Univ.AI

Ensemble Learning for Fire Calorimetry Database
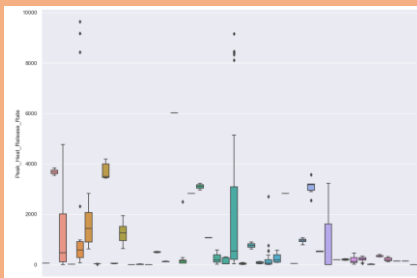## Ekanki, Srish, Bhaskar & Kuldeep

Goals and Motivation

- Ensemble techniques (Isolation Forest, Feature Bagging and XGBOD) to detect anomalies.
- Regression analysis for Peak Heat Release Rate. Model Interpretation.
- Aleotoric Error - Data Augmentation.
- Clustering / Ensemble methods for classifying fuel types.

Aleotoric Uncertainty

- Using Central Limit Theorem
- Feature follows a Gaussian distribution.
- Measurement uncertainty as standard deviation and average value as the mean.
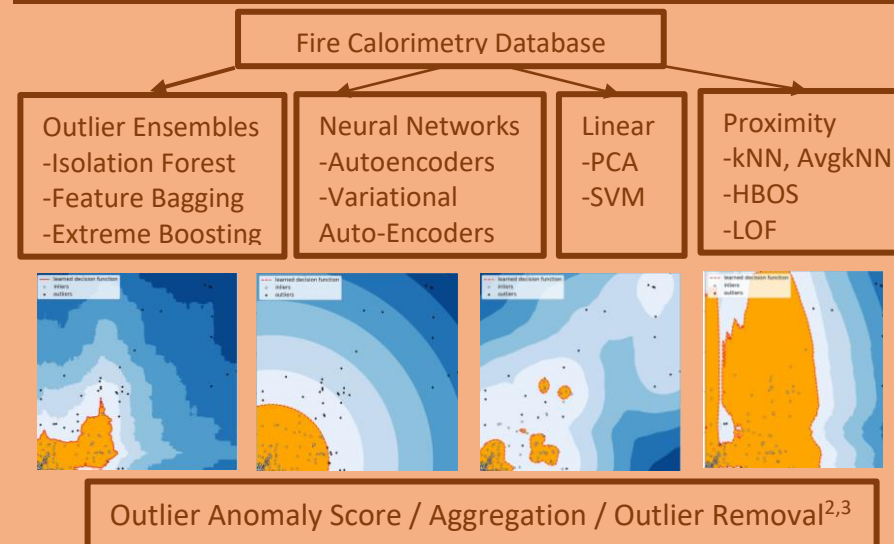- Perform Cholesky decomposition

References

1. *FCD Database,* http://www.nist.gov/fcd
2. *Outlier Ensembles, C Aggarwal, S Sathe, Springer 2017*
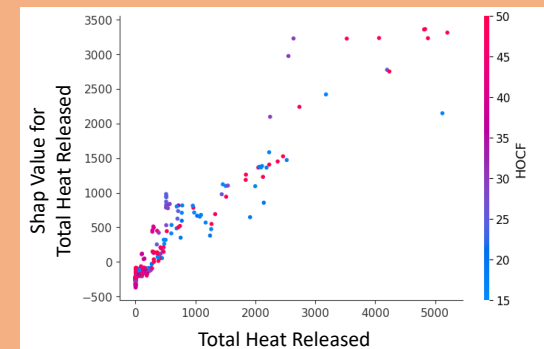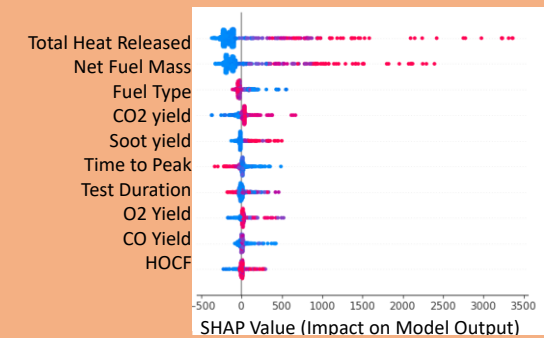3. *PyOD: A python toolbox for scalable outlier detection, JML 20 (2019) 1-7.*

NIST Fire Calorimetry Database (FCD)[1]

FCD consists of data augmented video and tabulated data from 1020 experiments. Each experiment is described with metadata, time dependent calculations based on dozens of sensors and quantified uncertainty.

Anomaly / Outlier Detection Flowchart

Fire Calorimetry Database

Outlier Ensembles	Neural Networks	Linear	Proximity
-Isolation Forest	-Autoencoders	-PCA	-kNN, AvgkNN
-Feature Bagging	-Variational	-SVM	-HBOS
-Extreme Boosting	Auto-Encoders		-LOF

Outlier Anomaly Score / Aggregation / Outlier Removal[2,3]

Interpretation Learning

Results and Analysis

Both original and augmented dataset yielded that HOCf and Ef are the more importance characteristic features.

Conclusions

-Shap Value and LIME are very useful tools for interpretation learning.
-Provide lots of insight into the model

Future Work & Impact

-Extension for data augmented videos and time-dependent output.
-Continue project as on-going research.