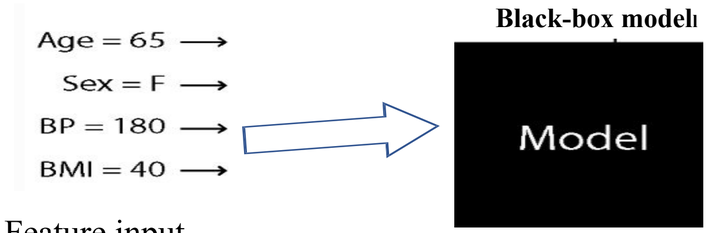
THE UNIVERSITY OF BRITISH COLUMBIA

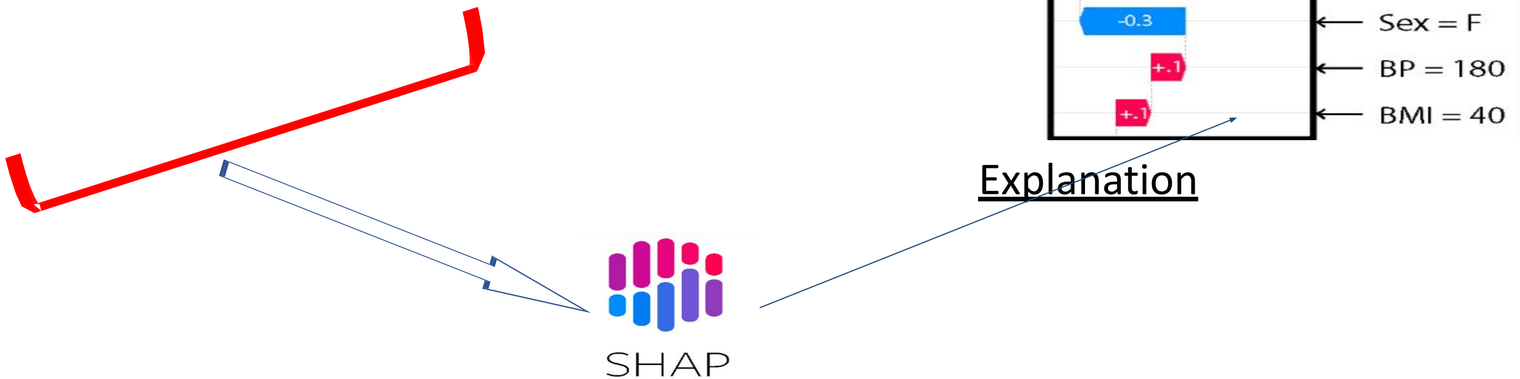
[redacted]

Ria Ji

Interpretable Soft Sensors using Extremely Randomized Trees and SHAP

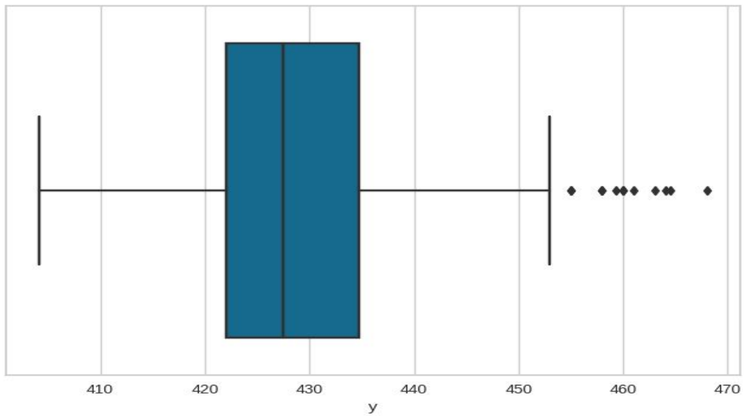
What is Shap(SHapley Additive exPlannations)?

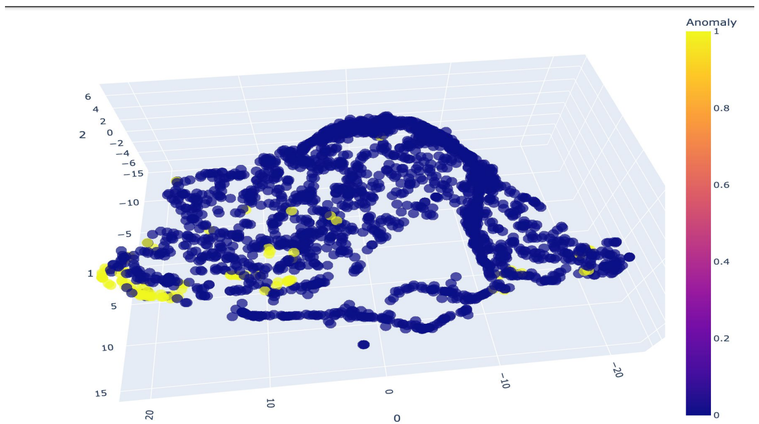




The impacts of this project is the ability to provide insights into the relationship between the input features and the target variable. This information can help in understanding the underlying process and can be used for further improvements. The model can also be used for monitoring and control purposes, for example, detecting anomalies or predicting outcomes in real-time.

data pre-process

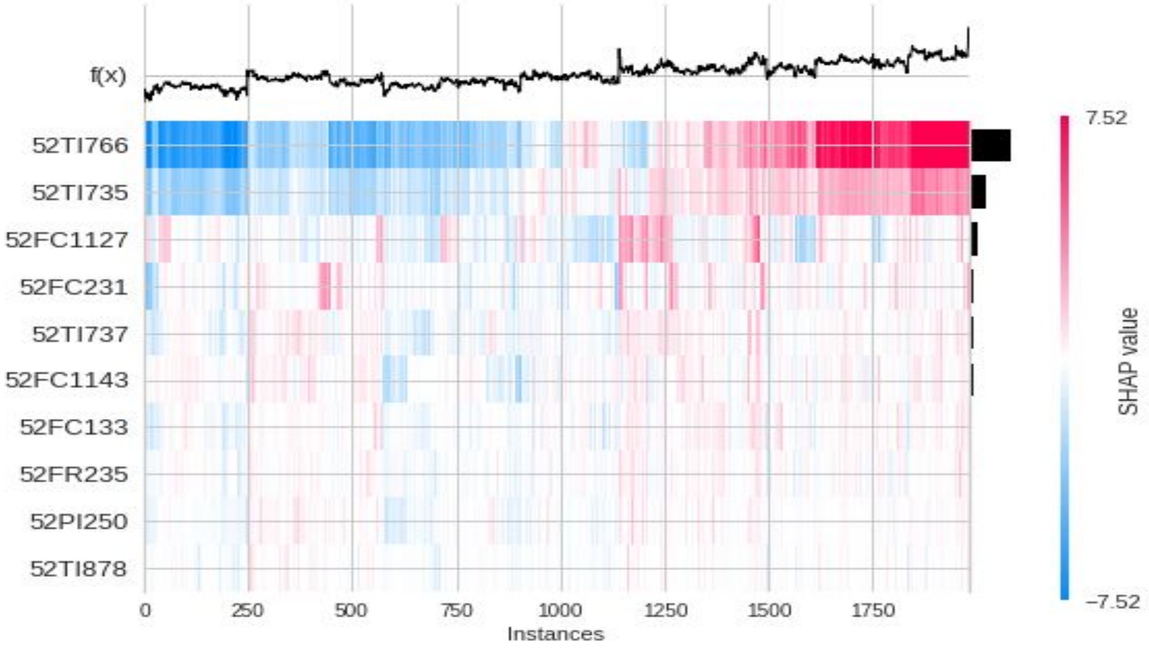




model selection

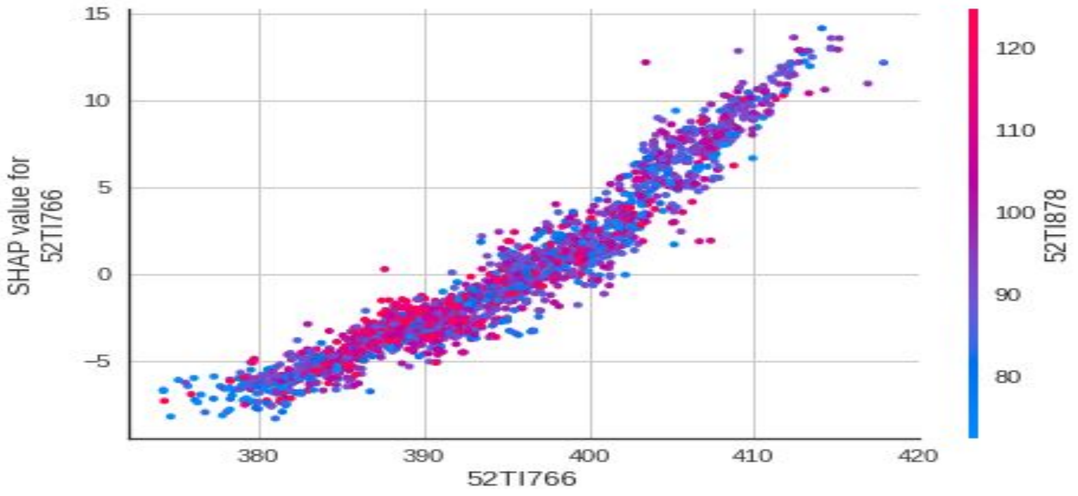
	RMSE	R^2
ET Regressor	3.8562	0.7932
Random Forest	4.0111	0.7771
Gradient Boosting Regressor	4.0301	0.7746
Huber Regressor	4.358	0.7367
Ridge Regression	4.4054	0.7311
Linear Regression	4.4067	0.7308
Neural networks (3 dense layers)	4.8609	0.6845
Lasso Regression	5.1631	0.6329
Elastic Net	5.3317	0.6093
Decision Tree Regressor	5.4937	0.5756

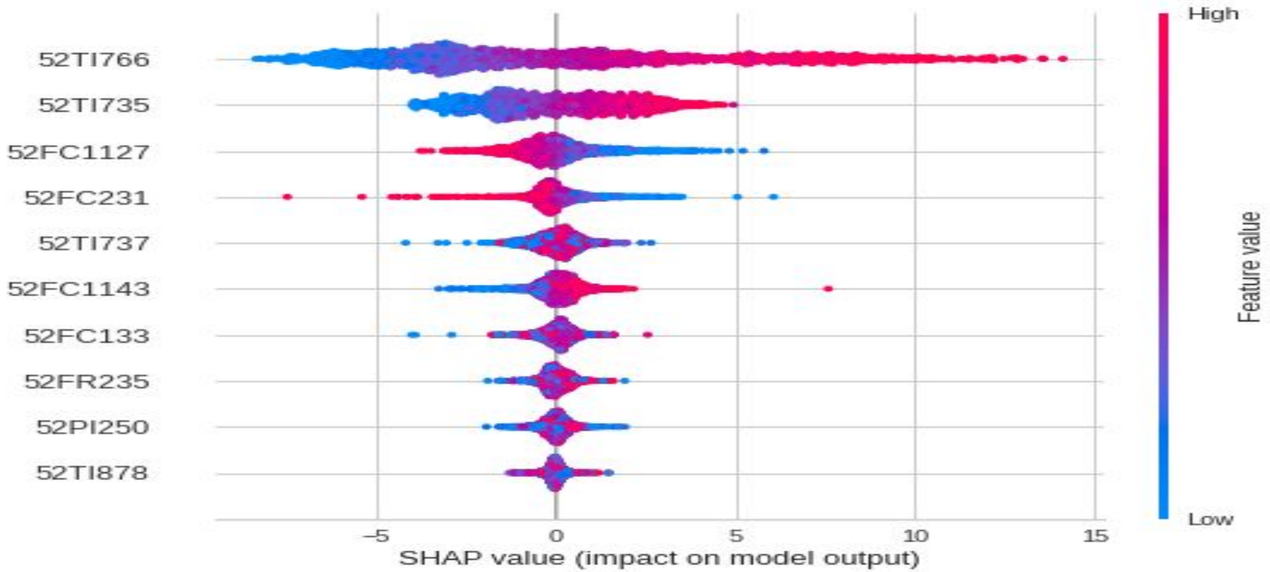
case study on real industrial process data using shap algorithm visualization-



Heatmap

Feature interactions





feature improtance

single prediction