ASME 2025 Student Hackathon

nTop

SURROGATE PHYSICS MODELING FOR HEAT EXCHANGER OPTIMIZATION: ACCELERATING DESIGN EXPLORATION VIA INVERSE DESIGN

A++ Team

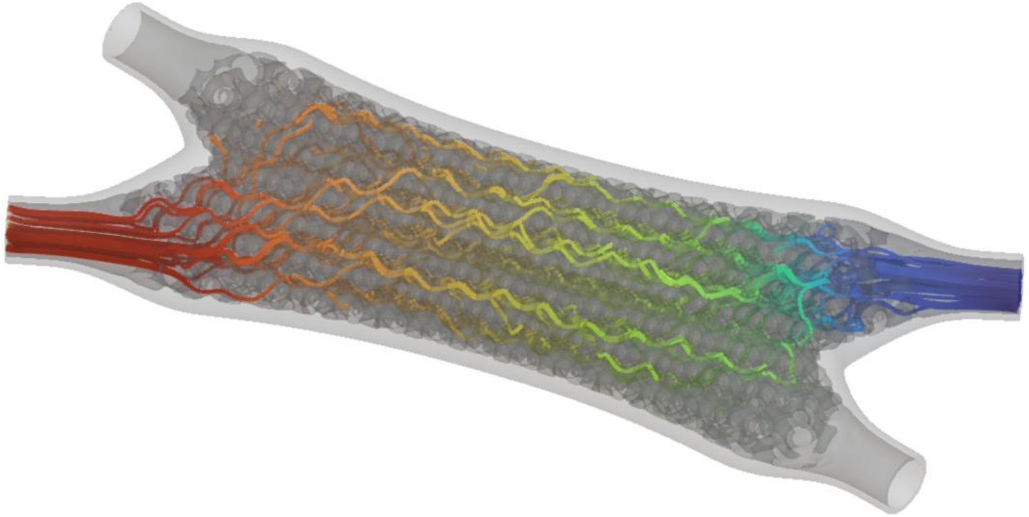
Amirhossein Iranmehr

Armin Hassanirad

UNIVERSITY OF ALBERTA

Overview

- Problem statement
- Review of previous works
- Primary Exploratory Data Analysis (EDA)
- Secondary EDA
- Model setup
- Model prediction
- Inverse design
- Potential applications and broader implications
- Key takeaways



UNIVERSITY OF ALBERTA

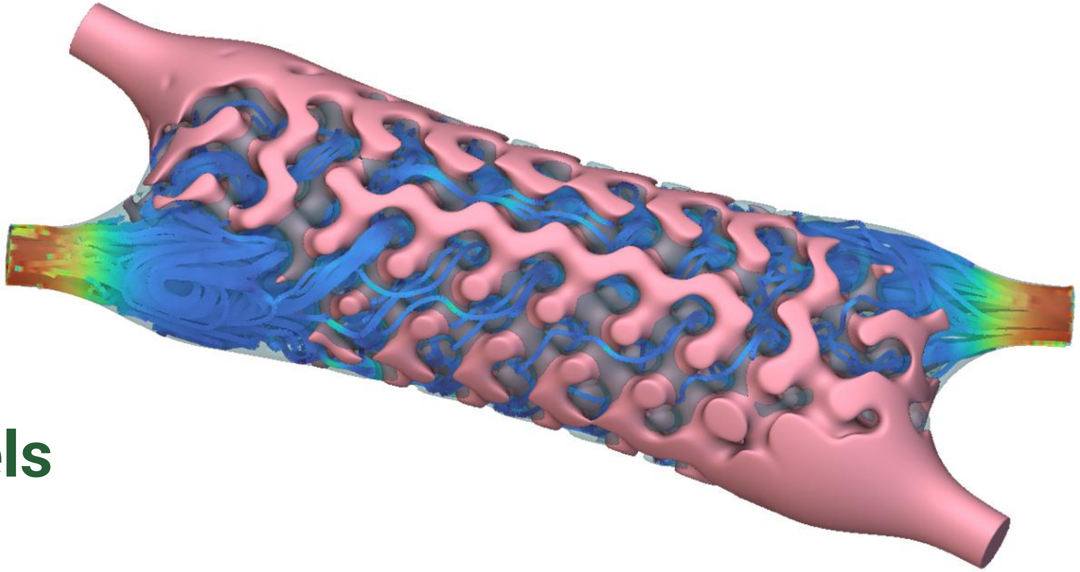
Problem Statement

➢HEXs Conventional Design

- CFD, high computational cost, and time-intensive
- Limited scalability

➢Design of HEXs via surrogate physics models

- Dataset provided by nTop
- Train a model to predict flow and geometry properties



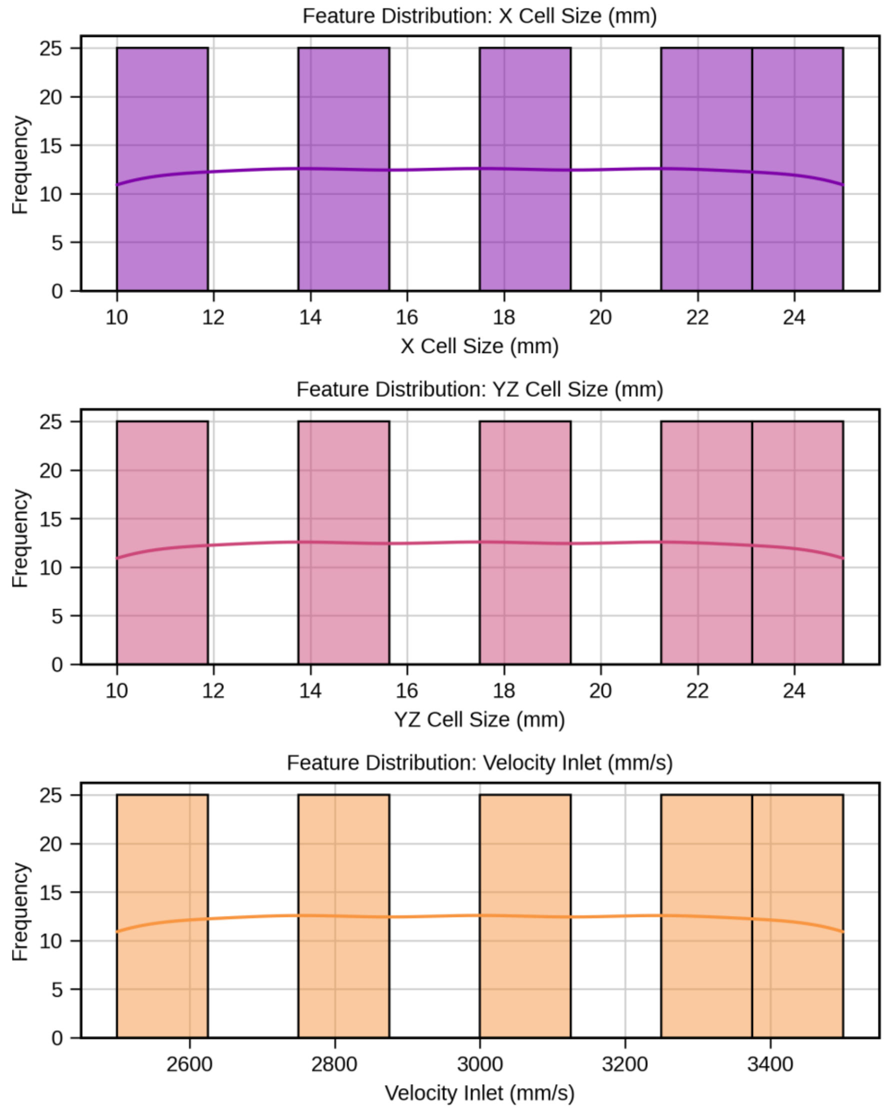
UNIVERSITY OF ALBERTA

Review of Previous Works

Study & Year	Method / Algorithm	Application	Key Outcome
Jeon et al. (2022)	Residual-based Physics-Informed Transfer Learning (RePIT)	CFD acceleration in industrial simulations	Reduced computational cost while maintaining accuracy
Kwon et al. (2020)	Random Forest	Predict convection heat transfer coefficients in cooling channels with rib roughness	Rapid multi-variable predictions without deep simulations
Krishnayatra et al. (2020)	k-Nearest Neighbor (k-NN)	Predict efficiency & effectiveness of axial finned-tube heat exchanger	High accuracy and effective geometry optimization tool

Primary Exploratory Data Analysis

- **Dataset Overview: 125 data points**
- **Data points appeared in distinct intervals**
- **Possible need for data augmentation**
- **Expand the physical domain knowledge of HEXS**
- **Provide sufficient data for bigger and more dense models (e.g., Neural Networks)**



UNIVERSITY OF ALBERTA

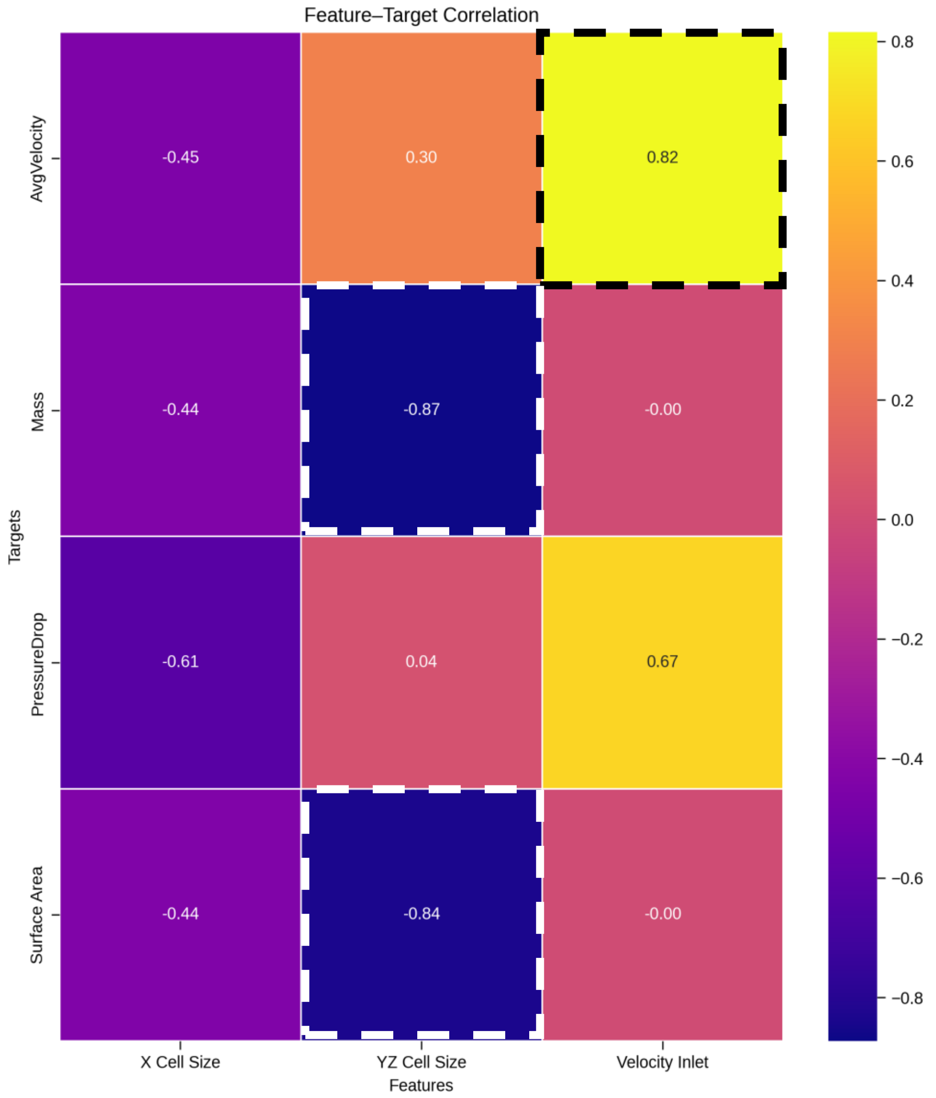
Data Augmentation

- Latin Hypercube Sampling to generate well-spread design points
- Target correlation analysis for proper sampling
- Sampling Design:

- 15 discrete levels for YZ cell size & inlet vel.

  (highest impact on targets)

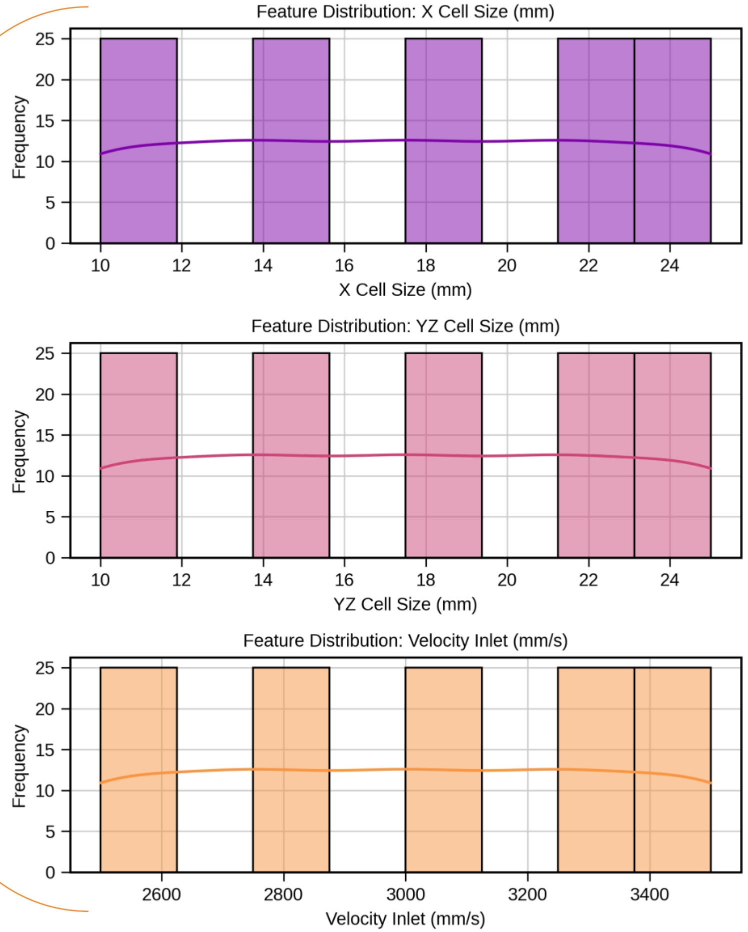
- 5 discrete levels X cell size
  (moderate impact on targets)

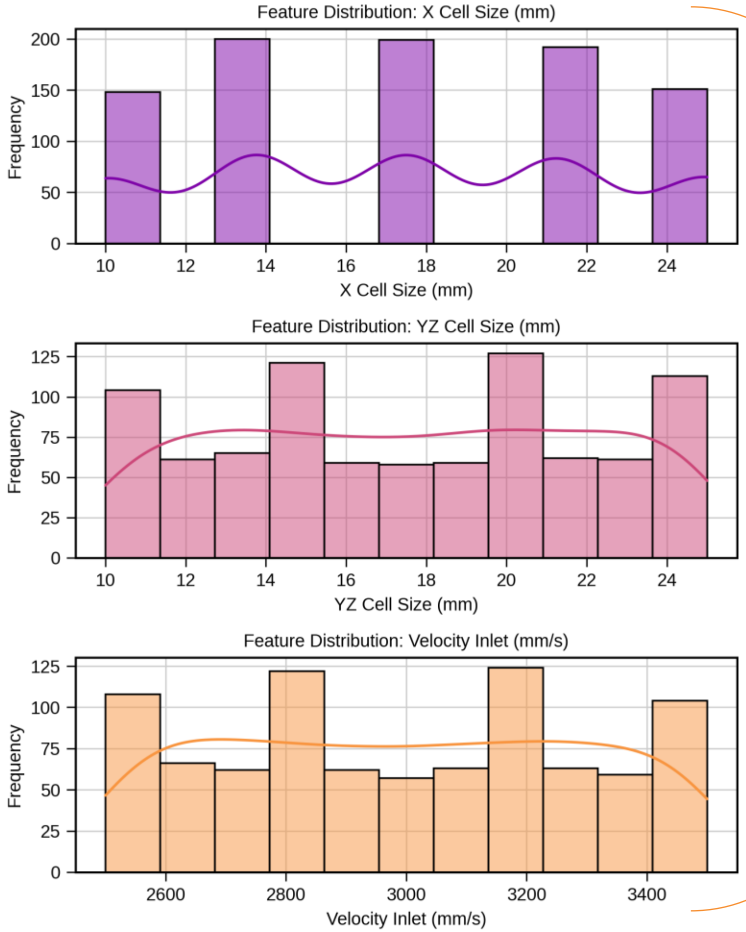


UNIVERSITY OF ALBERTA

Secondary EDA Comparing Features’ Distribution

initial dataset (125 datapoints)



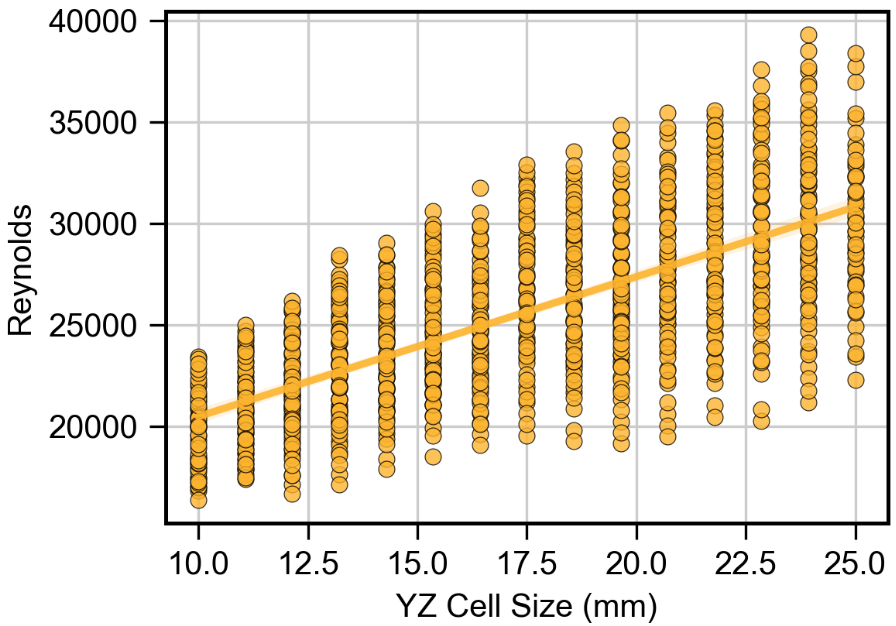


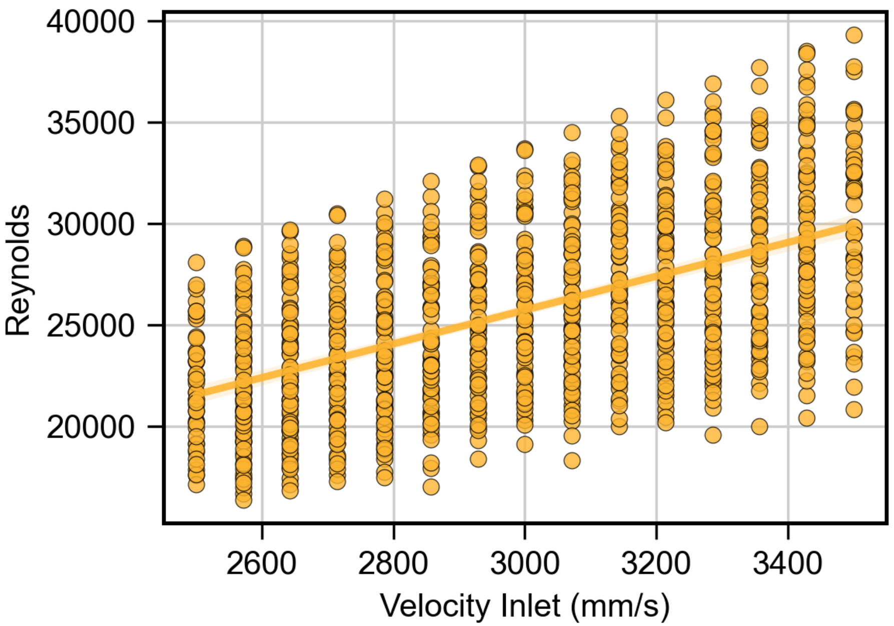
Augmented dataset (890 datapoints)

Reynolds Number a physics-informed analysis

- **Dataset preprocessing**
  - Cold fluid volume extraction from nTop
  - Calculating hydraulic diameter
  - Solving for Reynolds number

- **Reynolds number (Re) analysis**
  - Re increases linearly with inlet velocity (expected)
  - Scattered Re @ fixed inlet vel. Due to different cell sizes
  - Cold flow is <u>fully turbulent</u> in the Re range



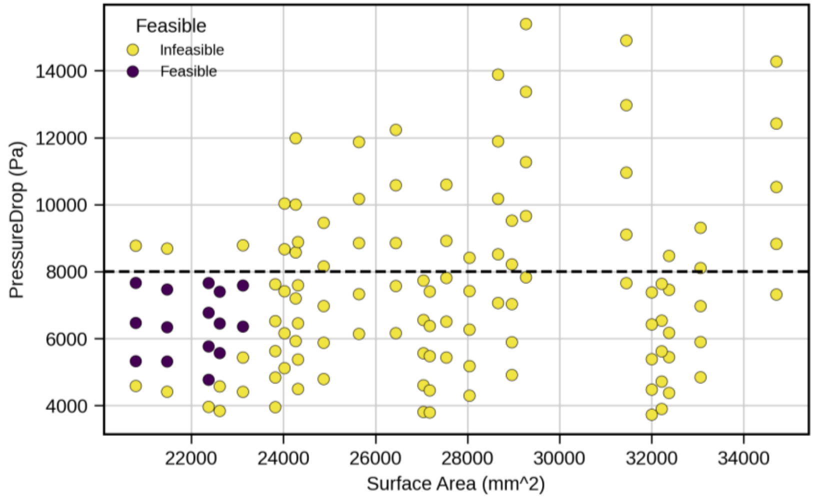


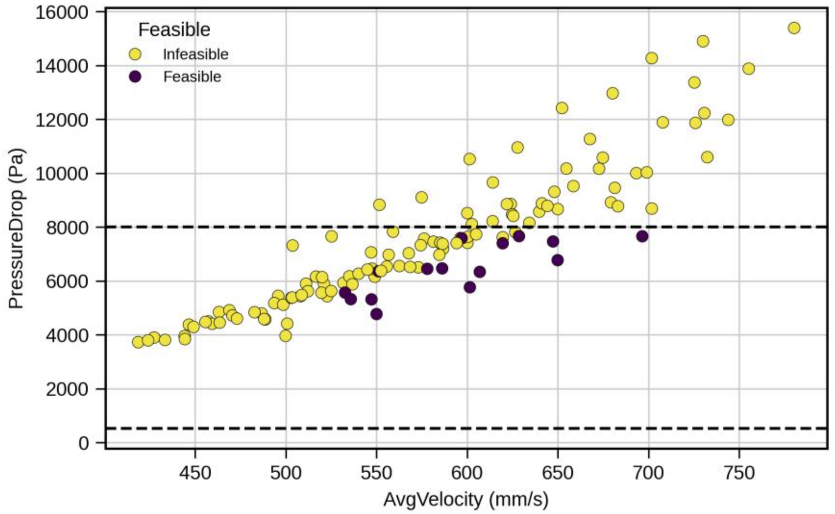
$$D_h = \frac{4V_S}{A_S}$$

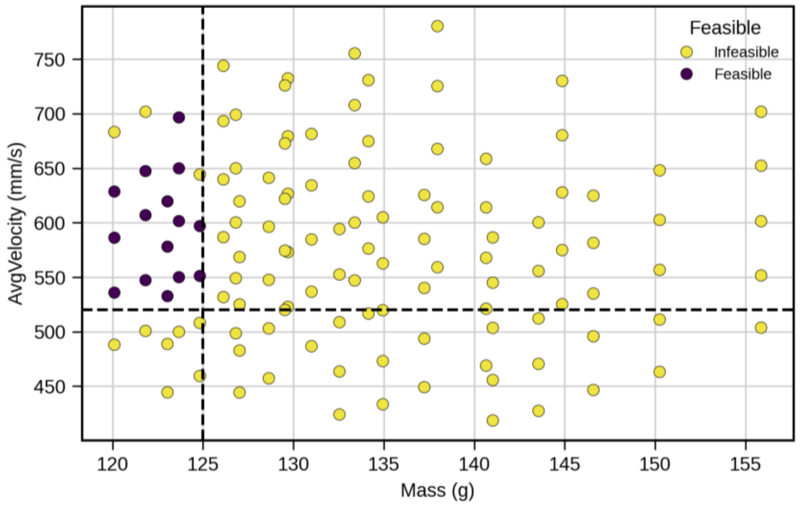
$$Re = \frac{\rho U D_h}{\mu}$$

UNIVERSITY OF ALBERTA

Primary EDA – Feasibility Plots

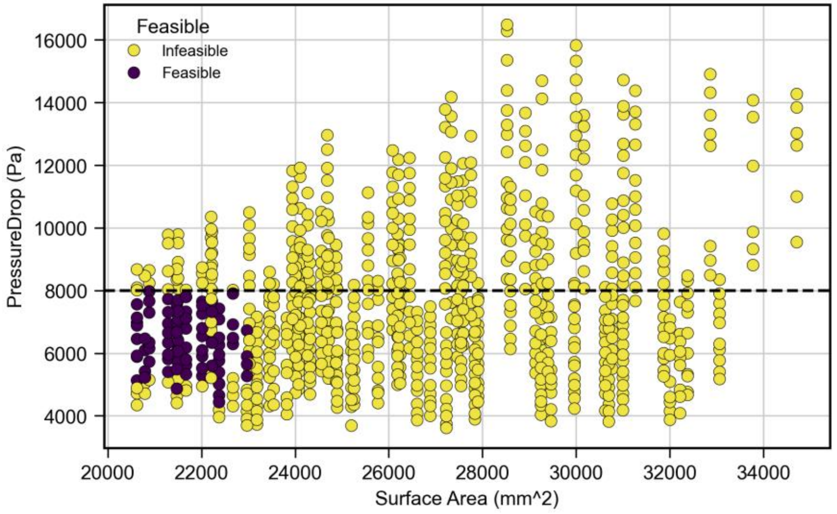


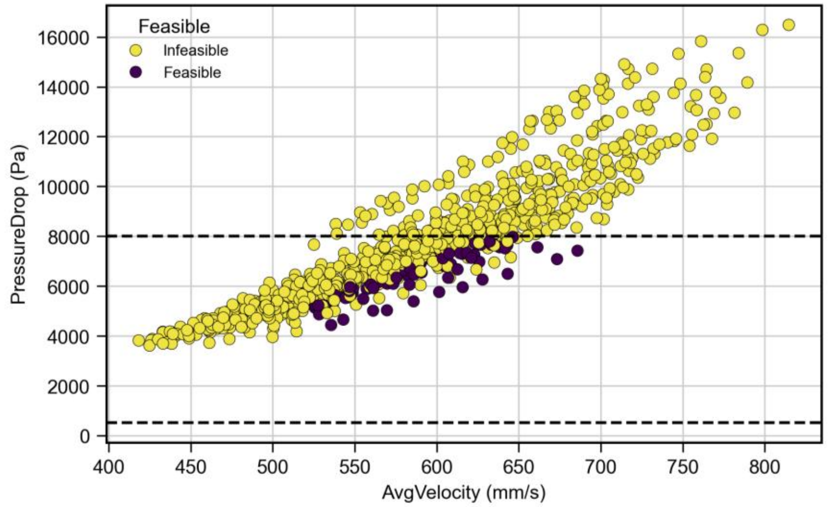


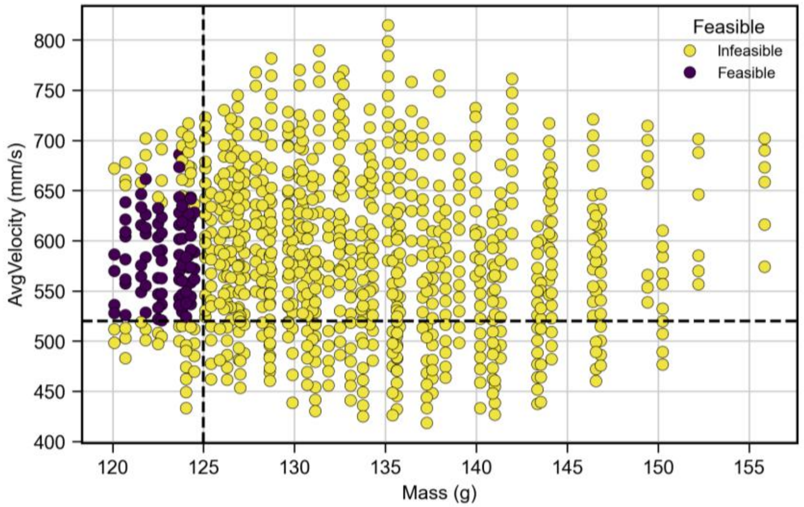


After data augmentation

After data augmentation

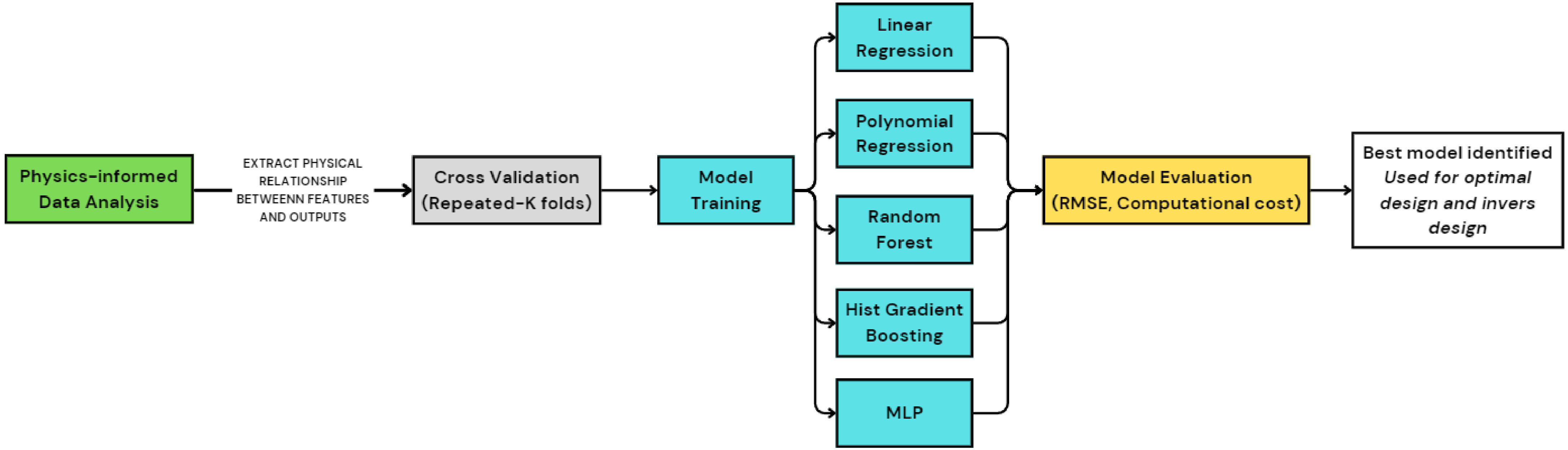






Models Setup-CV + Pipeline

- **Systematic Pipeline:** Preprocessing, scaling, and multi-output regression integrated.
- **Cross-Validation:** Repeated K-Fold (5 splits × 3 repeats = 15 folds) for robust training and reduced variance. It prevents overfitting and tunes the hyperparameters for each model.
- **Custom Scoring**: Defined target-specific RMSE plus a combined metric for fair comparison.



UNIVERSITY OF ALBERTA

Models Setup-Models + Evaluations

- **Models:**

Baselines: Linear Regression and Polynomial (degree 2) for linear and quadratic trends.

Ensembles: Random Forest and Histogram Gradient Boosting to capture complex nonlinearities.

Neural Models: MLP (128, 64) with ReLU, regularization, and early stopping for nonlinear mapping.

- Models were trained on 80% of the dataset (total: 890), and the other 20% was used to test the models.

- **Evaluation:**

- Statistical error metrics and visual comparisons of model predictions against the CFD results.
- The best model for accuracy and generality was selected through the comparison framework.
- This setup ensured that the surrogate model captured the underlying physics while.

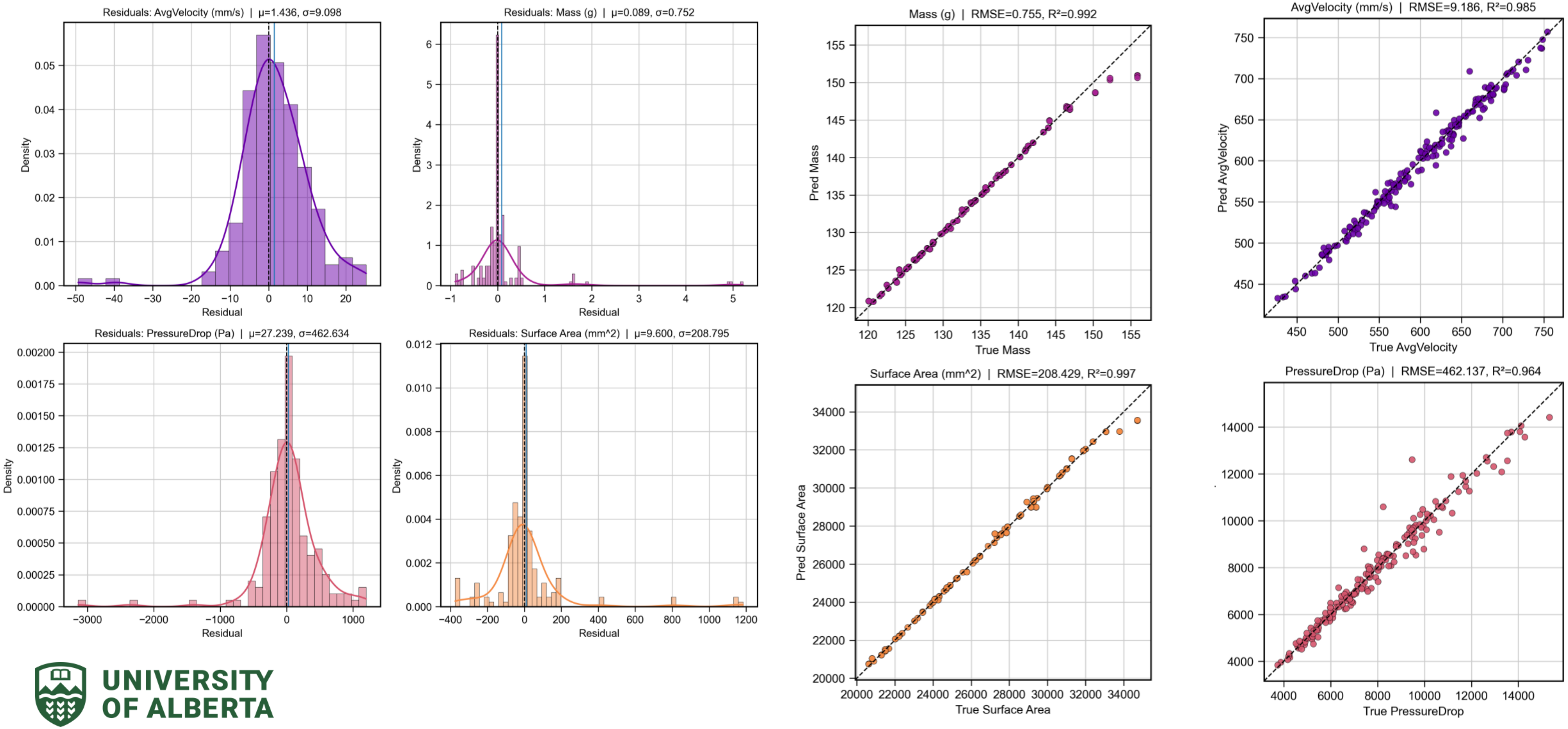
Models Comparison

- **HistGradientBoosting:** Best overall accuracy (Combined RMSE ≈ 471), strong across outputs.
- **Random Forest**: Comparable accuracy, lowest surface area and mass error, slightly higher combined RMSE.
- **Poly2+LR**: Fastest runtime (0.006 s), moderate accuracy — good speed/accuracy trade-off.
- **MLP (128,64)**: Captured nonlinearities but underperformed with high runtime (13 s).

Model	RMSE Pressure Drop	RMSE Avg Velocity	RMSE Surface Area	RMSE Mass	RunTime (s)
HistGradientBoosting	462.14	9.19	208.43	0.76	1.68
RandomForest	473.49	13.44	45.36	0.14	1.86
Poly2+LR	695.33	16.34	437.14	0.42	0.006
MLP (128,64)	746.04	15.51	373.48	0.41	12.93
LinearRegression	1015.01	19.00	976.78	1.72	0.005

UNIVERSITY OF ALBERTA

Model Predictions-HGB



UNIVERSITY OF ALBERTA

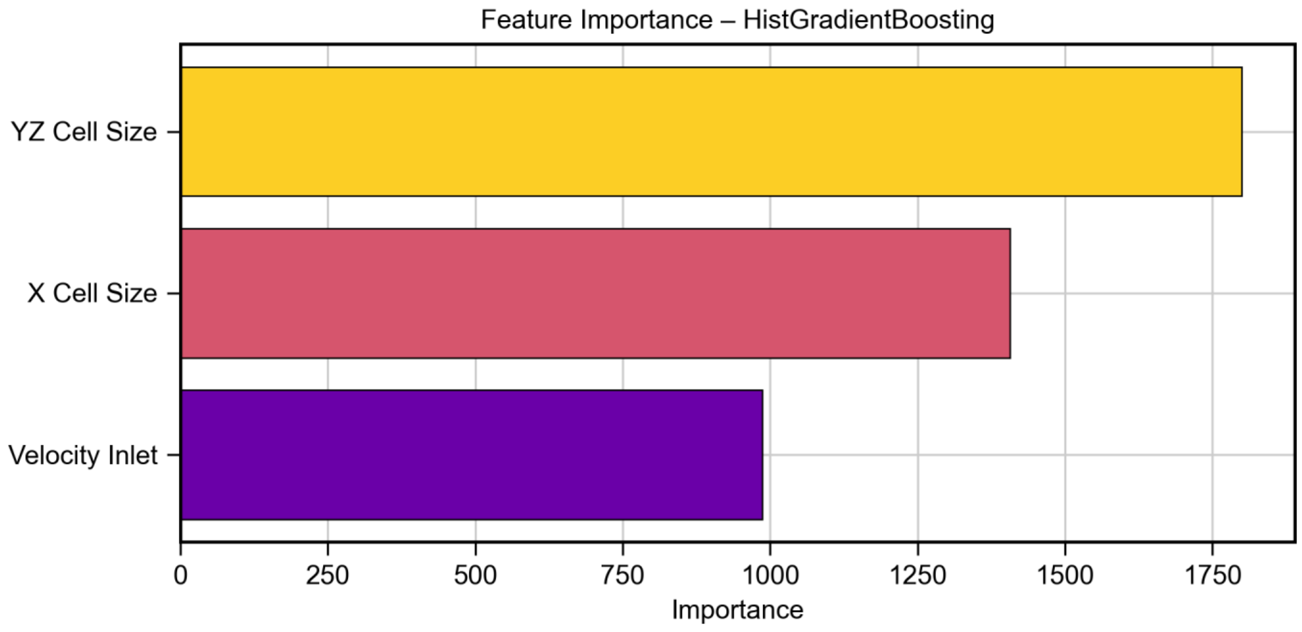
Model Prediction-Unseen Data/Cases

- ➢ It is necessary to evaluate the efficiency and practicality of the model against unseen data.
- ➢ Also, the feature importance for the HGB (best model) was analysed, which satisfied the reasons behind data augmentation. This should be noted when using inverse design as well.

Model Prediction Accuracy

- Tested on 8 random input cases.

Target Outputs	Model Prediction vs. nTop Error (%)
Average Velocity	2.73
Pressure Drop	7.53
Surface Area	2.93
Mass	2.86



Inverse Design

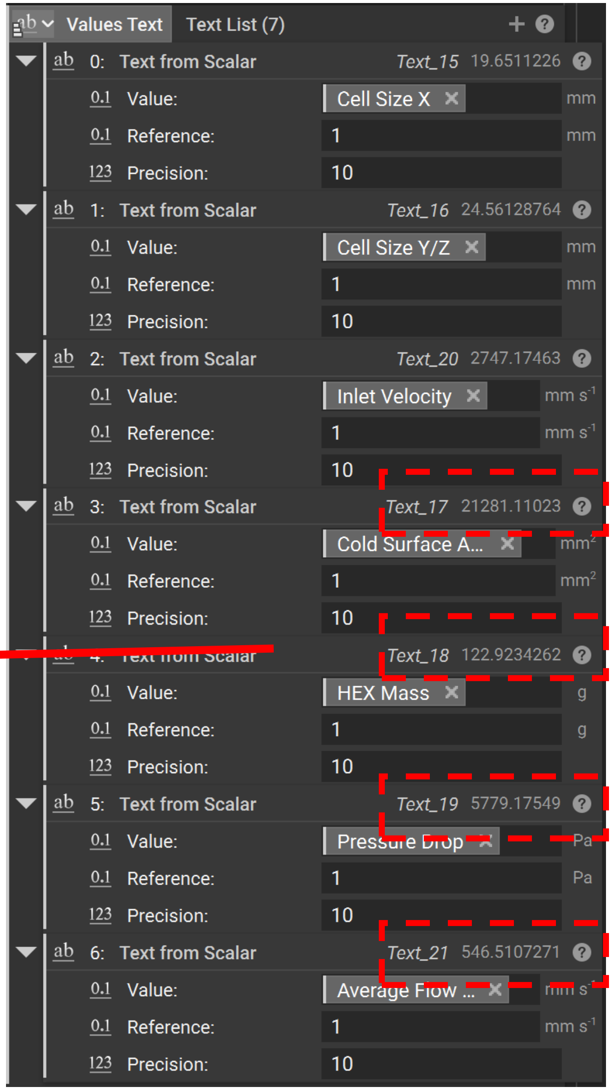
- ➢ A progressive Latin Hypercube Sampling (LHS) strategy was chosen to determine the optimal lattice design for the HEXs.
- ➢ Design space is iteratively zoomed in around the best feasible point, and the surface area is maximized while mass, pressure drop, and velocity constraints are enforced.

Optimal Design suggested

1. Cell X:19.6511226
2. CellYZ: 24.56128764
3. V-inlet: 2747.17462971

Target Outputs	HGB Predcition	nTop
Average Velocity	537.56	546.51
Pressure Drop	5124.33	5779.17
Surface Area	21516.2	21281.11
Mass	121.82	122.92

Mass < 125 grams
Pressure Drop < 8000 Pa
Avg Velocity > 520 mm/s^2

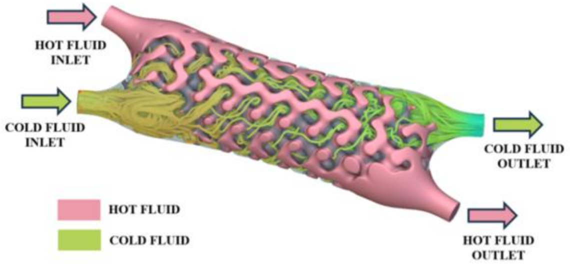


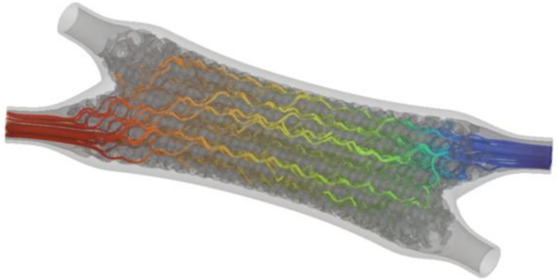
UNIVERSITY OF ALBERTA

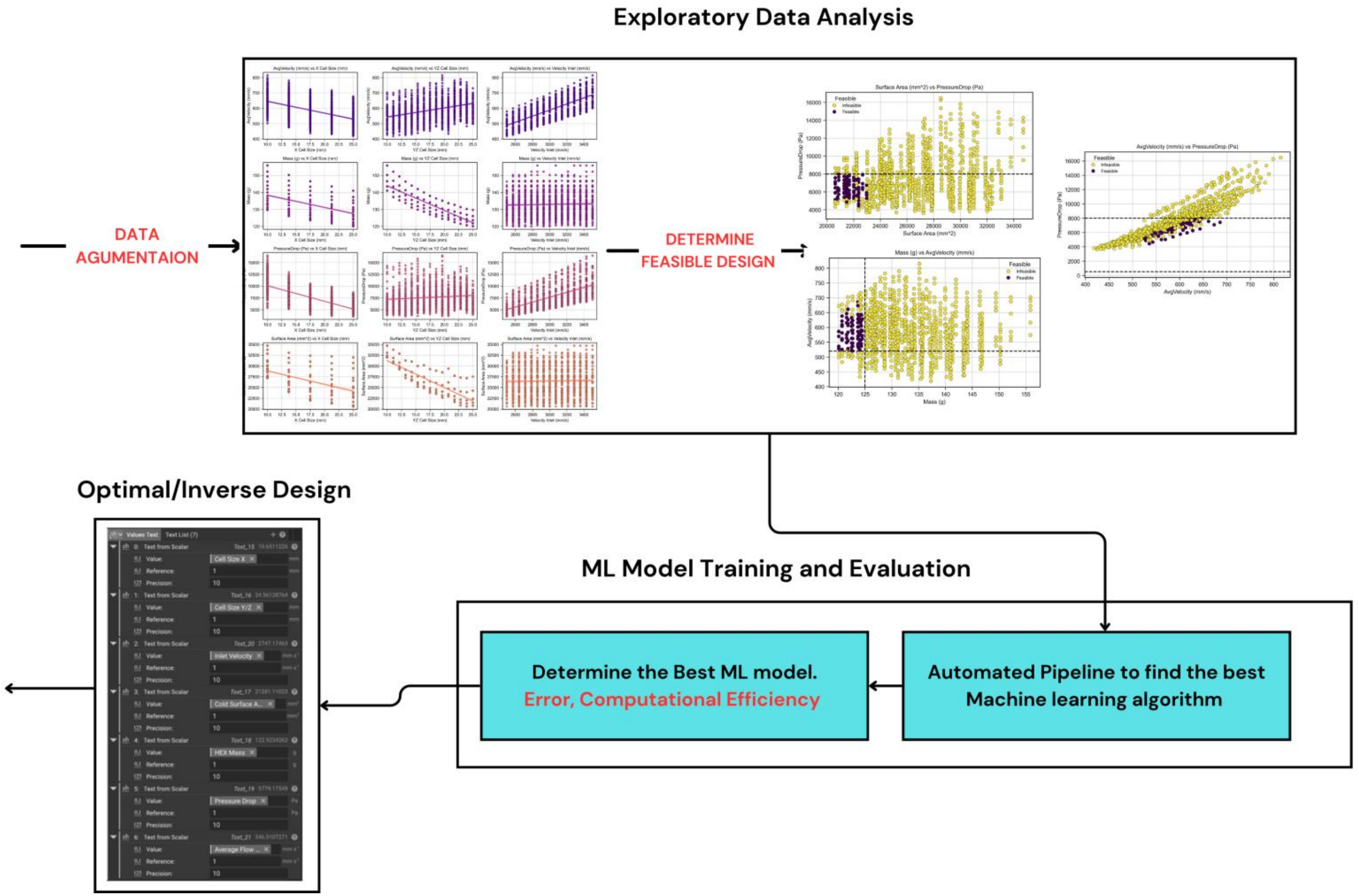
Potential applications and broader implications

Applications:

1. Coupled surrogate–CFD approach
2. Lattice-based structures' FEA analysis
3. Porous media simulations
4. Additive manufacturing process modeling







UNIVERSITY OF ALBERTA

Key takeaways-Future Works

- Physics-informed surrogate modeling successfully replaced computationally expensive CFD for heat exchanger optimization, cutting evaluation time from hours to milliseconds.
- Physics-informed exploratory data analysis (EDA) would further help us understand the domain of the training ML models and augment the data.
- Inverse design optimization identified optimal HEX geometry and flow parameters (e.g., X = 19.65 mm, YZ = 24.56 mm, velocity = 2747 mm/s) that maximize surface area while satisfying mass, velocity, and pressure constraints.
- Eight different sets of input parameters were utilized as unseen data to examine the model's performance compared to nTop.

Future works:

- Integrating physics-based empirical equations to enhance model accuracy
- In-depth analysis of data point efficacy and coverage to optimize prediction performance
- Expand the data to 5 to 10 thousand, ideal for training neural networks.

UNIVERSITY OF ALBERTA

Thanks!

Any Questions?