Imperial College London

IRP Project

15/Sep/2023

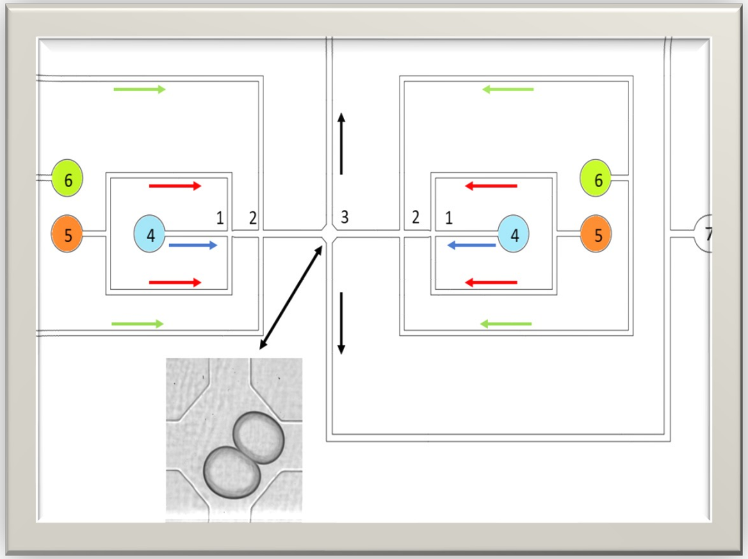
Explainable AI: understand the black box of predictive models with chemical engineering applications

Name: Jinwei Hu

Supervisors: Sibo Cheng, Rossella Arcucci

Student ID: 02261789

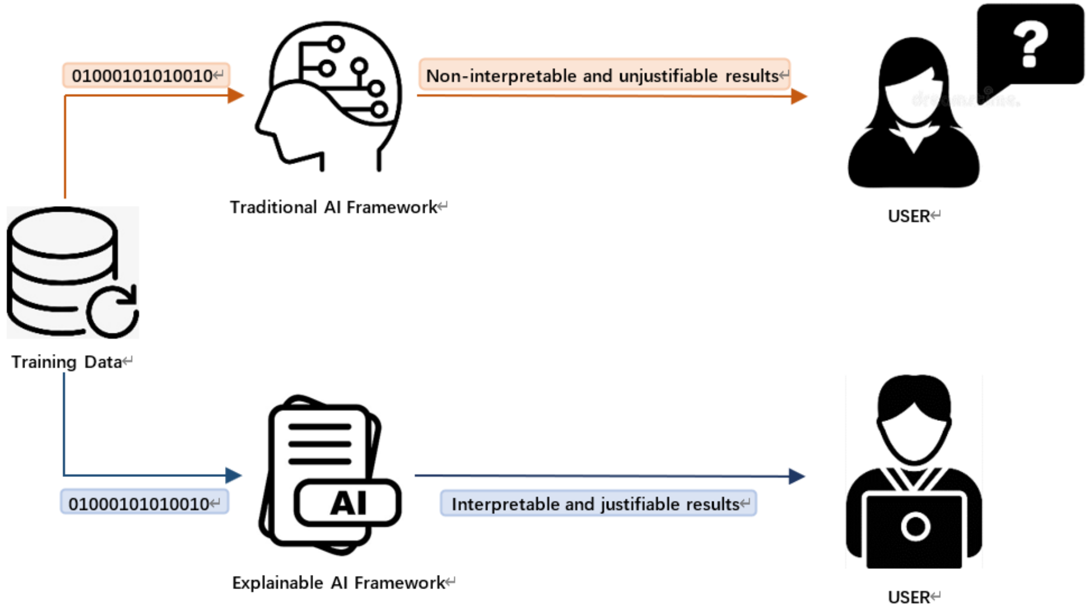
Integrating Machine Learning and Explainability for Studying Droplet Coalescence in Microfluidic System



- Objective: Predict & Control Droplet Coalescence & Explainability
- Droplet Coalescence Applications: Pharma, Food, Oil, Materials
- Methodology: ML Models + Explainability Tools
- Key Challenges: Non-linear Dynamics & Model Opacity
- Innovation: ML & Explainability on Droplet Coalescence in Microfluidic System

Research Motivation

- Complexity: Numerous parameters in fluid coalescence
- Variability: Different conditions yield diverse outcomes
- ML Opacity: Lack of transparency in predictions
- Trust: Need for credible, understandable results
- Necessity of XAI: ML + XAI for transparent, reliable solutions in chemical engineering

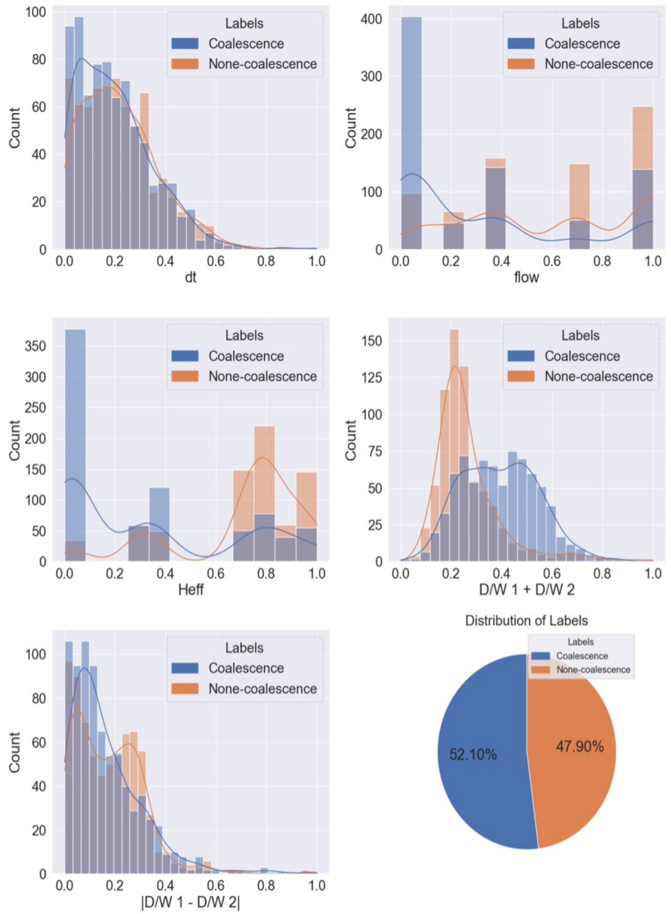


Imperial College London

Dataset Overview

- 1501 Samples
- Five Features & One Label
- Coalescence & Non-Coalescence
- Training & Testing
- Balance Ratio

	Coalescence	Non-coalescence	Balance Ratio (BR)	Total
Total dataset	782	719	1.09	1501
Training dataset	625	575	1.09	1200
Testing dataset	157	144	1.09	301



Imperial College London

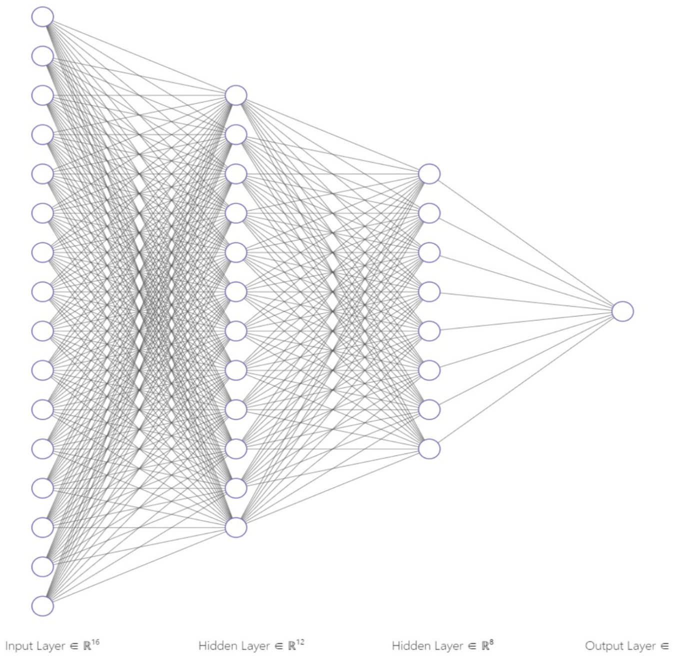
Data Preprocessing

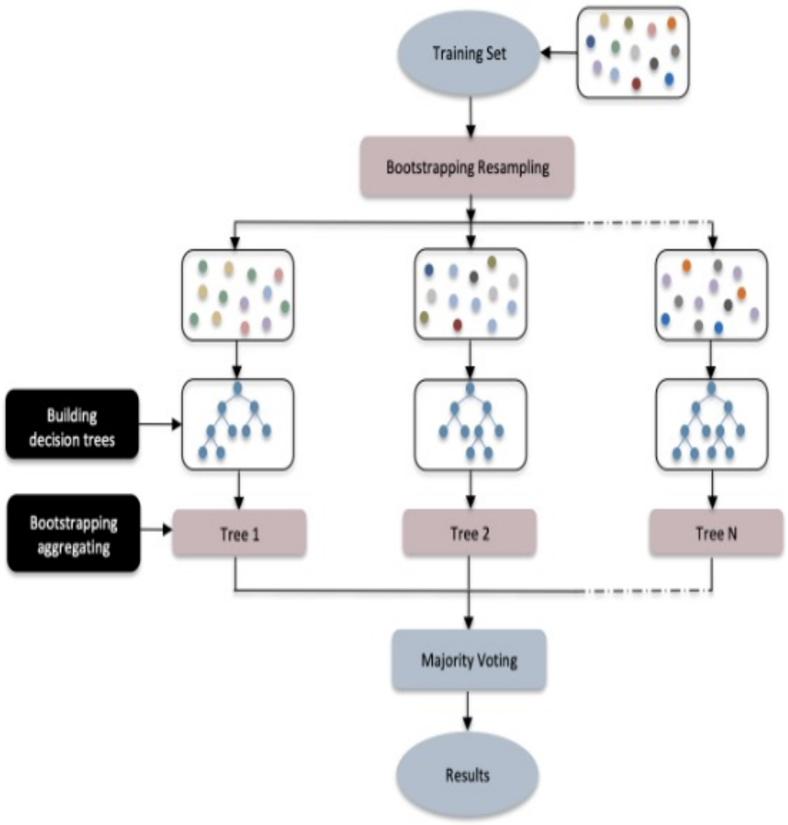
- Raw Data
- Min-Max Normalization: keep the relative relationships among the features

$$x_{\text{scaled}} = \frac{x - x_{\min}}{x_{\max} - x_{\min}}, \quad x \in \left\{x_{\frac{D}{L}1+\frac{D}{L}2}, x_{|\frac{D}{L}1-\frac{D}{L}2|}, x_{dt}, x_{\text{flow}}, x_{\text{Heff}}\right\}$$

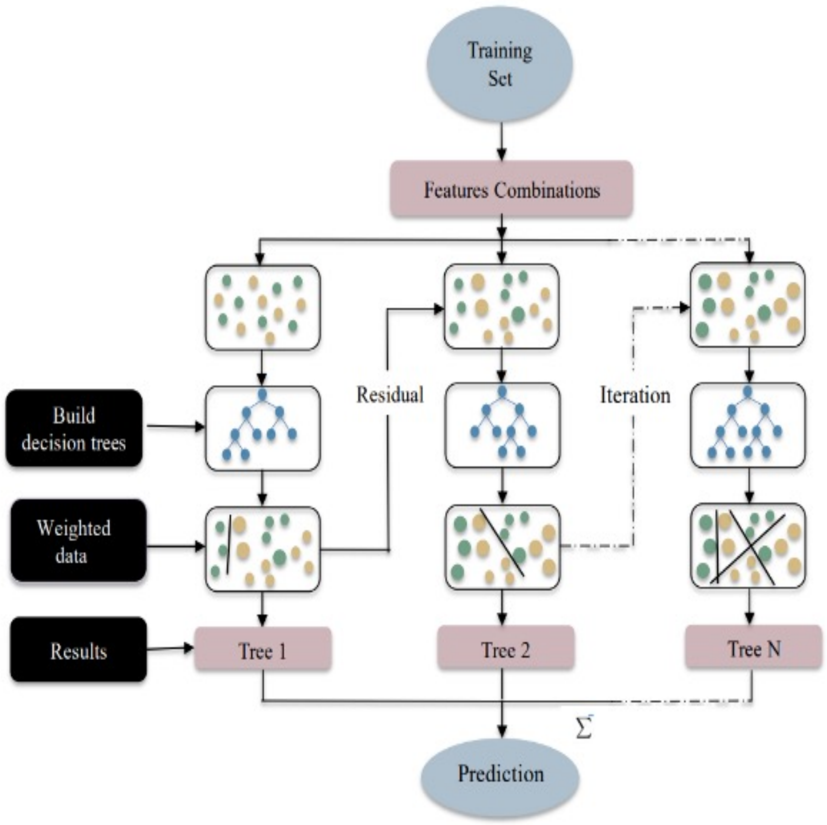
Predictive Models

- Random Forest
- XGBoost
- Multilayer Perceptron(MLP)





(a) Random Forest



(b) XGBoost

Imperial College London

XAI: SHapley Additive exPlanations(SHAP)

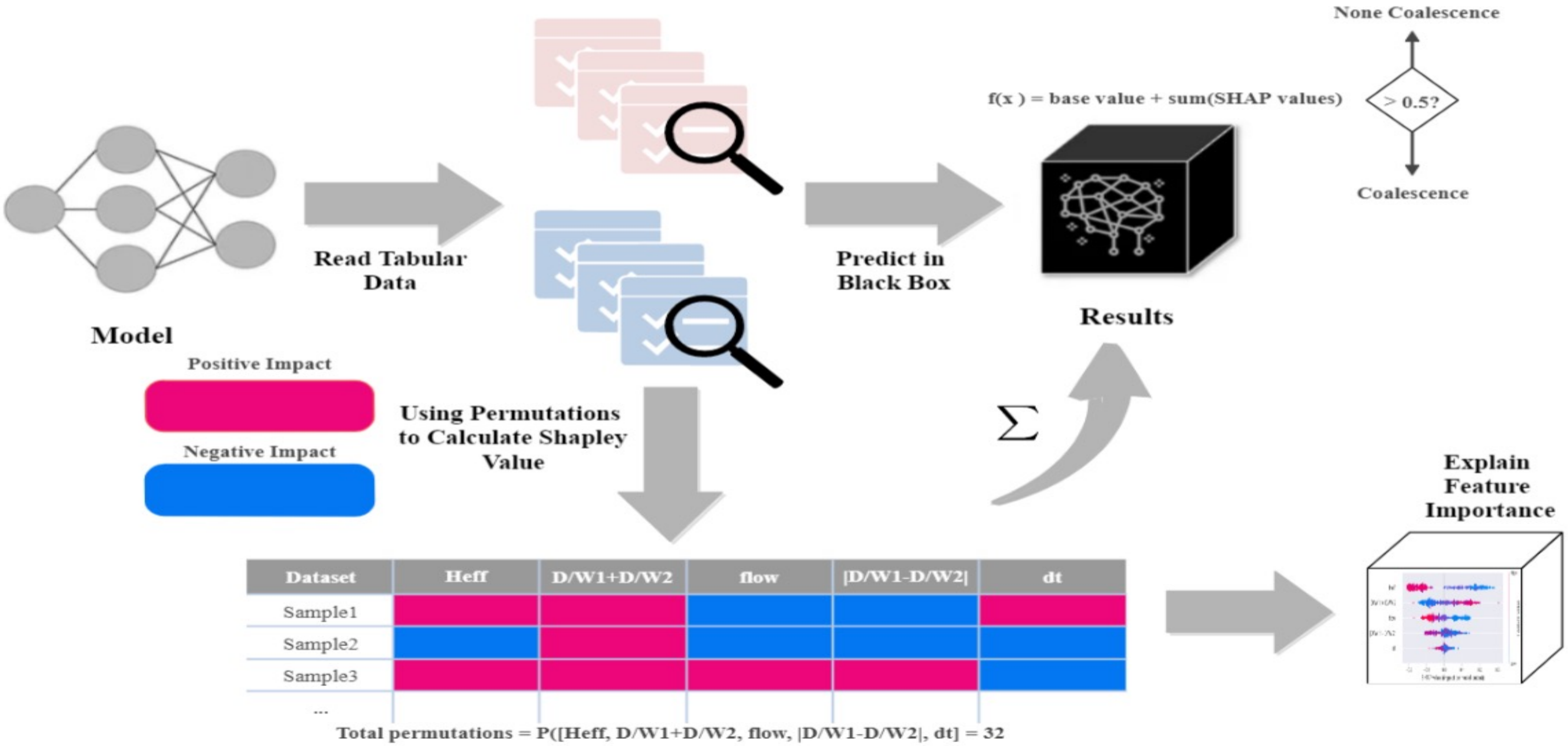


Figure 4: Visual Representation of SHapley Additive exPlanations (SHAP) in Chemical Applications

Imperial College London

Local Interpretable Model-agnostic Explanations(LIME)

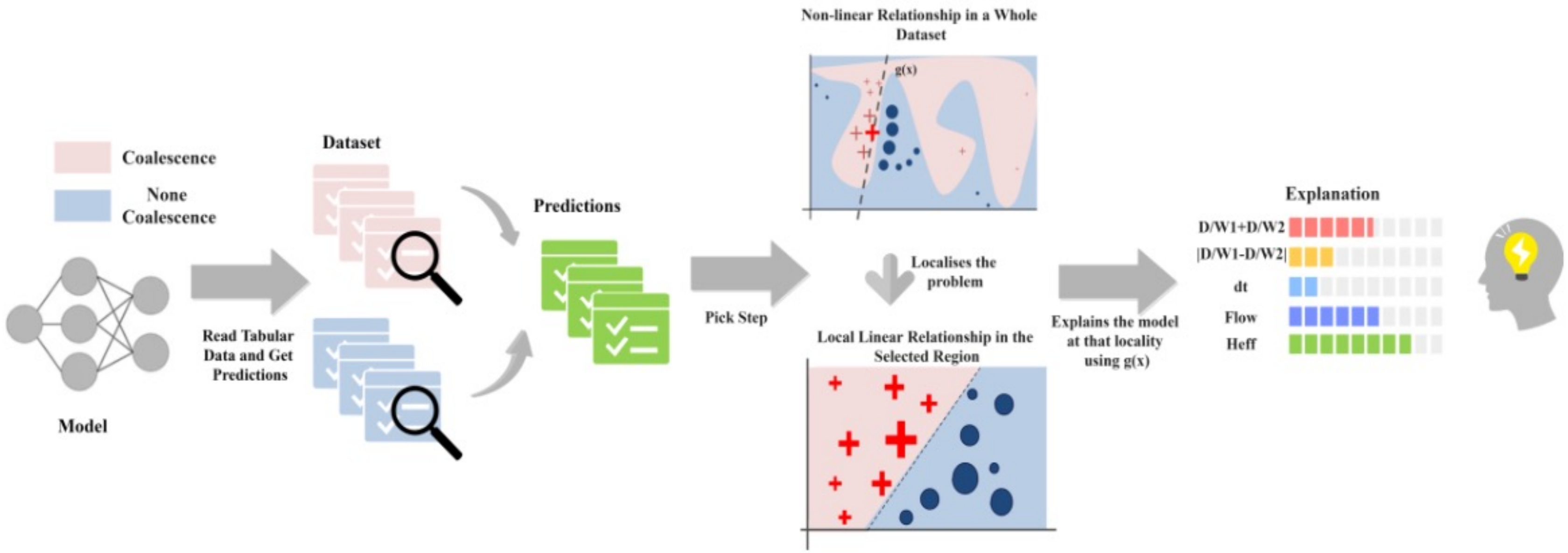
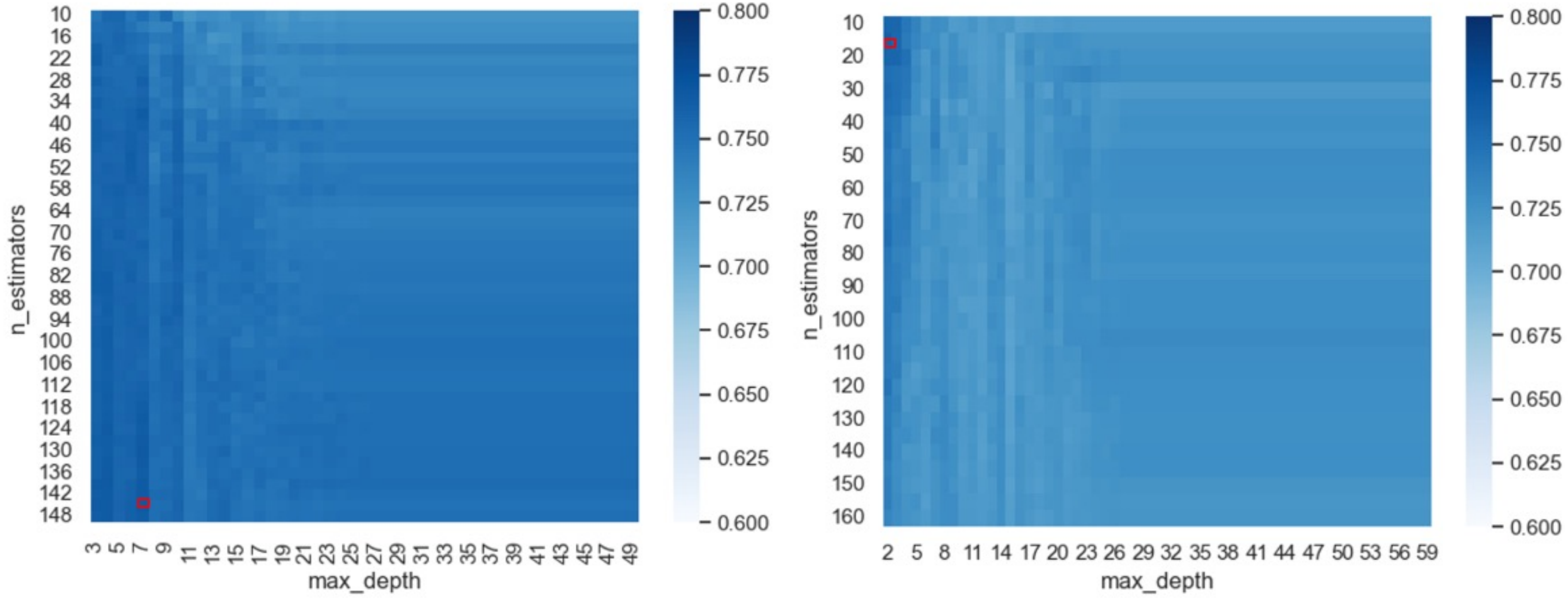


Figure 5: Visual Representation of Local Interpretable Model-agnostic Explanations (LIME) in Chemical Applications

Implementation Details

Random Forest: [n estimators = 145, max depth = 7]
XGBoost: [n estimators = 15, max depth = 2]



(a) Random Forest (b) XGBoost

Figure 6: Validation heatmaps for the tuning hyperparameters of tree-based predictive models. The selected hyperparameters are highlighted by the red frames.

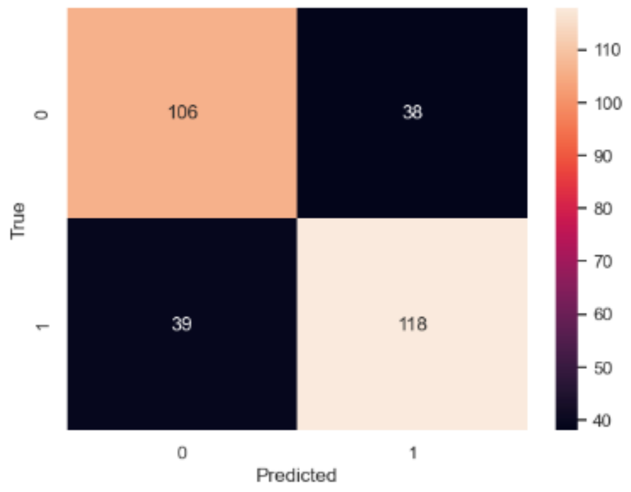
MLP Hyperparameters

Hyperparameter	Options
Activation Functions	**relu**, tanh, sigmoid, linear
Optimisers	**adam**, sgd, rmsprop
Learning Rates	**0.001**, 0.01, 0.1
L2 Rates	0.0, 0.001, **0.01**, 0.1
Dropout Rates	0.0, **0.1**, 0.2, 0.3
Epochs	10, 50, 100, 150, 200, 250, 300, 400, 450, 500, 800, 1000, 1500, 2000, **2500**, 3000

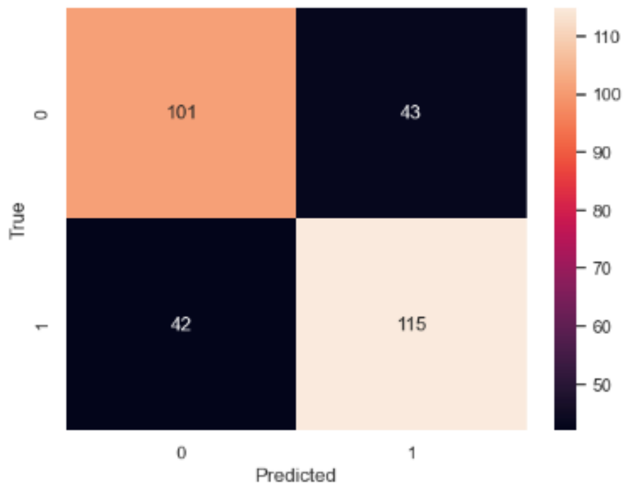
Imperial College London

Predictive Results

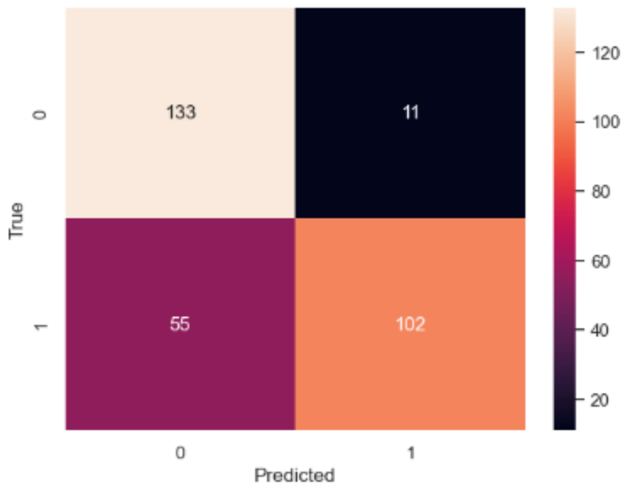
Model	Precision(%)	Recall(%)	F1-Score(%)	Accuracy(%)
Random Forest	75.64	**75.16**	75.40	74.42
XGBoost	72.78	73.25	73.02	71.76
MLP	**90.27**	64.97	**75.56**	**78.07**



(a) Random Forest



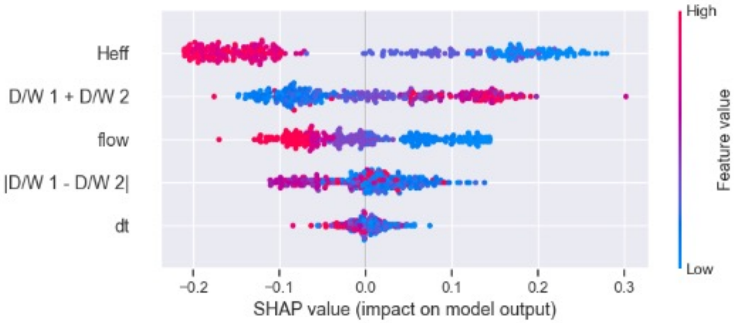
(b) XGBoost



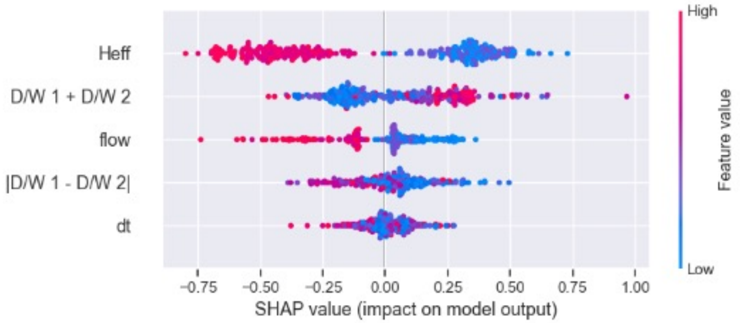
(c) MLP

Imperial College London

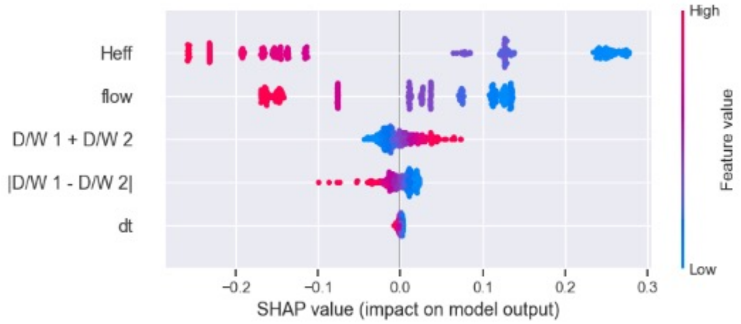
Global Interpretability



(a) Random Forest



(b) XGBoost



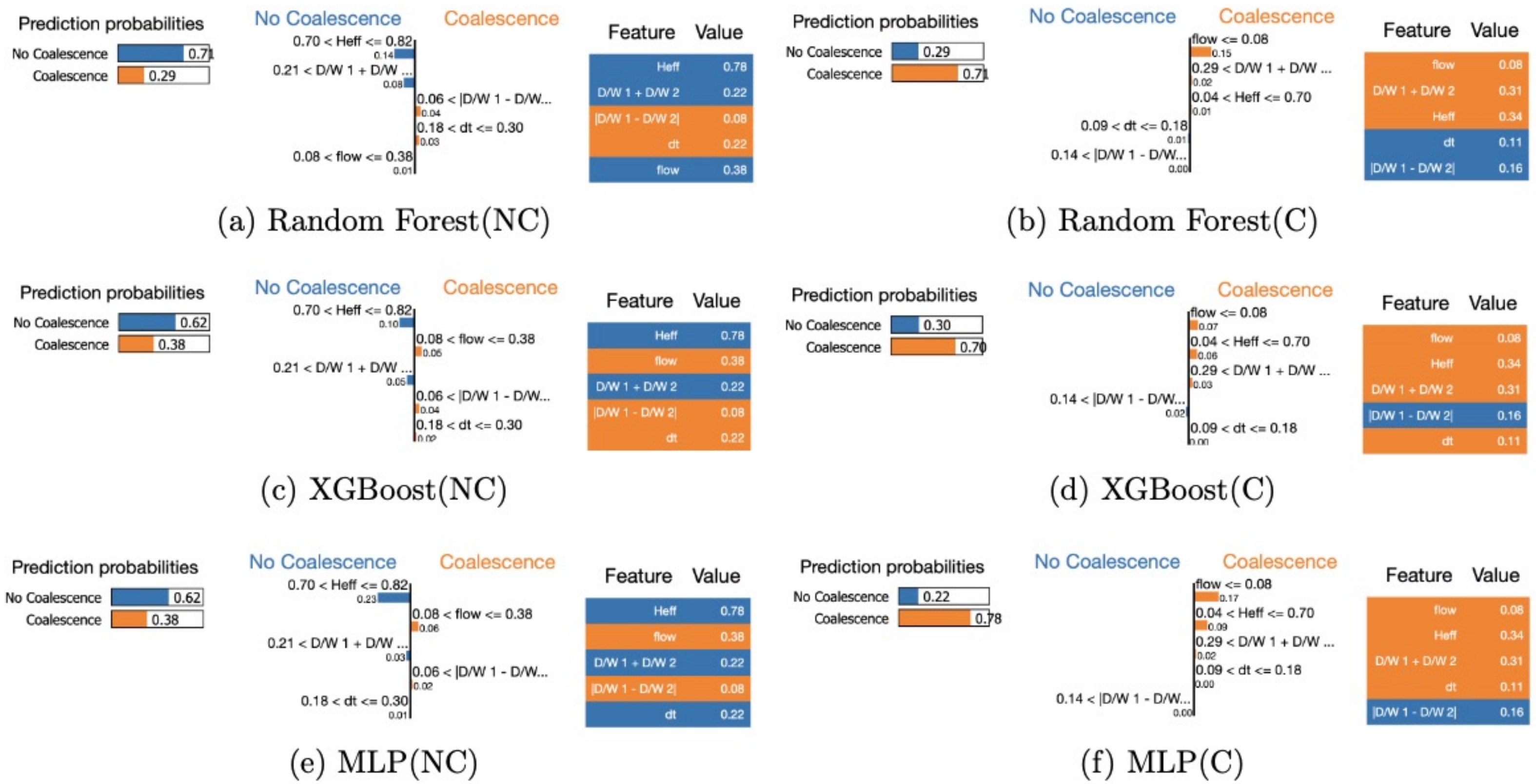
(c) MLP

Feature Ablation Testing

Model	Baseline	w/o Heff	w/o $\frac{D}{W}1 + \frac{D}{W}2$	w/o flow	w/o $\lvert\frac{D}{W}1 - \frac{D}{W}2\rvert$	w/o dt
Random Forest	**74.42%**	72.09%	70.10%	74.09%	72.09%	70.10%
XGBoost	71.76%	66.11%	72.43%	72.09%	72.43%	**74.09%**
MLP	**78.07%**	69.77%	75.08%	76.08%	76.74%	75.75%

Imperial College London

Local Interpretability



(a) Random Forest(NC) (b) Random Forest(C)

(c) XGBoost(NC) (d) XGBoost(C)

(e) MLP(NC) (f) MLP(C)

Figure 9: LIME Plot for two instances. NC means the actual label for instance is Non-Coalescence, and C means the actual label for instance is Coalescence

Future works

- Adapt to other fluid dynamics areas
- Advanced Deep learning Models
- Tailor-made interpretability frameworks

Thank you!