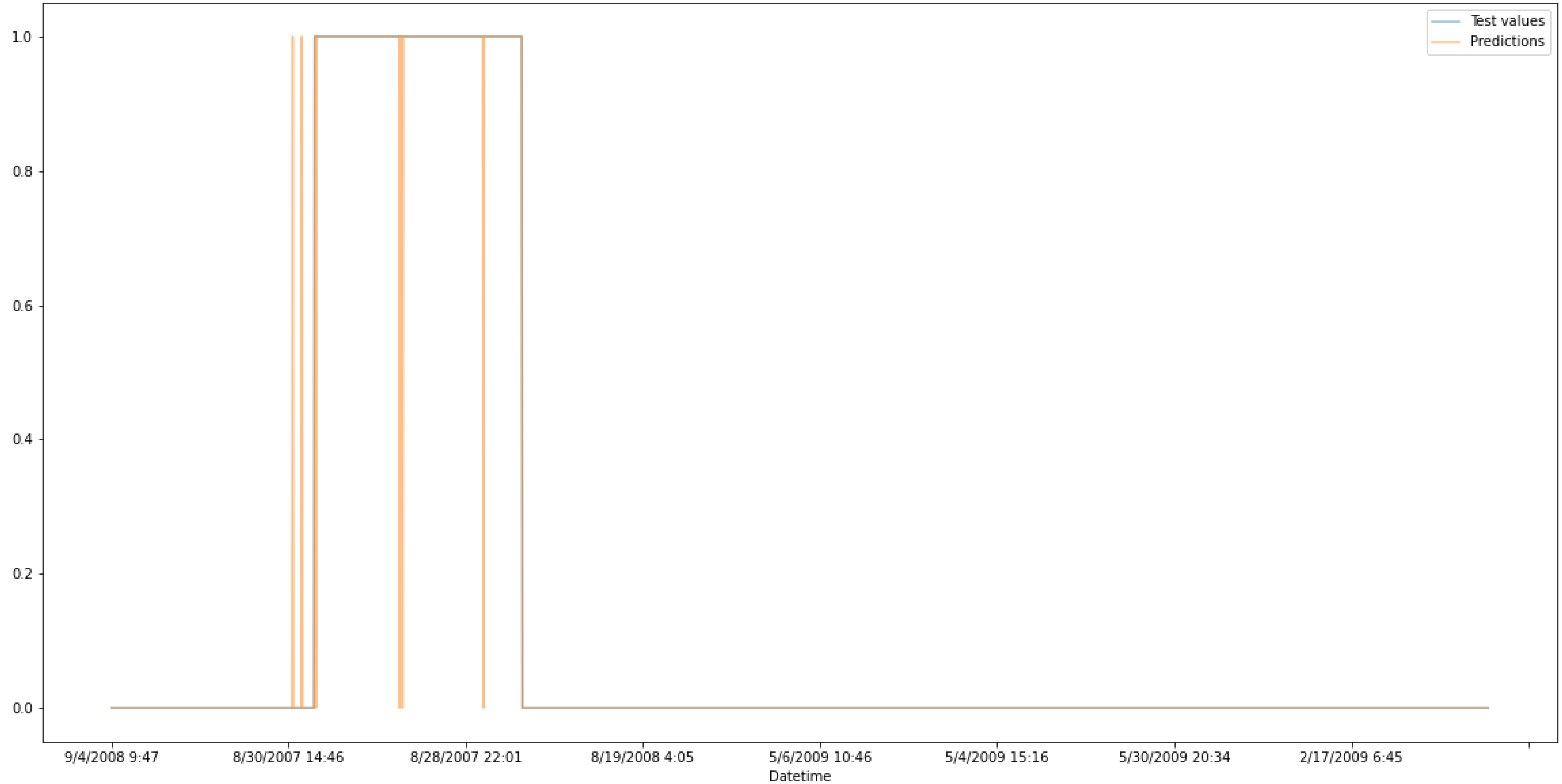
- Self-organizing maps with 2 clusters specified had an accuracy of around 57%
- Actual accuracy was unstable and varied between each run



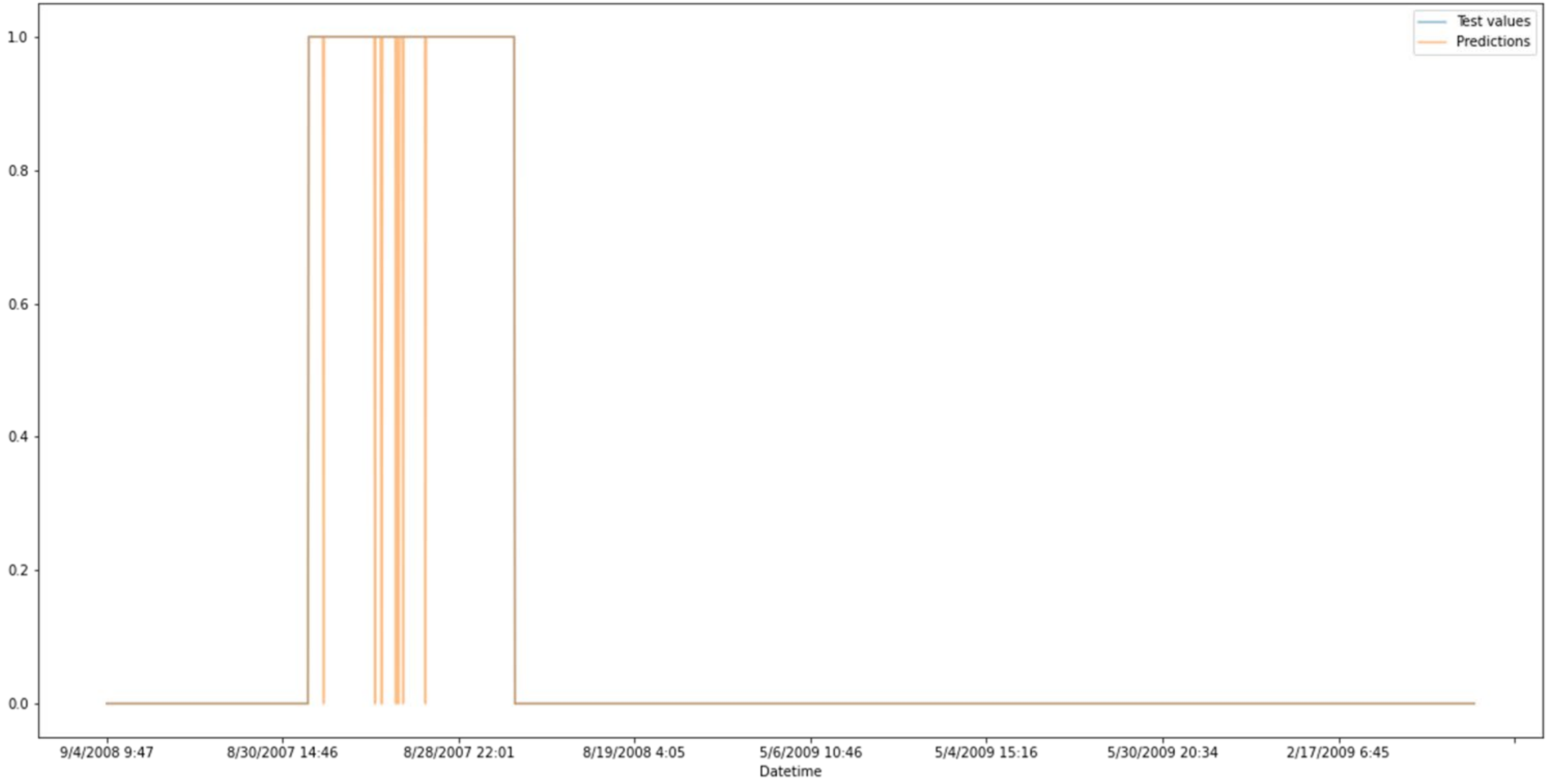
AI/ML Modeling



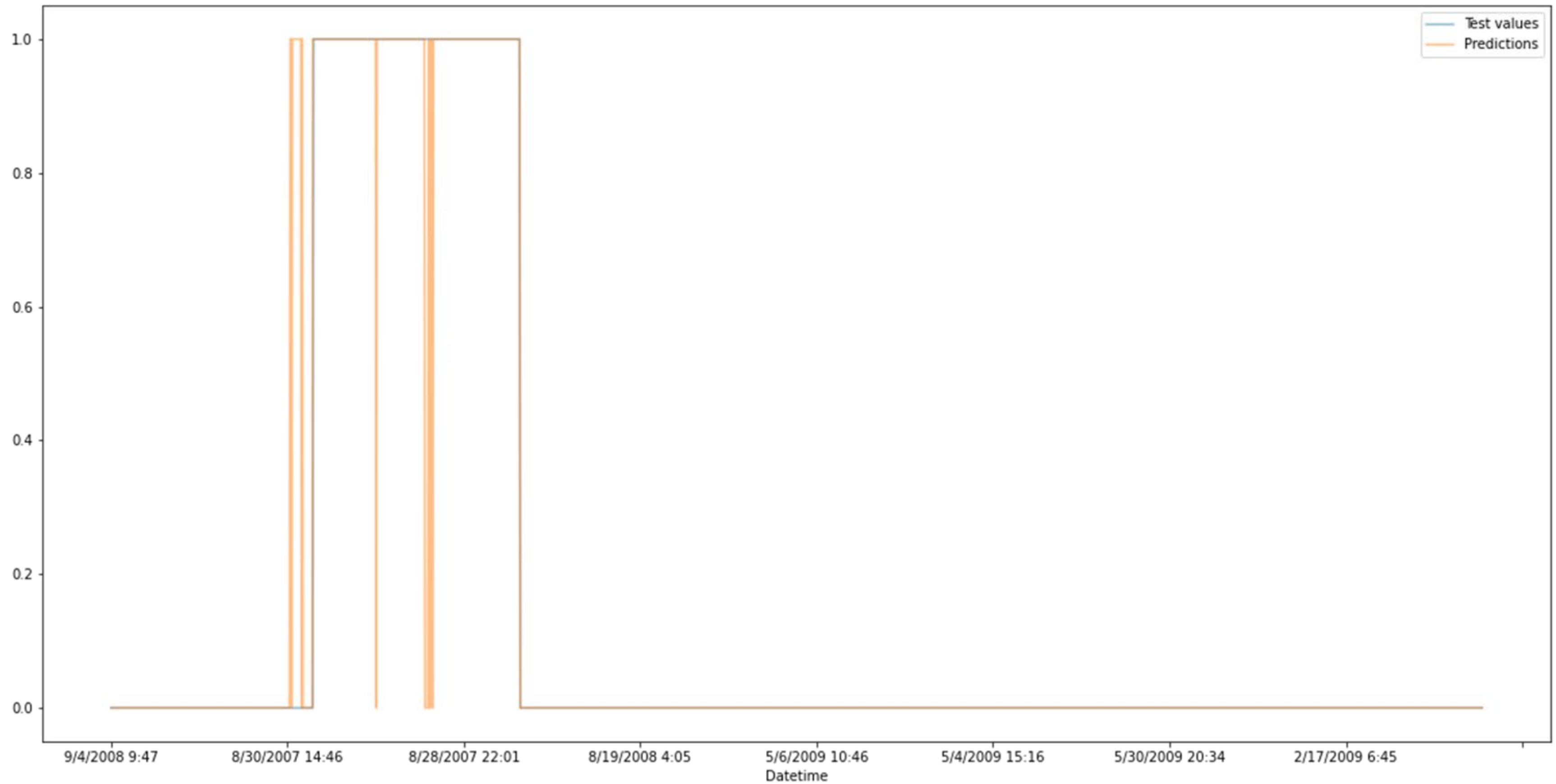
Classification results: decision trees



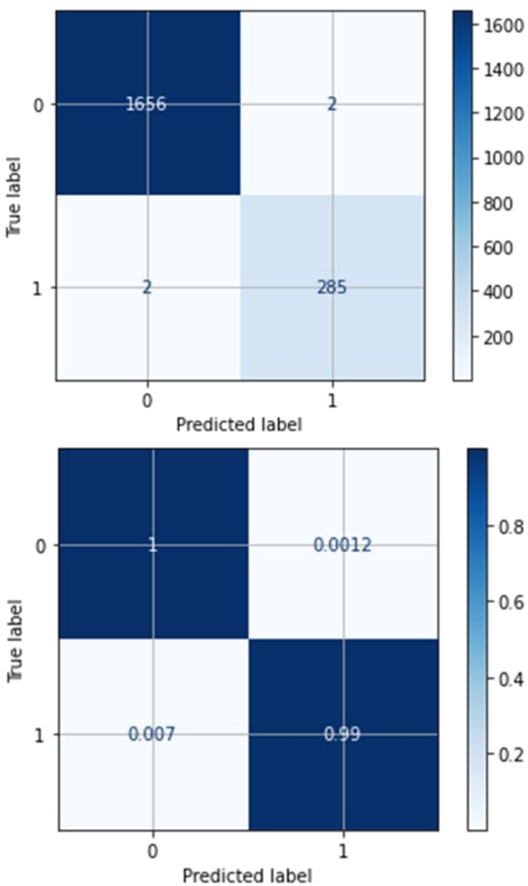
Classification results: kNN



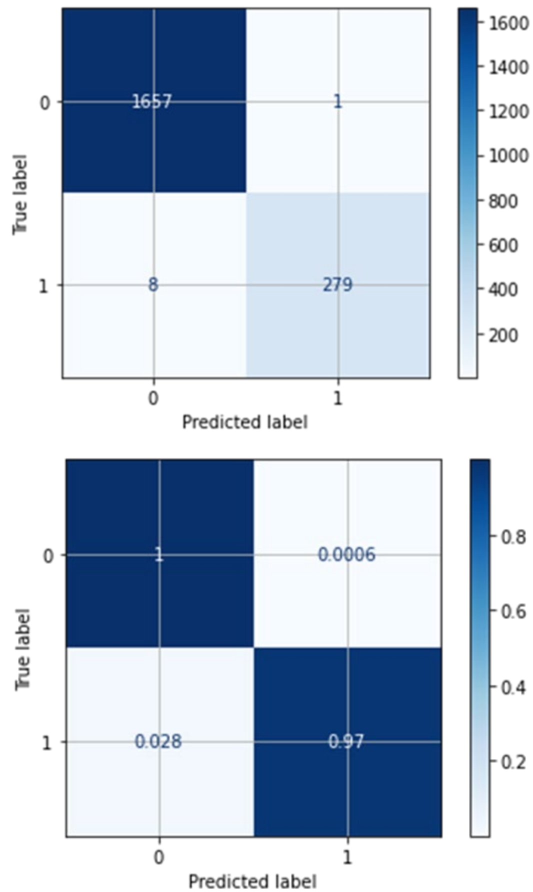
Classification results: MLP ANN



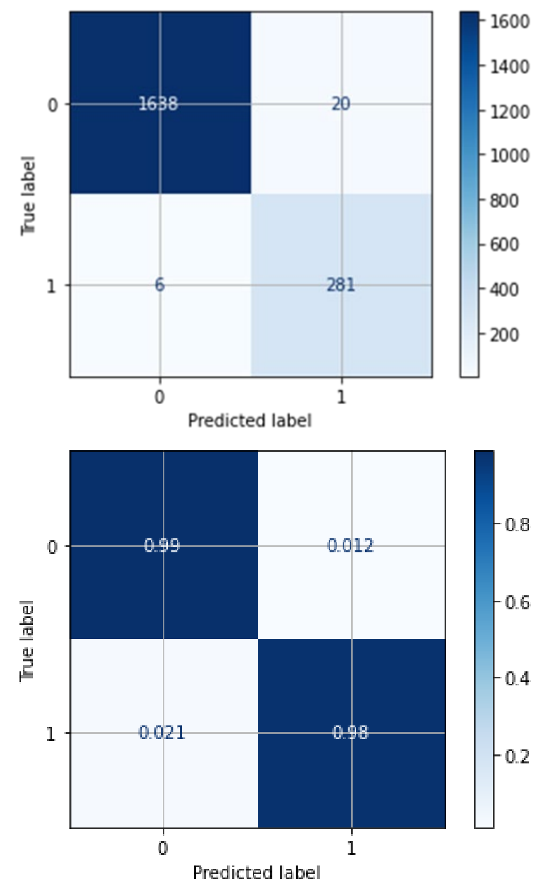
AI/ML Modeling (confusion matrices)



Decision trees
(validation data)

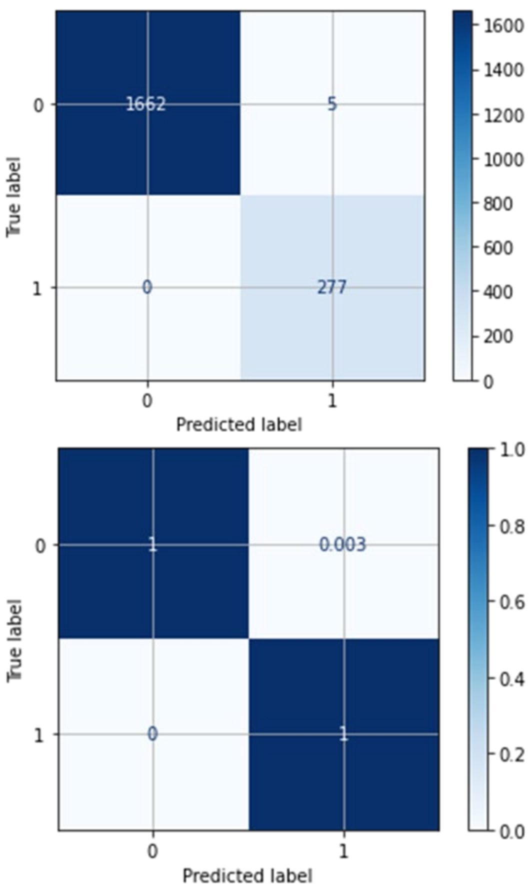


K-nearest neighbors
(validation data)

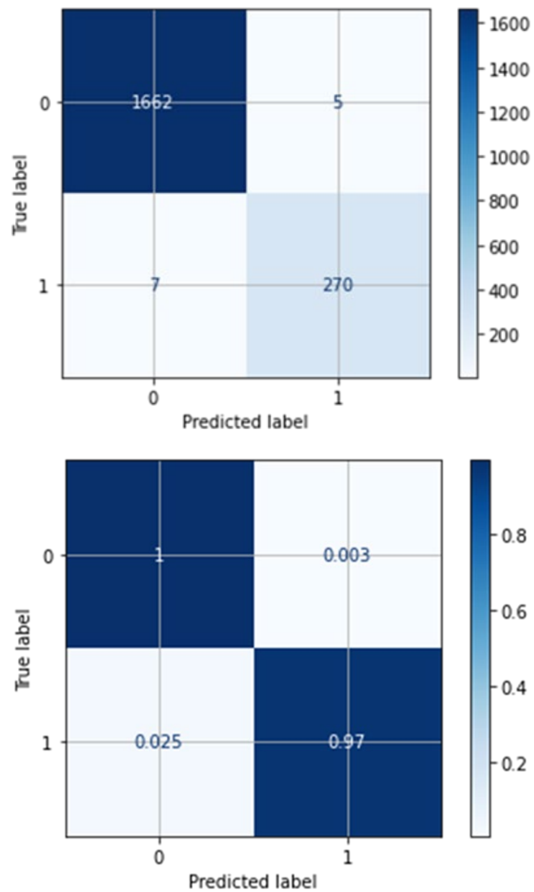


Multi-layer perceptron
(validation data)

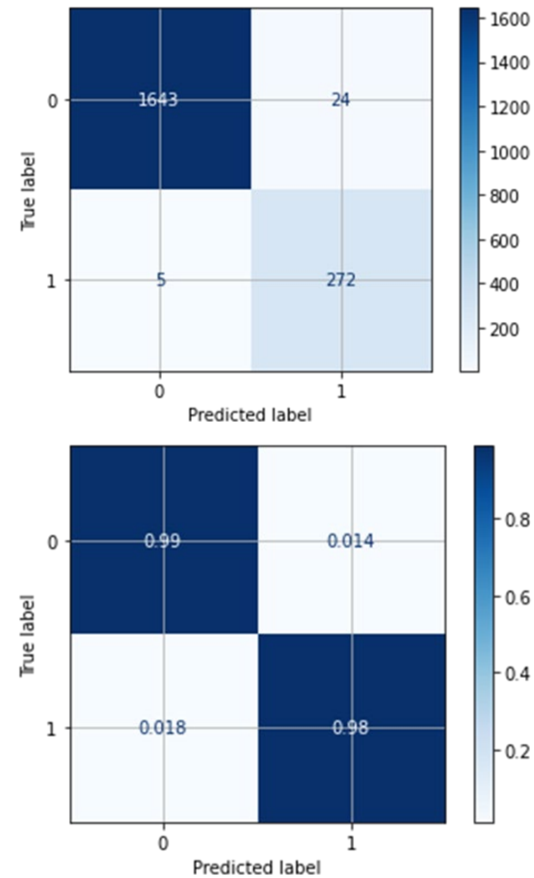
AI/ML Modeling (confusion matrices)



Decision trees
(test data)

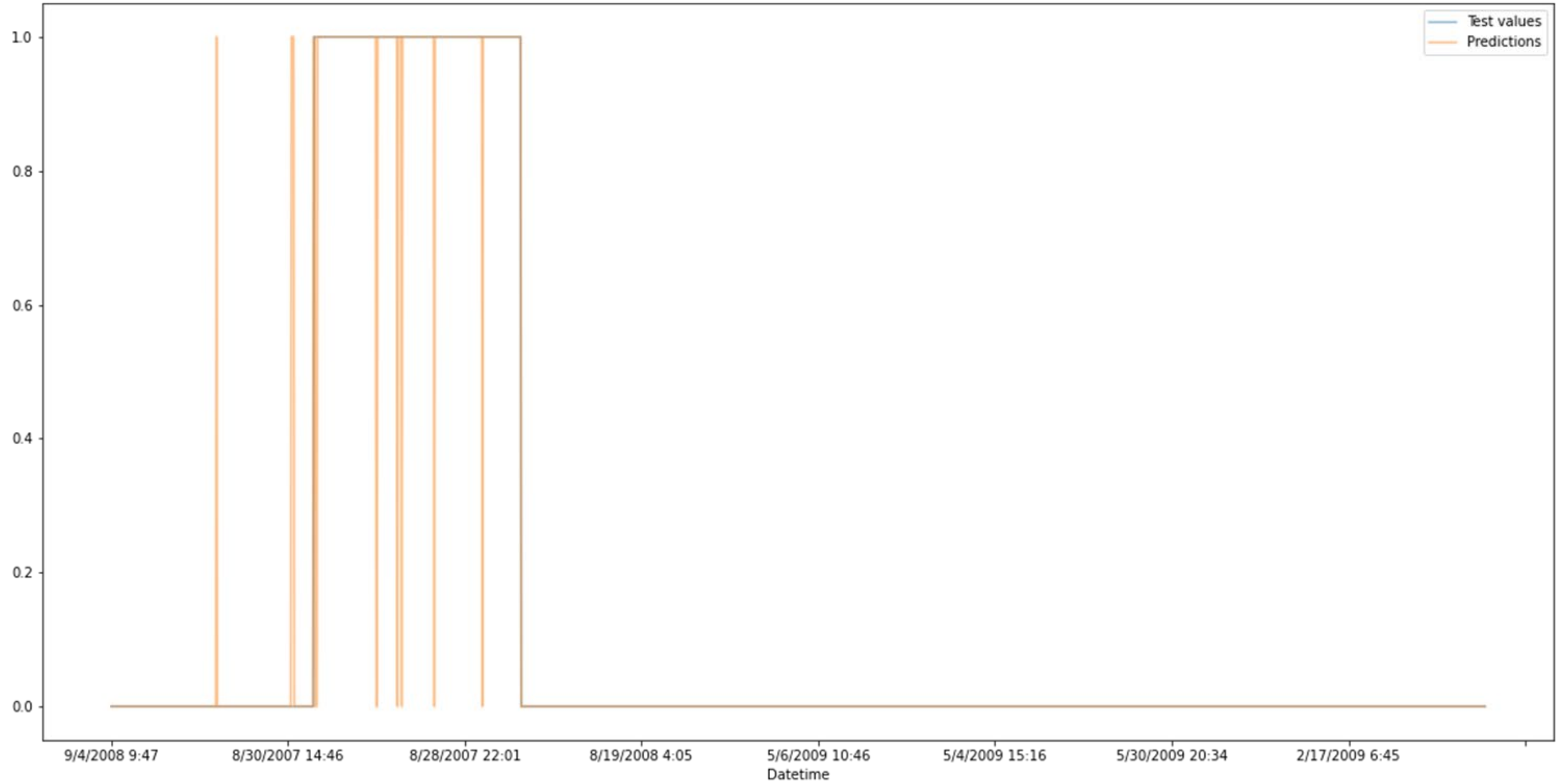


K-nearest neighbors
(test data)

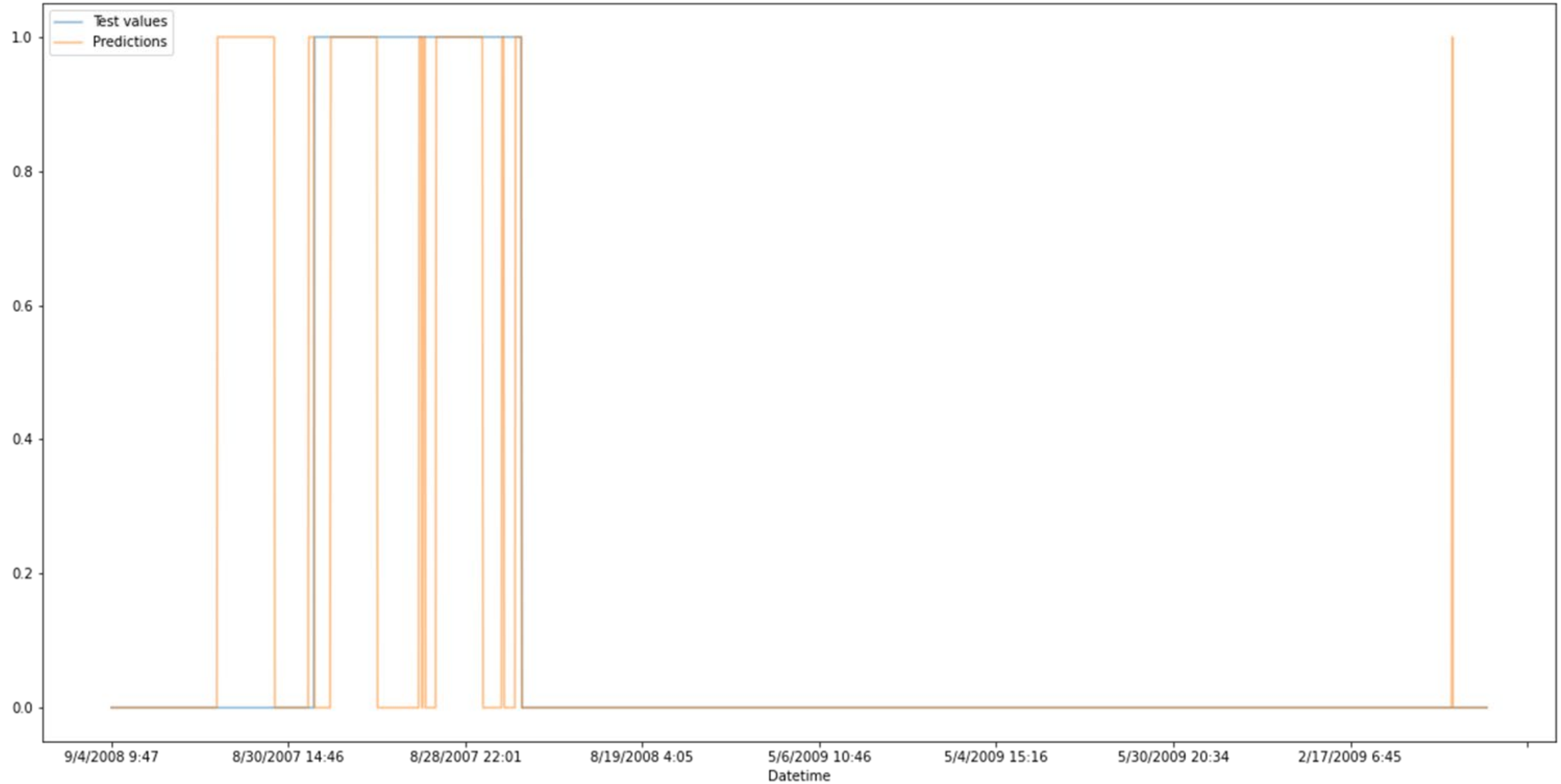


Multi-layer perceptron
(test data)

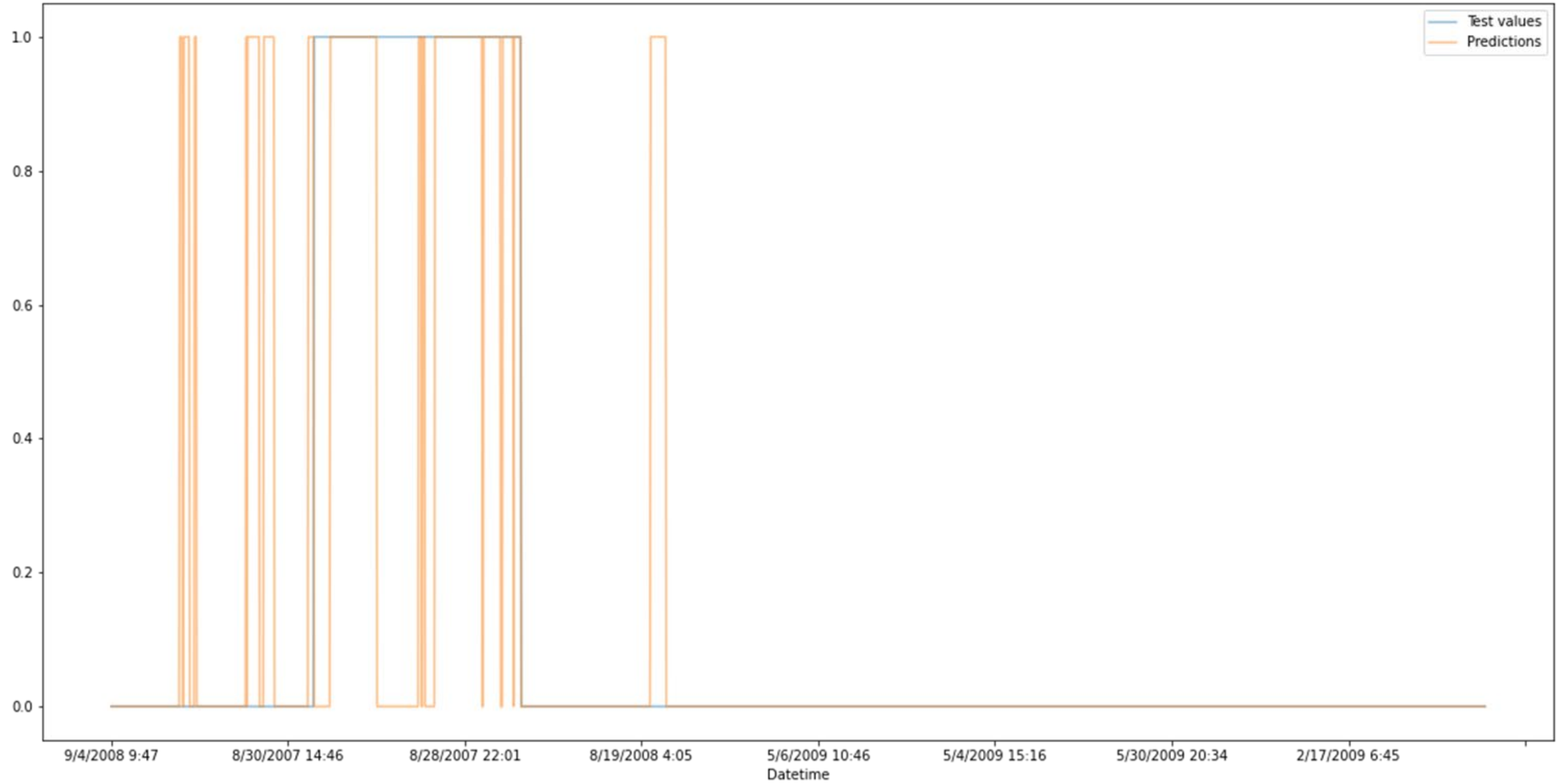
Regression results: decision trees



Regression results: linear regression

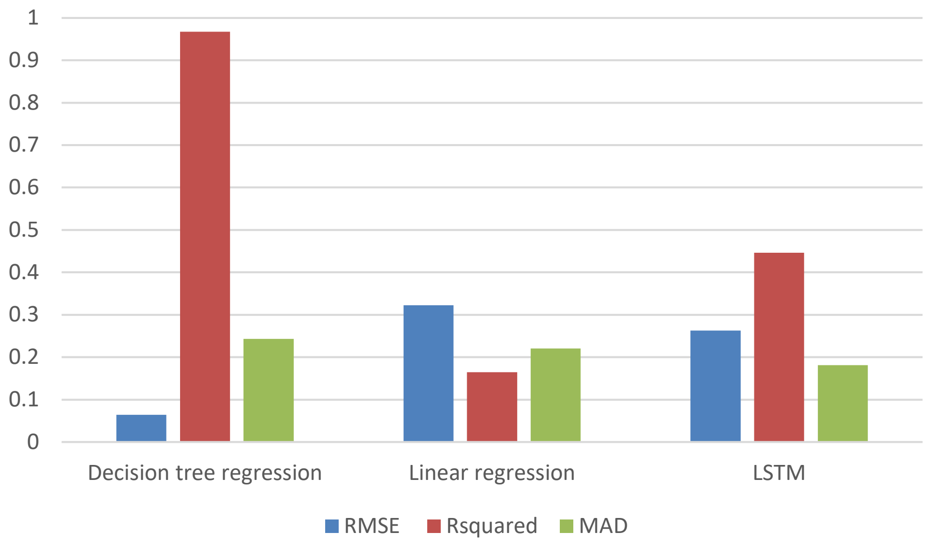


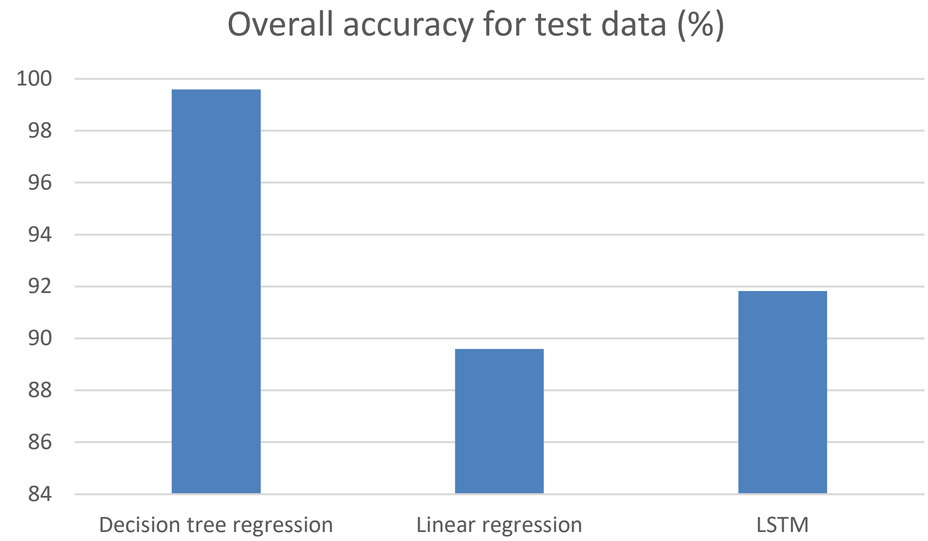
Regression results: LSTMs



AI/ML Modeling

- In regression, decision trees regression performed even better than the LSTM ANN





AI/ML Modeling

- Further testing: try models on simulated data for same system (AHU-A) with more failure modes and “noisier” data

Fault scenarios comparison

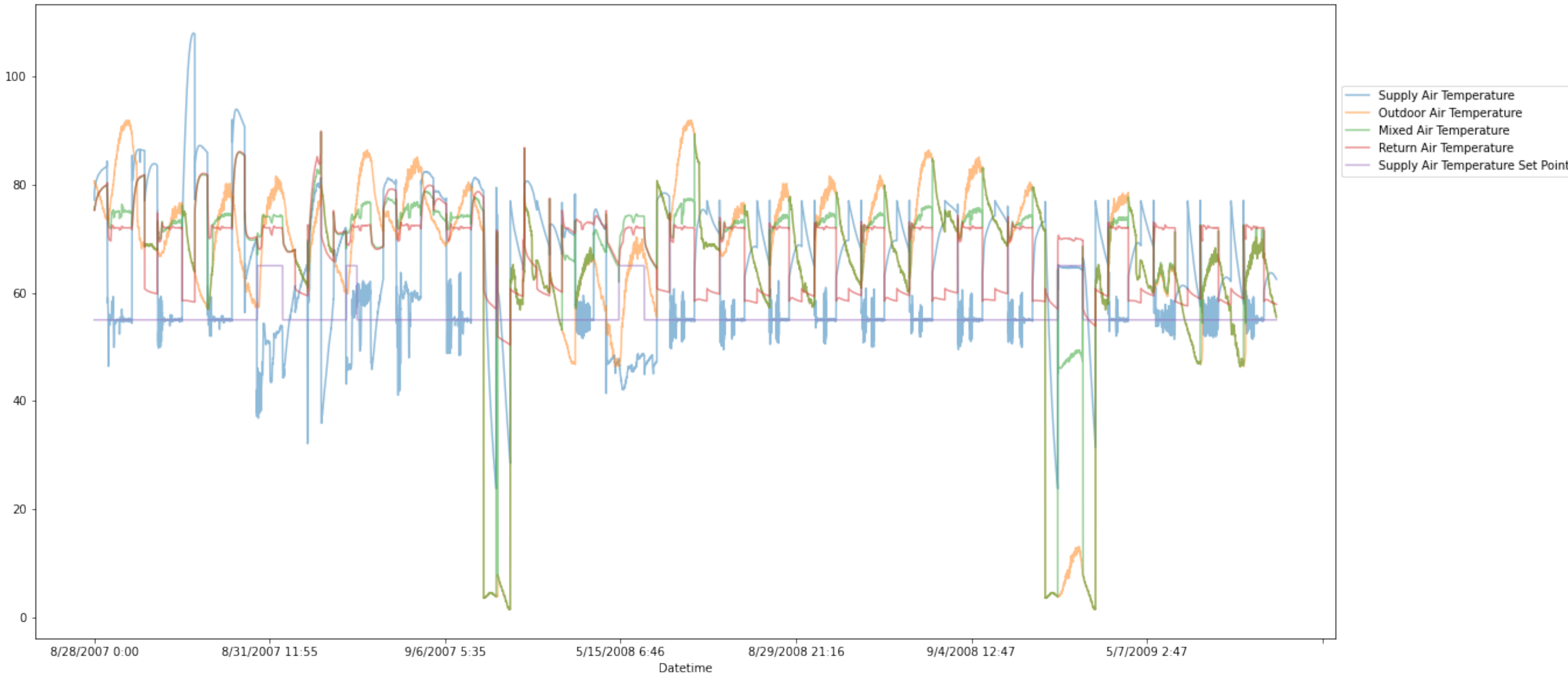
Table 2-2. Experimental input scenarios included in the dataset for MZVAV AHU

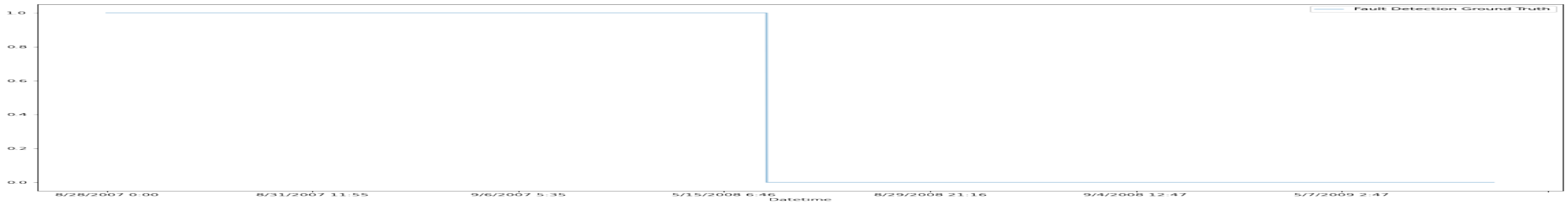
Input Scenarios			Method of fault imposition	Fault occurred time
Fault type		Fault intensity		
Valve of Heating Coil	Leaking	Stage 1: 0.4 GPM	Manually open heating coil bypass valve	8/28/2007
		Stage 2: 1.0 GPM		8/29/2007
		Stage 3: 2.0 GPM		8/30/2007
Unfaulted				8/19/2008 8/25/2008 9/4/2008 1/19/2009 2/16/2009 2/17/2009 5/3/2009 5/4/2009 5/5/2009 5/6/2009

Table 2-2. Simulation Input scenarios included in the dataset for MZVAV AHU

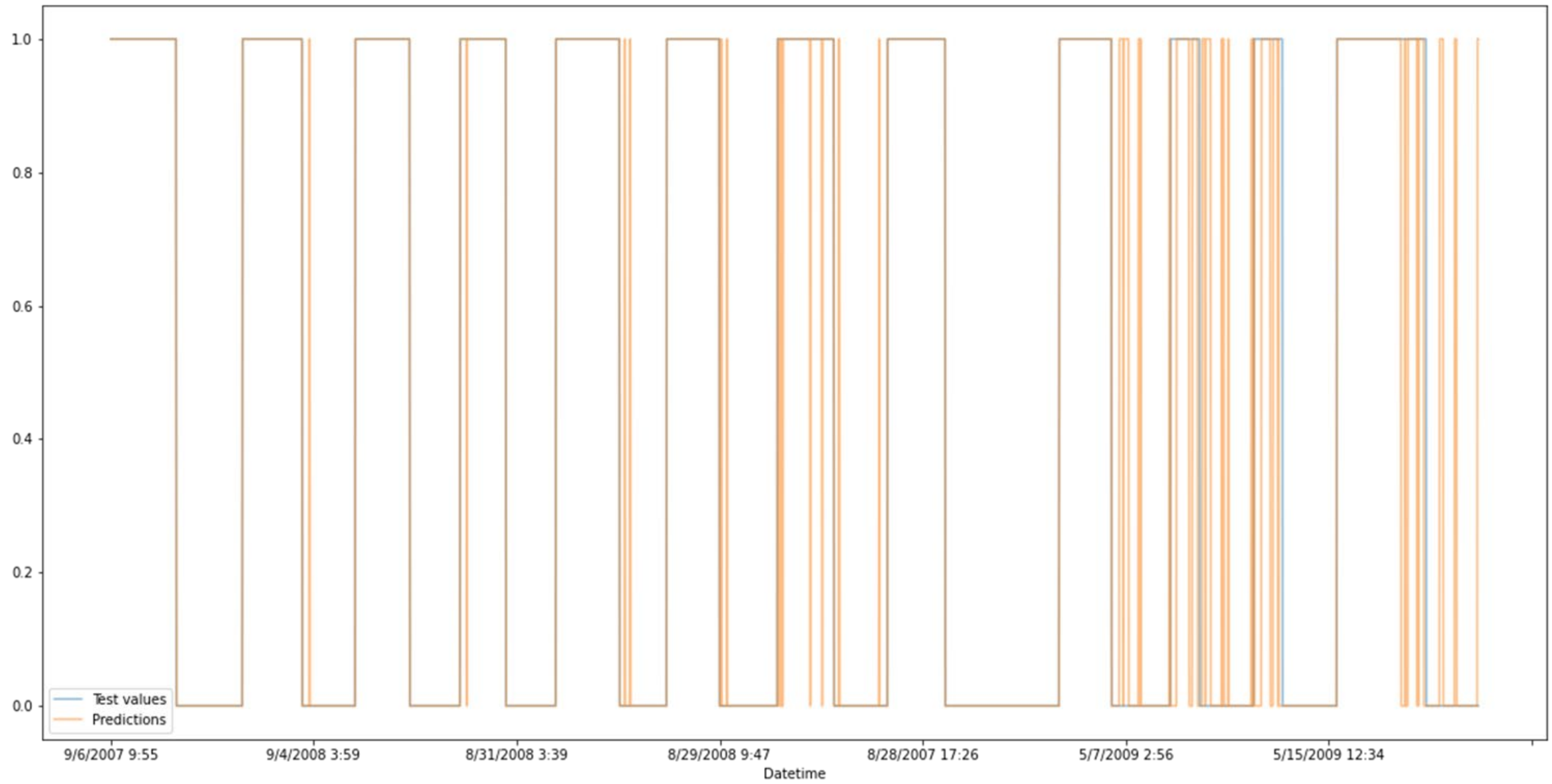
Input Scenarios			Method of fault imposition	Fault occurred time
Fault type		Fault intensity		
OA Damper	Stuck	Fully closed	Automated override of control signal values to indicate that OA damper is stuck.	2/12/2008 5/7/2008
		40% Open		5/8/2008
		45% Open		9/5/2007
		55% Open		9/6/2007
Valve of Heating Coil	Leaking	Stage 1: 0.4 GPM	Manually open heating coil bypass valve	8/28/2007
		Stage 2: 1.0 GPM		8/29/2007
		Stage 3: 2.0 GPM		8/30/2007
Valve of Cooling Coil	Stuck	Fully closed	Automated override of control signal values to indicate that cooing coil valve is stuck.	5/6/2008
		Fully open		8/31/2007 5/15/2008
		Partially open 15%		9/1/2007
		Partially open 65%		9/2/2007
Unfaulted				8/27/2008 8/28/2008 8/29/2008 8/30/2008 8/31/2008 9/1/2008 9/4/2008 9/5/2008 2/11/2009 5/6/2009 5/7/2009 5/8/2009 5/15/2009

AI/ML Modeling

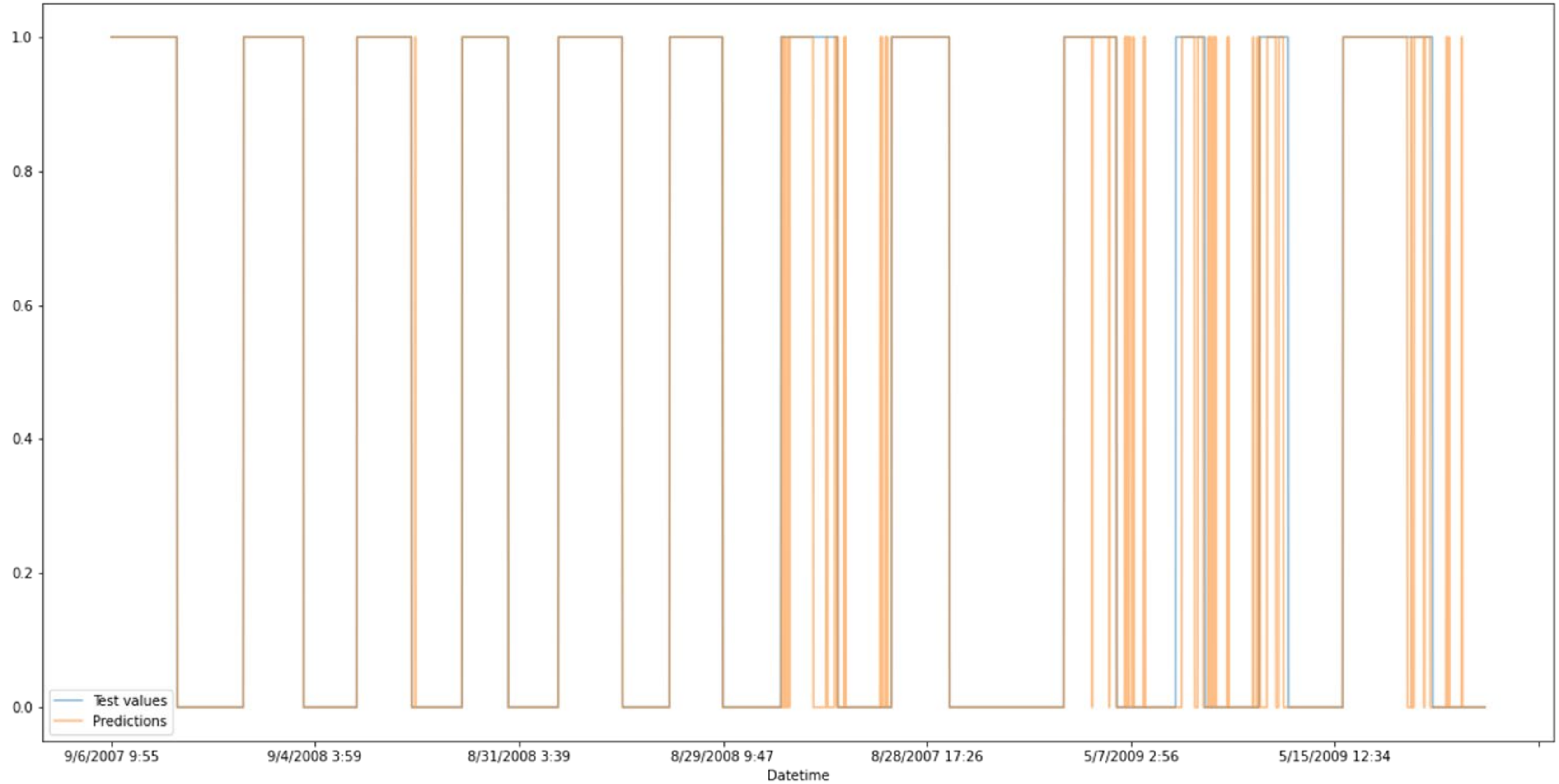




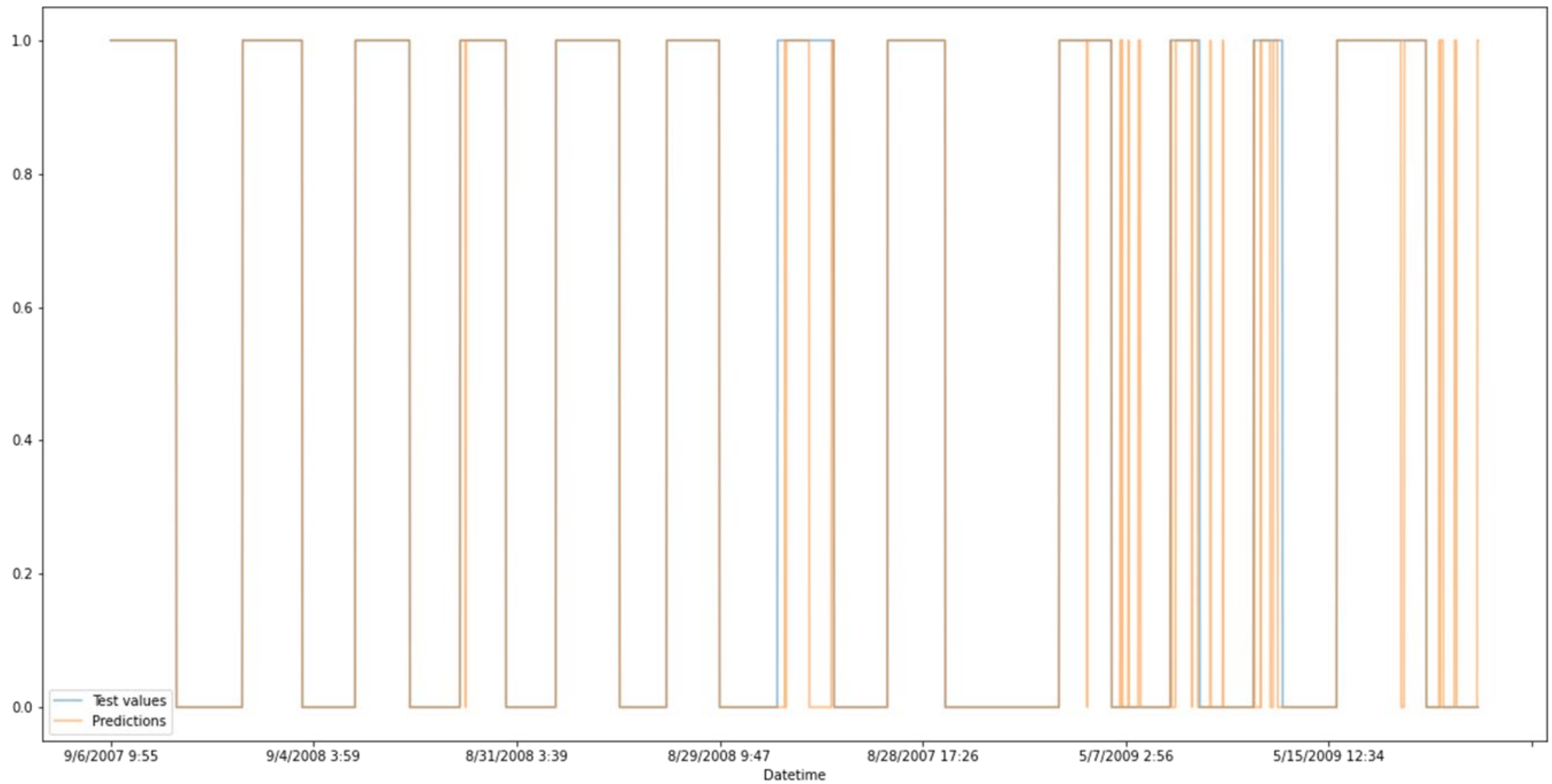
Classification results: decision trees



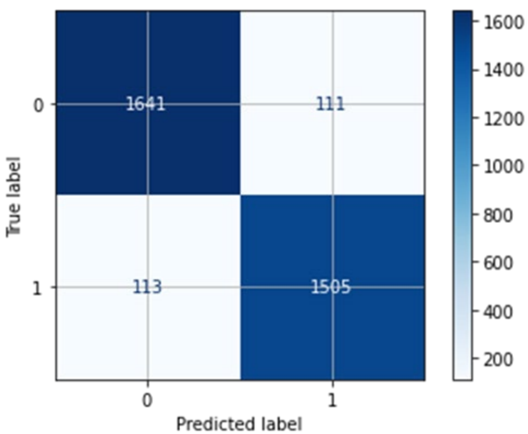
Classification results: kNN

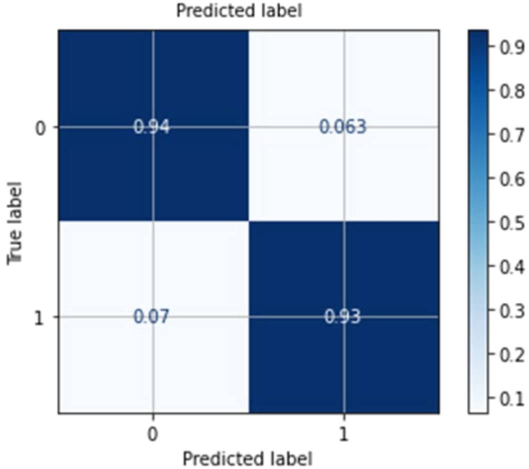


Classification results: MLP ANNs

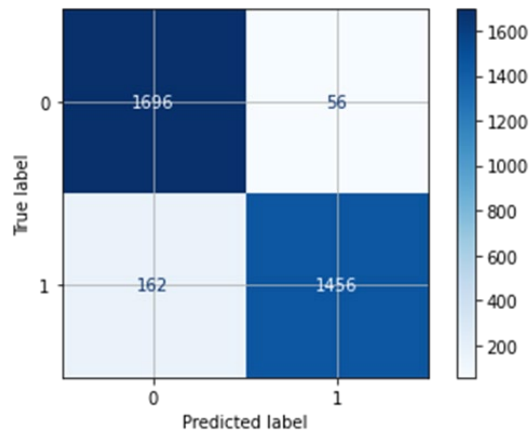


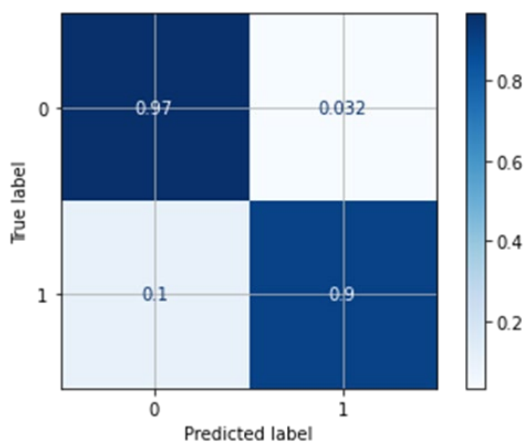
AI/ML Modeling (simulated test data)



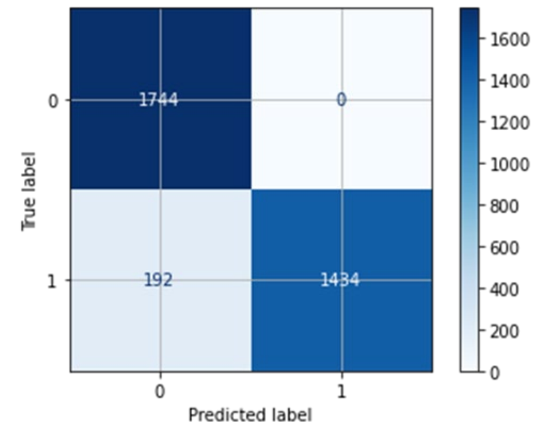


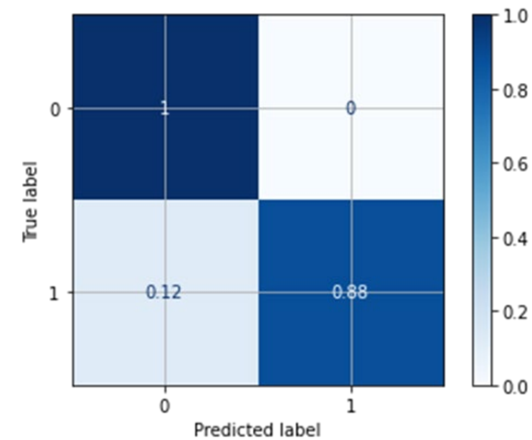
Decision trees
(test data)





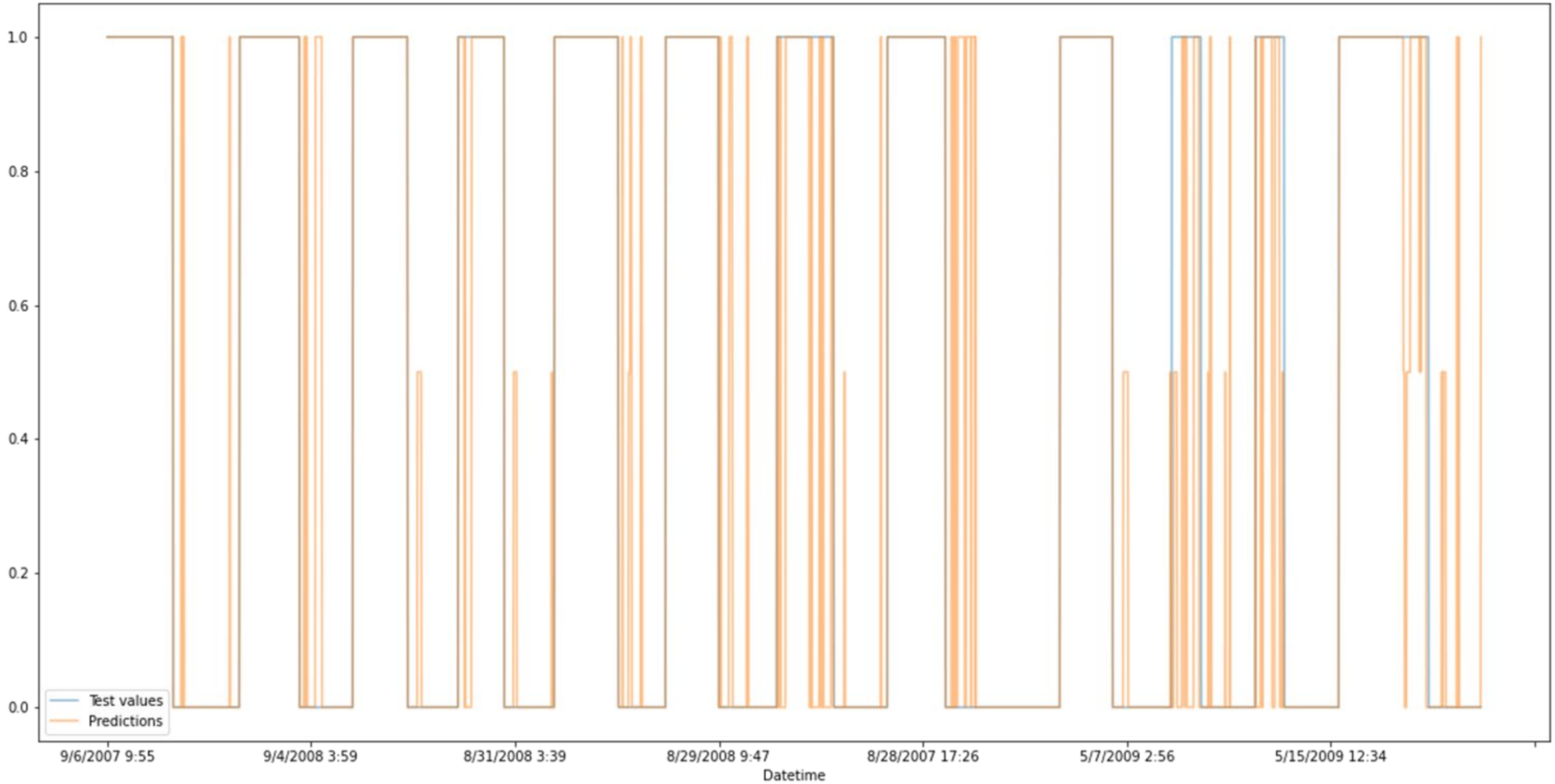
K-nearest neighbors
(test data)



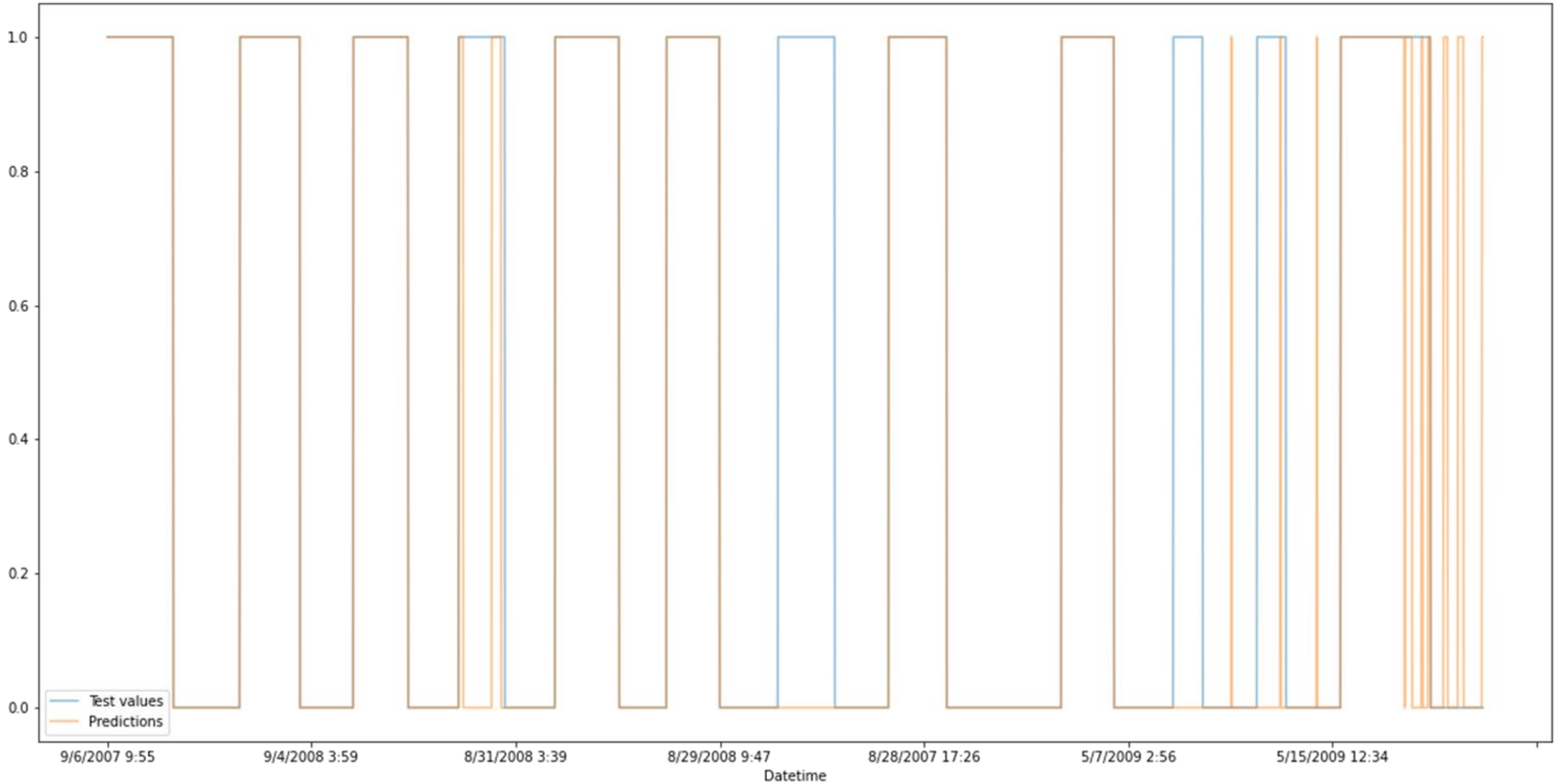


Multi-layer perceptron
(test data)

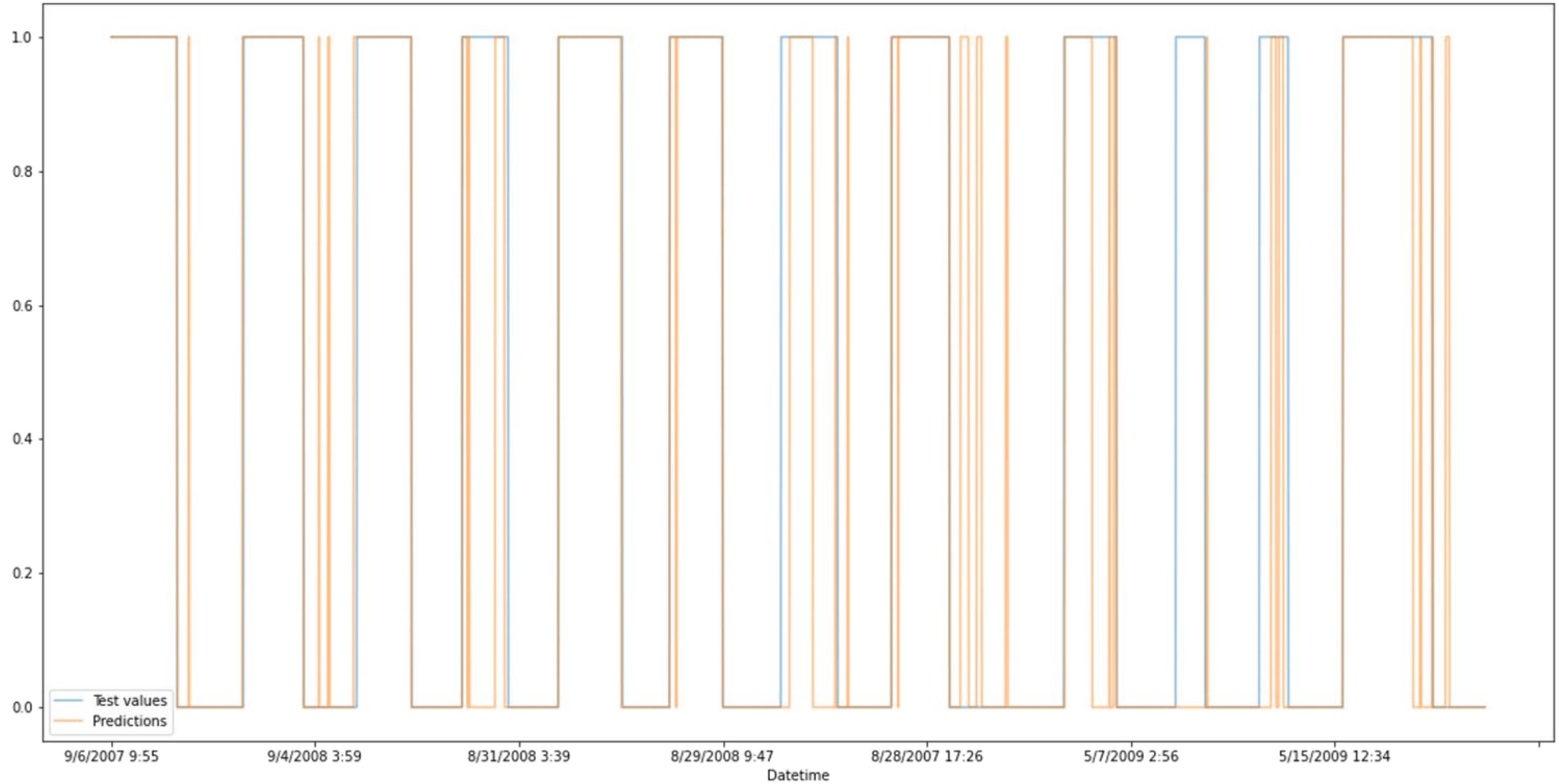
Regression results: decision trees



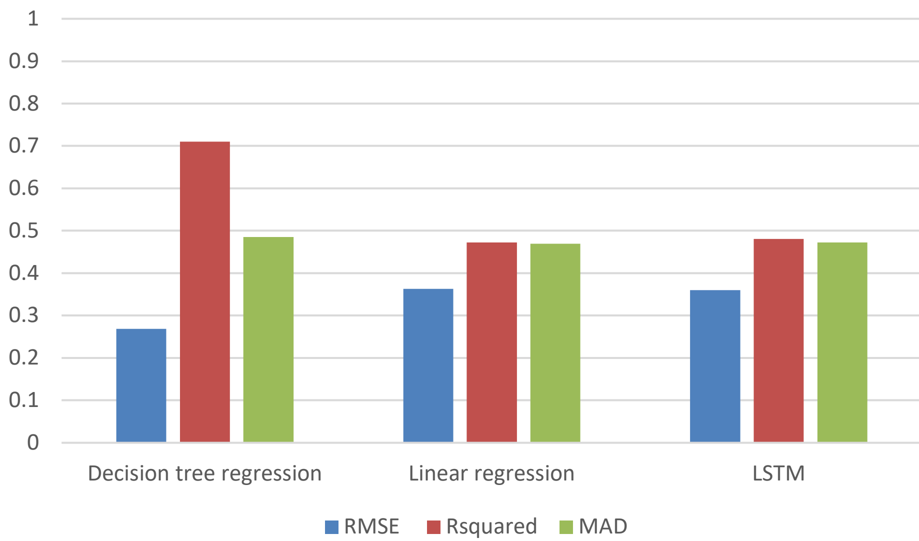
Regression results: linear regression

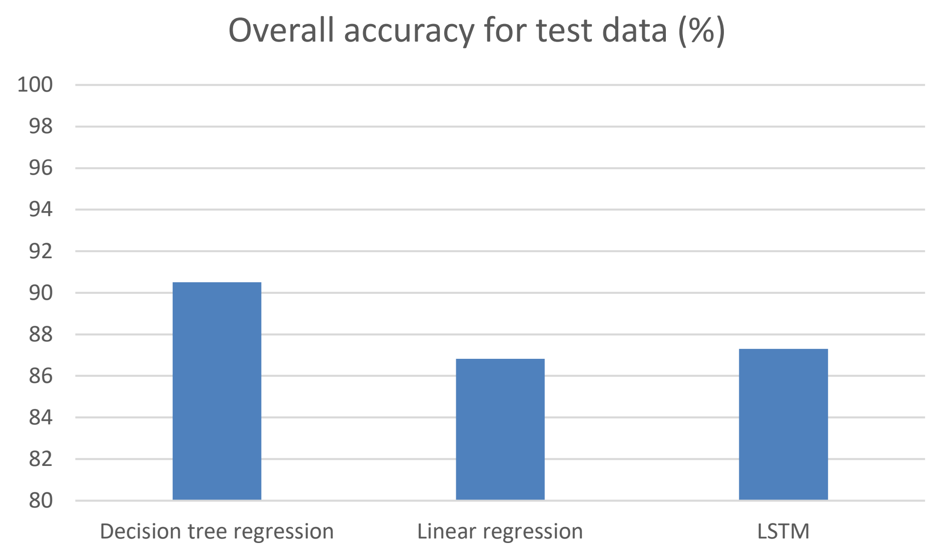


Regression results: LSTM ANN



AI/ML Modeling (simulated test data)





Conclusions

- Unsupervised learning methods did not yield significant insight into the dataset
- Supervised learning techniques (especially classification) yielded accurate outputs
- In both classification and regression, decision-tree based methods performed excellently, likely due to similarity with conditional programming
- Chosen inputs are sufficient to accurately determine faults
- In practice, the model will be fed data from AHU operations from the BMS via BACNET, and a prediction will be made for each timestamp.

Recommendations

- Focus on classification and regression methods for future analysis
- Attempt ensemble learning
- Explore more neural network options in place of weaker methods (e.g. MLP regression in place of linear regression)
- Non-neural network methods may be preferable where there is no significant accuracy increase (less time usage)
- Extend AI model to discriminate between specific types of faults and fault intensity instead of just fault/no fault
- Tweak parameters of LSTM model for greater accuracy
- Apply algorithm to a more diverse experimental dataset

References

- Du, Zhimin, Bo Fan, Xinqiao Jin, and Jinlei Chi. "Fault Detection and Diagnosis for Buildings and Hvac Systems Using Combined Neural Networks and Subtractive Clustering Analysis." *Building and Environment* 73 (2014/03/01/ 2014): 1-11. https://doi.org/https://doi.org/10.1016/j.buildenv.2013.11.021. https://www.sciencedirect.com/science/article/pii/S0360132313003417.
- Du, Zhimin, Xinqiao Jin, and Yunyu Yang. "Fault Diagnosis for Temperature, Flow Rate and Pressure Sensors in Vav Systems Using Wavelet Neural Network." *Applied Energy* 86, no. 9 (2009/09/01/ 2009): 1624-31. https://doi.org/https://doi.org/10.1016/j.apenergy.2009.01.015. https://www.sciencedirect.com/science/article/pii/S0306261909000233.
- Li, Shun, and Jin Wen. "Application of Pattern Matching Method for Detecting Faults in Air Handling Unit System." *Automation in Construction* 43 (2014/07/01/ 2014): 49-58. https://doi.org/https://doi.org/10.1016/j.autcon.2014.03.002. https://www.sciencedirect.com/science/article/pii/S0926580514000545.
- Lin, Guanjing, Hannah Kramer, and Jessica Granderson. "Building Fault Detection and Diagnostics: Achieved Savings, and Methods to Evaluate Algorithm Performance." *Building and Environment* 168 (2020/01/15/ 2020): 106505. https://doi.org/https://doi.org/10.1016/j.buildenv.2019.106505. https://www.sciencedirect.com/science/article/pii/S0360132319307176.
- Pourarian, Shokouh, Jin Wen, Daniel Veronica, Amanda Pertzborn, Xiaohui Zhou, and Ran Liu. "A Tool for Evaluating Fault Detection and Diagnostic Methods for Fan Coil Units." *Energy and Buildings* 136 (2017/02/01/ 2017): 151-60. https://doi.org/https://doi.org/10.1016/j.enbuild.2016.12.018. https://www.sciencedirect.com/science/article/pii/S0378778816317698.
- Yan, Rui, Zhenjun Ma, Yang Zhao, and Georgios Kokogiannakis. "A Decision Tree Based Data-Driven Diagnostic Strategy for Air Handling Units." *Energy and Buildings* 133 (2016/12/01/ 2016): 37-45. https://doi.org/https://doi.org/10.1016/j.enbuild.2016.09.039. https://www.sciencedirect.com/science/article/pii/S0378778816308854.

References (continued)

- Yu, Bing, Dolf H. C. Van Paassen, and Siamak Riahy. "General Modeling for Model-Based Fdd on Building Hvac System." *Simulation Practice and Theory* 9, no. 6 (2002/05/15/ 2002): 387-97. https://doi.org/https://doi.org/10.1016/S1569-190X(02)00062-X. https://www.sciencedirect.com/science/article/pii/S1569190X0200062X.
- Zhao, Yang, Tingting Li, Xuejun Zhang, and Chaobo Zhang. "Artificial Intelligence-Based Fault Detection and Diagnosis Methods for Building Energy Systems: Advantages, Challenges and the Future." *Renewable and Sustainable Energy Reviews* 109 (2019/07/01/ 2019): 85-101. https://doi.org/https://doi.org/10.1016/j.rser.2019.04.021. https://www.sciencedirect.com/science/article/pii/S1364032119302333.
- Zhao, Yang, Jin Wen, and Shengwei Wang. "Diagnostic Bayesian Networks for Diagnosing Air Handling Units Faults – Part Ii: Faults in Coils and Sensors." *Applied Thermal Engineering* 90 (2015/11/05/ 2015): 145-57. https://doi.org/https://doi.org/10.1016/j.applthermaleng.2015.07.001. https://www.sciencedirect.com/science/article/pii/S1359431115006584.
- Zhao, Yang, Jin Wen, Fu Xiao, Xuebin Yang, and Shengwei Wang. "Diagnostic Bayesian Networks for Diagnosing Air Handling Units Faults – Part I: Faults in Dampers, Fans, Filters and Sensors." *Applied Thermal Engineering* 111 (2017/01/25/ 2017): 1272-86. https://doi.org/https://doi.org/10.1016/j.applthermaleng.2015.09.121. https://www.sciencedirect.com/science/article/pii/S1359431116301193.
- Granderson, Jessica, and Guanjing Lin. 2019. "Inventory Of Data Sets For AFDD Evaluation". Lawrence Berkeley National LIbrary. https://data.openei.org/submissions/910.