Hands-on Experiences on AI Implementation

Module 11: AI Implementation Process

IDARE®

4 Simple Steps

Data Collection & Pre-Process

- Data Collection
- Perform Preliminary Data Preprocessing
- Get your 1st ML prediction result to have an overall assessment

Build Model Training-Testing

- Variable Selection or Feature Engineering
- ML Model Selection
- Parametric study
- Finalizing the model

Deploy for Production

- Production data collection
- Production mode setting
- Configuring, saving and distributing results
- Result post processing for decision making

Get Decision

- Result Visualization
- Scenario analysis for optimized decision

AI Implementation Process

IDARE®

Data | Training- Testing | Deploy | Decision

Data
- Hospital Patient
- BMD Weather
- City Vector Survey

Insight Brainstorming

Dengue Prediction

Data Ingestion

SME Driven Parametric Study

Data Preparation Cleanup

Anomaly Detection

Missing Value Replacement

Data Split

Test Data

Training Data

Validation Data

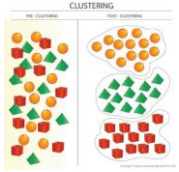
Feature Engineering

- New Variable Creation
- Variable elimination

Machine Learning
Statistical Validation

ML Model Selection

Before After



Parametric Study

docker

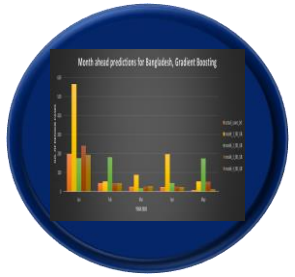
kubernetes

Model Deployment

Production Modes

Dengue Prediction

1. Year Ahead
2. Month Ahead
3. Week Ahead

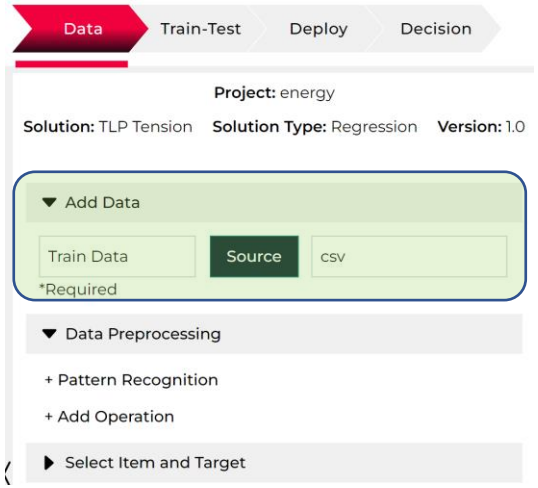


Tools:
In-house idareAI

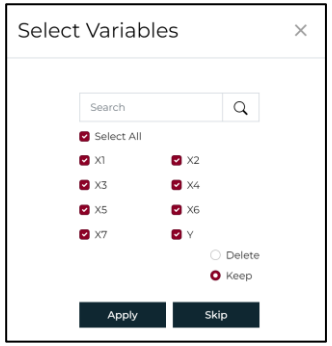
Decision

IDARE®

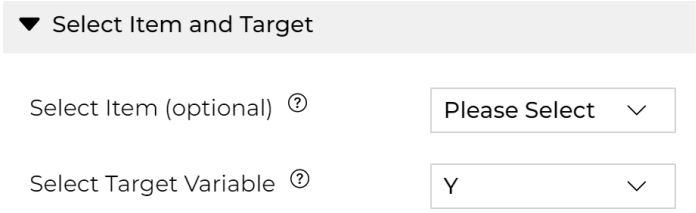
Data: Data Collection and Preprocess



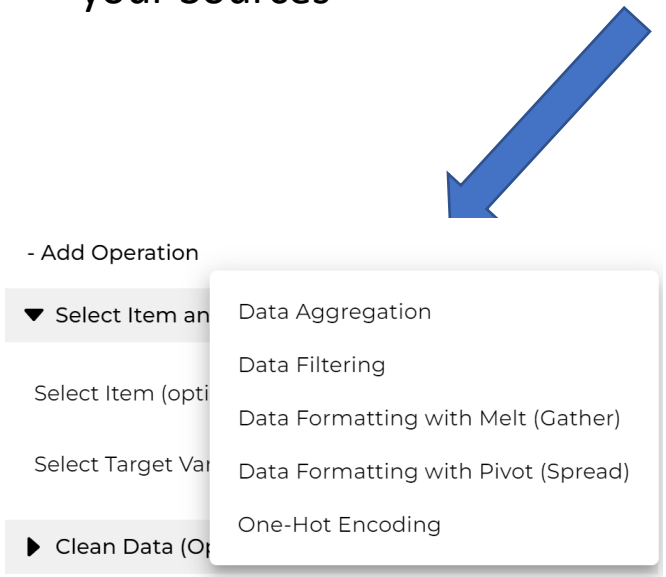
Read Data from your Sources



Remove Unnecessary Variable



- Select Item: if requires batch solution for different items i.e. location, assets, entity etc.
- Select Target: Choose the column or variable that needs to be predicted



Operate on Data

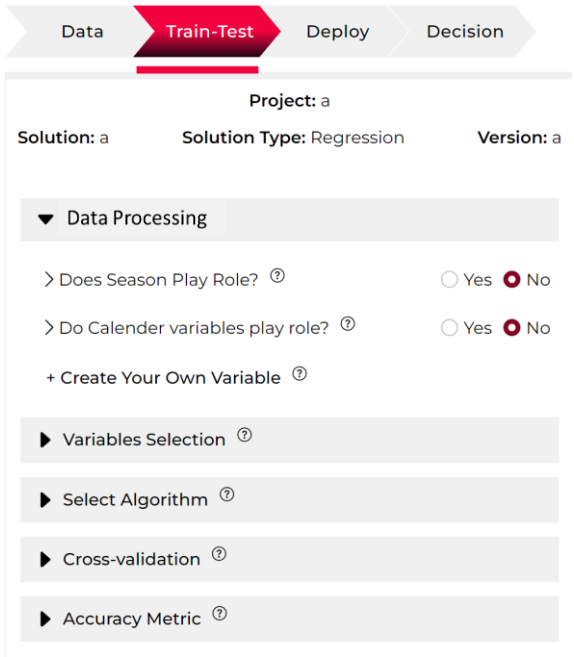
Train Test Split

X1	X2	X3	Y
A	D	9	30
B	D	3	30
C	D	4	40

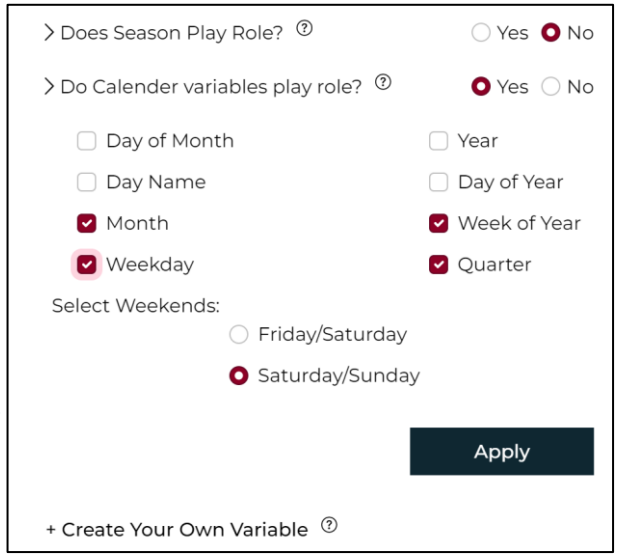
Feature Data

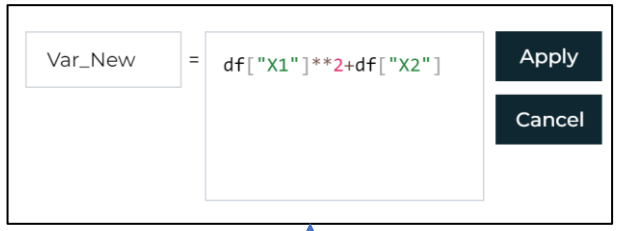
Model Building: Training Testing

IDARE®

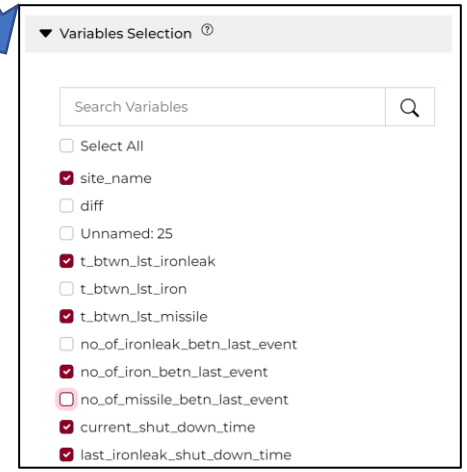


Work on your Variable

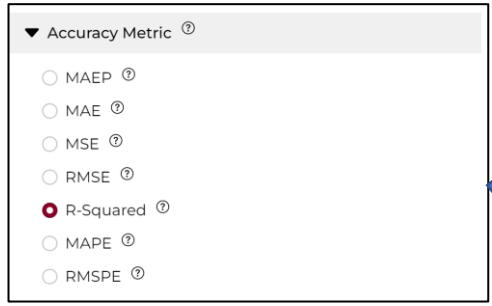




Validation

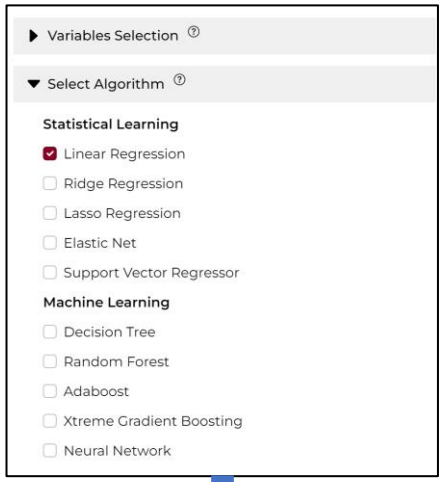


Select Unselect Variable



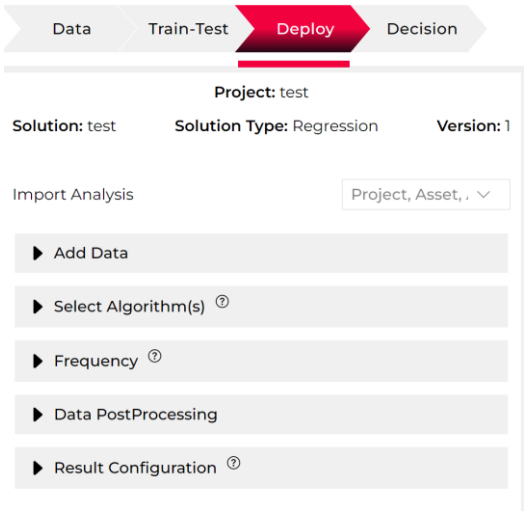
Choose Accuracy Metric

Choose Algorithm

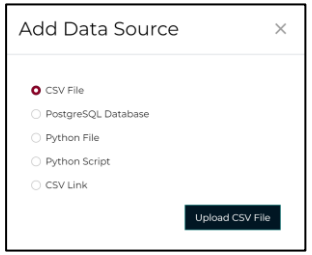


Finalize

Deploy: Prediction in Production

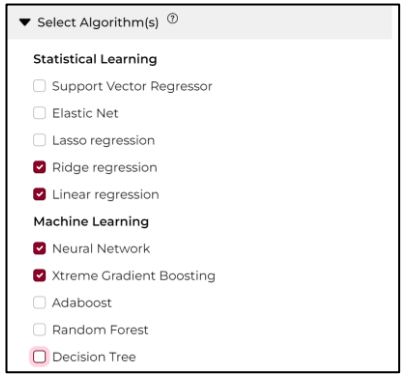


Read Production Data

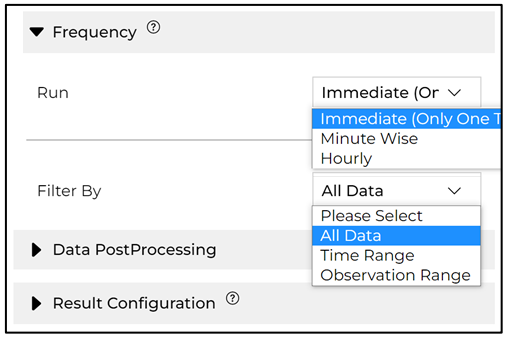


Perform all the data processing to create Feature Data

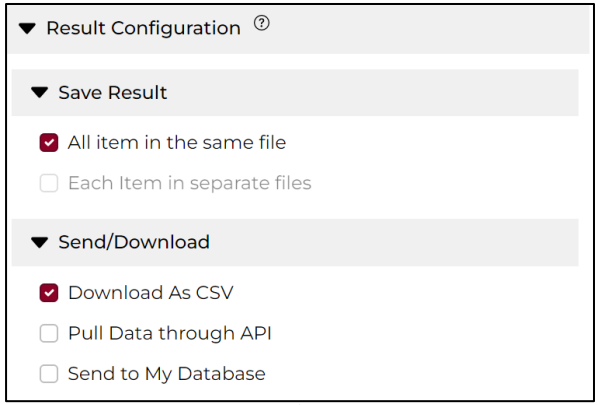
Select Trained ML



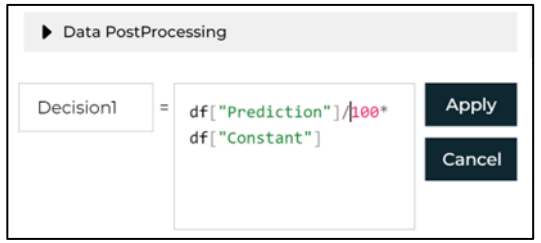
Set Production Mode Frequency



Configure retrieval of result data



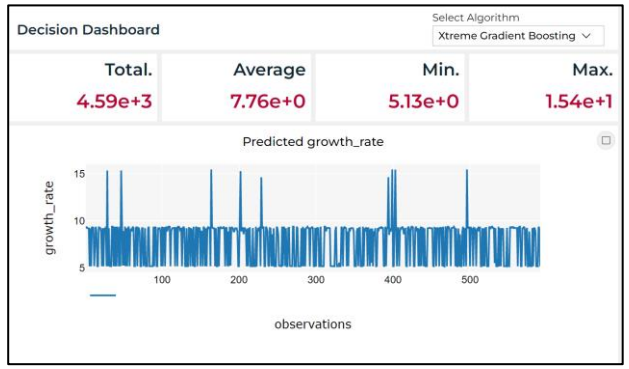
Create Decision Variable by Post-Processing Results



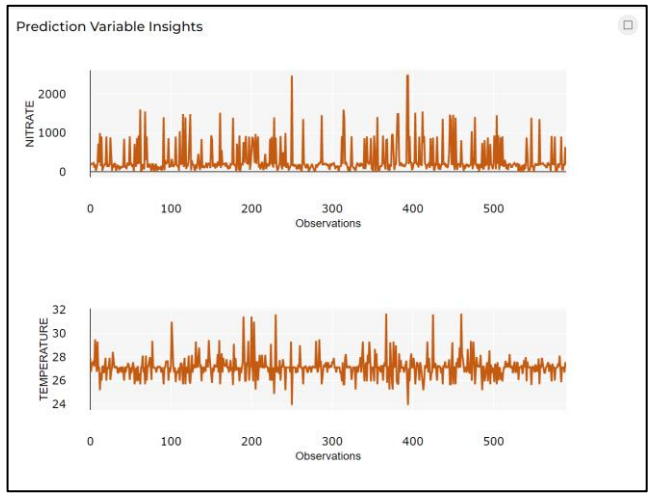
Predict

Decision

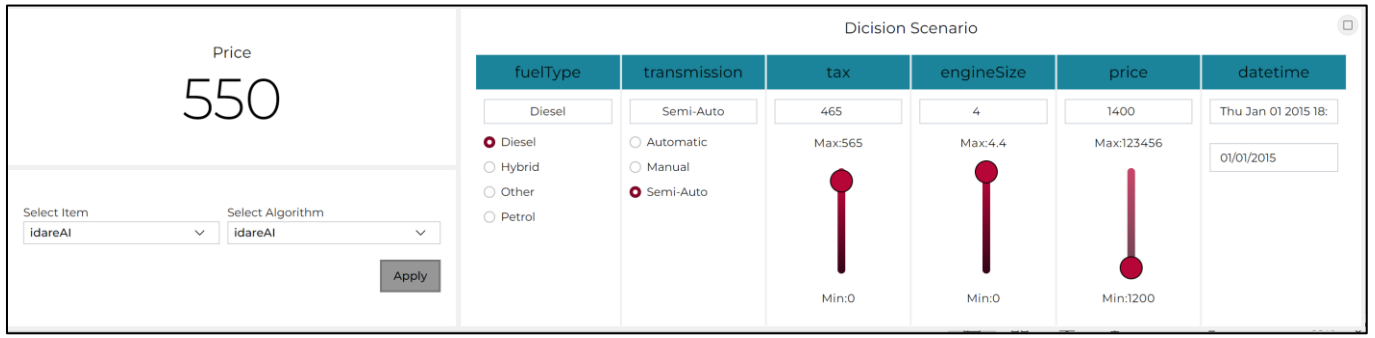
Visualize your Prediction Result

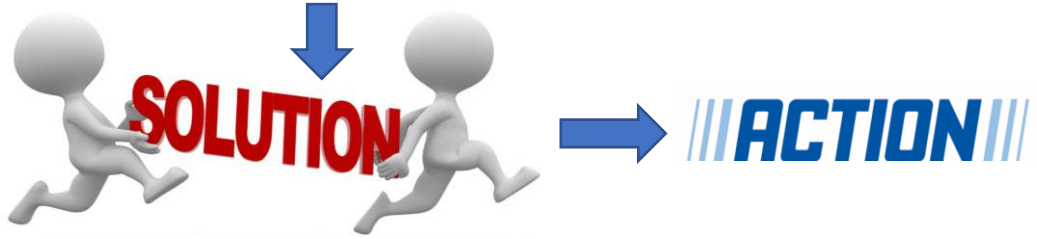


Visualize your data behind the Prediction



- Obtain Different Decision scenario by what if analysis
- Change different value of your production data to check what will be the prediction and associated decision variable





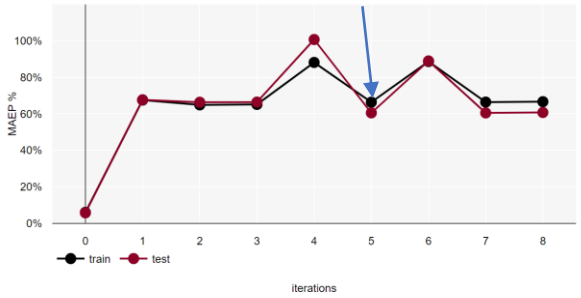
Successful Model Building: Variable/Feature Selection Strategy

IDARE®

Your target will be Lowest Error with Highest Stability with optimal bias and minimum variances.
Select the variable for

- the lowest error AND
- lowest differences between train and test error.

Min. Error with
Min. Diff betwn Train-Test



1 ML Only

Start with the most important variables from your domain knowledge

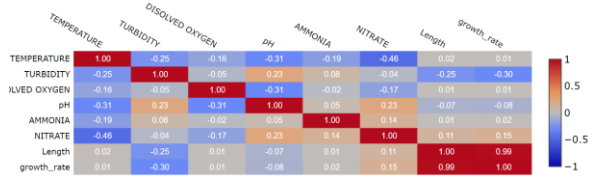
Train

Variables	Data Type	Missing Values Count	p value
TEMPERATURE	float	0	0
TURBIDITY	int	0	0
DISOLVED OXYGEN	float	0	0.27
pH	float	0	0
AMMONIA	float	0	0.08
NITRATE	int	0	0
Length	float	0	0
growth_rate	float	0	

- Add variables with p-value < 0.05
- Add variable with p-value > 0.05 if needed

1

Train

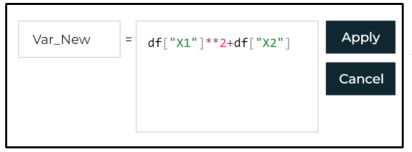


- Add variables Pearson correlation coefficient w.r.t Target more than 0.5
- Remove variable if a selected variables PCC w,r,t other variable > 0.3

2

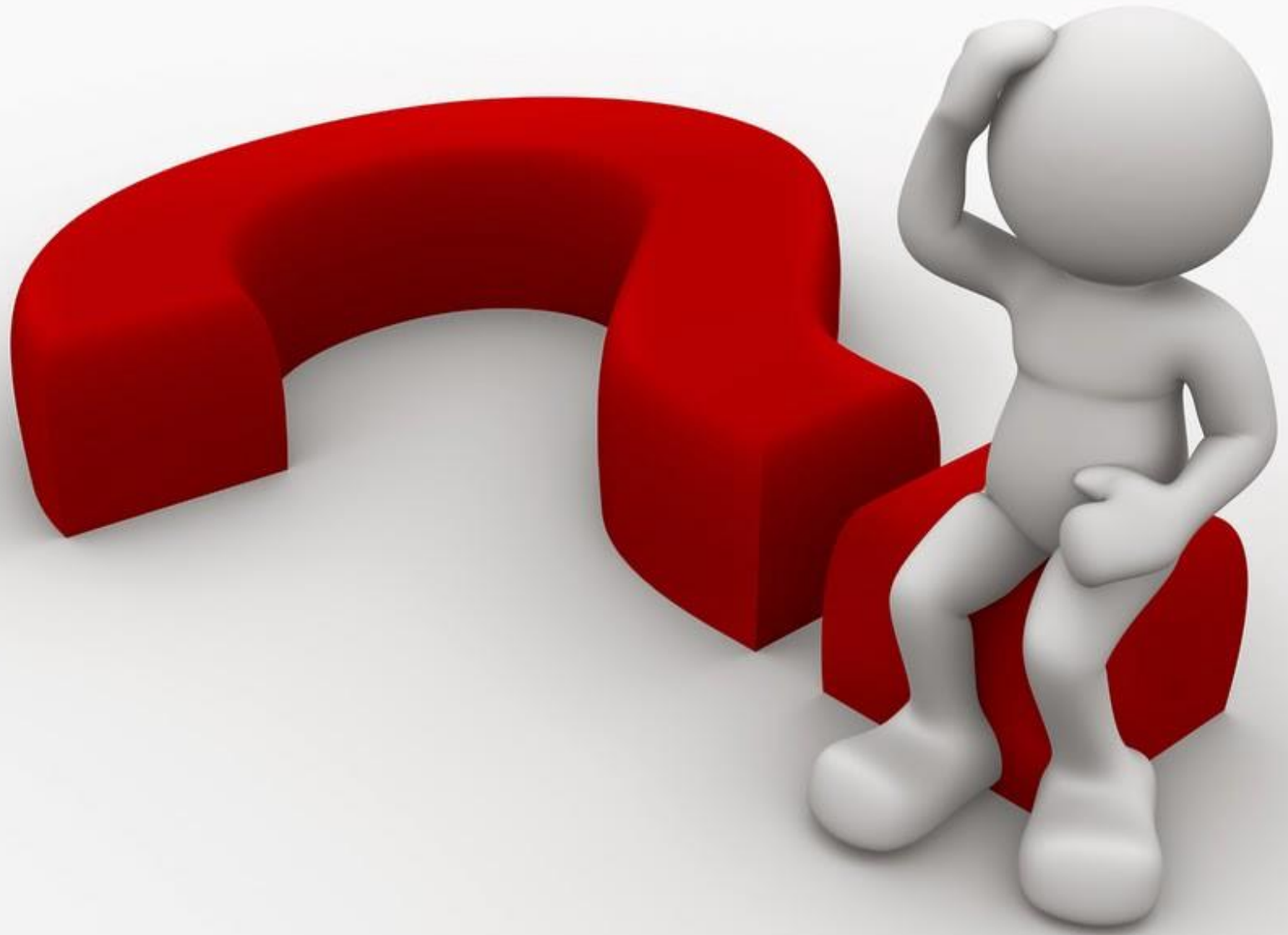
3

Train



Create Science Driven Variable

4

Repeat this process by many parametric study until reaching the desired accuracy

IDARE®