# A cat is a cat and a Euro is a Euro, not (?)[1]

## An assessment of AI/ML methods applied to Asset Allocation

Bernhard Pfaff
Invesco Asset Management GmbH, 60322 Frankfurt am Main
bernhard.pfaff@invesco.com

Data Science in Finance with R
WU, Vienna
19th - 20th September 2019

[1] Quote: 'a Euro is a Euro' by Willem F. Duisenberg, ECB Press Conference, 4th February 1999.

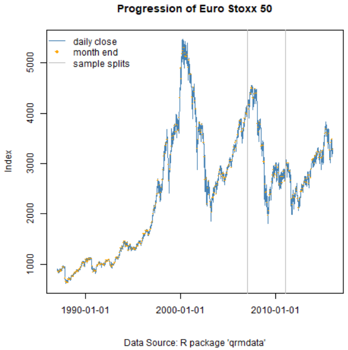
# Introduction

- First wave of ML-/AI methods applied to problems in finance dates back to the 1990s and was inspired by applications in the field of engineering (quality control), *e.g.*, Karlruher Ökonometrie Workshops, see Bol, Nakhaeizadeh, and Vollmer 1994; Bol, Nakhaeizadeh, and Vollmer 1996; Bol, Nakhaeizadeh, and Vollmer 2000.
- Second wave (attempt) of applying ML-/AI-method to problems in finance since the mid 2010s and was motivated by successful applications in image/video classification, NLP, *etc.*
- What has changed in the interim?
  1. Computational capacity.
  2. Count of data records/observations.
  3. Access to 'ready-made' software implementations, *e.g.* Chollet and Allaire 2018.
- Will this second wave last and form a new paradigm?[2]

[2] Kuhn 1970.

# The Challenge
## Euro Stoxx 50 Index

- Binary classification:
  Predict sign of subsequent monthly returns (aka as *label*).
- Use trailing daily observations for defining explanatory variables (aka as *features*).
- Sample Split:
  1. Training:
     31.01.1987 – 31.12.2006.
  2. Validating:
     31.01.2007 – 31.12.2010.
  3. Testing:
     31.01.2011 – 31.12.2015.



# The Challenge

## The Peer Models/Methods

1. Linear Discriminant Analysis: Fisher 1936
2. k-Nearest Neighbours: Hart 1968
3. Support Vector Machines: Vapnik and Lerner 1963; Vapnik and Chervonenkis 1974; Vapnik 1995; Vapnik, Golowich, and Smola 1997
4. Classification Trees: Breiman et al. 1984
5. Random Forest: Kleinberg 1990; Kleinberg 1996; Kleinberg 2000; Ho n.d.; Ho 1998; Breiman 2001
6. (Extreme) Gradient Boosting: Friedman 2001; Friedman 2002
7. Feed-Forward Neural Networks: Chollet and Allaire 2018; Goodfellow 2016; Prado 2018; Rosenblatt 1958; Russell 2010

# The Features

## Alternative Momentum Measures

1. Sign of (significant) slope coefficient in local trend regression; set of indicator values $\{-1, 0, 1\}$.
2. Two-part trend (last five and ten trading days in a month) based on simple moving averages; set of indicator values $\{-1, 0, 1\}$.
3. Relative strength index (RSI) with hurdles 30 and 70; set of indicator values $\{-1, 0, 1\}$.
4. Trend Detection Index (TDI) derived from last ten trading days in a month; set of indicator values $\{-1, 0, 1\}$.

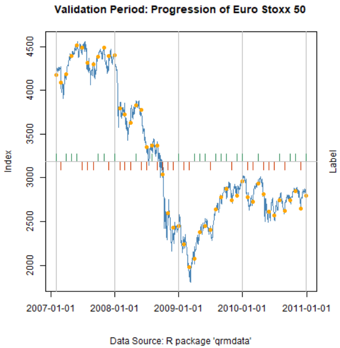
# The Features

## Alternative Risk Measures

1. Standard deviations of intra-month daily returns (last ten trading days vs. all observations);
   set of indicator values $\{-1, 1\}$.
2. Realized volatility of intra-month daily returns (last ten trading days vs. all observations);
   set of indicator values $\{-1, 1\}$.
3. Sharpe ratio (last ten trading days vs. all observations);
   set of indicator values $\{-1, 1\}$.
4. Sign of skewness of intra-month daily returns;
   set of indicator values $\{-1, 1\}$.
5. Sign of excess kurtosis of intra-month daily returns;
   set of indicator values $\{-1, 1\}$.

# Model Training
## Outline

- Each specified model is trained on the training set, and evaluated by its predictive accuracy on the validation set.
- An extending recursive training/shrinking validation set by 12 months is utilized.
- The model specification of a ML-method with the highest hit rate is used for the test set.



# Model Training

## Linear Discriminant Analysis

| LDA models | Validation Periods | | | | |
|---|---|---|---|---|---|
| | 48 Months | 36 Months | 24 Months | 12 Months | Average |
| Method = moment, DF = NA | 54.2 | 55.6 | 45.8 | 41.7 | 49.3 |
| Method = mle, DF = NA | 54.2 | 55.6 | 45.8 | 41.7 | 49.3 |
| Method = t, DF = 5 | 56.2 | 52.8 | 45.8 | 50.0 | 51.2 |
| Method = t, DF = 10 | 56.2 | 52.8 | 45.8 | 50.0 | 51.2 |
| Method = t, DF = 15 | 56.2 | 52.8 | 45.8 | 50.0 | 51.2 |
| Method = t, DF = 20 | 54.2 | 52.8 | 45.8 | 50.0 | 50.7 |

Table: Predicitve Accuracy of LDA on Validation Set.

# Model Training

## k-Nearest Neighbours

| KNN models | Validation Periods | | | | |
|---|---|---|---|---|---|
| | 48 Months | 36 Months | 24 Months | 12 Months | Average |
| k = 1 | 52.1 | 61.1 | 54.2 | 25.0 | 48.1 |
| k = 2 | 60.4 | 61.1 | 54.2 | 33.3 | 52.3 |
| k = 3 | 60.4 | 61.1 | 50.0 | 33.3 | 51.2 |
| k = 4 | 52.1 | 52.8 | 50.0 | 25.0 | 45.0 |
| k = 5 | 56.2 | 61.1 | 54.2 | 25.0 | 49.1 |
| k = 6 | 56.2 | 55.6 | 58.3 | 41.7 | 53.0 |
| k = 7 | 54.2 | 52.8 | 50.0 | 41.7 | 49.7 |
| k = 8 | 45.8 | 52.8 | 50.0 | 41.7 | 47.6 |
| k = 9 | 45.8 | 47.2 | 41.7 | 41.7 | 44.1 |
| k = 10 | 45.8 | 52.8 | 45.8 | 41.7 | 46.5 |

Table: Predicitve Accuracy of KNN on Validation Set.

# Model Training

## Support Vector Machine

| SVM models | Validation Periods | | | | |
|---|---|---|---|---|---|
| | 48 Months | 36 Months | 24 Months | 12 Months | Average |
| Kernel = linear, cost = 1 | 47.9 | 44.4 | 54.2 | 41.7 | 47.0 |
| Kernel = linear, cost = 10 | 47.9 | 44.4 | 54.2 | 41.7 | 47.0 |
| Kernel = linear, cost = 100 | 47.9 | 44.4 | 54.2 | 41.7 | 47.0 |
| Kernel = radial, cost = 1 | 47.9 | 44.4 | 41.7 | 33.3 | 41.8 |
| Kernel = radial, cost = 10 | 54.2 | 52.8 | 50.0 | 33.3 | 47.6 |
| Kernel = radial, cost = 100 | 52.1 | 58.3 | 54.2 | 33.3 | 49.5 |
| Kernel = polynomial, cost = 1 | 52.1 | 55.6 | 45.8 | 33.3 | 46.7 |
| Kernel = polynomial, cost = 10 | 54.2 | 55.6 | 37.5 | 33.3 | 45.1 |
| Kernel = polynomial, cost = 100 | 54.2 | 52.8 | 37.5 | 33.3 | 44.4 |
| Kernel = sigmoid, cost = 1 | 47.9 | 47.2 | 58.3 | 50.0 | 50.9 |
| Kernel = sigmoid, cost = 10 | 33.3 | 25.0 | 66.7 | 66.7 | 47.9 |
| Kernel = sigmoid, cost = 100 | 33.3 | 30.6 | 29.2 | 58.3 | 37.8 |

Table: Predicitve Accuracy of SVM on Validation Set.

# Model Training

## Classification Tree

| CTR models | Validation Periods | | | | |
|---|---|---|---|---|---|
| | 48 Months | 36 Months | 24 Months | 12 Months | Average |
| SplitCrit = information, minsize = 2 | 43.8 | 47.2 | 45.8 | 33.3 | 42.5 |
| SplitCrit = information, minsize = 4 | 43.8 | 47.2 | 41.7 | 33.3 | 41.5 |
| SplitCrit = information, minsize = 6 | 45.8 | 47.2 | 41.7 | 33.3 | 42.0 |
| SplitCrit = information, minsize = 8 | 45.8 | 47.2 | 45.8 | 41.7 | 45.1 |
| SplitCrit = information, minsize = 10 | 45.8 | 44.4 | 41.7 | 33.3 | 41.3 |
| SplitCrit = gini, minsize = 2 | 45.8 | 47.2 | 45.8 | 33.3 | 43.1 |
| SplitCrit = gini, minsize = 4 | 43.8 | 47.2 | 41.7 | 33.3 | 41.5 |
| SplitCrit = gini, minsize = 6 | 45.8 | 47.2 | 41.7 | 33.3 | 42.0 |
| SplitCrit = gini, minsize = 8 | 45.8 | 47.2 | 45.8 | 41.7 | 45.1 |
| SplitCrit = gini, minsize = 10 | 45.8 | 44.4 | 41.7 | 33.3 | 41.3 |

Table: Predicitve Accuracy of CTR on Validation Set.

# Model Training

## Random Forest

| RFM models | Validation Periods | | | | |
|---|---|---|---|---|---|
| | 48 Months | 36 Months | 24 Months | 12 Months | Average |
| Node size = 1, mtry = 2 | 52.1 | 52.8 | 45.8 | 33.3 | 46.0 |
| Node size = 1, mtry = 3 | 45.8 | 50.0 | 45.8 | 33.3 | 43.8 |
| Node size = 1, mtry = 4 | 45.8 | 50.0 | 45.8 | 33.3 | 43.8 |
| Node size = 1, mtry = 5 | 43.8 | 47.2 | 45.8 | 33.3 | 42.5 |
| Node size = 1, mtry = 6 | 45.8 | 47.2 | 41.7 | 33.3 | 42.0 |
| Node size = 1, mtry = 7 | 43.8 | 47.2 | 45.8 | 33.3 | 42.5 |
| Node size = 10, mtry = 2 | 54.2 | 58.3 | 45.8 | 33.3 | 47.9 |
| Node size = 10, mtry = 3 | 50.0 | 58.3 | 45.8 | 33.3 | 46.9 |
| Node size = 10, mtry = 4 | 52.1 | 58.3 | 45.8 | 33.3 | 47.4 |
| Node size = 10, mtry = 5 | 52.1 | 52.8 | 45.8 | 33.3 | 46.0 |
| Node size = 10, mtry = 6 | 50.0 | 52.8 | 45.8 | 33.3 | 45.5 |
| Node size = 10, mtry = 7 | 50.0 | 55.6 | 45.8 | 33.3 | 46.2 |
| Node size = 20, mtry = 2 | 58.3 | 52.8 | 37.5 | 33.3 | 45.5 |
| Node size = 20, mtry = 3 | 56.2 | 50.0 | 41.7 | 33.3 | 45.3 |
| Node size = 20, mtry = 4 | 58.3 | 52.8 | 41.7 | 33.3 | 46.5 |
| Node size = 20, mtry = 5 | 56.2 | 52.8 | 41.7 | 33.3 | 46.0 |
| Node size = 20, mtry = 6 | 56.2 | 52.8 | 41.7 | 33.3 | 46.0 |
| Node size = 20, mtry = 7 | 56.2 | 55.6 | 41.7 | 33.3 | 46.7 |

Table: Predicitve Accuracy of RFM on Validation Set.

# Model Training

## Extreme Gradient Boosting

| XGB models | Validation Periods | | | | |
|---|---|---|---|---|---|
| | 48 Months | 36 Months | 24 Months | 12 Months | Average |
| Depths = 2, eta = 0.25 | 50.0 | 47.2 | 45.8 | 50.0 | 48.3 |
| Depths = 2, eta = 0.5 | 50.0 | 47.2 | 45.8 | 50.0 | 48.3 |
| Depths = 2, eta = 0.75 | 50.0 | 47.2 | 45.8 | 33.3 | 44.1 |
| Depths = 2, eta = 1 | 50.0 | 44.4 | 45.8 | 33.3 | 43.4 |
| Depths = 3, eta = 0.25 | 50.0 | 47.2 | 41.7 | 33.3 | 43.1 |
| Depths = 3, eta = 0.5 | 50.0 | 47.2 | 50.0 | 33.3 | 45.1 |
| Depths = 3, eta = 0.75 | 47.9 | 44.4 | 45.8 | 33.3 | 42.9 |
| Depths = 3, eta = 1 | 47.9 | 44.4 | 41.7 | 41.7 | 43.9 |
| Depths = 4, eta = 0.25 | 43.8 | 44.4 | 45.8 | 41.7 | 43.9 |
| Depths = 4, eta = 0.5 | 47.9 | 44.4 | 41.7 | 41.7 | 43.9 |
| Depths = 4, eta = 0.75 | 52.1 | 47.2 | 41.7 | 58.3 | 49.8 |
| Depths = 4, eta = 1 | 45.8 | 52.8 | 50.0 | 50.0 | 49.7 |
| Depths = 5, eta = 0.25 | 43.8 | 50.0 | 41.7 | 41.7 | 44.3 |
| Depths = 5, eta = 0.5 | 43.8 | 47.2 | 45.8 | 41.7 | 44.6 |
| Depths = 5, eta = 0.75 | 56.2 | 47.2 | 41.7 | 50.0 | 48.8 |
| Depths = 5, eta = 1 | 43.8 | 50.0 | 45.8 | 41.7 | 45.3 |

Table: Predicitve Accuracy of XGB on Validation Set.

# Model Training

## Neural Network

| RFM models | Validation Periods | | | | |
|---|---|---|---|---|---|
| | 48 Months | 36 Months | 24 Months | 12 Months | Average |
| 1st hidden = 2, 2nd hidden = 0 | 52.1 | 55.6 | 41.7 | 50.0 | 49.8 |
| 1st hidden = 3, 2nd hidden = 0 | 39.6 | 44.4 | 50.0 | 58.3 | 48.1 |
| 1st hidden = 4, 2nd hidden = 0 | 58.3 | 36.1 | 33.3 | 50.0 | 44.4 |
| 1st hidden = 5, 2nd hidden = 0 | 52.1 | 52.8 | 58.3 | 41.7 | 51.2 |
| 1st hidden = 2, 2nd hidden = 1 | 45.8 | 44.4 | 54.2 | 75.0 | 54.9 |
| 1st hidden = 3, 2nd hidden = 1 | 50.0 | 44.4 | 54.2 | 41.7 | 47.6 |
| 1st hidden = 4, 2nd hidden = 1 | 62.5 | 44.4 | 41.7 | 58.3 | 51.7 |
| 1st hidden = 5, 2nd hidden = 1 | 56.2 | 55.6 | 50.0 | 58.3 | 55.0 |
| 1st hidden = 2, 2nd hidden = 2 | 54.2 | 47.2 | 41.7 | 50.0 | 48.3 |
| 1st hidden = 3, 2nd hidden = 2 | 43.8 | 58.3 | 50.0 | 41.7 | 48.4 |
| 1st hidden = 4, 2nd hidden = 2 | 47.9 | 44.4 | 45.8 | 58.3 | 49.1 |
| 1st hidden = 5, 2nd hidden = 2 | 43.8 | 41.7 | 54.2 | 50.0 | 47.4 |

Table: Predicitve Accuracy of FNN on Validation Set.

# Model Training

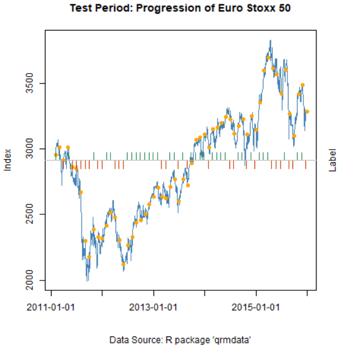
## Assessment

- Classification accuracy of models roughly equivalent with coin-tossings.
- No clear improvement of correct predictions with increasing training data set.
- Results depend on models' specifications and are sensitive to those hyperparameters.
- Methods delivered in general poor results, whence evaluated on last twelve month of validation period ('sideway market environment'). In these market phases directional guidance is required the most, but ML-models did not excel (whip-saw effect).
- Lack of structural stability in data generating process might be an explaination for the rather sobering result of ML during the validation period.

# Model Testing

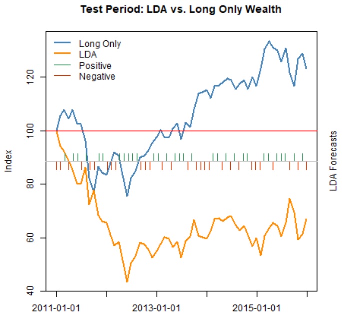
## Outline

- Utilize 'optimal' model specifications determined from validation stage in test period.
- Re-estimate models based on training and validation data.
- Compute one-step ahead predictions.
- Evaluate wealth progressions by entering at the beginning of each month a position with exposure equivalent to one hundred monetary units and close position at month's end.



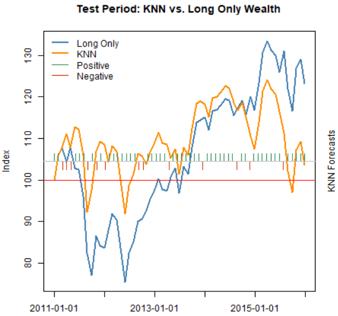
# Model Testing

## Linear Discriminant Analysis




| Performance | LDA | LongOnly |
|---|---|---|
| Return (annualized) | −7.67 | 4.26 |
| Risk (annualized) | 26.04 | 16.74 |
| Sharpe Ratio | −0.29 | 0.25 |
| Maximum Drawdown | 56.49 | 30.04 |
| Average Drawdown | 56.49 | 9.07 |
| Hit Rate (percentage) | 45.00 | |
| False Positive (percentage) | 48.39 | |
| False Negative (percentage) | 62.07 | |

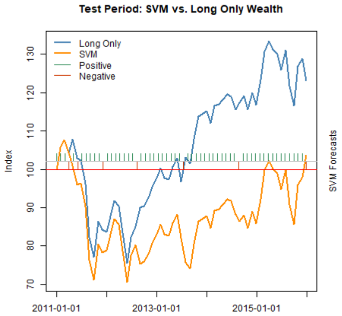
# Model Testing
## k-Nearest Neighbours




| Performance | KNN | LongOnly |
|---|---|---|
| Return (annualized) | 0.69 | 4.26 |
| Risk (annualized) | 15.88 | 16.74 |
| Sharpe Ratio | 0.04 | 0.25 |
| Maximum Drawdown | 21.82 | 30.04 |
| Average Drawdown | 11.81 | 9.07 |
| Hit Rate (percentage) | 48.33 | |
| False Positive (percentage) | 46.81 | |
| False Negative (percentage) | 69.23 | |

# Model Testing

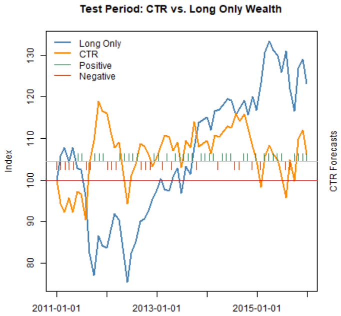
## Support Vector Machine




| Performance | SVM | LongOnly |
|---|---|---|
| Return (annualized) | 0.74 | 4.26 |
| Risk (annualized) | 19.54 | 16.74 |
| Sharpe Ratio | 0.04 | 0.25 |
| Maximum Drawdown | 34.43 | 30.04 |
| Average Drawdown | 34.43 | 9.07 |
| Hit Rate (percentage) | 55.00 | |
| False Positive (percentage) | 43.40 | |
| False Negative (percentage) | 57.14 | |

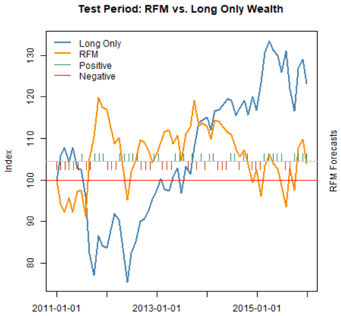
# Model Testing

## Classification Tree




| Performance | CTR | LongOnly |
|---|---|---|
| Return (annualized) | 1.20 | 4.26 |
| Risk (annualized) | 16.49 | 16.74 |
| Sharpe Ratio | 0.07 | 0.25 |
| Maximum Drawdown | 20.85 | 30.04 |
| Average Drawdown | 15.22 | 9.07 |
| Hit Rate (percentage) | 45.00 | |
| False Positive (percentage) | 48.65 | |
| False Negative (percentage) | 65.22 | |

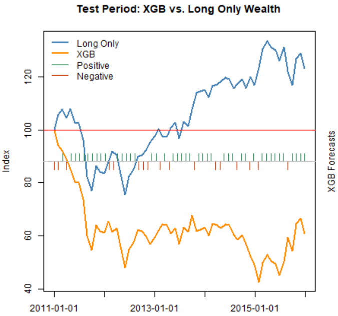
# Model Testing
## Random Forest




| Performance | RFM | LongOnly |
|---|---|---|
| Return (annualized) | 0.78 | 4.26 |
| Risk (annualized) | 16.52 | 16.74 |
| Sharpe Ratio | 0.05 | 0.25 |
| Maximum Drawdown | 22.06 | 30.04 |
| Average Drawdown | 15.36 | 9.07 |
| Hit Rate (percentage) | 45.00 | |
| False Positive (percentage) | 48.39 | |
| False Negative (percentage) | 62.07 | |

# Model Testing

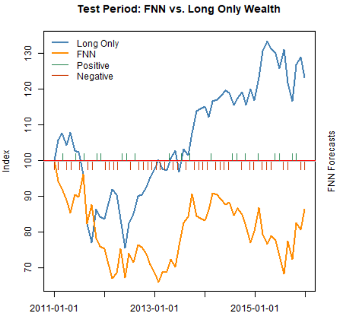
## Extreme Gradient Boosting




| Performance | XGB | LongOnly |
|---|---|---|
| Return (annualized) | −9.45 | 4.26 |
| Risk (annualized) | 28.53 | 16.74 |
| Sharpe Ratio | −0.33 | 0.25 |
| Maximum Drawdown | 57.39 | 30.04 |
| Average Drawdown | 57.39 | 9.07 |
| Hit Rate (percentage) | 38.33 | |
| False Positive (percentage) | 53.49 | |
| False Negative (percentage) | 82.35 | |

# Model Testing

## Neural Network




| Performance | FNN | LongOnly |
|---|---|---|
| Return (annualized) | −2.90 | 4.26 |
| Risk (annualized) | 21.02 | 16.74 |
| Sharpe Ratio | −0.14 | 0.25 |
| Maximum Drawdown | 33.97 | 30.04 |
| Average Drawdown | 33.97 | 9.07 |
| Hit Rate (percentage) | 40.00 | |
| False Positive (percentage) | 55.00 | |
| False Negative (percentage) | 62.50 | |

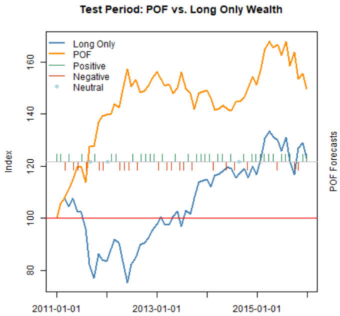
# Model Testing

## Neural Network: Comments

- Activation by rectified linear unit; full batch optimization (rmsprop); no dropout in layers.
- Results depend on chosen optimization engine.
- Results can be not reproducible, even if `keras::use_session_with_seed()` is used (in particular, for settings like mini-batch, stochstic gradient decent, random dropout in layers and non-fixed validation data).
- The peril of data augmentation, *e.g.* by adding Gaussian noise: Beforehand, models should better generalize in testing period, but how much jitter to add? If to small, then no big generalization gains are made, if to high, the labels in the training/validation set might not be any longer in sync with randomly created records.

# Model Testing

## 'Alternative facts: these classy features are all great!' Majority voting of features[3]




| Performance | POF | LongOnly |
|---|---|---|
| Return (annualized) | 8.43 | 4.26 |
| Risk (annualized) | 11.58 | 16.74 |
| Sharpe Ratio | 0.73 | 0.25 |
| Maximum Drawdown | 10.72 | 30.04 |
| Average Drawdown | 6.04 | 9.07 |
| Hit Rate (percentage) | 64.91 | |
| False Positive (percentage) | 32.35 | |
| False Negative (percentage) | 39.13 | |

[3] Bates and Granger 1969.

# Model Testing

## Performance Summary

| Performance | LDA | KNN | SVM | CTR | RFM | XGB | FNN | POF | LongOnly |
|---|---|---|---|---|---|---|---|---|---|
| Return (annualized) | −7.67 | 0.69 | 0.74 | 1.20 | 0.78 | −9.45 | −2.90 | 8.43 | 4.26 |
| Risk (annualized) | 26.04 | 15.88 | 19.54 | 16.49 | 16.52 | 28.53 | 21.02 | 11.58 | 16.74 |
| Sharpe Ratio | −0.29 | 0.04 | 0.04 | 0.07 | 0.05 | −0.33 | −0.14 | 0.73 | 0.25 |
| Maximum Drawdown | 56.49 | 21.82 | 34.43 | 20.85 | 22.06 | 57.39 | 33.97 | 10.72 | 30.04 |
| Average Drawdown | 56.49 | 11.81 | 34.43 | 15.22 | 15.36 | 57.39 | 33.97 | 6.04 | 9.07 |
| Hit Rate (percentage) | 45.00 | 48.33 | 55.00 | 45.00 | 45.00 | 38.33 | 40.00 | 64.91 | |
| False Positive (percentage) | 48.39 | 46.81 | 43.40 | 48.65 | 48.39 | 53.49 | 55.00 | 32.35 | |
| False Negative (percentage) | 62.07 | 69.23 | 57.14 | 65.22 | 62.07 | 82.35 | 62.50 | 39.13 | |

# Model Testing

## Lifting the Veil of Secrecy

- Signs and magnitudes of Kendall correlations between predicators and label give hindsight to differing relations.
- Pooled forecasts by majority votes yield a more robust outcome compared 'trained' ML-/AI models.

| Variable | All | Train | Valid | Test |
|---|---|---|---|---|
| EuroStoxx | 1.00 | 1.00 | 1.00 | 1.00 |
| Kurtosis | 0.02 | 0.04 | −0.07 | 0.00 |
| RSI | −0.11 | −0.10 | −0.06 | −0.23 |
| RvRisk | 0.00 | −0.05 | −0.04 | 0.23 |
| SdRisk | −0.01 | −0.05 | −0.13 | 0.23 |
| SharpeRatio | −0.01 | −0.01 | −0.08 | 0.07 |
| Skewness | 0.00 | 0.01 | −0.09 | 0.03 |
| SlopeCoef | 0.08 | 0.09 | 0.07 | 0.03 |
| TrendDetection | 0.07 | 0.07 | −0.04 | 0.18 |
| TwoPart | −0.02 | 0.02 | −0.15 | −0.07 |

# Summary & Conclusion

- ML-/AI methods have been applied to a binary classification problem.
- All tested methods showed inferior results compared to either a long-only and/or pooled majority forecasts of the predictors.
- Potential reason for this empirical finding is a lack of structural stability across training, validation and testing periods.
- In this sense: ‘a euro is not a euro’.
- So, ‘Will this second wave last . . . ?’: Probably not in the field of empirical social science, but in fairly stable and structured domains *e.g.* in which ‘a cat is still a cat’.

# References: Literature I


Bates, J.M. and C.W.J. Granger (1969). "The Combination of Forecasts". In: *Operations Research Quarterly* 20.4, pp. 451–468.

Bol, Georg, Gholamreza Nakhaeizadeh, and Karl-Heinz Vollmer, eds. (1994). *Ergebnisse des 4. Karlsruher Ökonometrie-Workshops*. Wissenschaftliche Beiträge 93. Heidelberg: Physica-Verlag. ISBN: 978-3-7908-0748-6. DOI: 10.1007/978-3-642-46948-0.

— eds. (1996). *Ergebnisse des 5. Karlsruher Ökonometrie-Workshops*. Wissenschaftliche Beiträge 125. Heidelberg: Physica-Verlag. ISBN: 978-3-7908-0925-1. DOI: 10.1007/978-3-642-61489-7.

— eds. (2000). *Ergebnisse des 7. Karlsruher Ökonometrie-Workshops*. Wissenschaftliche Beiträge 174. Heidelberg: Physica-Verlag. ISBN: 978-3-7908-1284-8. DOI: 10.1007/978-3-642-57656-0.

Breiman, L. (Oct. 2001). "Random Forests". In: *Machine Learning* 45.1, pp. 5–32. DOI: 10.1023/A:1010933404324. URL: https://doi.org/10.1023/A:1010933404324.

Breiman, L. et al. (1984). *Classification and Regression Trees*. Monterey, CA: Wadsworth and Brooks.

Chollet, F. and J.J. Allaire (2018). *Deep Learning with R*. Birmingham: Manning Publications Company. ISBN: 978-1-617-29554-6.

Fisher, R. A. (1936). "THE USE OF MULTIPLE MEASUREMENTS IN TAXONOMIC PROBLEMS". In: *Annals of Eugenics* 7.2, pp. 179–188. DOI: 10.1111/j.1469-1809.1936.tb02137.x. URL: http://dx.doi.org/10.1111/j.1469-1809.1936.tb02137.x.

Friedman, Jerome H. (2001). "Greedy Function Approximation: A Gradient Boosting Machine". In: *Annals of Statistics* 29.5, pp. 1189–1232.

— (Feb. 2002). "Stochastic Gradient Boosting". In: *Computational Statistics & Data Analysis* 38.4, pp. 367–378. ISSN: 0167-9473. DOI: 10.1016/S0167-9473(01)00065-2. URL: http://dx.doi.org/10.1016/S0167-9473(01)00065-2.

Goodfellow, I. (2016). *Deep learning*. Cambridge, Massachusetts: The MIT Press. ISBN: 9780262035613.

Hart, P. (1968). "The condensed nearest neighbor rule (Corresp.)". In: *IEEE Transactions on Information Theory* 14.3, pp. 515–516. DOI: 10.1109/tit.1968.1054155. URL: http://dx.doi.org/10.1109/TIT.1968.1054155.

Ho, T.K. (n.d.). "Random Decision Forests". In: Proceedings of the 3rd International Conference on Document Analysis and Recognition. Montreal, QC, Canada, pp. 278–282.

— (1998). "The Random Subspace Method for Constructing Decision Forests". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 20.8, pp. 832–844. DOI: 10.1109/34.709601. URL: http://dx.doi.org/10.1109/34.709601.

# References: Literature II


Kleinberg, E. (1990). "Stochastic Discrimination". In: *Annals of Mathematics and Artificial Intelligence* 1.1–4, pp. 207–239.

— (1996). "An Overtraining-Resistant Stochastic Modeling Method for Pattern Recognition". In: *Annals of Statistics* 24.6, pp. 2319–2349.

— (2000). "On the Algorithmic Implementation of Stochastic Discrimination". In: *IEEE Transactions on PAMI* 22.5, pp. 473–490.

Kuhn, Thomas S. (1970). *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press, pp. xii, 210.

Prado, M.L. (2018). *Advances in Financial Machine Learning*. 1st. New York: John Wiley & Sons. ISBN: 978-1-119-48208-6.

Ripley, B. (2018). *tree: Classification and Regression Trees*. R package version 1.0-39. URL: https://CRAN.R-project.org/package=tree.

Rosenblatt, Frank (1958). "The perceptron: a probabilistic model for information storage and organization in the brain". In: *Psychological Review* 65.6, pp. 386–408. URL: https://www.ling.upenn.edu/courses/cogs501/Rosenblatt1958.pdf.

Russell, Stuart (2010). *Artificial intelligence : a modern approach*. Upper Saddle River, NJ: Prentice Hall. ISBN: 9780136042594.

Seligman, M. (2017). *Rborist: Extensible, Parallelizable Implementation of the Random Forest Algorithm*. R package version 0.1-8. URL: https://CRAN.R-project.org/package=Rborist.

Vapnik, V. (1995). *The Nature of Statistical Learning Theory*. New York: Springer.

Vapnik, V. and A. Chervonenkis (1974). *Theory of Pattern Recognition [in Russian]*. (German Translation: W. Wapnik and A. Tscherwonenkis, Theorie der Zeichenerkennung, Akademie-Verlag, Berlin, 1979). Moscow: Nauka.

Vapnik, V., S. Golowich, and A. Smola (1997). "Support vector method for function approximation regression estimation and signal processing". In: *Advances in Neural Information Processing Systems*. Ed. by M. Mozer, M. Jordan, and T. Petsche. Vol. 9. Cambridge, MA: MIT Press, pp. 281–287.

Vapnik, V. and A. Lerner (1963). "Pattern recognition using generalized portrait method". In: *Automation and Remote Control* 24.

# References: R packages I


Allaire, J.J. and F. Chollet (2018). *keras: R Interface to 'Keras'*. R package version 2.2.4. URL: https://CRAN.R-project.org/package=keras.

Allaire, J.J. and Y. Tang (2018). *tensorflow: R Interface to 'TensorFlow'*. R package version 1.10. URL: https://CRAN.R-project.org/package=tensorflow.

Ardia, David and Kris Boudt (2015). "Testing equality of modified Sharpe ratios". In: *Finance Research Letters* 13, pp. 97–104. DOI: 10.1016/j.frl.2015.02.008.

— (2018). "The peer performance ratios of hedge funds". In: *Journal of Banking and Finance* 87, pp. 351–368. DOI: 10.1016/j.jbankfin.2017.10.014.

Chen, Tianqi et al. (2019). *xgboost: Extreme Gradient Boosting*. R package version 0.90.0.2. URL: https://CRAN.R-project.org/package=xgboost.

David, Ardia and Kris Boudt (2018). *PeerPerformance: Luck-corrected peer performance analysis in R*. version 2.2.1. URL: https://github.com/ArdiaD/PeerPerformance.

E., LeDell; et al. (2019). *h2o: R Interface for 'H2O'*. R package version 3.22.1.1. URL: https://CRAN.R-project.org/package=h2o.

Greenwell, Brandon et al. (2019). *gbm: Generalized Boosted Regression Models*. R package version 2.1.5. URL: https://CRAN.R-project.org/package=gbm.

Harrell Jr, Frank E, with contributions from Charles Dupont, and many others. (2019). *Hmisc: Harrell Miscellaneous*. R package version 4.2-0. URL: https://CRAN.R-project.org/package=Hmisc.

Hothorn, T., P. Bühlmann, et al. (2006). "Survival Ensembles". In: *Biostatistics* 7.3, pp. 355–373.

Hothorn, T., K. Hornik, and A. Zeileis (2006). "Unbiased Recursive Partitioning: A Conditional Inference Framework". In: *Journal of Computational and Graphical Statistics* 15.3, pp. 651–674.

Karatzoglou, A. et al. (2004). "kernlab – An S4 Package for Kernel Methods in R". In: *Journal of Statistical Software* 11.9, pp. 1–20. URL: http://www.jstatsoft.org/v11/i09/.

Liaw, A. and M. Wiener (2002). "Classification and Regression by randomForest". In: *R News* 2.3, pp. 18–22. URL: https://CRAN.R-project.org/doc/Rnews/.

Meyer, D. et al. (2018). *e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien*. R package version 1.7-0. URL: https://CRAN.R-project.org/package=e1071.

# References: R packages II


Peterson, Brian G. and Peter Carl (2019). *PerformanceAnalytics: Econometric Tools for Performance and Risk Analysis*. R package version 1.5.3. URL: https://CRAN.R-project.org/package=PerformanceAnalytics.

R Core Team (2019). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. URL: https://www.R-project.org/.

Rong, X. (2014). *deepnet: deep learning toolkit in R*. R package version 0.2. URL: https://CRAN.R-project.org/package=deepnet.

Strobl, C., A. Boulesteix, T. Kneib, et al. (2008). "Conditional Variable Importance for Random Forests". In: *BMC Bioinformatics* 9.307. URL: http://www.biomedcentral.com/1471-2105/9/307.

Strobl, C., A. Boulesteix, A. Zeileis, et al. (2007). "Bias in Random Forest Variable Importance Measures: Illustrations, Sources and a Solution". In: *BMC Bioinformatics* 8.25. URL: http://www.biomedcentral.com/1471-2105/8/25.

Therneau, T. and B. Atkinson (2018). *rpart: Recursive Partitioning and Regression Trees*. R package version 4.1-13. URL: https://CRAN.R-project.org/package=rpart.

Ulrich, Joshua (2018). *TTR: Technical Trading Rules*. R package version 0.23-4. URL: https://CRAN.R-project.org/package=TTR.

Venables, W.N. and B.D. Ripley (2002). *Modern Applied Statistics with S*. 4th. ISBN 0-387-95457-0. New York: Springer. URL: http://www.stats.ox.ac.uk/pub/MASS4.

Weihs, C. et al. (2005). "klaR Analyzing German Business Cycles". In: *Data Analysis and Decision Support*. Ed. by D. Baier, R. Decker, and L. Schmidt-Thieme. Berlin: Springer-Verlag, pp. 335–343.

Zeileis, A., T. Hothorn, and K. Hornik (2008). "Model-Based Recursive Partitioning". In: *Journal of Computational and Graphical Statistics* 17.2, pp. 492–514.