# Learning Side and Size

**Advanced Topics in Financial Machine Learning**

# We believe that the scientific method is the best way to approach investment management.

Our core focus is on the implementation of academic research within buy-side asset management. We are first and foremost a research group which productionalizes tools in the form of libraries.
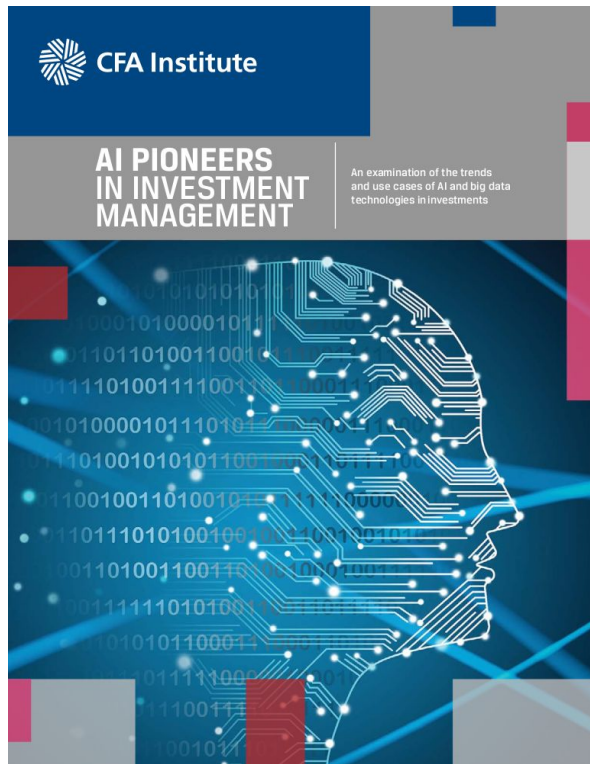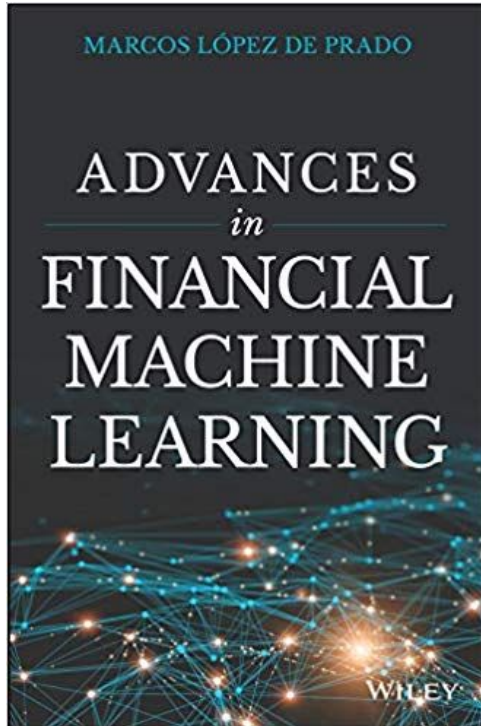
MLFINLAB          NOTEBOOKS          PRESENTATIONS          MEETUP GROUP

# Case Studies



1. Enhancing Trading Strategy and **Execution** with Machine Learning: Man AHL.
2. Generating **Signals** for Quant Models with Machine Learning: New York Life Investments.
3. Refining Equity Trading **Volume Prediction** with Deep Learning: State Street Corporation.
4. Leveraging AI/Alternative **Data Analysis in Sell-Side** Research: Goldman Sachs
5. Dissecting **Earnings Conference Calls** with AI and Big Data: American Century.
6. AI and Big Data Assist in **Debt Portfolio Management**: China Life Asset Management and China Securities Credit Investment.
7. Applying AI and Big Data Technologies in the Filing and Processing of Insurance Claims and Assessing Corporate Risk: Ping An.
8. **Sentiment Analysis**: Bloomberg.
9. Building the Data Science Team: Schroders.
10. Special Focus: **Enhancing the MPT** Efficient Frontier with Machine Learning.
11. Special Focus: Using **Intelligent Searches** to Collect and Process Information.

GitHub

# Advances in Financial Machine Learning



MARCOS LÓPEZ DE PRADO

**ADVANCES** *in* **FINANCIAL MACHINE LEARNING**

WILEY

[www.quantresearch.org](www.quantresearch.org)

QUANTCON 2018

April 26-28 / NYC / #QUANTCON

Quantopian's fourth annual quantitative finance and algorithmic trading conference.

QUANTCON.COM

Grand Hyatt New York, 109 East 42nd Street, New York City NY

THE JOURNAL OF PORTFOLIO MANAGEMENT
QUANT OF THE YEAR 2019

# Journal of Financial Data Science

Academic Journal:

1. A Backtesting Protocol in the Era of Machine Learning
2. **Neural Networks in Finance:** *Design and Performance*
3. Enhancing Time-Series Momentum Strategies Using Deep Neural Networks
4. Time-Series Momentum: *A Monte Carlo Approach*
5. Extracting Signals from High-Frequency Trading with Digital Signal Processing Tools
6. Industry Return Predictability: *A Machine Learning Approach*
7. A Machine Learning Approach to Risk Factors: *A Case Study Using the Fama–French–Carhart Model*
8. Big Data in Portfolio Allocation: *A New Approach to Successful Portfolio Optimization*
9. A Practical Approach to Advanced Text Mining in Finance
10. Dynamic Replication and Hedging: *A Reinforcement Learning Approach*

GitHub

# Table of Contents

# Overview

Often we have the side of the position but are left asking how large should the size of our bet be, or if we should bet at all.

Meta-Labeling helps us to address situations where the primary model is likely to fail and thus reduce our position size, however in an event with a high probability of success it up weights position sizes. In this way it helps to both filter out false positives and size positions.
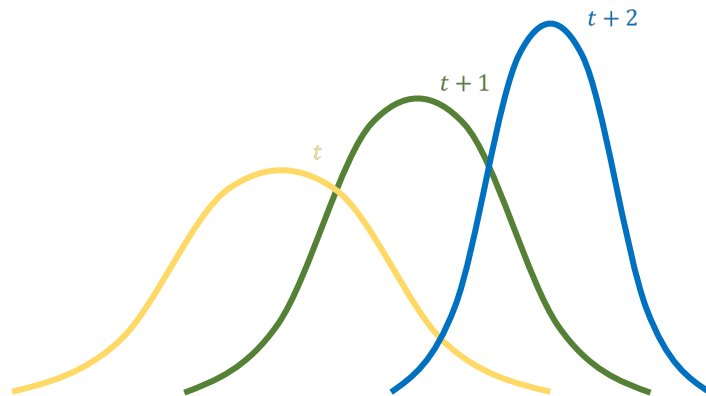
# Problem: non-stationarity

Probability distribution shifts in time, consequently parameters such as mean, variance, and covariance also change over time.

The underlying data generating process f(x) is changing through time.

**Most ML models assume that the data is generated by an independent and identically distributed (IID) processes.**



Stuart Reid, Dynamic Systems, IndabaX, 2018

# Problem: non-stationarity

Fundamentals change:
- Open outcry -> electronic
- A single exchange -> multiple across various locations.
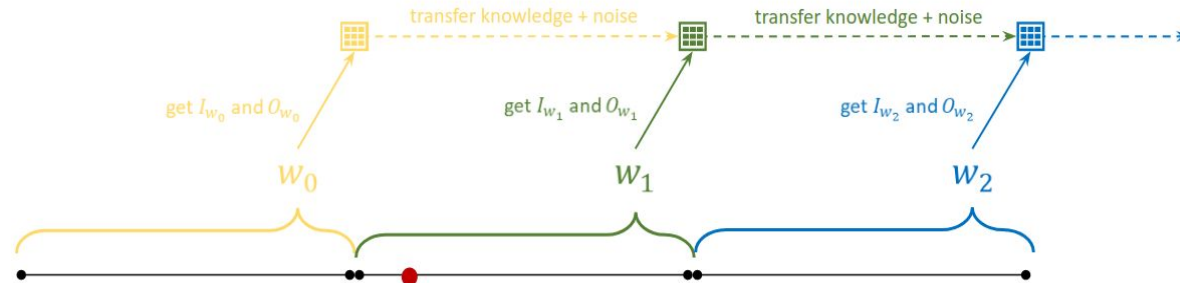- Introduction of HFT

Regimes:
- Trending
- Mean reverting
- Volatility clustering
- Random walk
- Recession



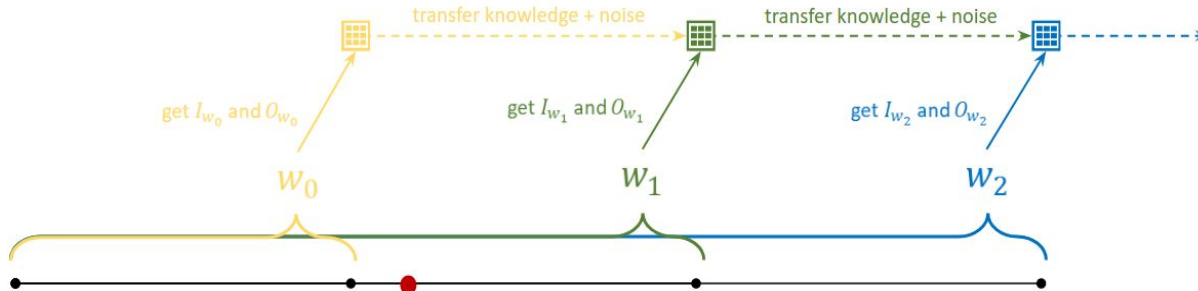**Daily Percentage Changes**
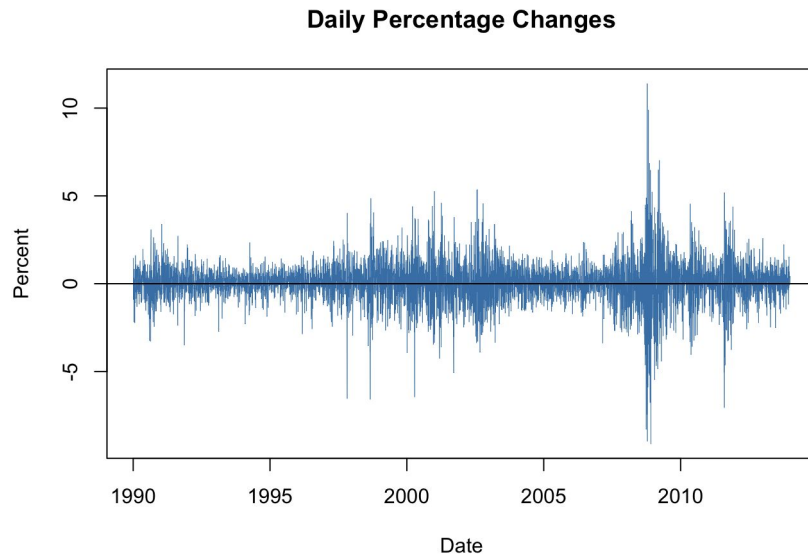
# Problem: Structural Break / Regime Shift
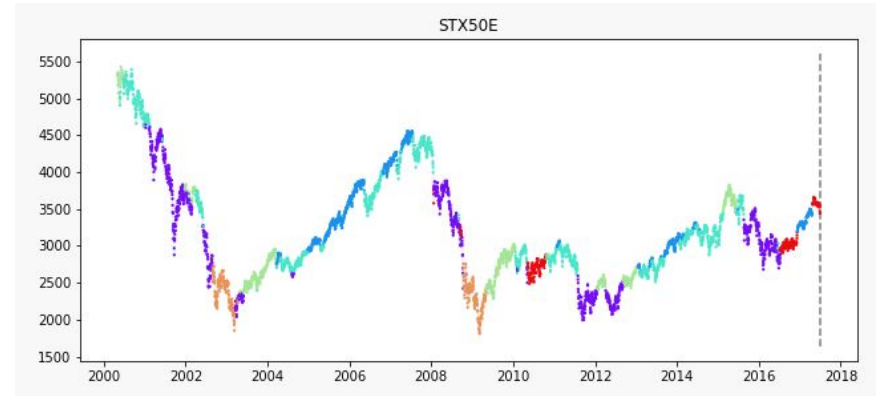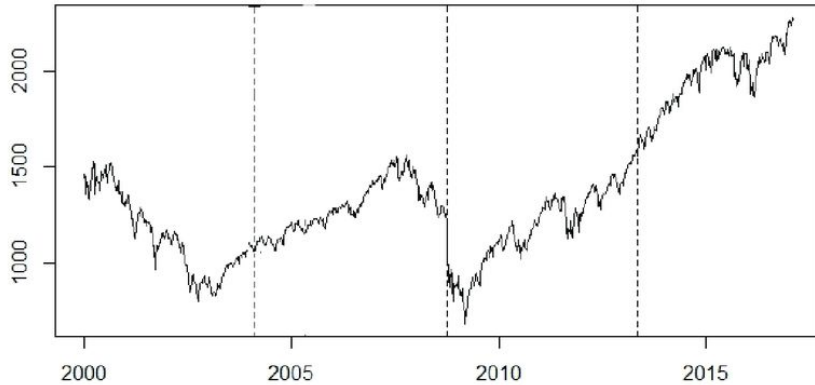
# Solution 1: Online Machine Learning



Fixed window size
-     Pessimistic data sampling
- +     Adapts to change quickly
-     Less data to train on
- +     Faster (less data)
-     Inefficient data usage

Increasing window size
-     Optimistic data sampling
-     Adapts to change slowly
- +     More data to learn from
-     Slower (more data)
-     Most data is *irrelevant* ~ poor model performance

Stuart Reid, Dynamic Systems, IndabaX, 2018
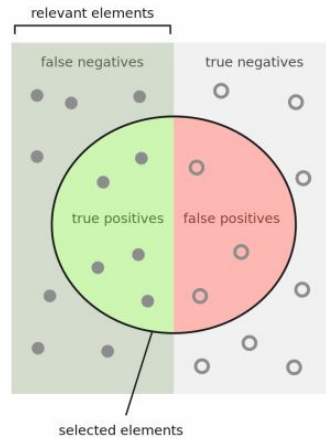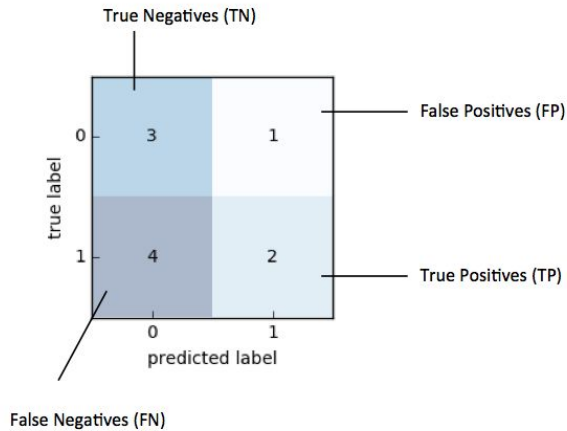
# Solution 2: Meta Labeling

To trade or not to trade!
- Meta-Labels: Takes the side from the primary model [-1, 1] and labels it as correct or incorrect.
- Train a secondary model to determine if we should trade the signal or not.
- Features:
  - Primary model features (Market state)
  - Features indicative of false positives
  - Additional market information
- Primary model can be, discretionary trader, technical rules, classic quant, ml model.
- Trade off between recall and precision. (Want more correct trades).
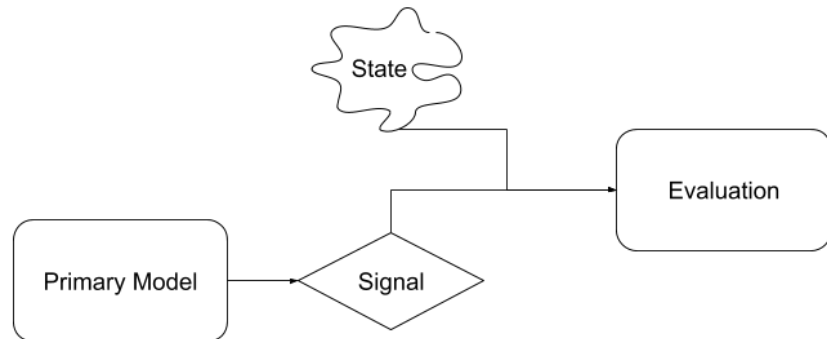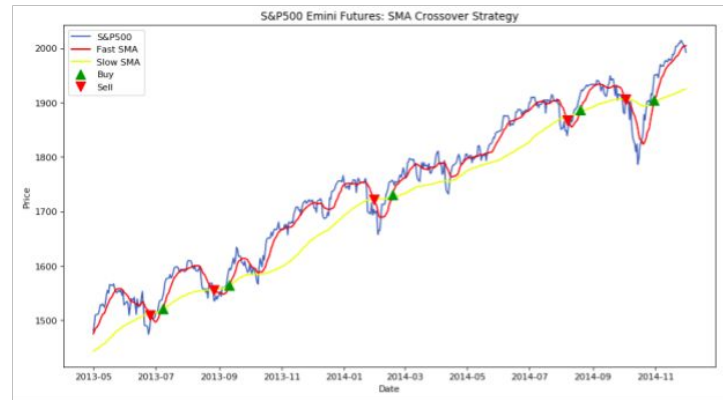
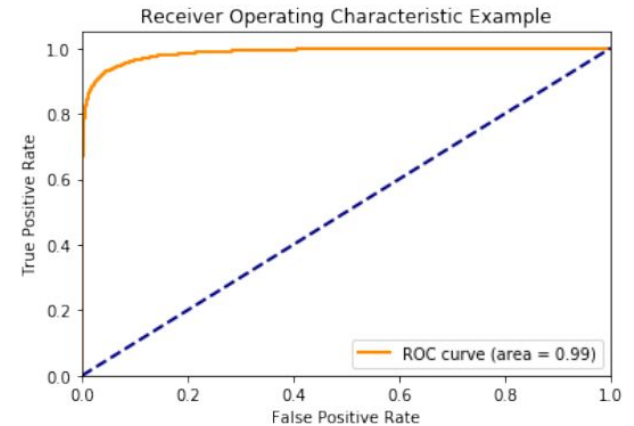# Strategy Framework
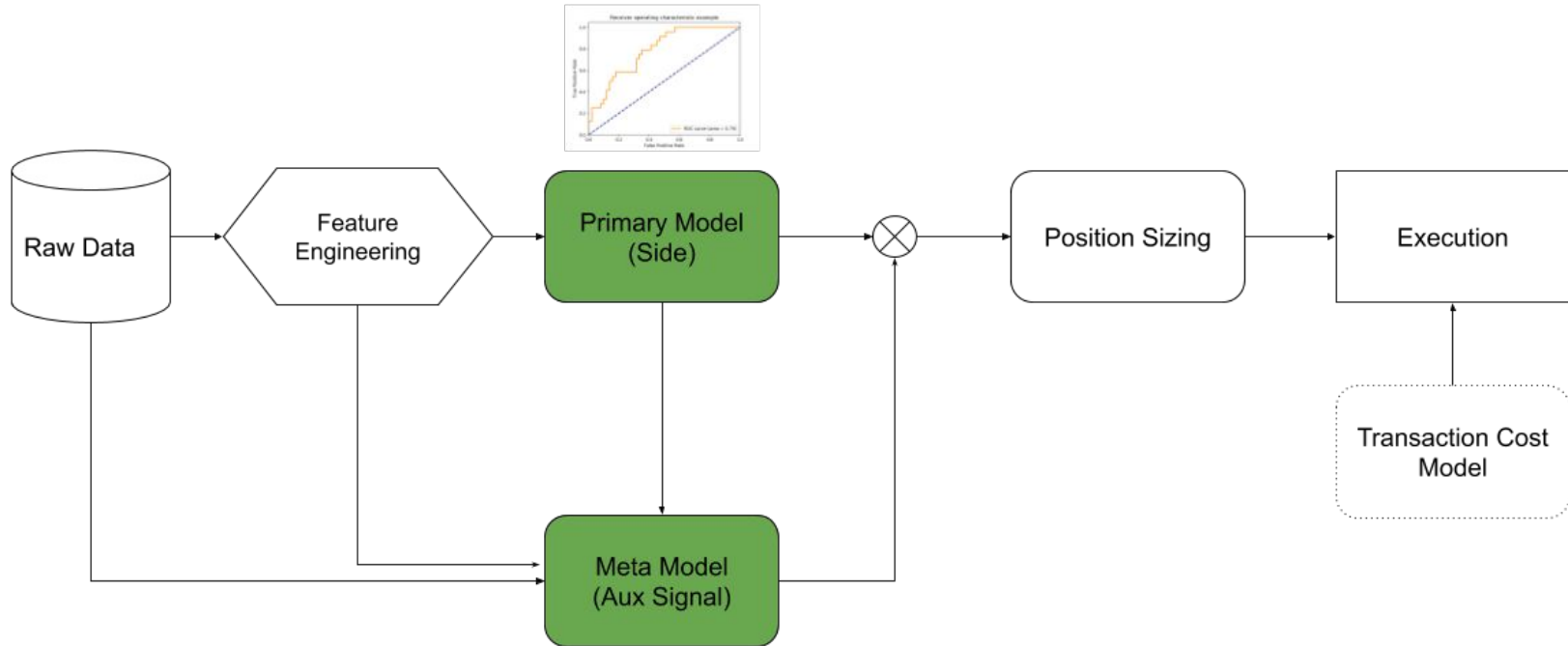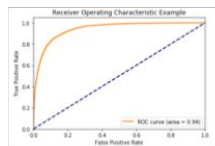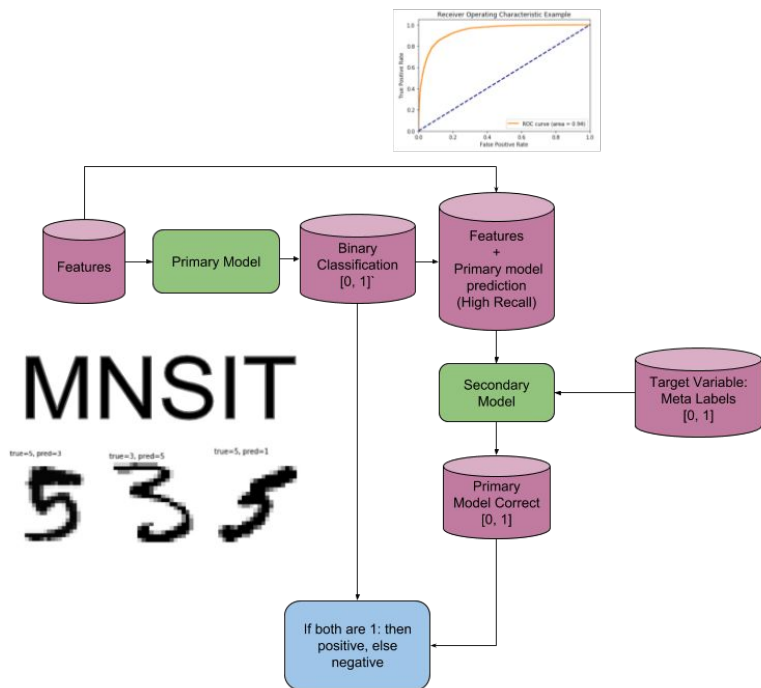
# Important Classification Metrics

# Toy Example: MNIST



Base Model Metrics:

|  | precision | recall | f1-score | support |
|---|---|---|---|---|
| False | 0.95 | 0.94 | 0.94 | 892 |
| True | 0.94 | 0.96 | 0.95 | 1010 |
| micro avg | 0.95 | 0.95 | 0.95 | 1902 |
| macro avg | 0.95 | 0.95 | 0.95 | 1902 |
| weighted avg | 0.95 | 0.95 | 0.95 | 1902 |

Confusion Matrix
[[700 192]
 [ 11 999]]
Accuracy:  0.8933

Meta Label Metrics:

|  | precision | recall | f1-score | support |
|---|---|---|---|---|
| False | 0.95 | 0.96 | 0.95 | 892 |
| True | 0.96 | 0.95 | 0.96 | 1010 |
| micro avg | 0.96 | 0.96 | 0.96 | 1902 |
| macro avg | 0.96 | 0.96 | 0.96 | 1902 |
| weighted avg | 0.96 | 0.96 | 0.96 | 1902 |

Confusion Matrix
[[857  35]
 [ 47 963]]
Accuracy:  0.9569

# Meta Labeling: Trading Example



|  | precision | recall | f1-score | support |
|---|---|---|---|---|
| 0 | 0.00 | 0.00 | 0.00 | 749 |
| 1 | 0.17 | 1.00 | 0.29 | 151 |
| micro avg | 0.17 | 0.17 | 0.17 | 900 |
| macro avg | 0.08 | 0.50 | 0.14 | 900 |
| weighted avg | 0.03 | 0.17 | 0.05 | 900 |

Confusion Matrix
[[   0 749]
 [   0 151]]

Accuracy
0.16777777777777778

|  | precision | recall | f1-score | support |
|---|---|---|---|---|
| 0 | 0.85 | 0.68 | 0.75 | 749 |
| 1 | 0.20 | 0.41 | 0.27 | 151 |
| micro avg | 0.63 | 0.63 | 0.63 | 900 |
| macro avg | 0.53 | 0.54 | 0.51 | 900 |
| weighted avg | 0.74 | 0.63 | 0.67 | 900 |

Confusion Matrix
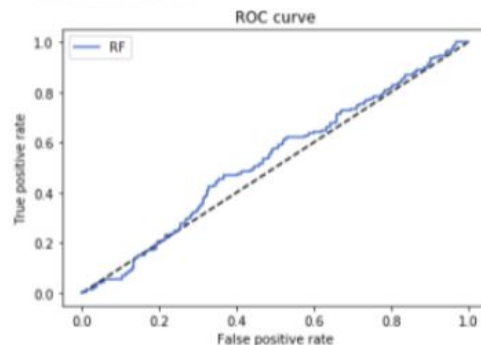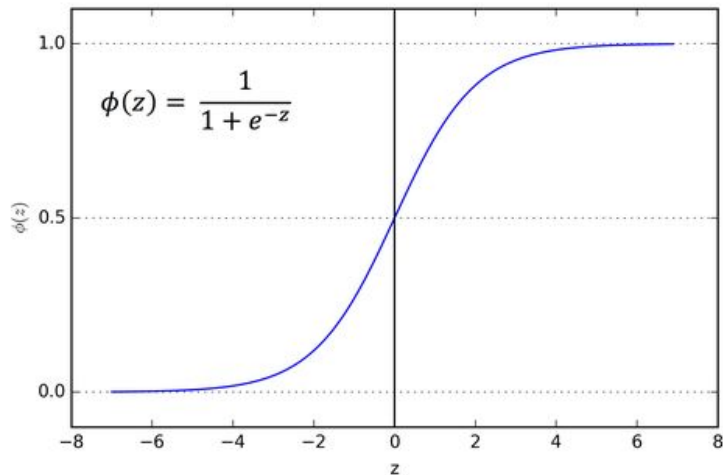[[506 243]
 [ 89  62]]

Accuracy
0.6311111111111111

Table 1: Out-of-sample (2018-01-04 : 2019-01-28)

|  | Primary Model | Meta Model |
|---|---|---|
| Annual return |  |  |
| Cumulative returns | 19.7% | 39.6% |
| Annual volatility | 95.0% | 56.7% |
| Sharpe ratio | 0.65 | 0.82 |
| Calmar ratio | 0.29 | 0.96 |
| Max drawdown | -61.9% | -36.8% |
| Daily value at risk | -11.7% | -7.0% |

# Meta Model Output



$$\phi(z) = \frac{1}{1 + e^{-z}}$$
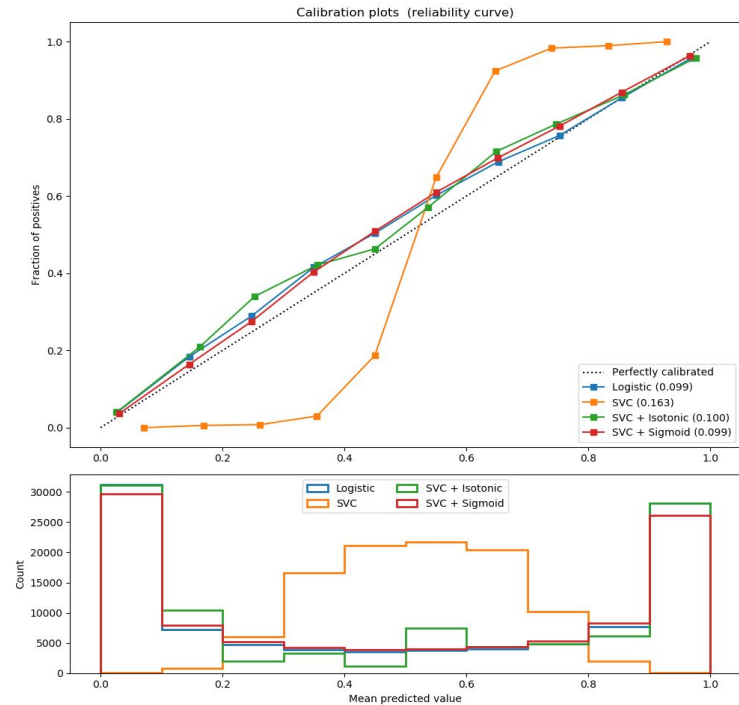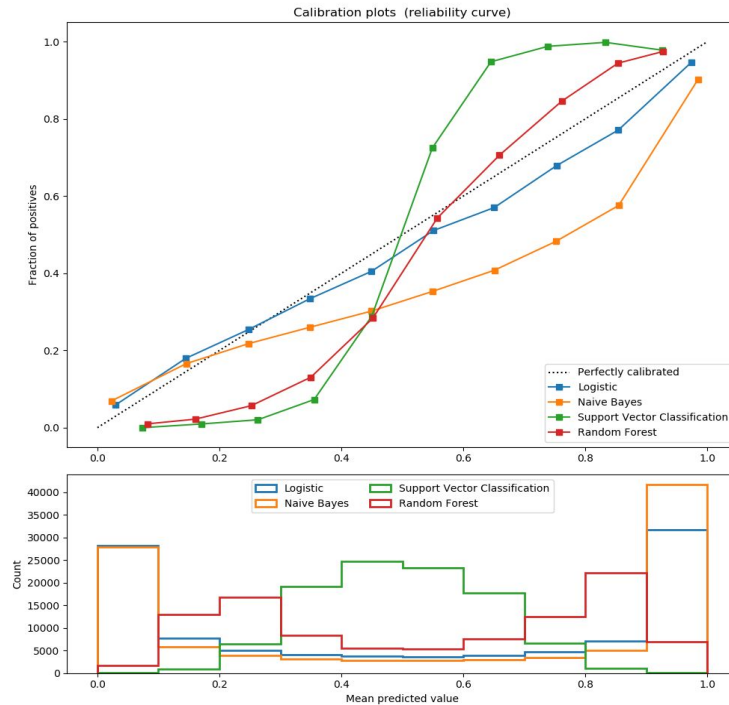
For instance, a well calibrated (binary) classifier should classify the samples such that among the samples to which it gave a predict_proba value close to 0.8, approximately 80% actually belong to the positive class.
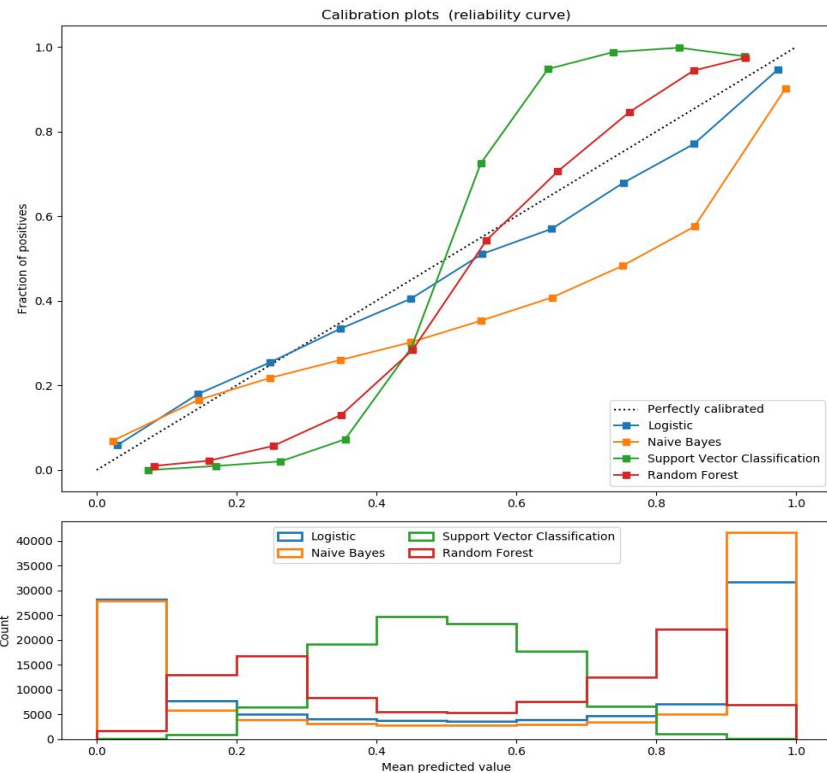
# Probability Calibration

# Position Sizing: Kelly Criterion



$$f^* = \frac{\beta p - q}{\beta} \qquad f^* = p - q$$
$$f^* = 2p - 1$$

Where:
- f* = optimal bet size
- Beta = odds (win amount / lose amount)
- p = probability of success
- q = probability of failure (1-p)

Additional papers:
- A New Interpretation of Information Rate
- Understanding the Kelly Capital Growth Investment Strategy
- How Does the Fortune's Formula Kelly Capital Growth Model Perform?
- A Response to Professor Paul A. Samuelson's Objections to Kelly Capital Growth Investing
- Good and bad properties of the Kelly criterion

**The Investment Opportunities**

| Win Probability | Odds | Prob. of Selection in Simulation | Kelly Bets |
|---|---|---|---|
| 0.570 | 1–1 | 0.1 | 0.140 |
| 0.380 | 2–1 | 0.3 | 0.070 |
| 0.285 | 3–1 | 0.3 | 0.047 |
| 0.228 | 4–1 | 0.2 | 0.035 |
| 0.190 | 5–1 | 0.1 | 0.028 |

**Final Wealth Statistics by Kelly Fraction: Ziemba-Hausch [1986] Model**

| Statistic | 1.0k | 0.75k | 0.50k | 0.25k | 0.125k |
|---|---|---|---|---|---|
| Max | 318854673 | 4370619 | 1117424 | 27067 | 6330 |
| Mean | 524195 | 70991 | 19005 | 4339 | 2072 |
| Min | 4 | 56 | 111 | 513 | 587 |
| St. Dev. | 8033178 | 242313 | 41289 | 2951 | 650 |
| Skewness | 35 | 11 | 13 | 2 | 1 |
| Kurtosis | 1299 | 155 | 278 | 9 | 2 |
| > 5 × 10 | 1981 | 2000 | 2000 | 2000 | 2000 |
| $10^2$ | 1965 | 1996 | 2000 | 2000 | 2000 |
| > 5 × 102 | 1854 | 1936 | 1985 | 2000 | 2000 |
| > $10^3$ | 1752 | 1855 | 1930 | 1957 | 1978 |
| > $10^4$ | 1175 | 1185 | 912 | 104 | 0 |
| > $10^5$ | 479 | 284 | 50 | 0 | 0 |
| > $10^6$ | 111 | 17 | 1 | 0 | 0 |

How Does the Fortune's Formula Kelly Capital Growth Model Perform?



Chapter 9

THE KELLY CRITERION IN BLACKJACK SPORTS BETTING, AND THE STOCK MARKET[1]

EDWARD O. THORP
Edward O. Thorp and Associates, Newport Beach, CA 92660, USA

**Contents**

# Position Sizing with Meta Labeling

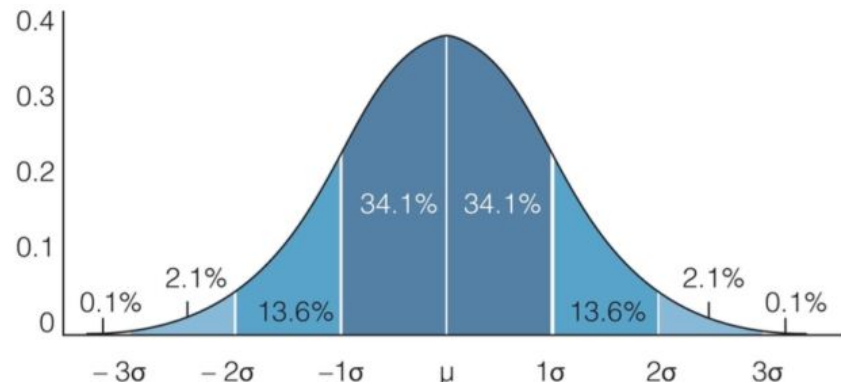**1** Transform meta-model outputs.

Transformed outputs assumed to be normally distributed.

$$z = \frac{p[x = 1] - \frac{1}{2}}{\sqrt{p[x = 1](1 - p[x = 1])}}$$

$$z = \frac{x - \bar{x}}{\sigma}$$

Where:
- p[x] = probability that label x takes place.
- z = test statistic
- x is element of {-1, 1}
- Z is standard normal distribution

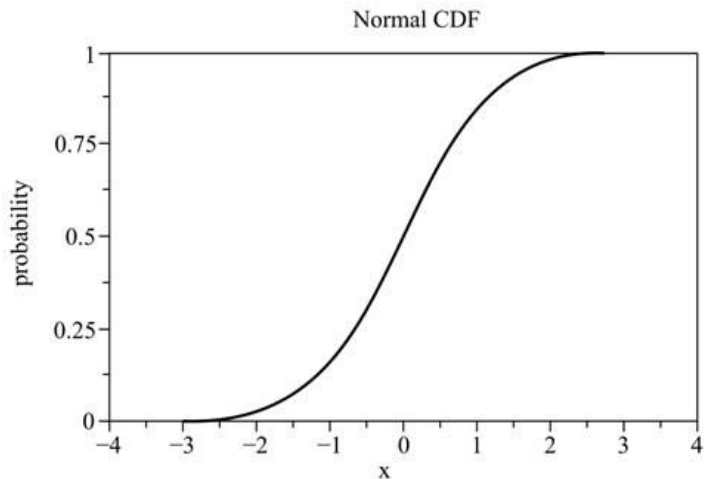# Position Sizing with Meta Labeling

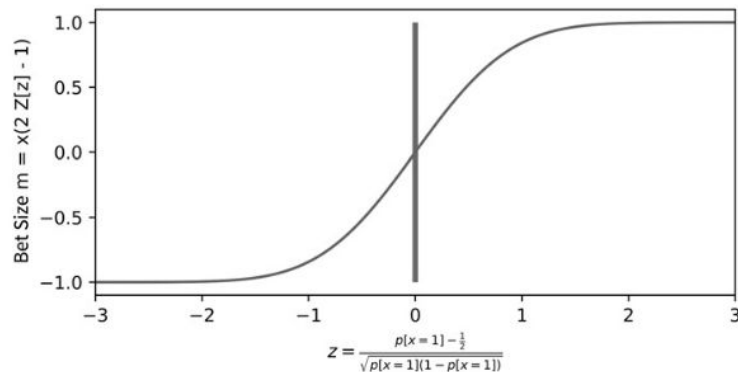**2**  Compute the CDF of Z:

Z[.] is the CDF of Z



Normal CDF

**3**  Compute Bet Size:

$$f = side(2Z[z] - 1) \qquad f^* = 2p - 1$$



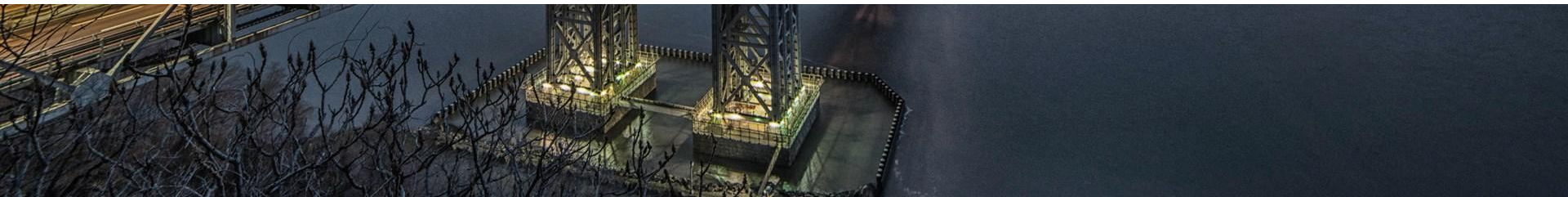$$z = \frac{p[x=1] - \frac{1}{2}}{\sqrt{p[x=1](1-p[x=1])}}$$

# Next Steps

- The concepts of position sizing and meta-labeling are both addressed in the textbook Advances in Financial Machine Learning.
- A great additional resource is the Journal of Financial Data Science and the Journal of Portfolio Management.
- If these concepts interest you, there is room available in our research group. You will get to work with high quality tick data and contribute to open-source.

MARCOS LÓPEZ DE PRADO

ADVANCES
*in*
FINANCIAL
MACHINE
LEARNING

WILEY
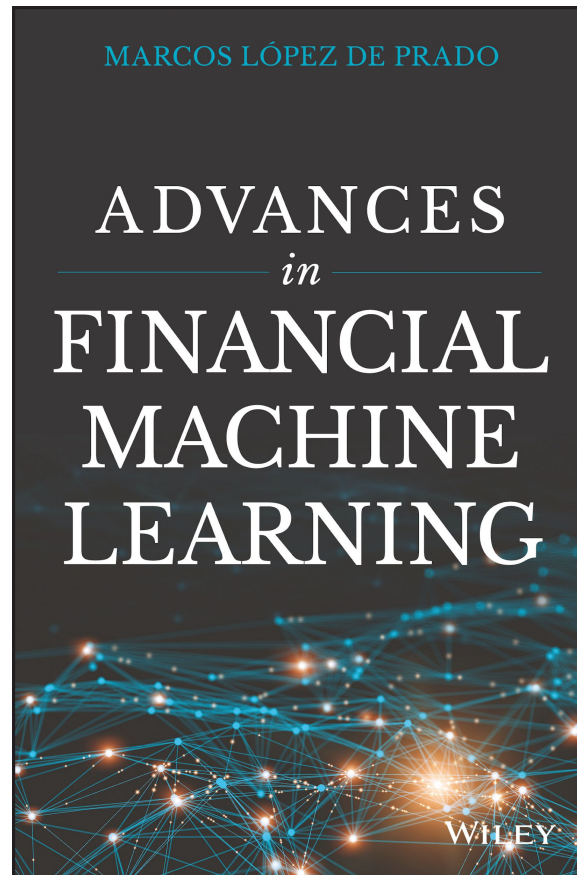
# Conclusion

- Meta-Labeling helps to address the problem of non-stationarity and structural breaks by down weighting position sizes in strategies that have a low probability of success in a given market state.
- The Kelly Criterion can be used to determine optimal position sizes.

Thank You