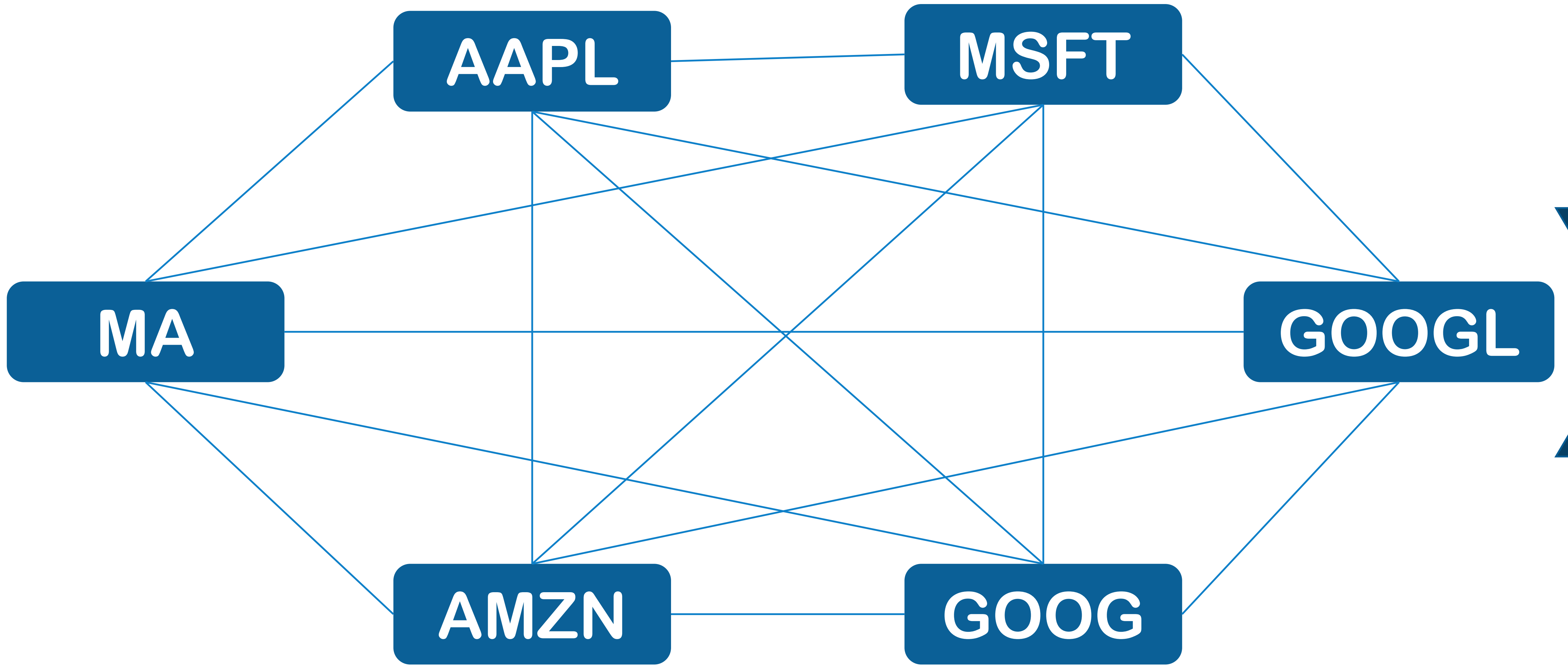
Market Intelligence Overview

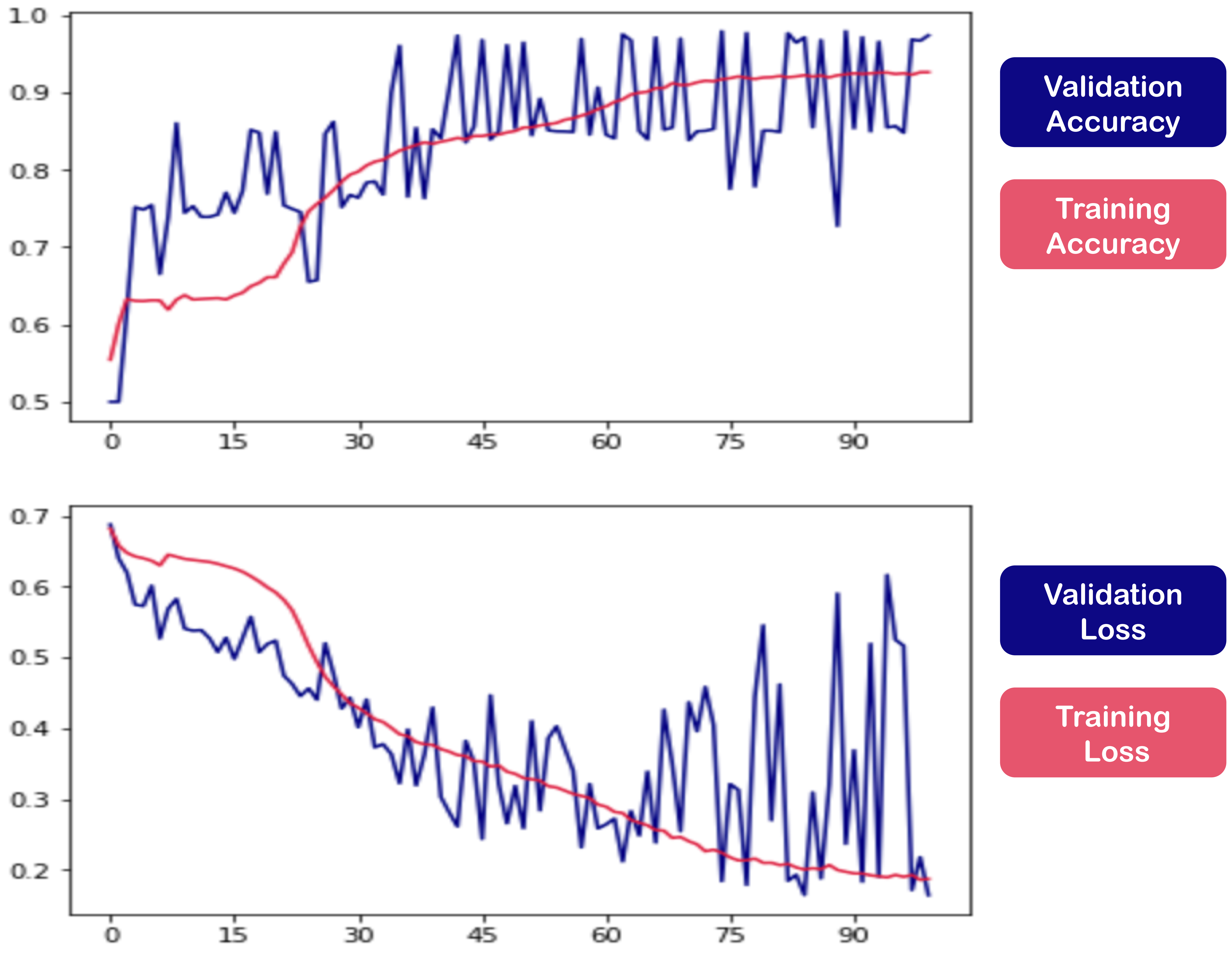
Donggeun Kim

0. Motivation

Can AML predict extreme market outcomes such as the 2008 Financial Crisis?

1. Overview





From feature selection to AML pipelines

2.1 Feature Selection – Bayesian Network

Track most relevant stocks by means of edge contributions or search for the strongest cointegrating relationship.

2.2 Feature Selection – Discerning Explanatory Variables

Figures on diagonals represent return distributions of LNT, AEP and AEE, separated by labels based on realization of LNT returns

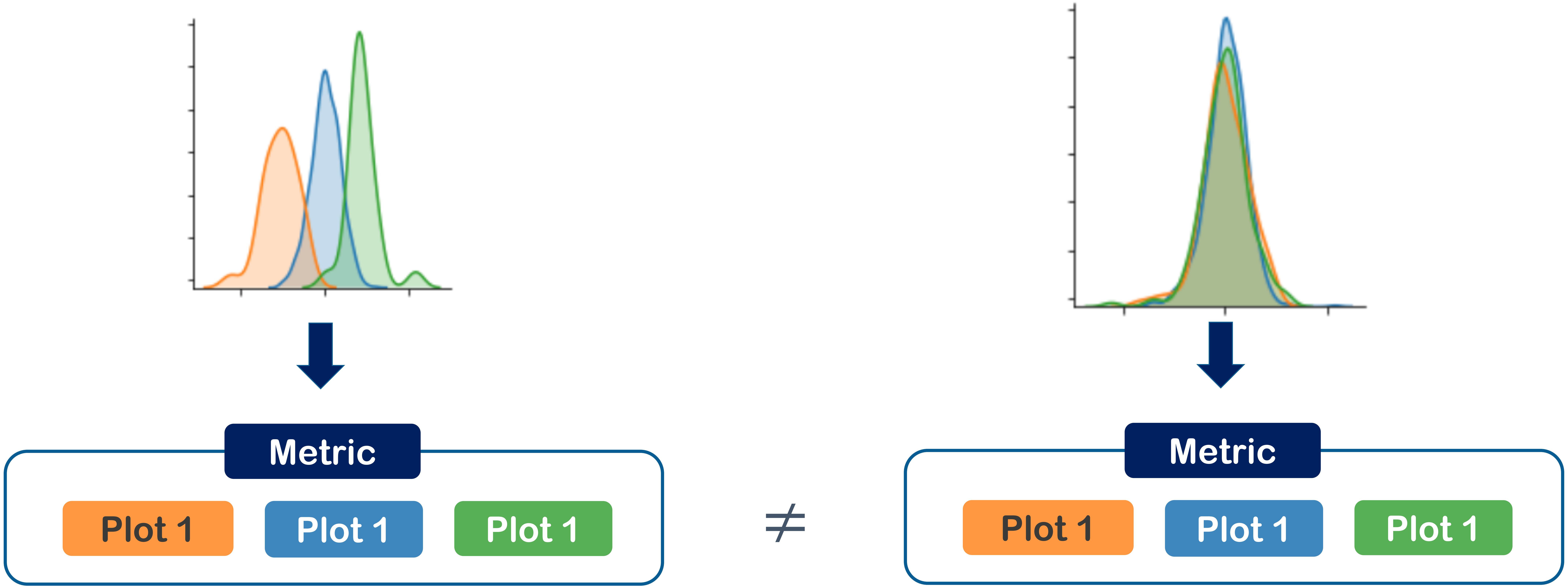
2.2 Feature Selection – Discerning Explanatory Variables

Characteristic of discerning distributions: Distributions are well separated by target labels

2.2 Feature Selection – Non-Discerning Explanatory Variables

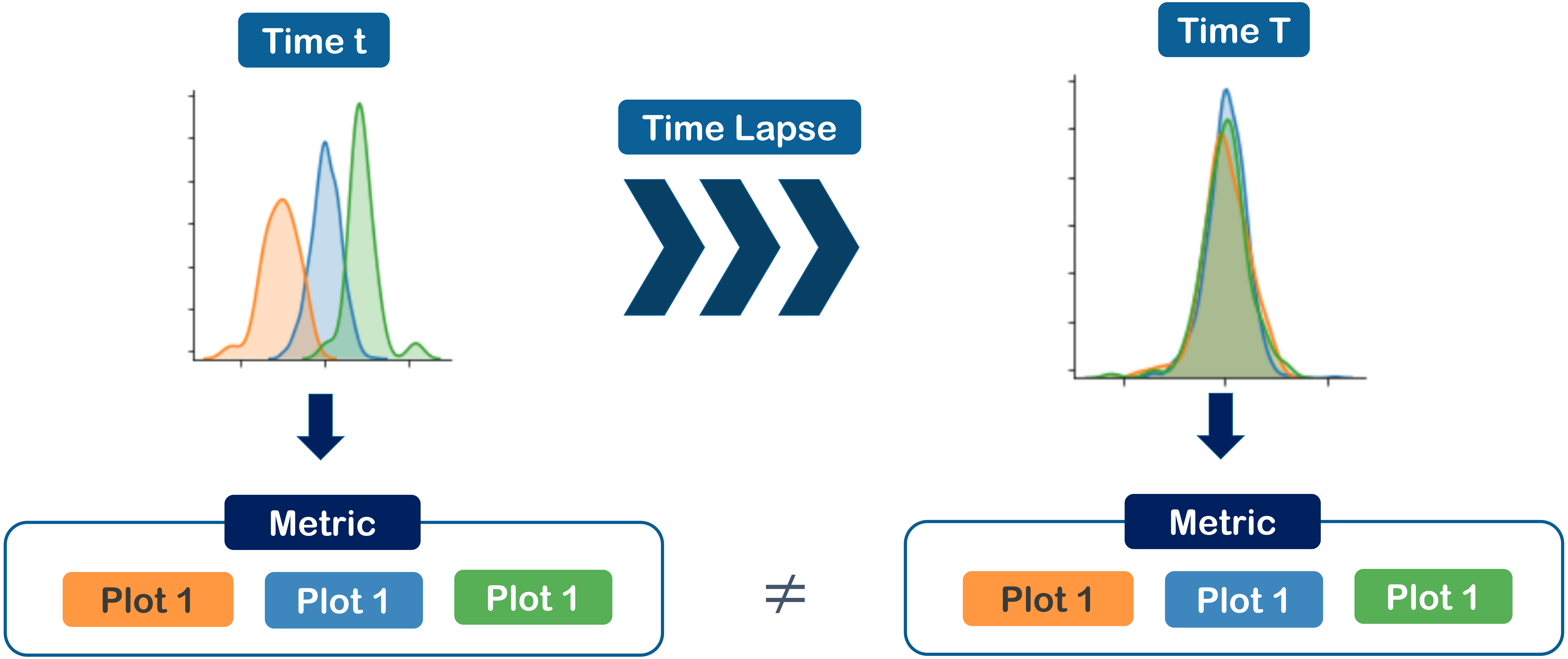
Characteristic of Non-Discerning Explanatory Variables: Distributions cannot be isolated by target labels
Less contributions from these variables

2.3 Feature Selection – Auto-selection



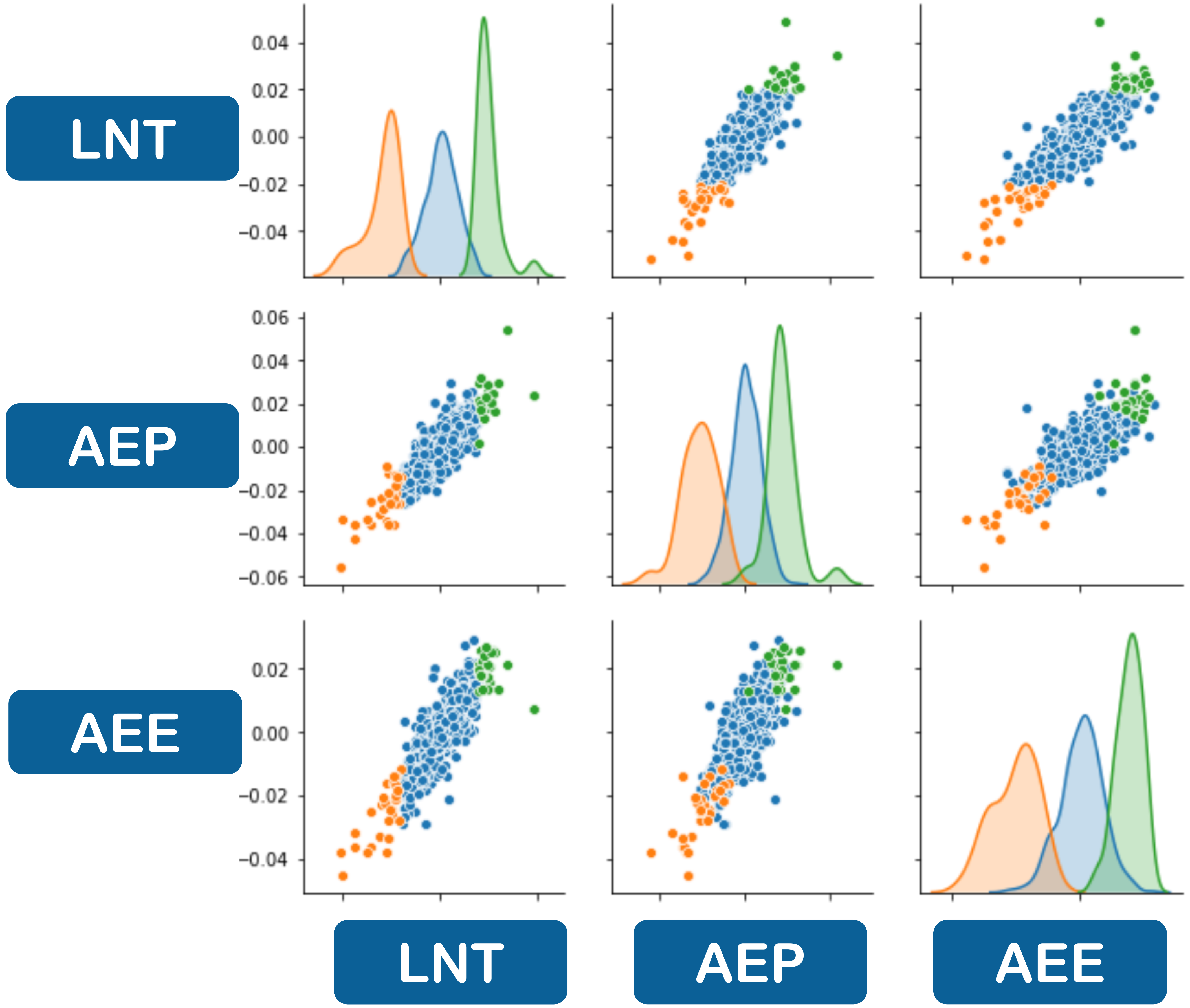
Auto-selection should prefer differentiated target-label-induced distributions such as plot 1 relative to plot 2.
Candidate metrics include Kullback-Leibler divergence or Two-sample Kolmogorov-Smirnov test.
Not referring to a metric induced by a metric space

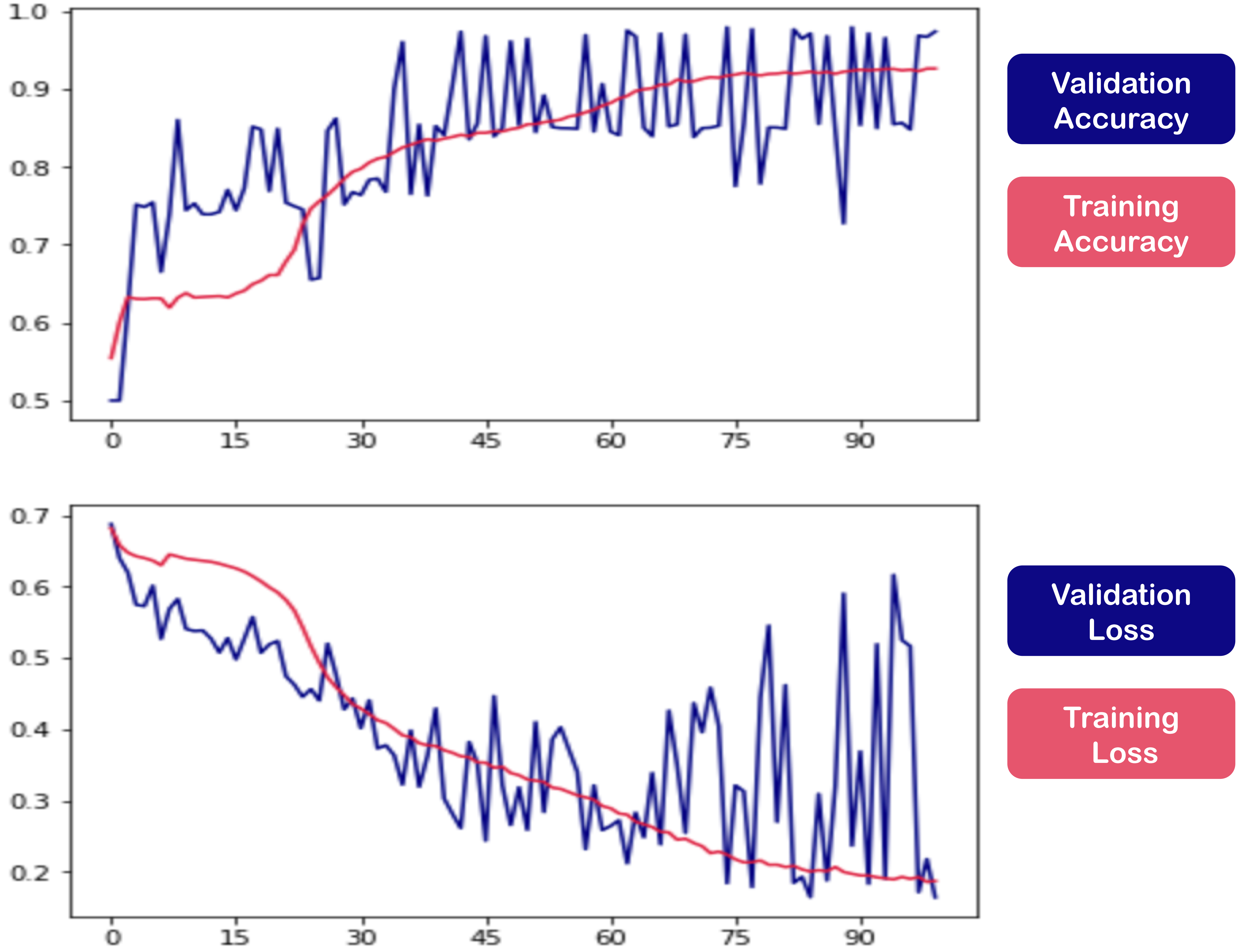
2.4 Feature Selection – Non-stationarity



Nonstationary data may result in changes in target-label-induced distributions, violating modeling assumptions.
Use metrics to compare recent realization of distributions with distributions used during training process.
For example, Kullback-Leibler divergence computes information gain by making use of recently realized distributions.

3. Modeling



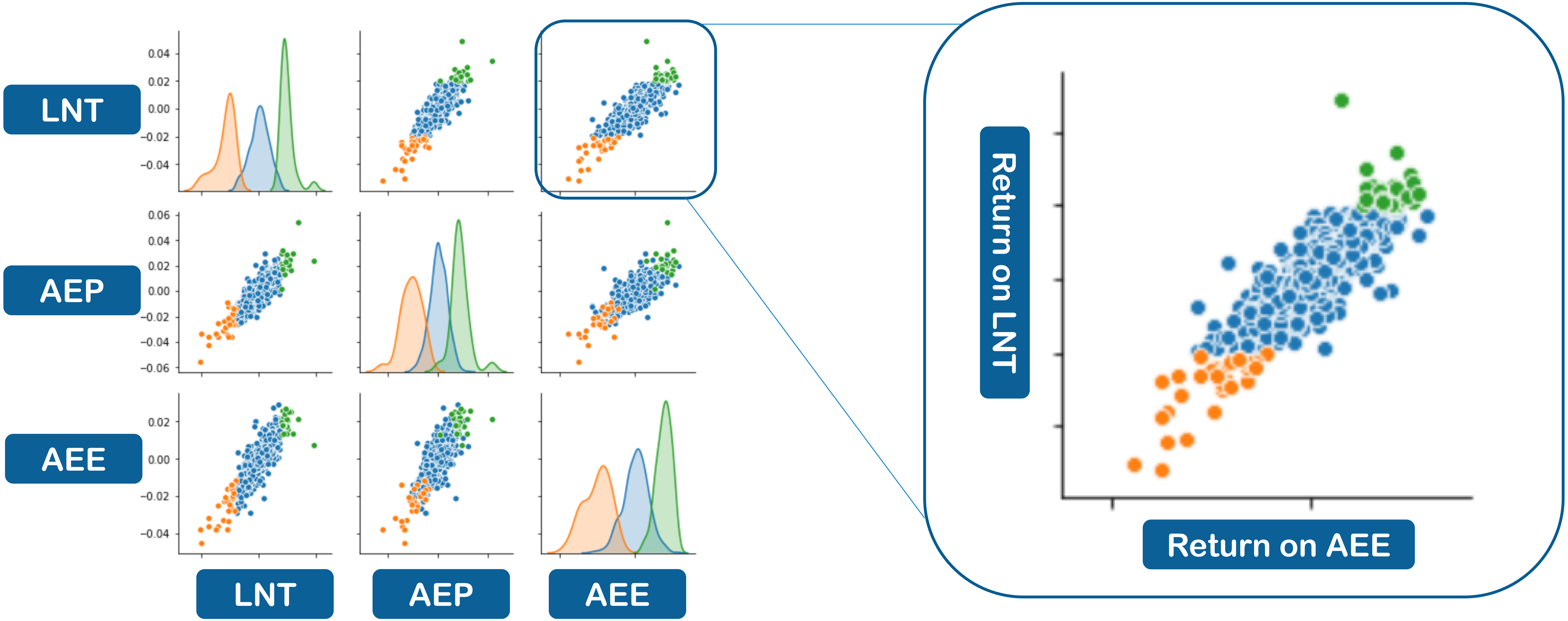


Use the best set of features for a variety of modeling approaches

3.1 Modeling – Group Return Prediction

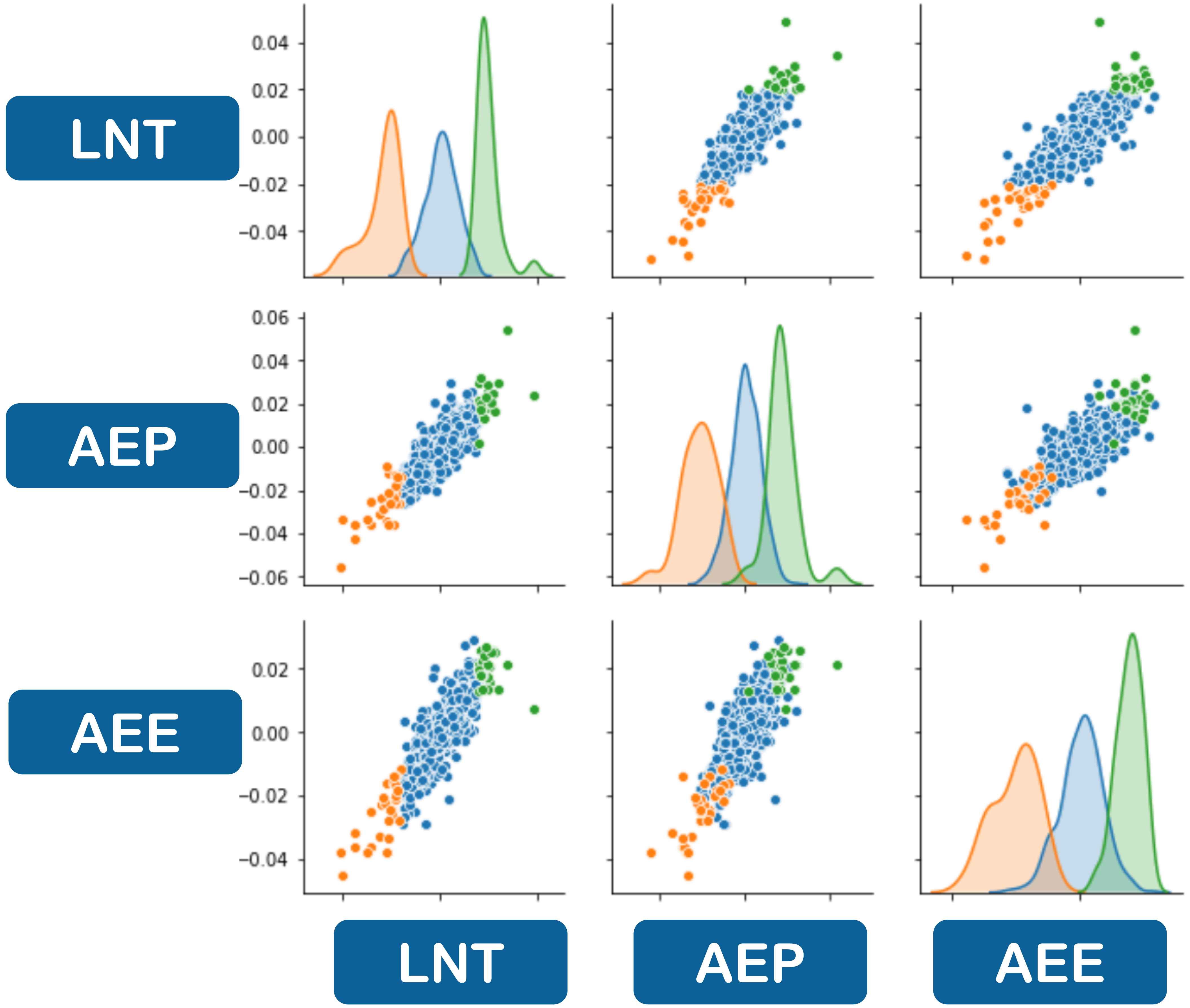
Observation 0:
When LNT return is less than -2%, return on AEP is always negative.
When LNT return is greater than 2%, return on AEP is always positive

3.1 Modeling – Group Return Prediction



Observation 1: high correlations for stocks within this particular group

3.1 Modeling – Group Return Prediction



Stock Specific Models

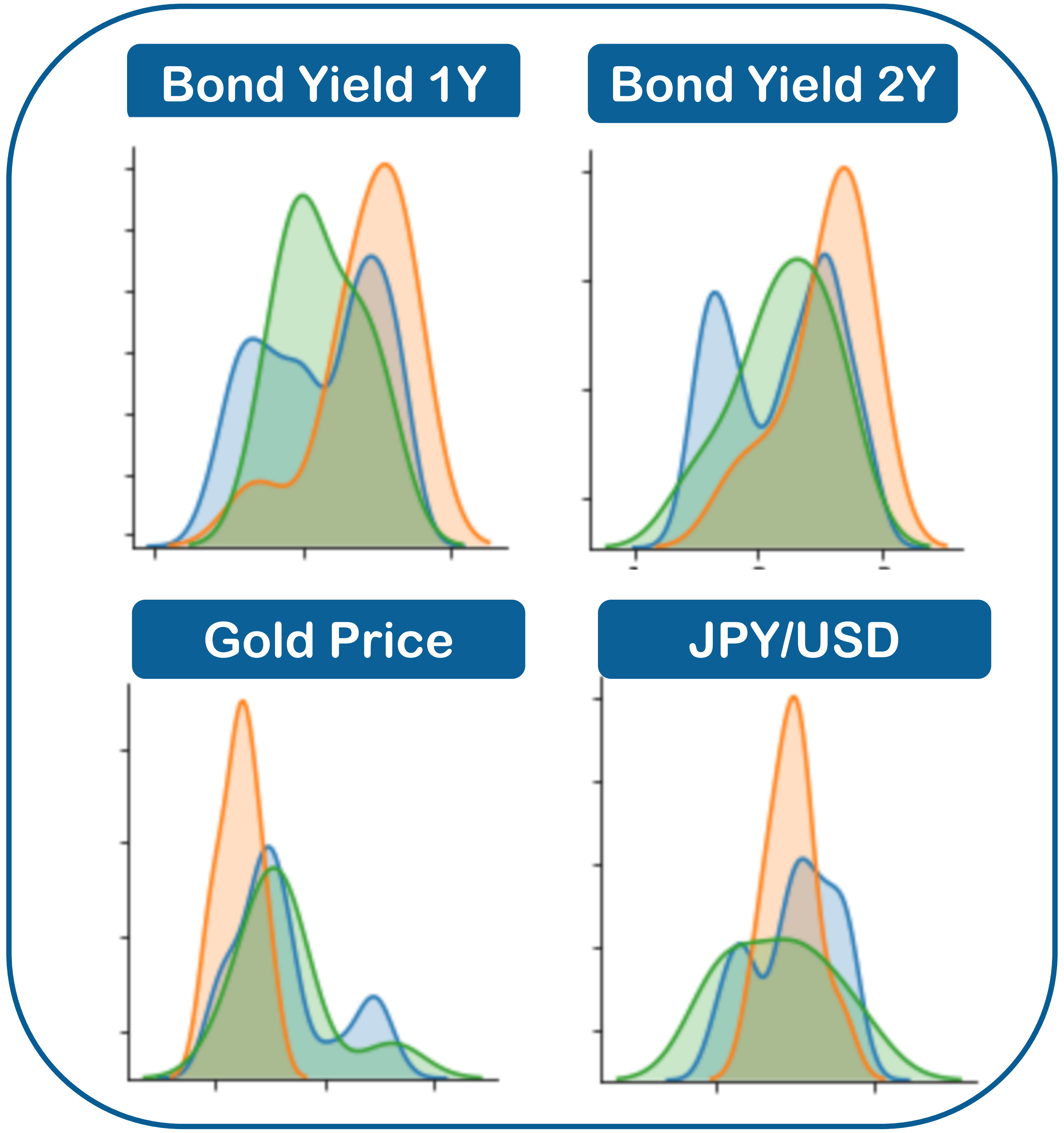
Model 1	LNT Return Prediction
Model 2	AEP Return Prediction
Model 3	AEE Return Prediction

Universal Model

Model 1	LNT, AEP, AEE Return Prediction

Observations give possibility of stock-specific modeling approaches AND locally universal modeling frameworks

3.2 Modeling – Predict Universal Labels



Universal Model for LNT, AEP, AEE Return Prediction

Model 1	Decision Tree
Model 2	AdaBoost
Model 3	QDA
Model 4	LSTM

Observations give possibility of stock-specific modeling approaches AND locally universal modeling frameworks

3.2 Modeling – Predict Universal Labels

Universal Model for LNT, AEP, AEE Return Prediction

Model 1	Decision Tree
Model 2	AdaBoost
Model 3	QDA
Model 4	LSTM

Predict Positive Return for the Group	
Precision	0%
Accuracy	49.71%
Predicted Positive Count	105

Predict Negative Return for the Group	
Precision	94.05%
Accuracy	61.74%
Predicted Negative Count	1682

When distribution can be well separated by target labels, there should exist an effective model

3.2 Modeling – Predict Universal Labels

Universal Model for LNT, AEP, AEE Return Prediction

Model 1	Decision Tree
Model 2	**AdaBoost**
Model 3	QDA
Model 4	LSTM

Predict Positive Return for the Group	
Precision	100%
Accuracy	62.74%
Predicted Positive Count	1536

Predict Negative Return for the Group	
Precision	82.47%
Accuracy	59.68%
Predicted Negative Count	1883

When distribution can be well separated by target labels, there should exist an effective model

3.2 Modeling – Predict Universal Labels

Universal Model for LNT, AEP, AEE Return Prediction

Model 1	Decision Tree
Model 2	AdaBoost
Model 3	**QDA**
Model 4	LSTM

Predict Positive Return for the Group	
Precision	0%
Accuracy	50.55%
Predicted Positive Count	0

Predict Negative Return for the Group	
Precision	76.80%
Accuracy	76.10%
Predicted Negative Count	6142

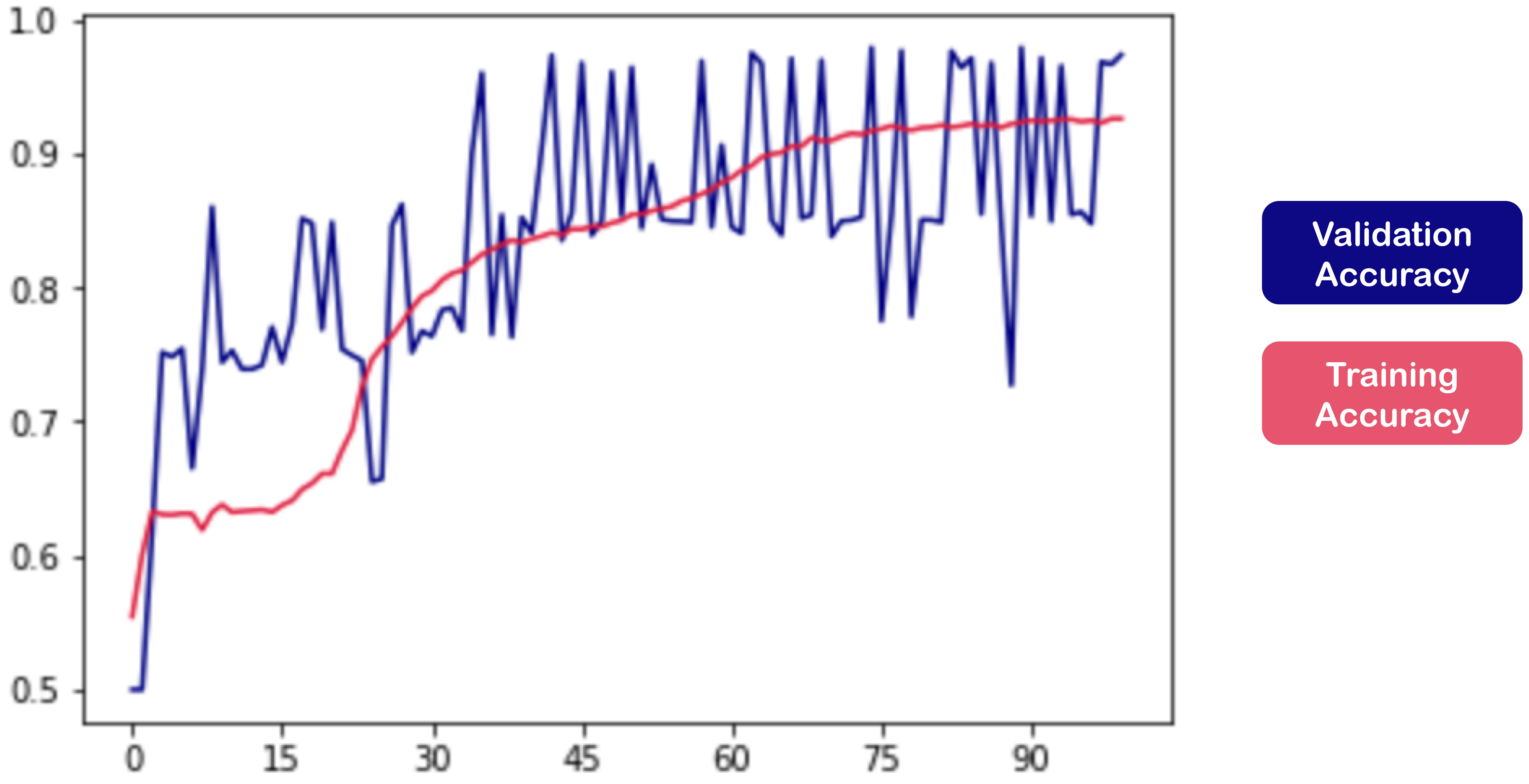
When distribution can be well separated by target labels, there should exist an effective model

3.2 Modeling – Predict Universal Labels

Universal Model for LNT, AEP, AEE Return Prediction

Model 1	Decision Tree
Model 2	AdaBoost
Model 3	QDA
Model 4	**LSTM**

Predict Positive Return for the Group



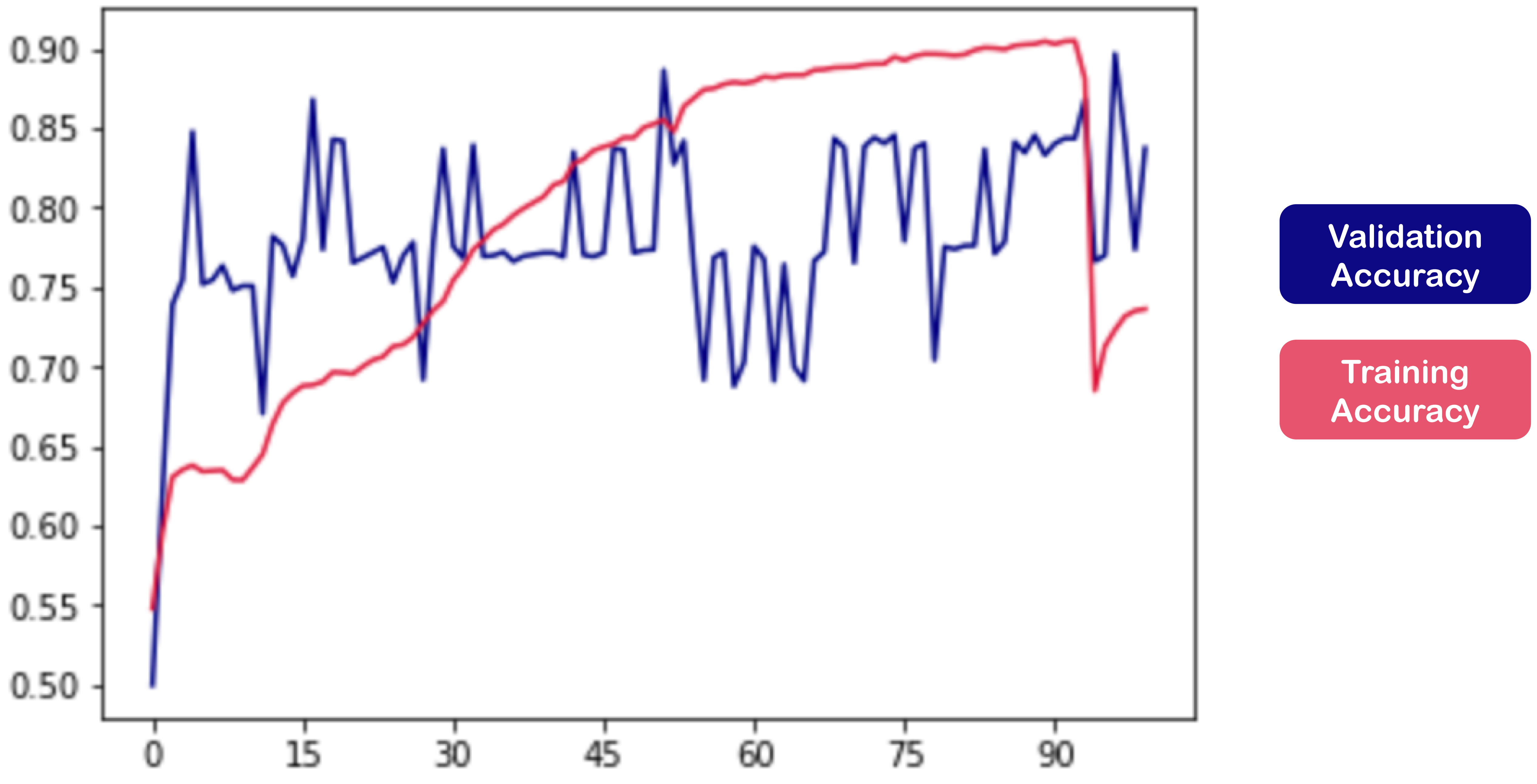
When distribution can be well separated by target labels, there should exist an effective model

3.2 Modeling – Predict Universal Labels

Universal Model for LNT, AEP, AEE Return Prediction

Model 1	Decision Tree
Model 2	AdaBoost
Model 3	QDA
Model 4	**LSTM**

Predict Negative Return for the Group



When distribution can be well separated by target labels, there should exist an effective model