



# Supervised Learning for Stock Trading

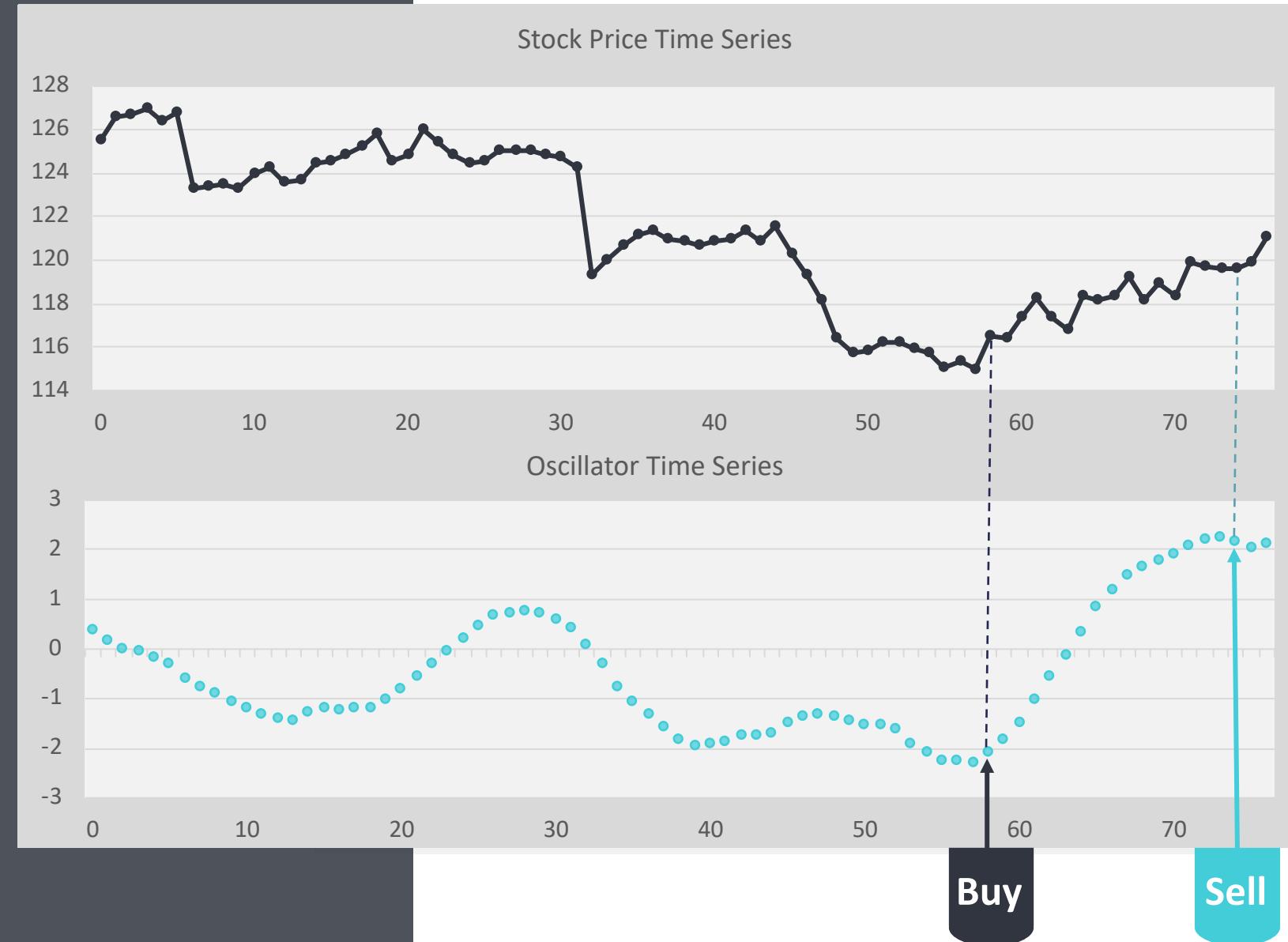
Seahorse Strategies

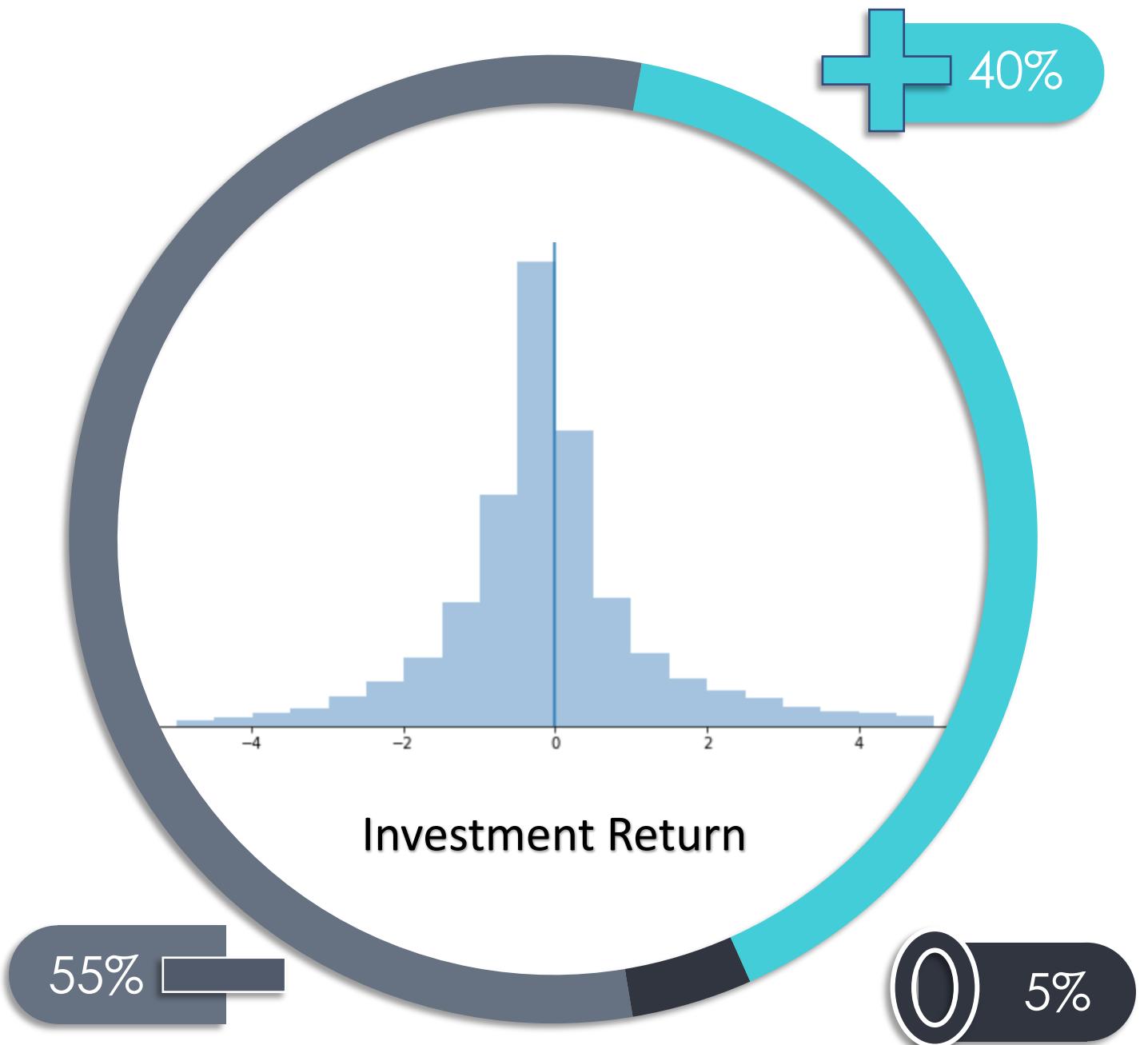
**Team Member:** Simon Chiu, Gilbert Lei, Fan Wu, Linyang Yu

**Mentor:** Rodolfo Lourenzutti

# Intro to Seahorse Strategies

- Toronto stock trading firm
- Oscillator (Tech. indicator)
- Gives buy and sell signals
- 30-minute time interval





# Seahorse's Needs



Increase Return From Investment



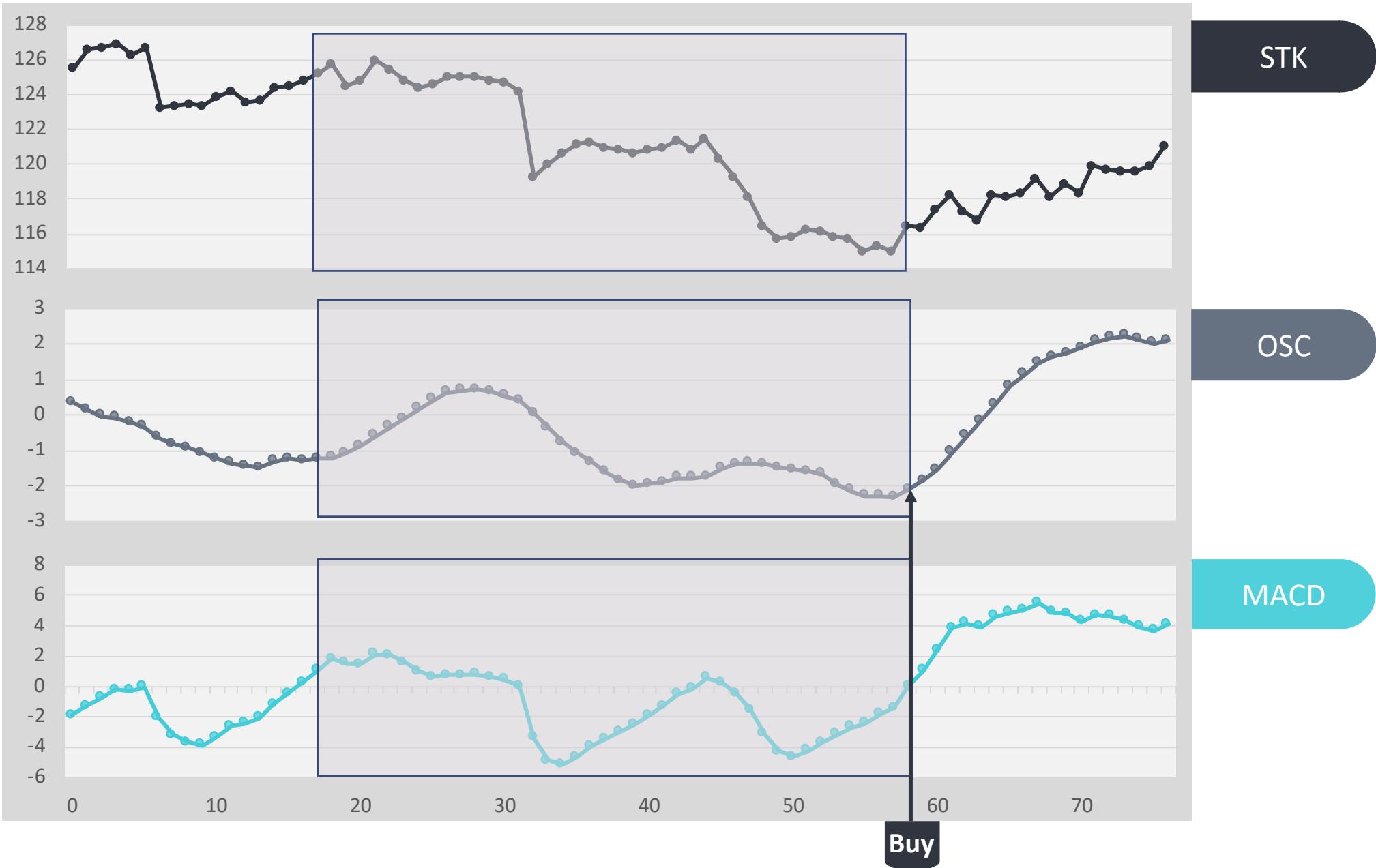
Minimize Human Judgement for  
Investment Decisions



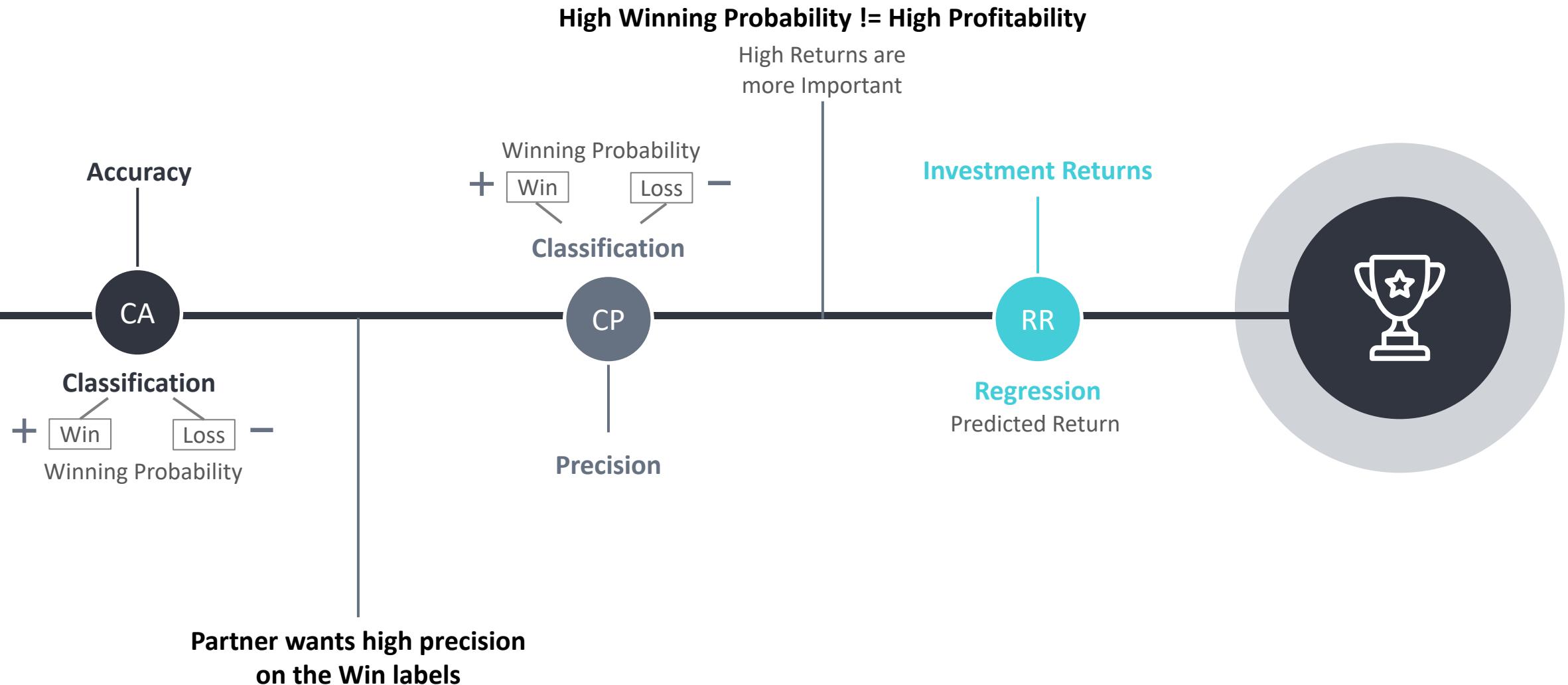
More Profitable Opportunities



# Original Features



# Objectives



# Our Footprints

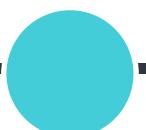
Classification



Classification

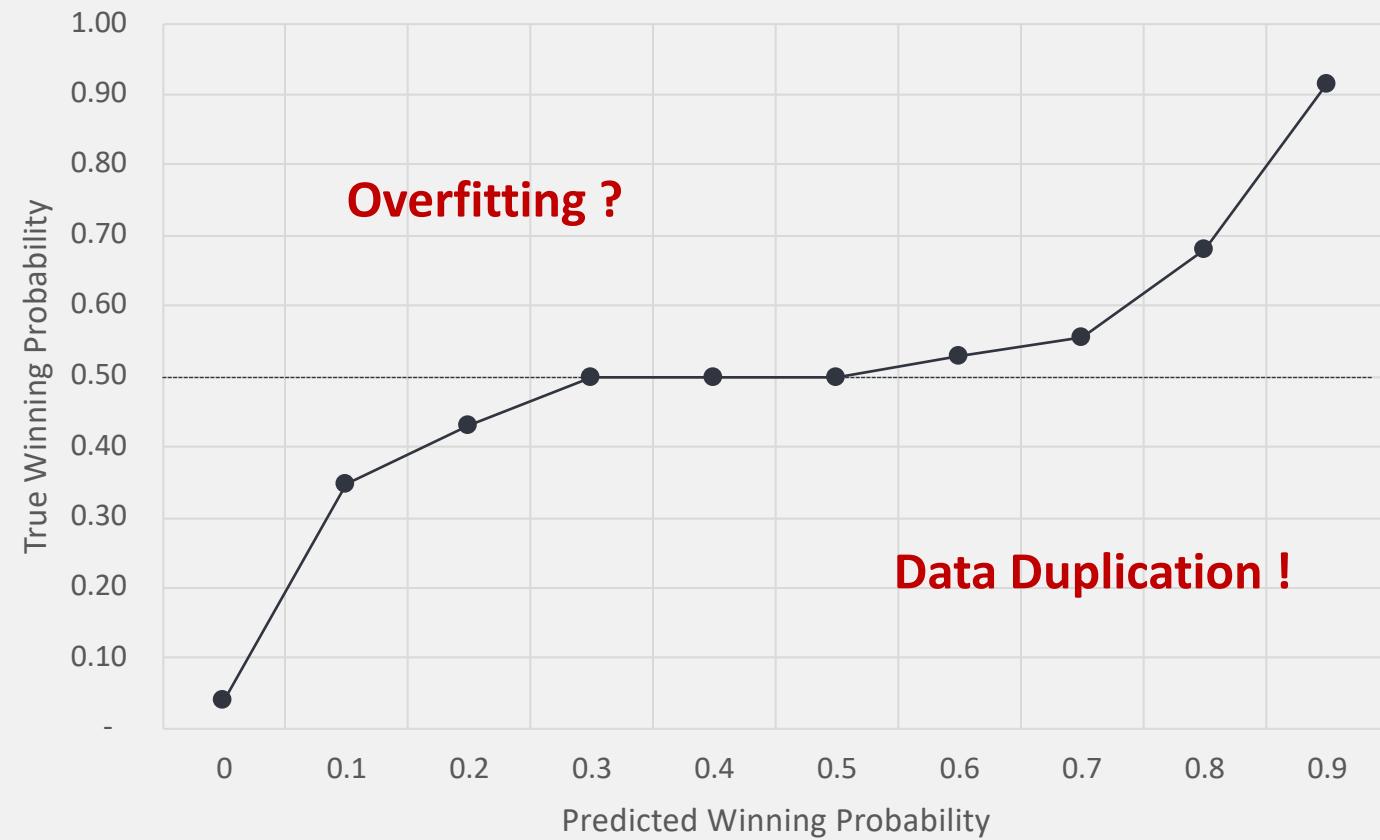


Regression



## Basic Model Exploring:

- Logistic Regression
- Random Forest
- SVC
- KNN
- Bernoulli Naïve Bayes



# Our Footprints

Classification



Classification

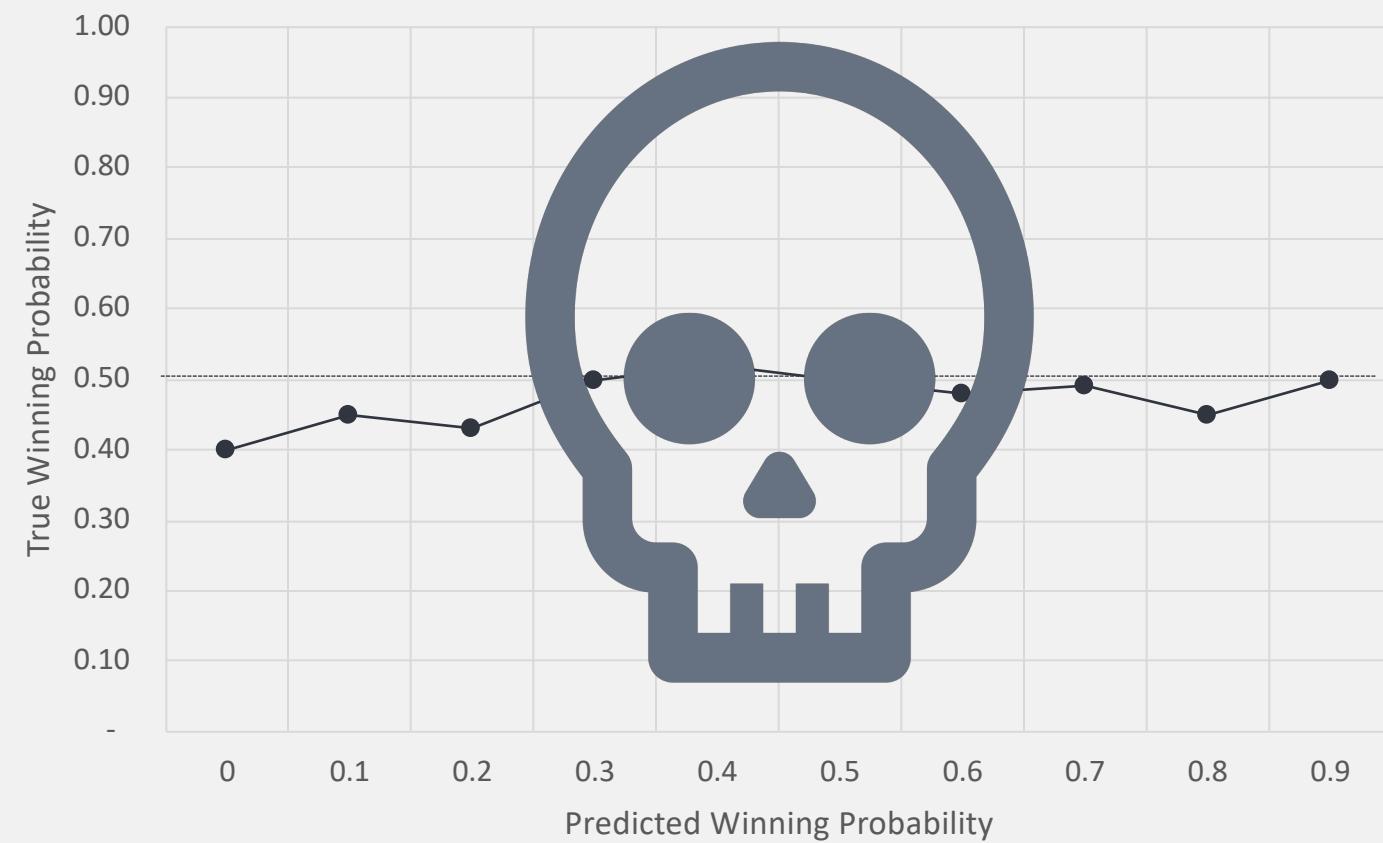


Regression



## Basic Model Exploring:

- Logistic Regression
- Random Forest
- SVC
- KNN
- Bernoulli Naïve Bayes



# Our Footprints

Classification



Classification

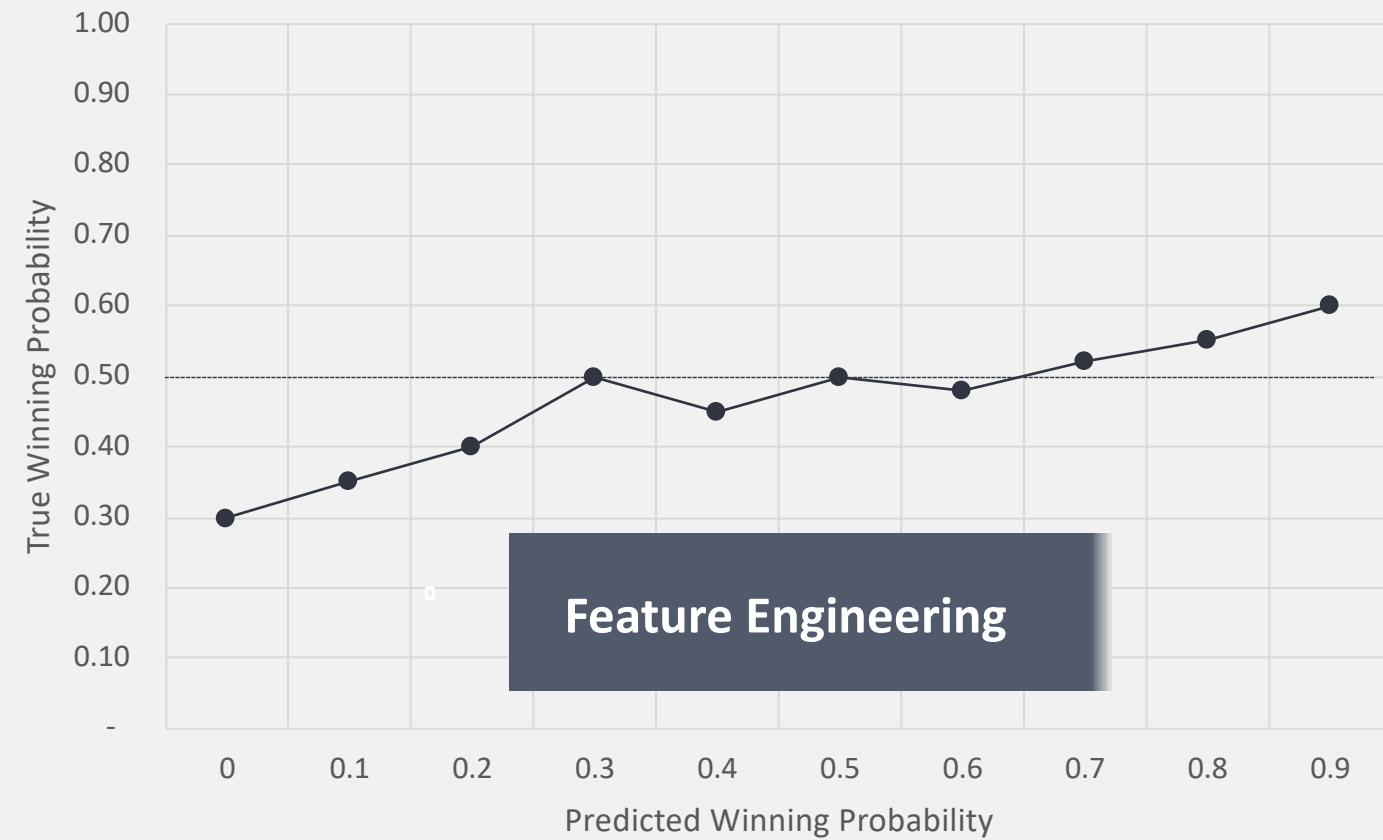


Regression



## Basic Model Exploring:

- Logistic Regression
- Random Forest
- SVC
- KNN
- Bernoulli Naïve Bayes



# Our Footprints

Classification



Classification



Regression



## Basic Model Exploring:

- Logistic Regression
- Random Forest
- SVC
- KNN
- Bernoulli Naïve Bayes

## Feature Engineering

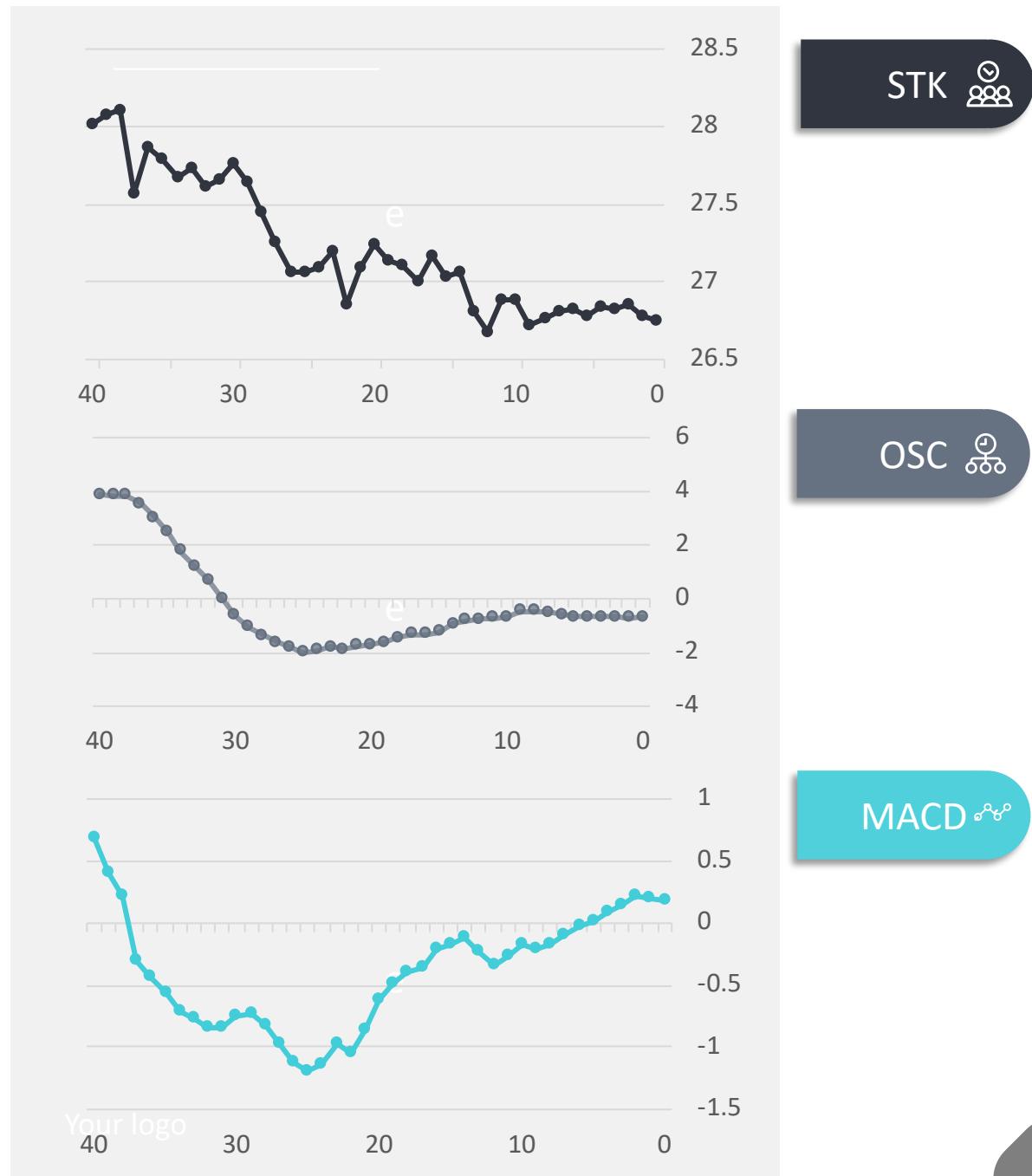
## LightGBM Regression



## Continue Exploring:

- LightGBM Classifier
- CNN → Ensembled CNN
- RNN for time series

# Feature Engineering

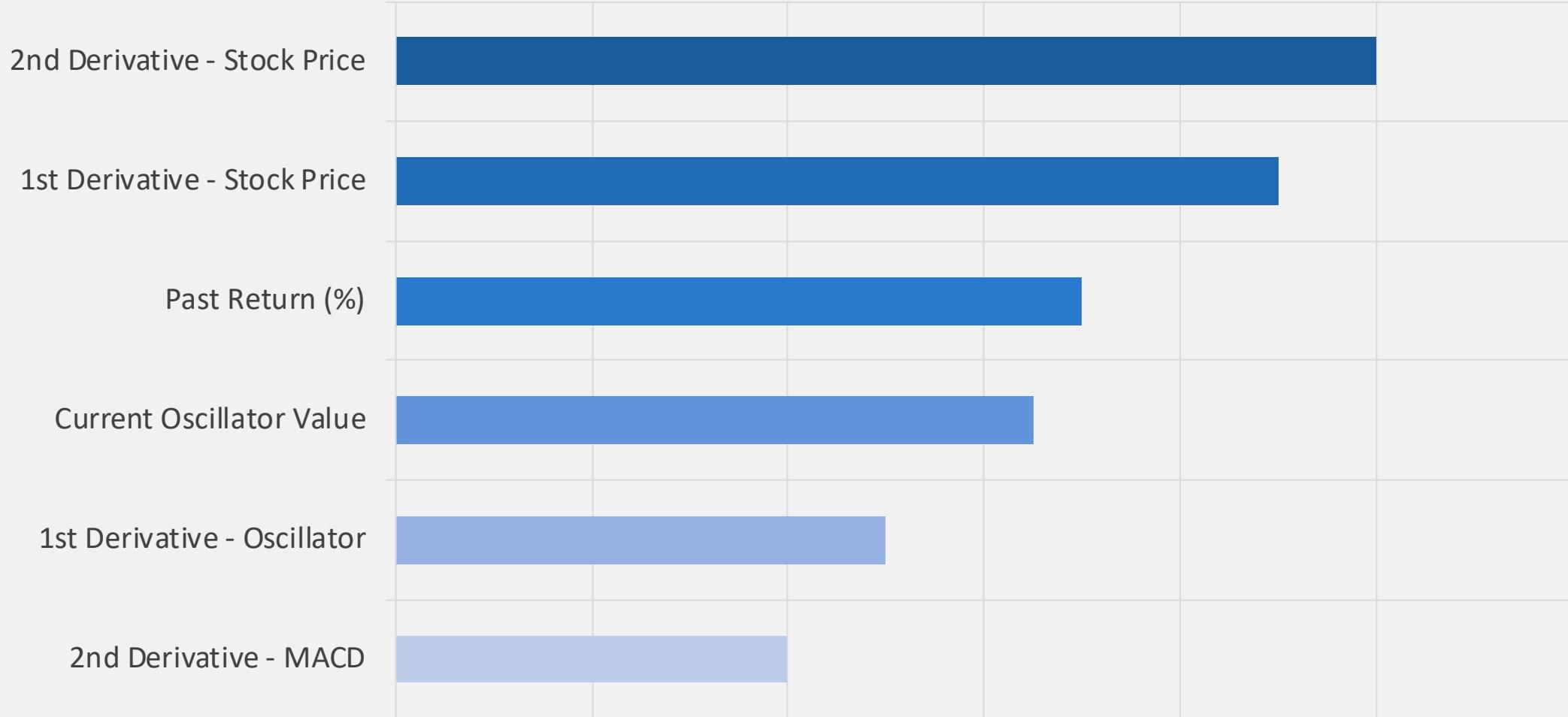


STK 🧑

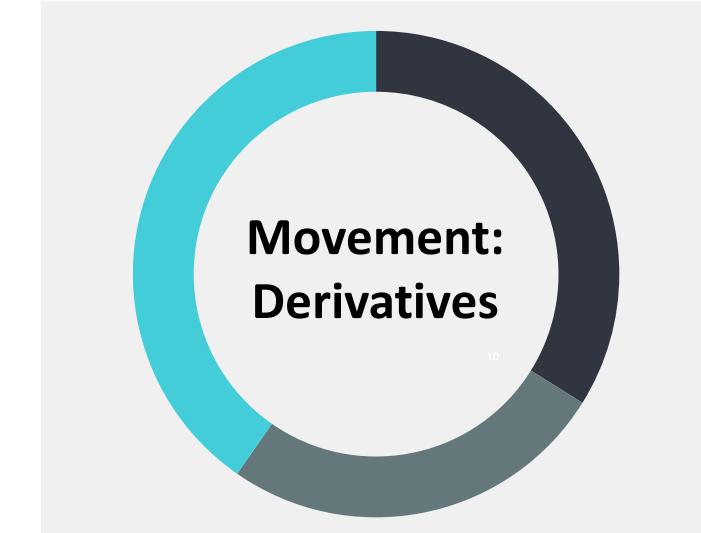
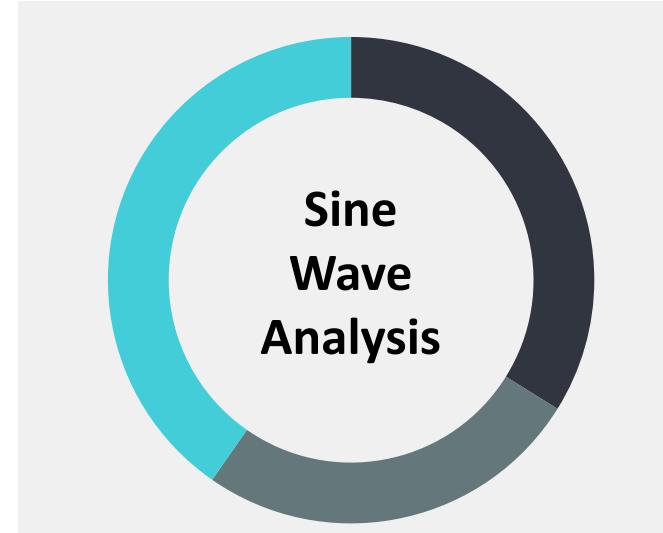
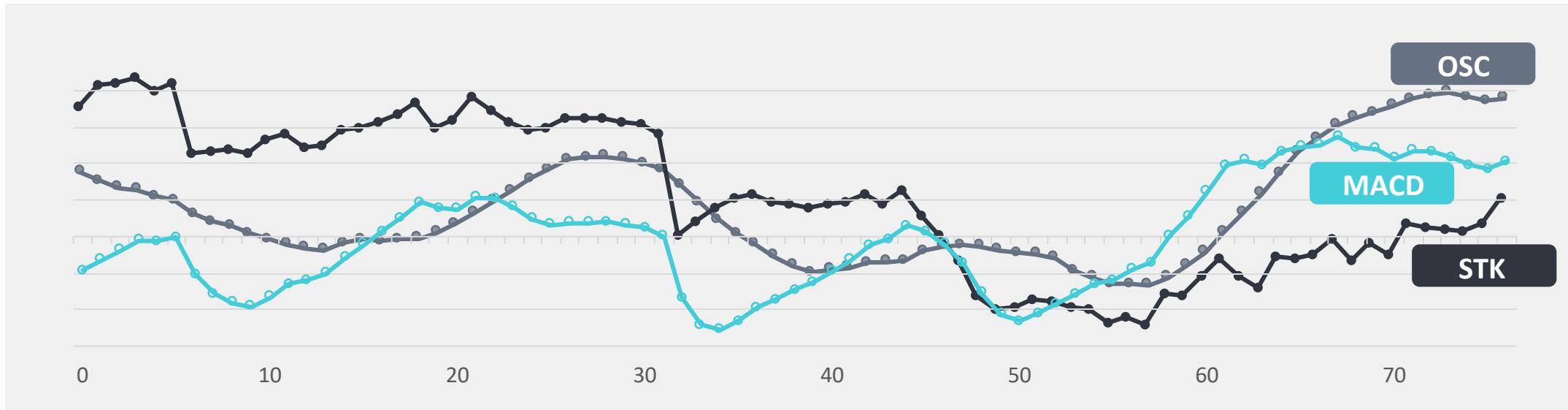
OSC ♂

MACD ⚡⚡

# Top Features



# Features Engineering

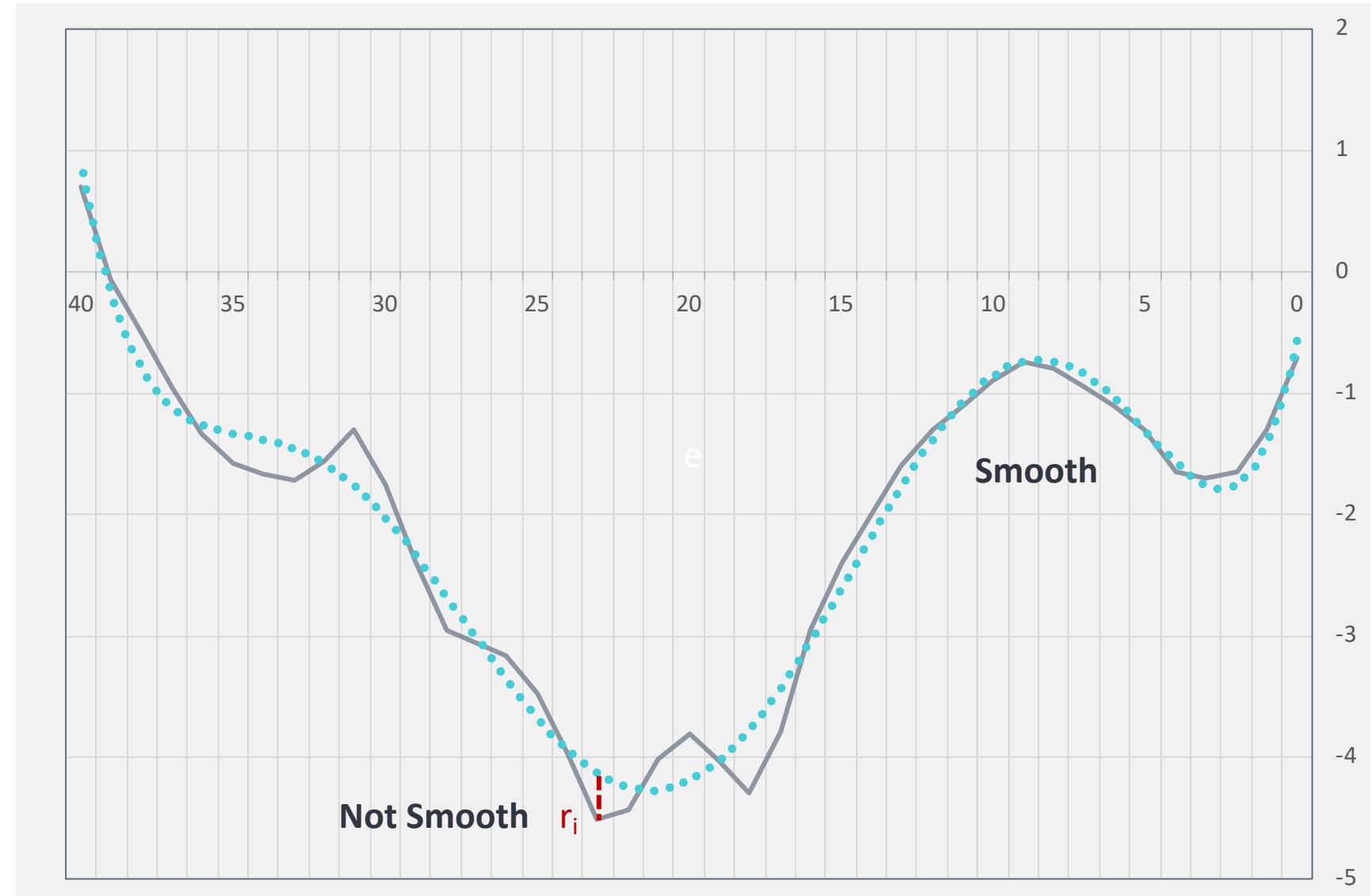


# Smoothness



Residual Method

$$S = \sum_{i=1}^n |r_i|$$

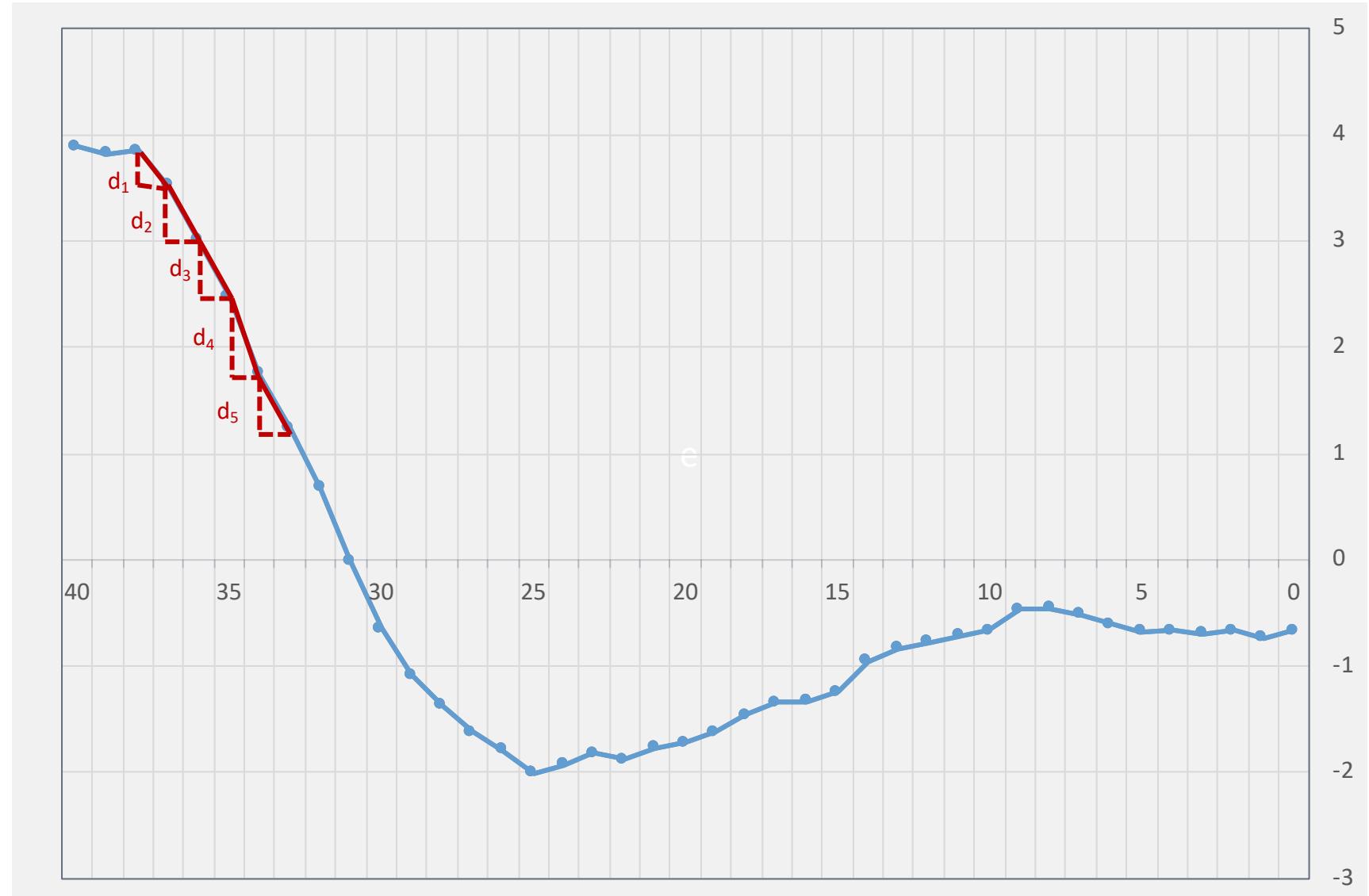


# Smoothness

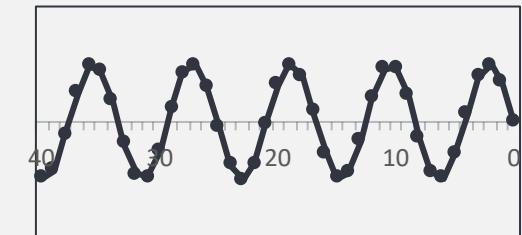
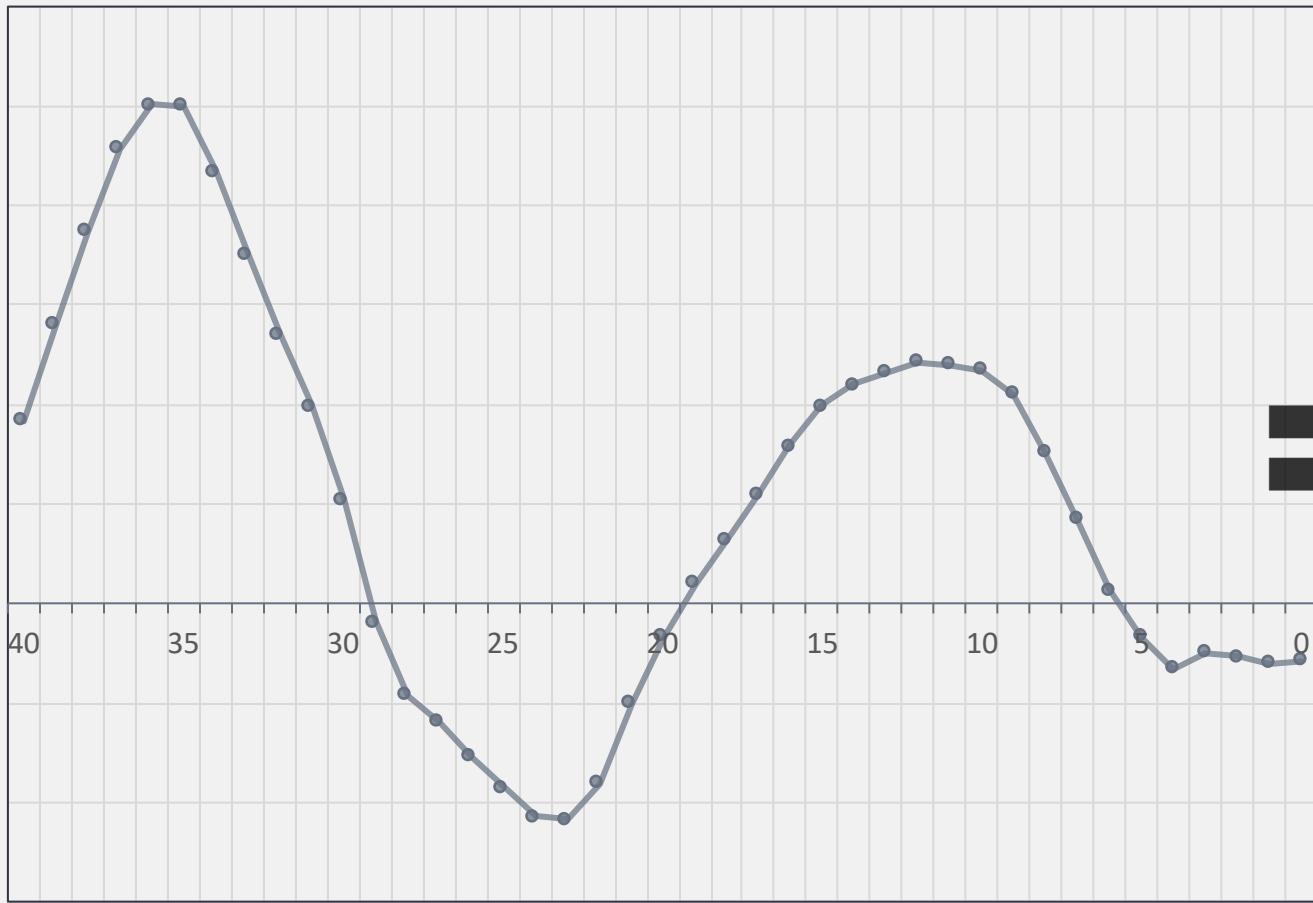


Differencing Method

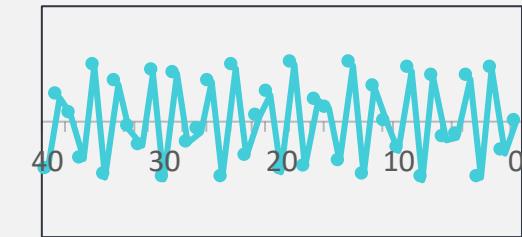
$$S = \text{Var}(d_1, \dots, d_n)$$



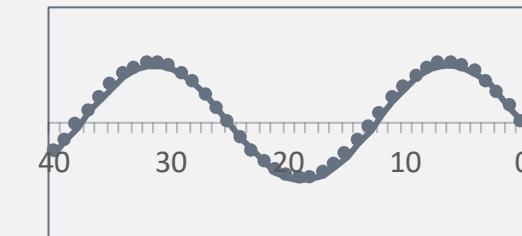
# Sine Wave Analysis: Power Spectral Density



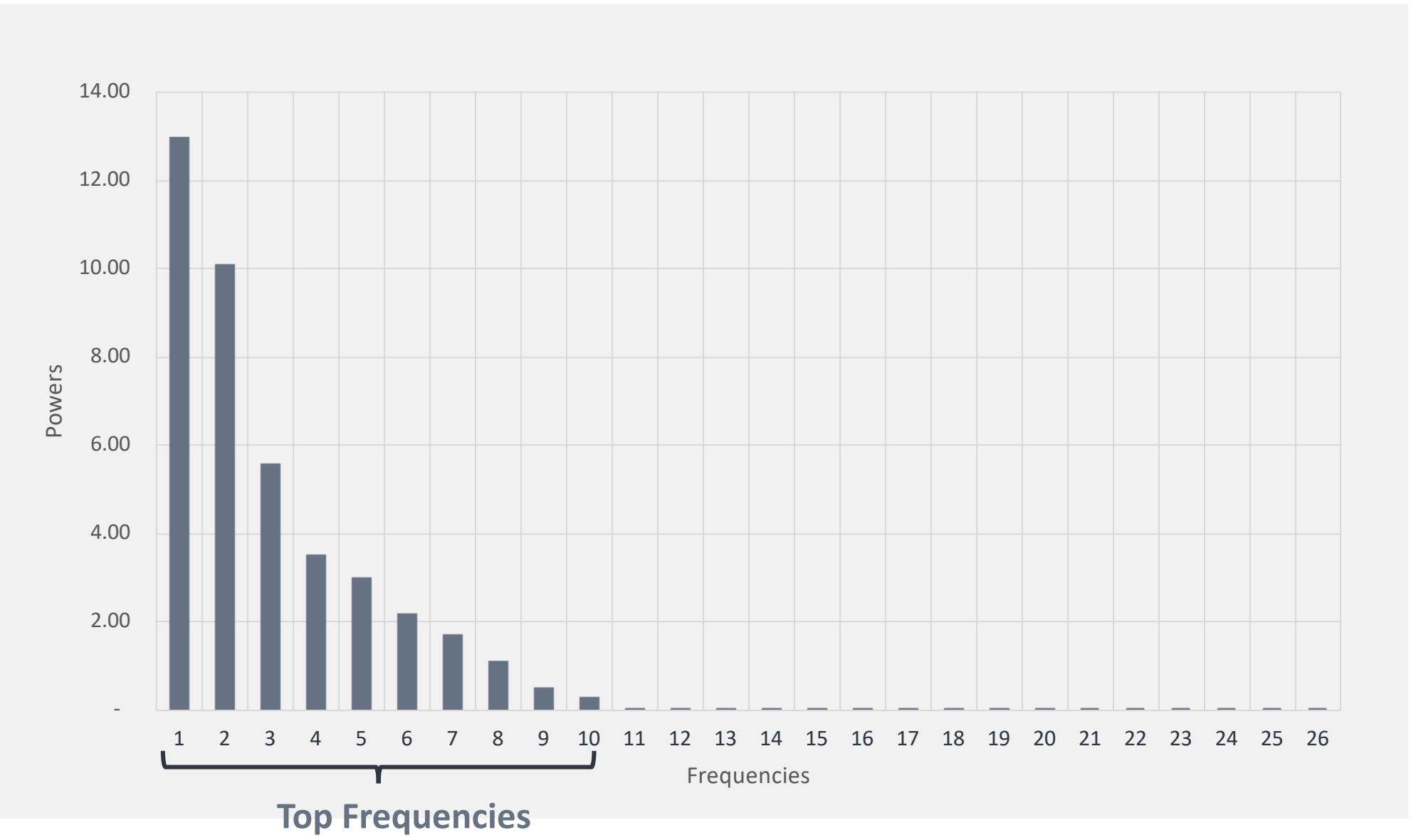
+



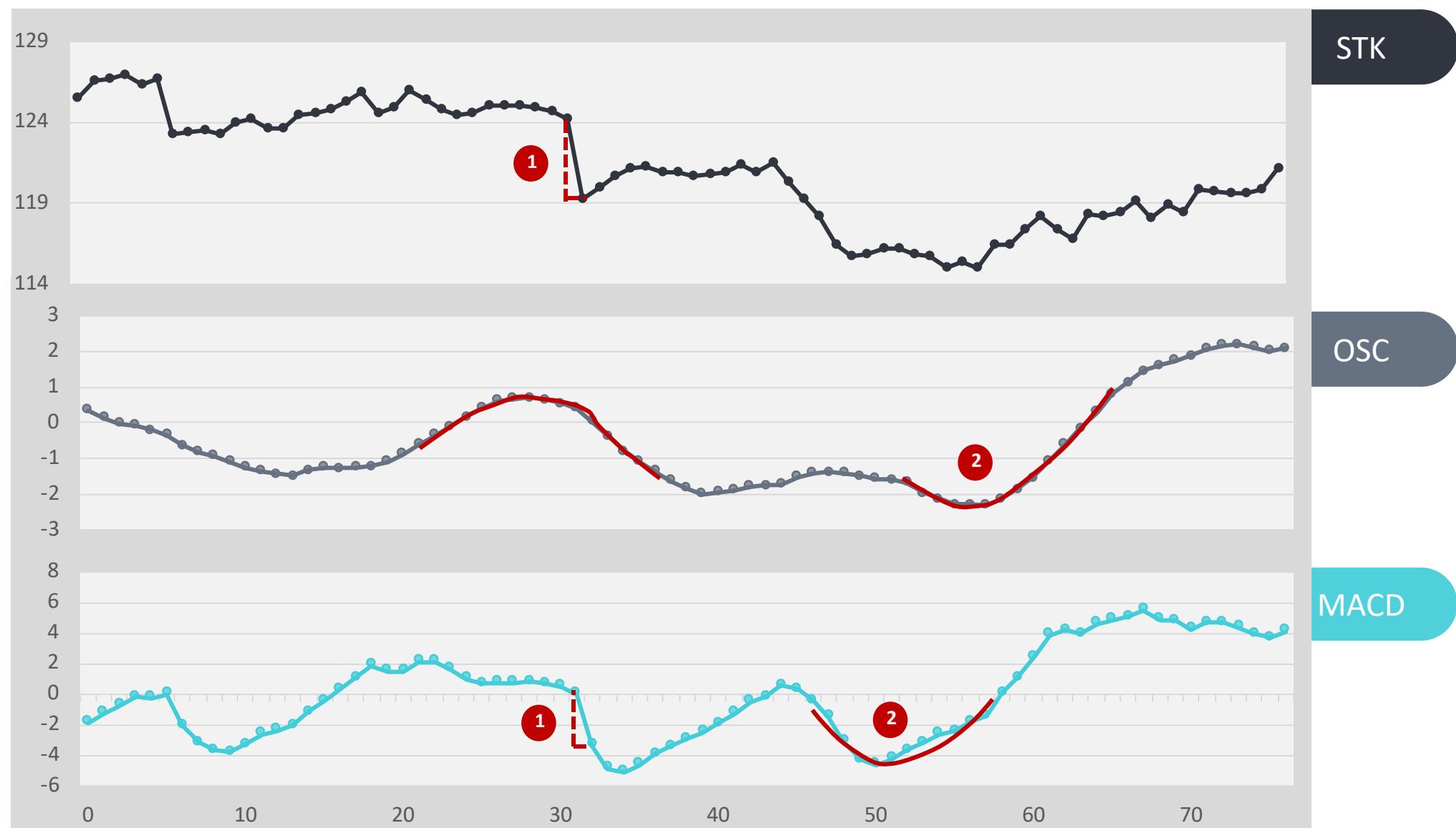
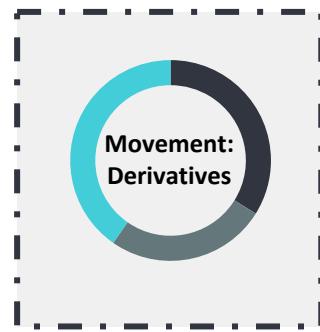
+



# Power Spectral Density



# Relative Derivatives (1<sup>st</sup> & 2<sup>nd</sup>)

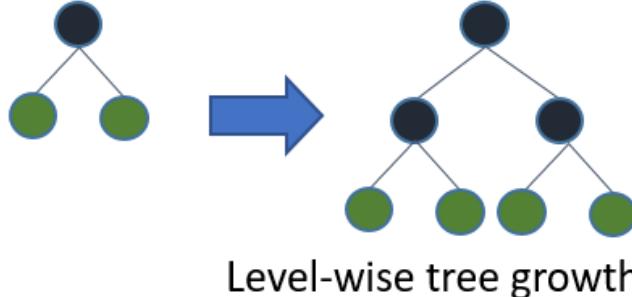


# Algorithm

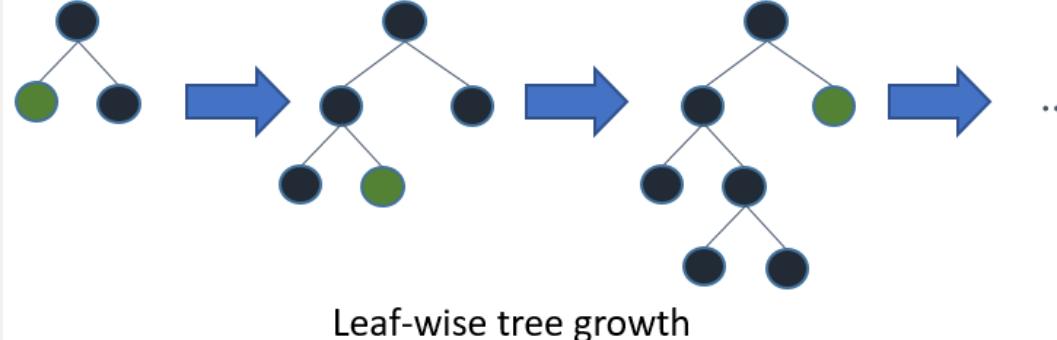


# LightGBM Regressor

XGBoost



LightGBM



Gradient Boosting Decision Tree like XGBoost, but use Leaf-wise Tree Growth instead



Fast



Categorical/  
Numerical  
Inputs

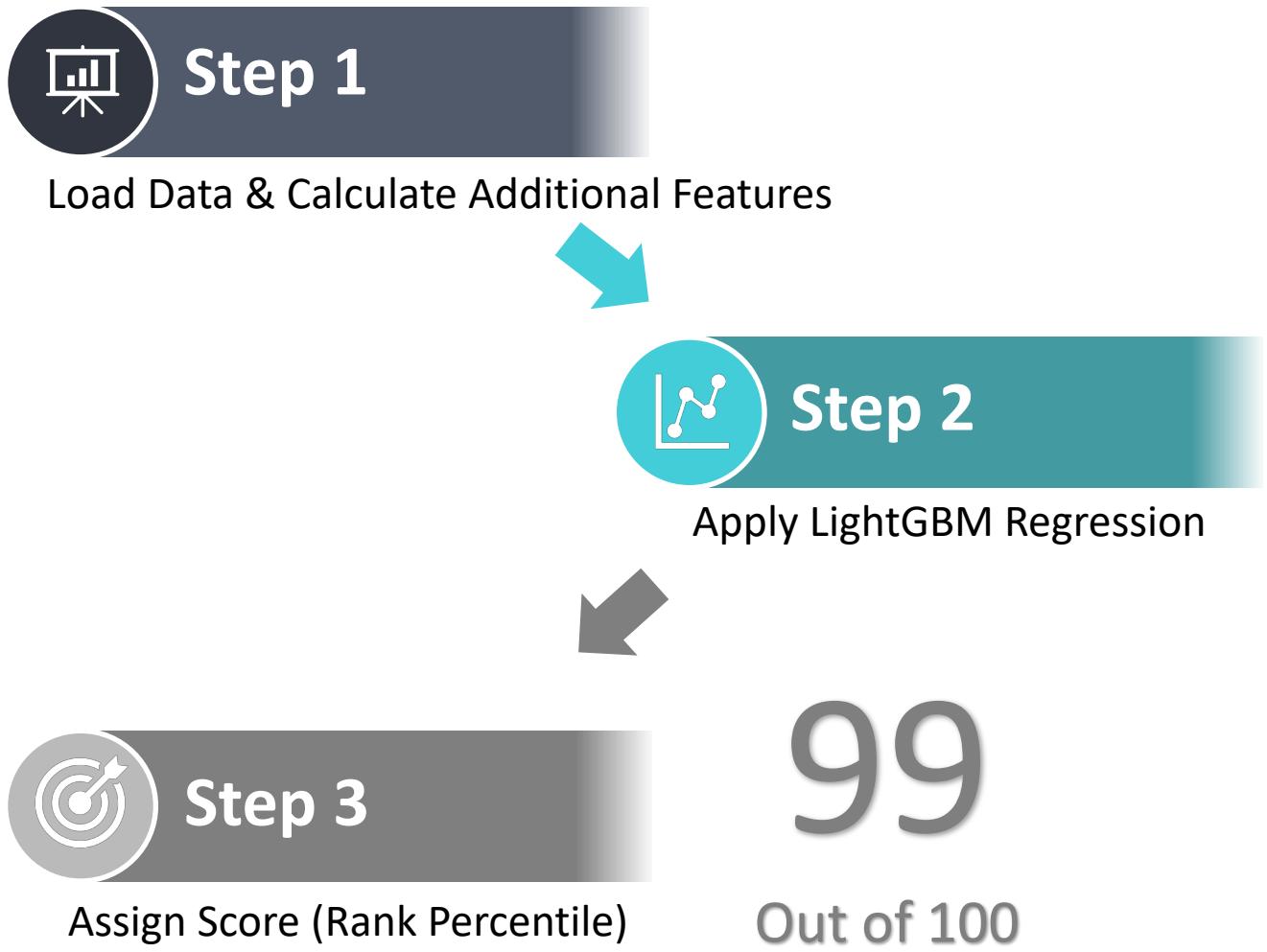
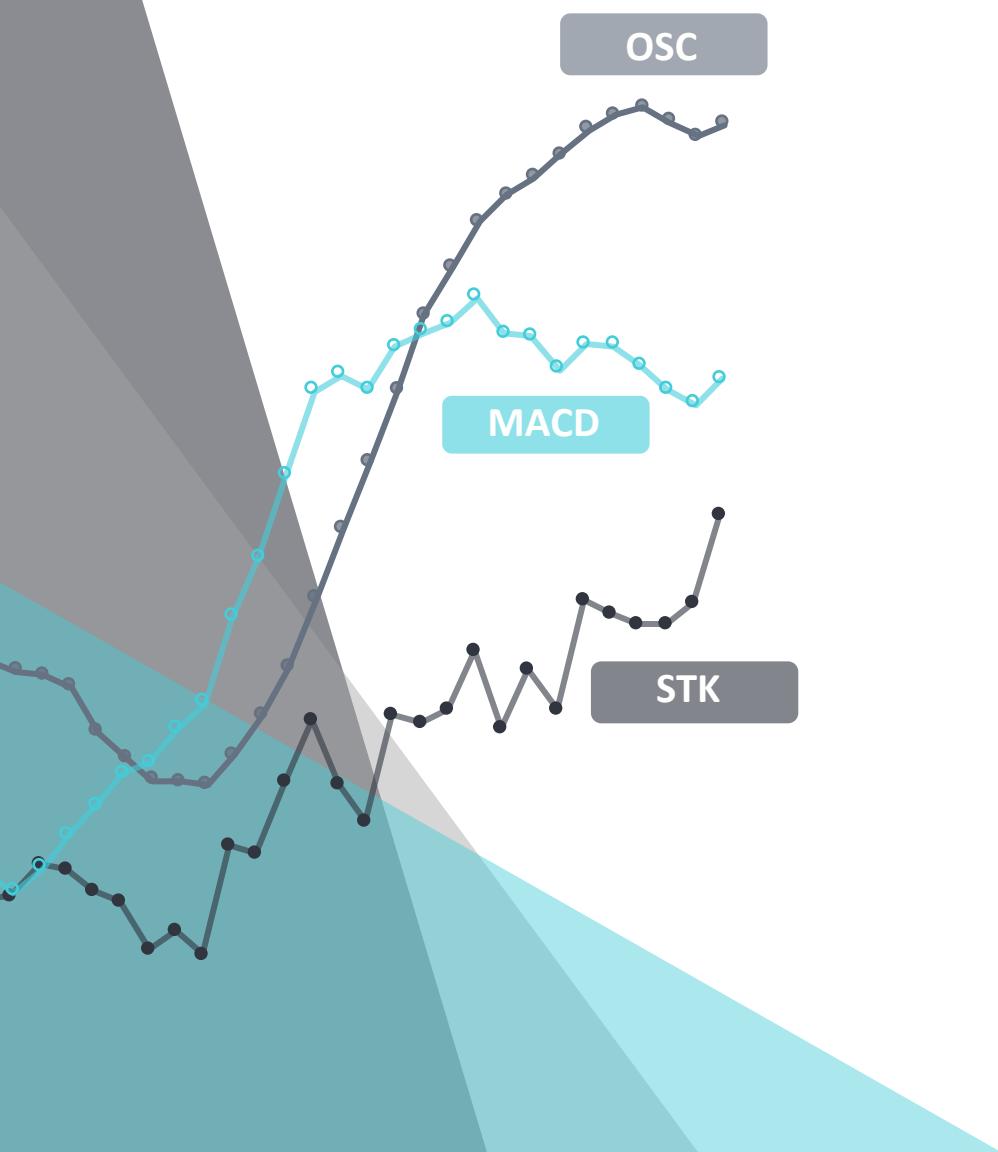


Small Dataset

# Final Product



# Final Product



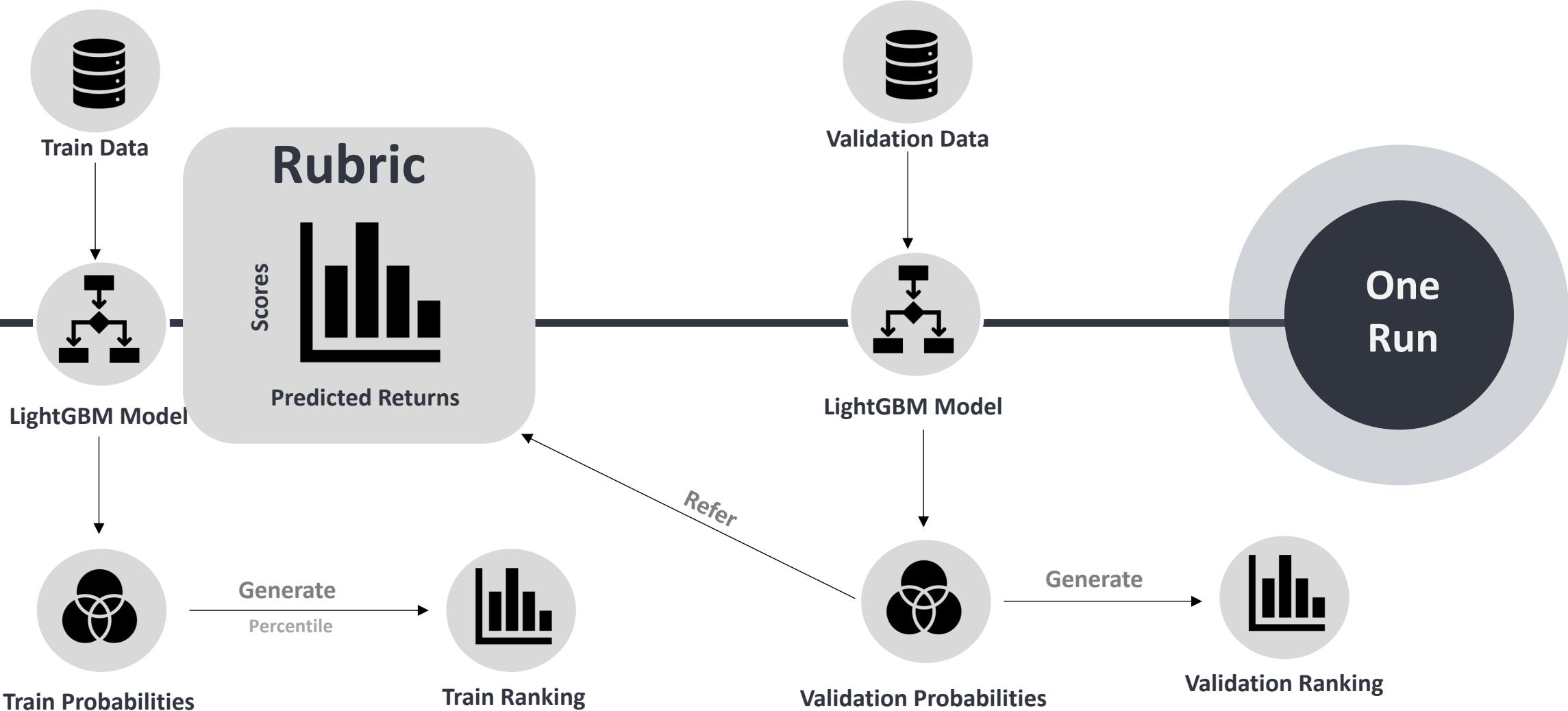
# Why Scoring?

Solution: Score 0 -100

Score ↑  
Return ↑



# Methodology – Single Run



# Results (Single Run)

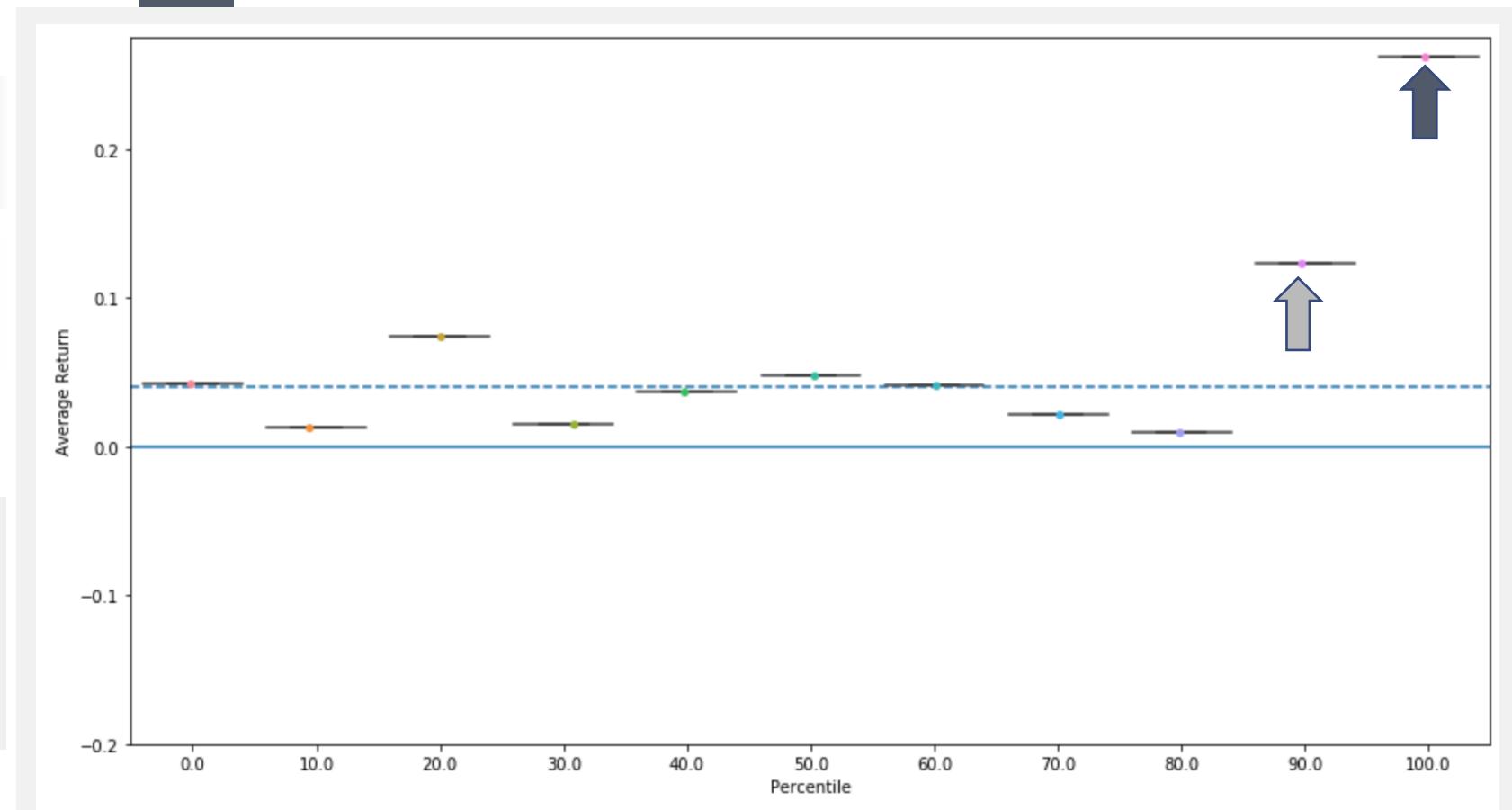
100      **~6X**  
Better (0.26%)

90      **~3X**  
Better (0.12%)

Performance  
increases:  
**/trade**

Dash Line:

Average Return of the Original Data (0.04%)



# Results (Cross-Validation)

100    **~5x\***  
Better

90    **~2.5x\***  
Better

\*: Long term average

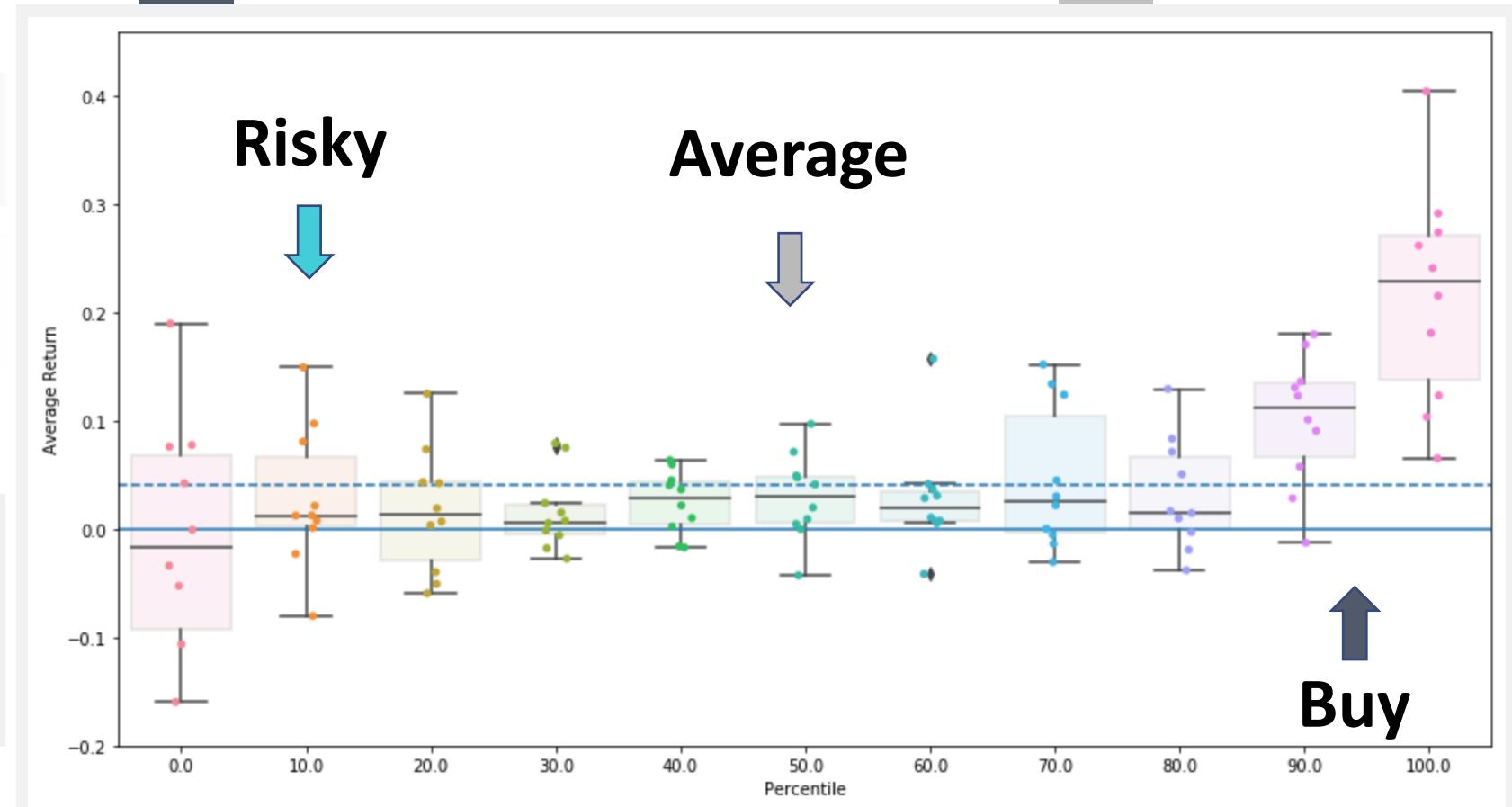
Performance increases:  
**/trade**

Dash Line:

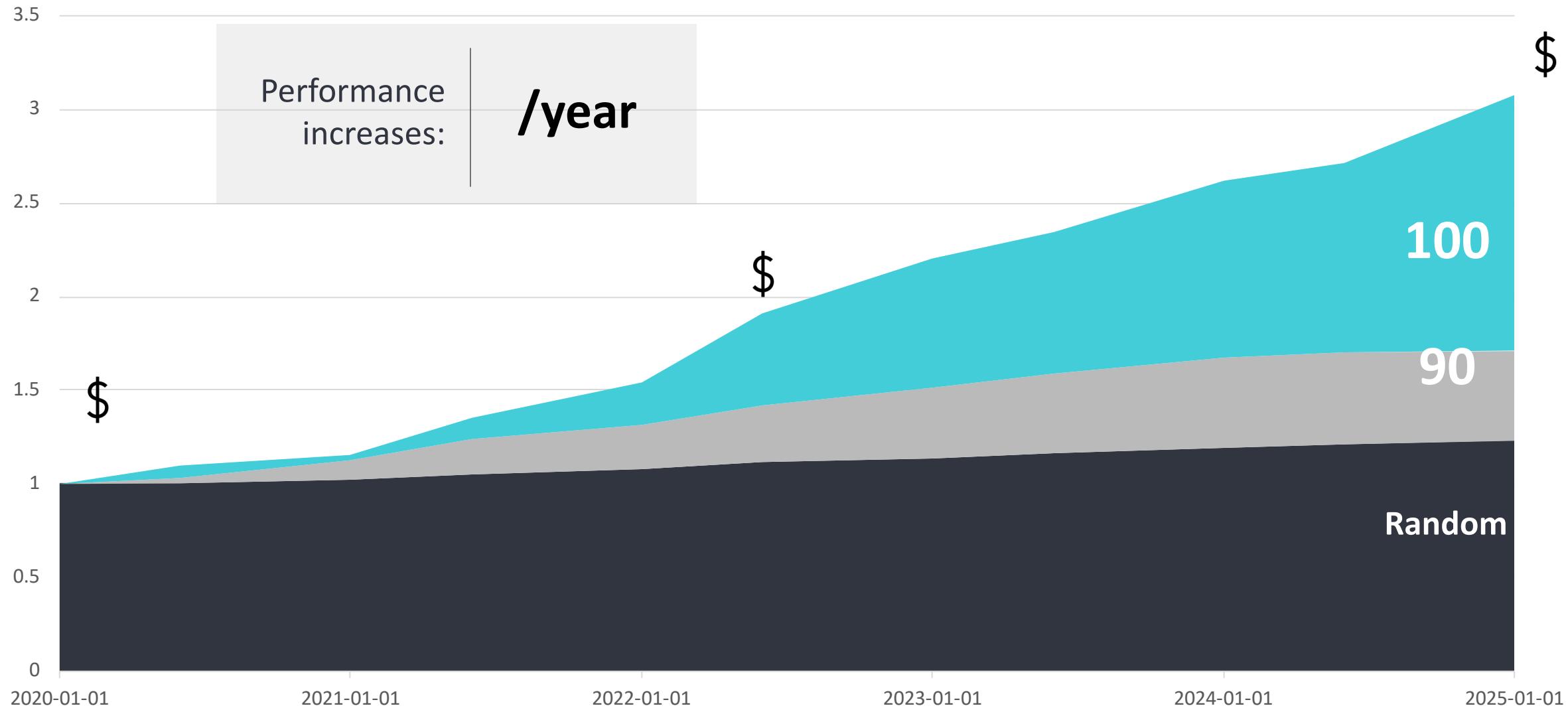
Average Return of the Original Data

CV:

10 Cross-Validation



# 5-Year Profit Simulation



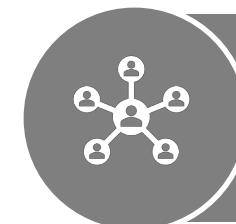
# Future Work



**Model Monthly Renew**



**Explore More Insightful  
Features**



**Ensembled Model**



THANK YOU