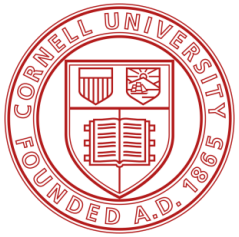


Advances in Financial Machine Learning: Lecture 2/10

Prof. Marcos López de Prado

Seeing Beyond The Hype

- Financial ML offers the opportunity to gain insight from data:
  - Modelling non-linear relationships in a high-dimensional space
  - Analyzing unstructured data (asynchronous, categorical)
  - Learning patterns with complex interactions (hierarchical, non-parametric)
  - Focusing on predictability over parametric adjudication
  - Controlling for overfitting (early-stopping, cross-validation)
- At the same time, **Finance is not a plug-and-play subject** as it relates to machine learning.
  - Modelling financial series is harder than driving cars or recognizing faces.
  - **A ML algorithm will always find a pattern, even if there is none!**
- In this presentation, we review a few important financial ML applications.

What is Machine Learning?

What Is Machine Learning?

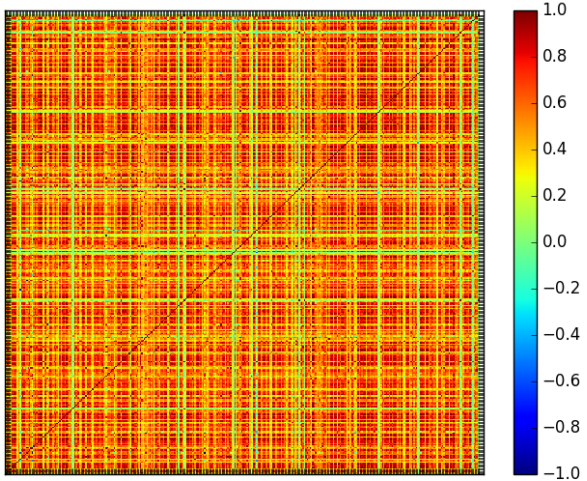
"An ML algorithm learns complex patterns in a high-dimensional space without being specifically directed."

Advances in Financial Machine Learning (2018, p.15)

Let's break this statement into its components.

- **"learns … without being specifically directed"**: Unlike with other empirical tools, researchers do not impose a particular structure on the data. Instead, researchers let the data speak.
- **"learns complex patterns"**: The ML algorithm may find a pattern that cannot be easily represented with a finite set of equations.
- **"learns … in a high-dimensional space"**: Solutions often involve a large number of variables and the interactions between them.

Suppose that you have a 1000x1000 correlation matrix...



What Is Machine Learning?

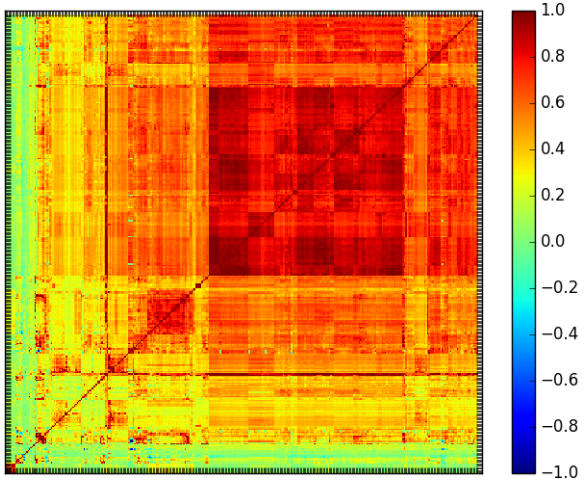
"An ML algorithm learns complex patterns in a high-dimensional space without being specifically directed."

Advances in Financial Machine Learning (2018, p.15)

Let's break this statement into its components.

- **"learns … without being specifically directed"**: Unlike with other empirical tools, researchers do not impose a particular structure on the data. Instead, researchers let the data speak.
- **"learns complex patterns"**: The ML algorithm may find a pattern that cannot be easily represented with a finite set of equations.
- **"learns … in a high-dimensional space"**: Solutions often involve a large number of variables and the interactions between them.

Suppose that you have a 1000x1000 correlation matrix... A clustering algorithm finds that there are 3 blocks: Highly correlated, low correlated, uncorrelated.



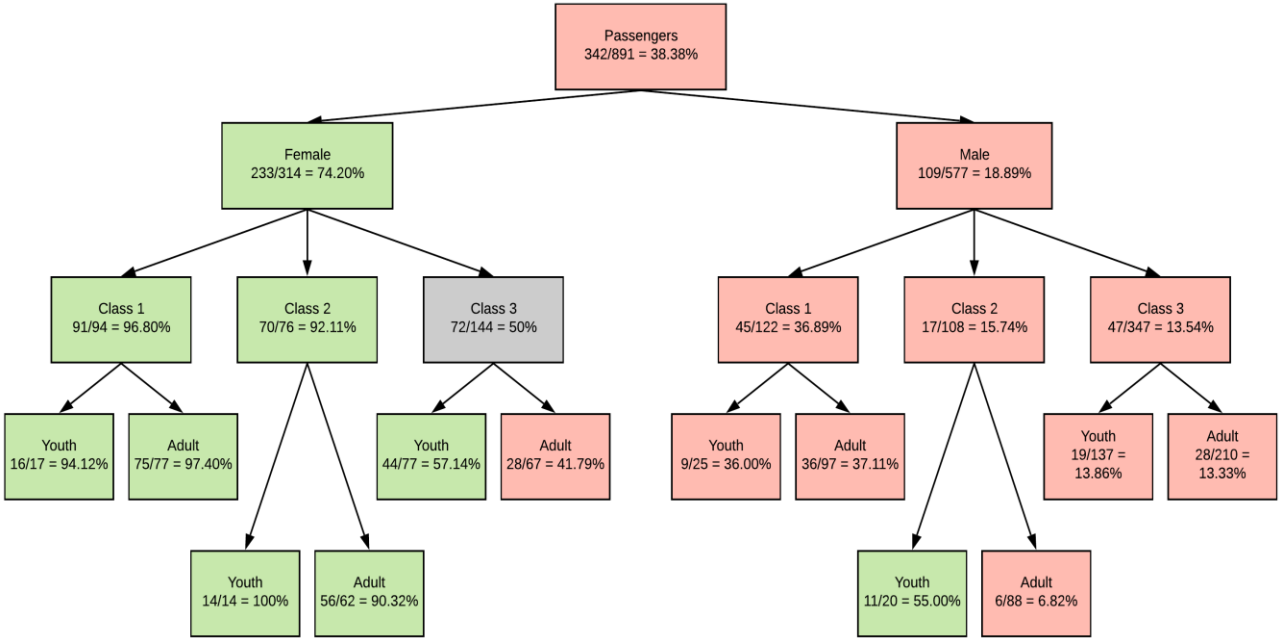
A Simple Example: The Survival Rate At The Titanic

How would you predict the probability that a particular passenger at the Titanic survived?

An ML algorithm will find that three variables are relevant: *Gender*, *ticket class*, and *age*.

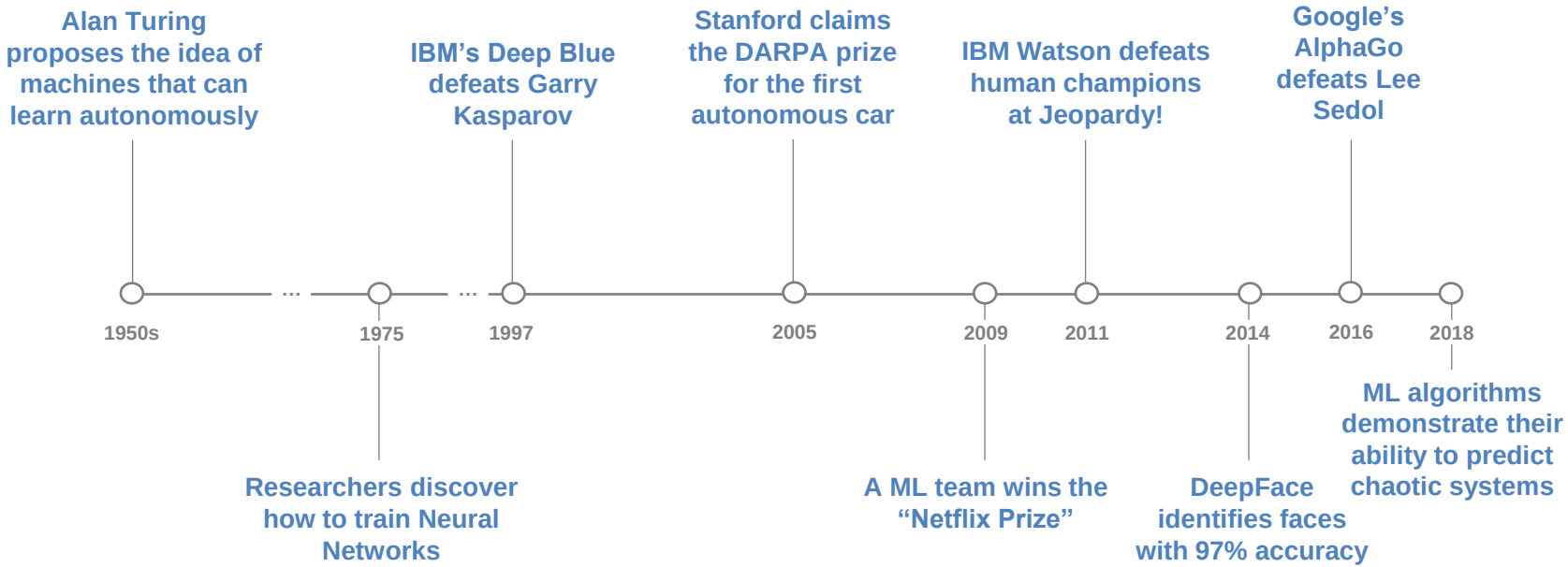
The algorithm will also find that the there is a hierarchical structure in the data.

For example, for the purpose of surviving, being female is more important than being young or having a first class ticket.



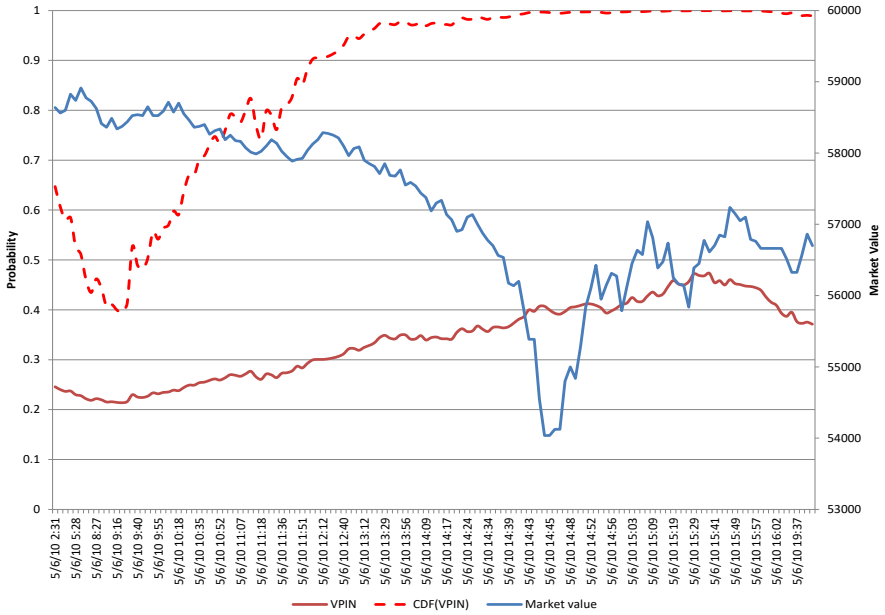
Source: Lam, E. and C. Tang. For illustrative purposes only. Available at https://bit.ly/2UjVmOI

Timeline



Can ML Predict Black Swans?

A **black swan** is an extreme event that has not been observed before. E.g., the "flash crash" of May 6 2010.



VPIN detected extreme levels of persistent order flow imbalance two hours prior to the crash

The official investigation into the flash crash found that the likely cause was an order to sell 75,000 E-mini S&P 500 futures contracts at a high participation rate.

That large order caused a persistent imbalance in the order flow, which triggered a cascade of stop-outs across market makers, until nobody stood on the bid.

Imbalanced order flow is the norm, with various degrees of persistency. The 10% sudden drop in prices was a black swan, but the causes were known to microstructure theory.

Conclusion: **Black swans can be predicted *by theory,* even if they cannot be predicted *by algorithms***.

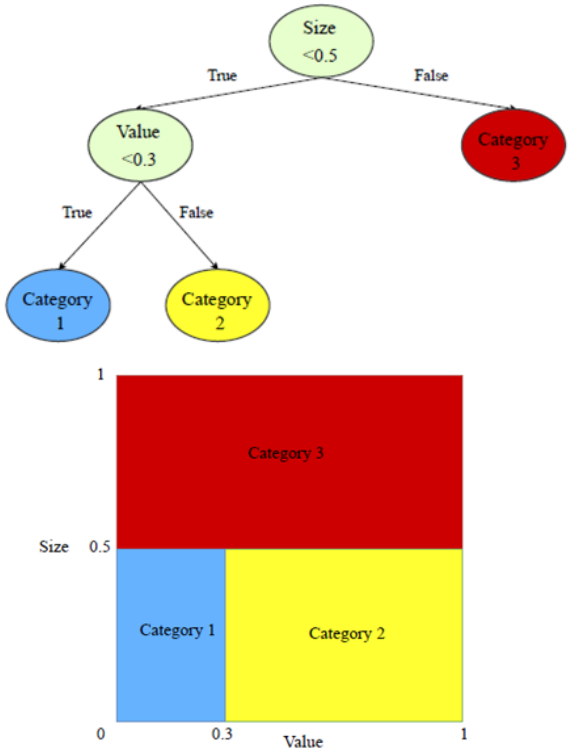
Corollary: Use ML for developing theories, and let the theories make the predictions (not the algorithms).

Simons on Machine Learning

Current Applications of Financial ML

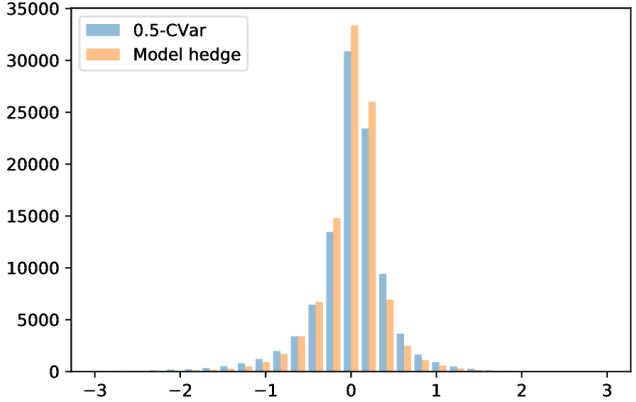
1. Price Prediction

- ML methods allow the modelling of complex relations among the 6 or 7 widely accepted economic factors, including
  - Non-linear relations
  - Threshold relations
  - Hierarchical relations
  - Categorical variables
  - Unknown specification
  - Interaction effects
  - Control variables
- Econometric methods fail to recognize complex relationships, hence leading to inferior results.

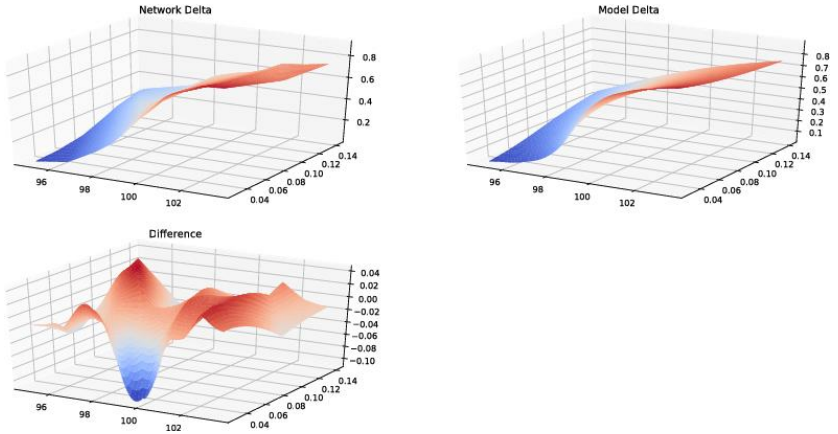


2. Hedging

- Analytical hedging is problematic in presence of market frictions, such as transaction costs, market impact, liquidity constraints, risk limits, etc.
- Reinforcement learning approaches are Greek-free and model free. They are purely empirical, with very few theoretical assumptions.
  - These models consider many more variables and data points when making hedging decisions, and can generate more accurate hedges at greater speeds.

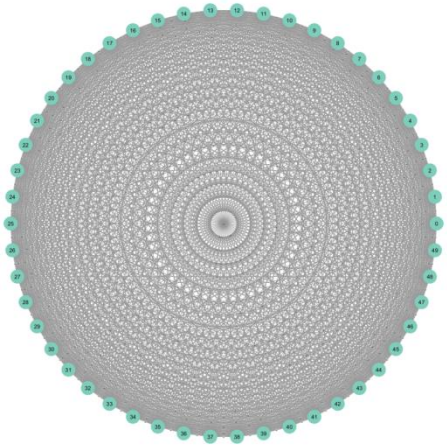


Comparison of model hedge and deep hedge associated to 50%-expected shortfall criterion



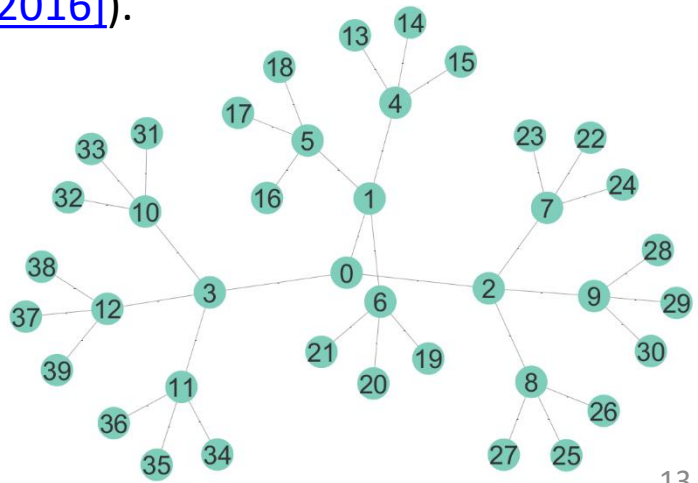
3. Portfolio Construction / Risk Analysis

- Most firms continue to allocate trillions of dollars using mean-variance portfolio optimization (MVO). *"The most expensive piece of beautiful math in history."*
- It is widely known that MVO underperforms the naïve allocation out-of-sample (De Miguel et al. [2009]).
- In contrast, ML solutions outperform MVO (and 1/N) out-of-sample, with gains in Sharpe ratio that exceed 31% (López de Prado [2016]).



Covariance-based models require the independent estimation of $N(N+1)/2$ variables.

ML models need only $N-1$ *hierarchical* estimates, making them more robust and reliable.



4. Structural Breaks / Outlier Detection

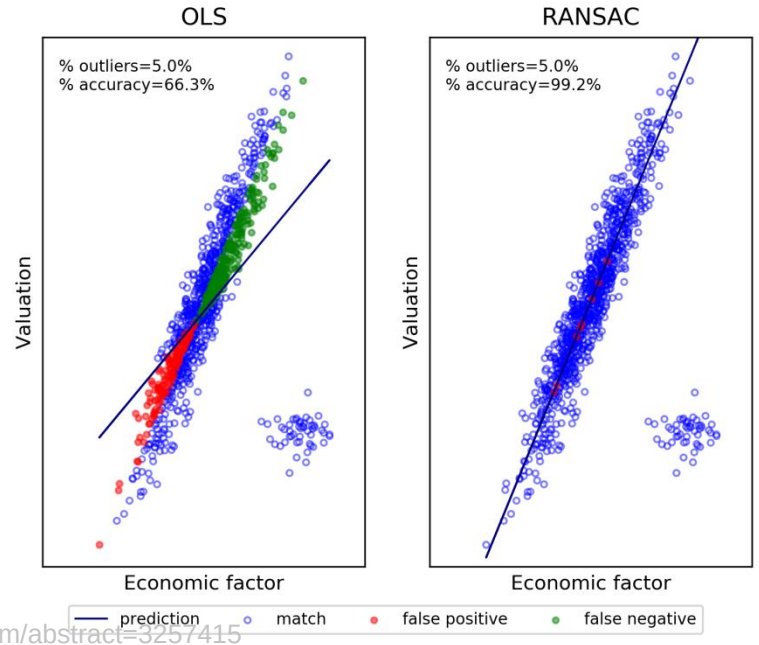
Cross-sectional studies are particularly sensitive to the presence of outliers. Even a small percentage of outliers can cause a very large percentage of wrong signals: Buys that should be sells (false positives), and sells that should be buys (false negatives).

In this plot we run a regression on a cross-section of securities, where a very small percentage (only 5%) are outliers:

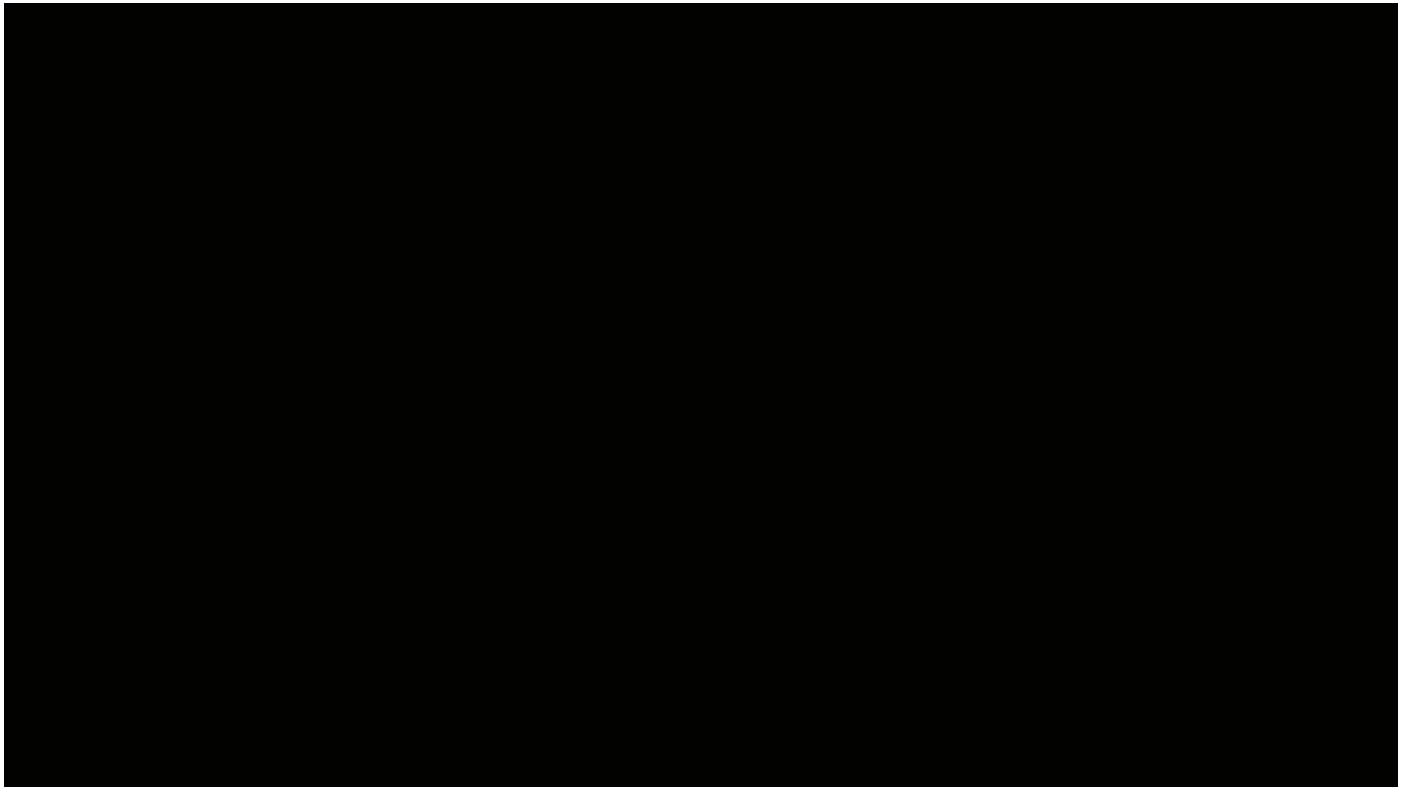
- The **red dots** are securities that are expensive, but the regression wrongly classified as cheap.
- The **green dots** are securities that are cheap, but the regression wrongly classified as expensive.

With only 5% of outliers, the cross-sectional regression produced a 34% classification error. In contrast, RANSAC's classification error was 1%, involving borderline cases.

Whenever you suspect the presence of outliers in your data, consider applying RANSAC or similar ML methods.

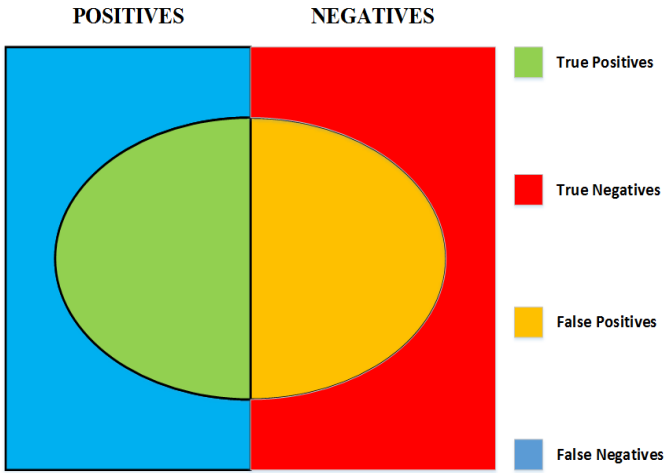


4. Structural Breaks / Outlier Detection



5. Bet Sizing / Alpha Capture

- Suppose that you have a model for making a buy-or-sell decision:
  - You just need to learn the size of that bet, which includes the possibility of no bet at all (zero size).
  - This is a situation that practitioners face regularly. We often know whether we want to buy or sell a product, and the only remaining question is how much money we should risk in such bet.
  - Meta-labeling: Label the outcomes of the primary model as 1 (gain) or 0 (loss).



- Meta-labeling builds a secondary ML model that learns how to use a primary exogenous model.
- The secondary model does not learn the *side*. It only learns the *size*.
- We can maximize the F1-score:

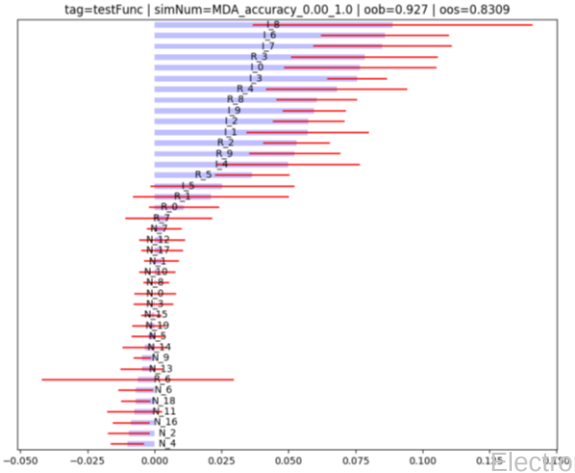
$$F1 = 2\frac{precision \cdot recall}{precision + recall}$$

$$precision = \frac{TP}{TP + FP}$$

$$recall = \frac{TP}{TP + FN}$$

6. Feature Importance

- ML algorithms identify patterns in a high dimensional space.
- These patterns associate features with outcomes.
- The nature of the relationship can be extremely complex, however we can always study what features are more important.
  - E.g., even if a ML algorithm may not derive an analytical formula for Newton's Gravitational Law, it will tell us that *mass* and *distance* are the key features.



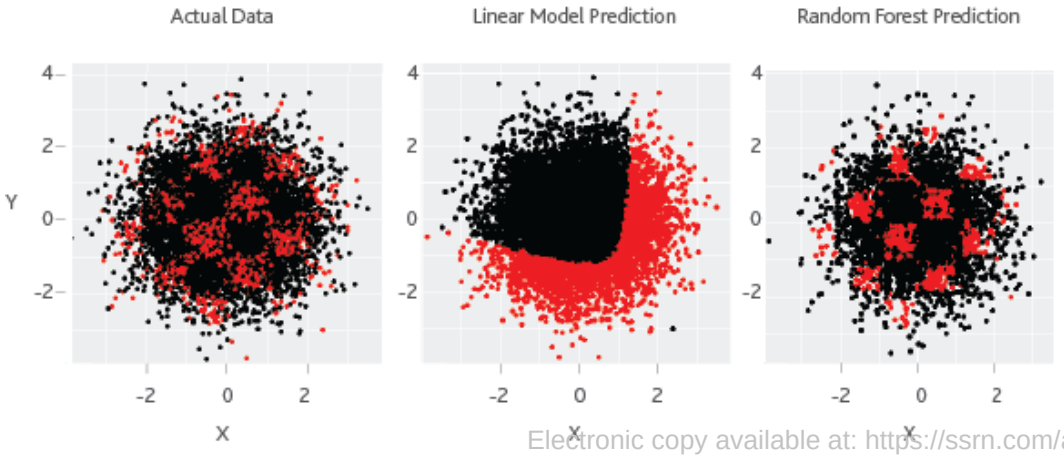
In traditional statistical analysis, key features are often missed as a result of the model's misspecification.

In ML analysis, we give up closed-form specifications in exchange for identifying what variables are important for forecasting.

Once we know *what* are the factors at play, we can develop a theory of *how*.

7. Credit Ratings, Analyst Recommendations

- Stock analysts apply a number of models and heuristics to produce credit and investment ratings.
- These decisions are not entirely arbitrary, and correspond to a complex logic that cannot be represented with a simple set of formulas or a well-defined procedure.
- Machine learning algorithms have been successful at replicating a large percentage of recommendations produced by bank analysts and credit rating agencies.



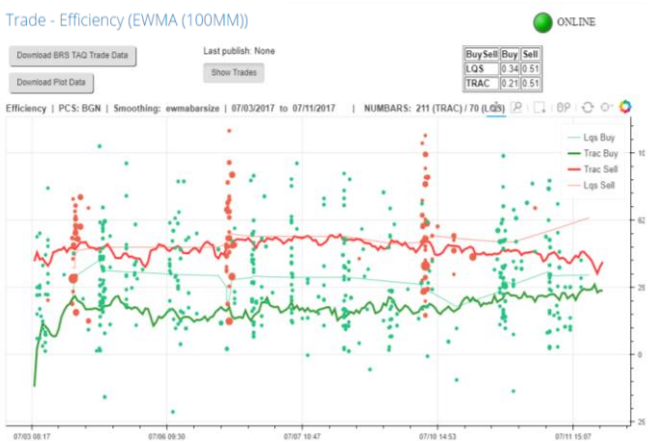
In this example by Moody's, the left figure shows a scatter plot of bonds as a function of two features (X,Y), where defaults are colored in red. The middle plot shows that traditional econometric methods fail at modelling this complex, non-linear relationship. The right plot shows that a very simple ML algorithm performs well.

8. Unstructured Data

- In the plot below, an algorithm has identified news articles containing information relevant to Tesla (TSLA US Equity).
  - **Blue bars**: Daily count of the total number of articles. The average is 458 articles/day, with a maximum of ~5000.
  - **Green bars**: Daily count of articles expressing a positive sentiment.
  - **Red bars**: Daily count of articles expressing a negative sentiment.

9. Execution

- Credit instruments are
  - traded over-the-counter
  - relatively illiquid (they may not trade for days and weeks)
- Kernel-based methods identify “similar” trades based on their common features.
  - The set of common trades enables us to derive theoretical prices.
  - If we buy a bond at a price higher than subsequent “similar” bonds, we can bust the trade.

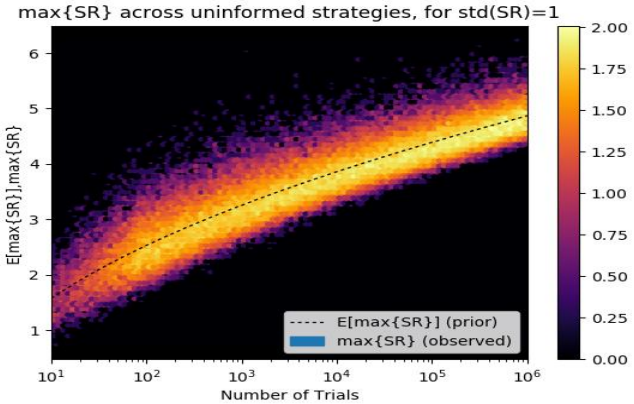


This plot shows the trade efficient of buys (green) and sales (red):

- A **buy** has efficiency 0 when it prints at the quoted offer, and it has efficiency 100 when it prints at the quoted bid.
- A **sale** has efficiency 0 when it prints at the quoted bid, and it has efficiency 100 when it prints at the quoted offer.
- Both have efficiency 50 at the mid.

In this example, the rebalancing of the portfolio has been profitable, as it has captured about 1/3 of the bid-ask spread (approx. 50 bps in price).

10. Detection of False Investment Strategies

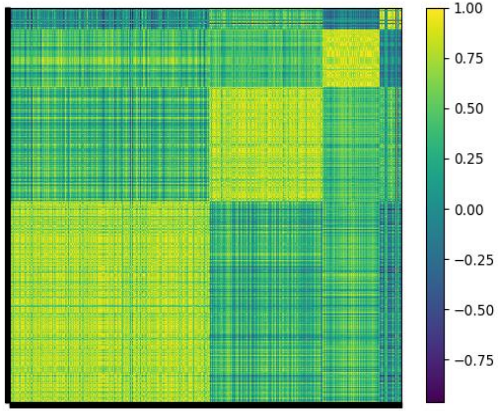


The y-axis displays the distribution of the maximum Sharpe ratios (max{SR}) for a given number of trials (x-axis). A lighter color indicates a higher probability of obtaining that result, and the dash-line indicates the expected value.

For example, after only 1,000 independent backtests, the expected maximum Sharpe ratio (E[max{SR}]) is 3.26, even if the true Sharpe ratio of the strategy is zero!

Most quantitative firms invest in false discoveries.

Solution: Deflate the Sharpe ratio by the number and variance of trials.



Stats	Cluster 0	Cluster 1	Cluster 2	Cluster 3
Strat Count	3265	1843	930	347
aSR	1.5733	1.4907	2.0275	1.0158
SR	0.0974	0.0923	0.1255	0.0629
Skew	-0.3333	-0.4520	-0.4194	0.8058
Kurt	11.2773	6.0953	7.4035	14.2807
T	2172	2168	2174	2172
StartDt	2010-01-04	2010-01-04	2010-01-04	2010-01-04
EndDt	2018-05-01	2018-04-25	2018-05-03	2018-05-01
Freq	261.0474	261.0821	261.1159	261.0474
sqrt(V[SR_k])	0.0257	0.0256	0.0256	0.0257
E[max SR_k]	0.0270	0.0270	0.0270	0.0270
DSR	0.9993	0.9985	**1.0000**	0.9558

The selected strategy belongs to Cluster 2. After taking into account the number and variance of trials involved in the discovery, the probability that $SR > 0$ is virtually 1. Hence, the backtest is unlikely to be overfit.

The Perils of Financial ML

The “spilled samples” problem (1/2)

- Most non-financial ML researchers can assume that observations are drawn from IID processes. For example, you can obtain blood samples from a large number of patients, and measure their cholesterol.
- Of course, various underlying common factors will shift the mean and standard deviation of the cholesterol distribution, but the samples are still independent: There is one observation per subject.
- Suppose you take those blood samples, and someone in your laboratory spills blood from each tube to the following 9 tubes to their right.
  - That is, tube 10 contains blood for patient 10, but also blood from patients 1 to 9. Tube 11 contains blood from patient 11, but also blood from patients 2 to 10, and so on.

The “spilled samples” problem (2/2)

- Now you need to determine the features predictive of high cholesterol (diet, exercise, age, etc.), without knowing for sure the cholesterol level of each patient.
- That is the equivalent challenge that we face in financial ML.
  - Labels are decided by outcomes.
  - Outcomes are decided over multiple observations.
  - Because labels overlap in time, we cannot be certain about what observed features caused an effect.

My friend Luna can recognize faces, like Google or FaceBook. She is not so good at investing, and Google’s ML would probably fail miserably if applied to financial markets.

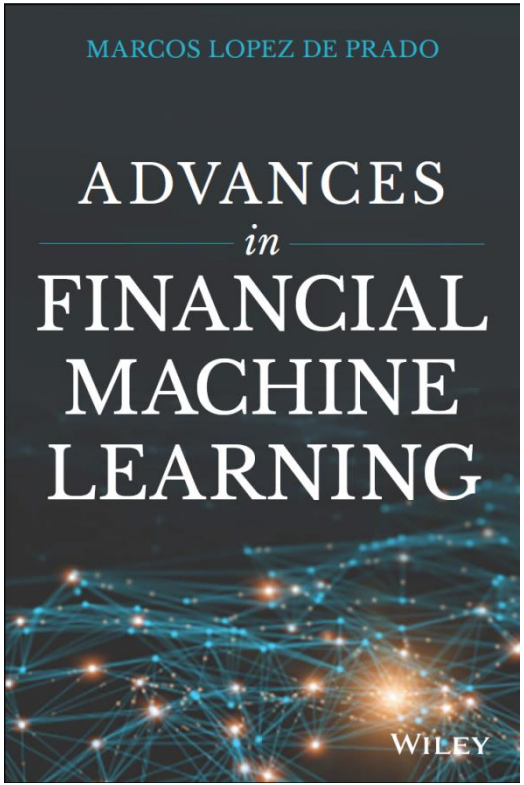
Finance is not a plug-and-play subject as it relates to ML. There are no “West Coast” solutions to “East Coast” problems.

Financial ML as a specific subject

- Financial series exhibit properties that are inconsistent with standard ML assumptions. **A ML algo will always find a pattern, even if there is none!**

PROBLEM	A SOLUTION
Non-stationarity with long memory	Fractional differentiation
Variable information arrival rate	Order imbalance bars
Outcomes span multiple observations	Triple barrier method, with uniqueness weighting
Regime switches	Structural-break methods
Dependence, serial and cross-sectional	K-fold CV with purging, embargoing
Single path for backtesting	Combinatorial cross-validation
Low signal/noise ratio. Backtest overfitting	Deflated Sharpe ratio by controlling for the number of trials

For Additional Details



The first wave of quantitative innovation in finance was led by Markowitz optimization. Machine Learning is the second wave and it will touch every aspect of finance. López de Prado's Advances in Financial Machine Learning is essential for readers who want to be ahead of the technology rather than being replaced by it.

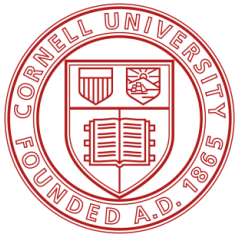
— Prof. **Campbell Harvey**, Duke University. Former President of the American Finance Association.

Financial problems require very distinct machine learning solutions. Dr. López de Prado's book is the first one to characterize what makes standard machine learning tools fail when applied to the field of finance, and the first one to provide practical solutions to unique challenges faced by asset managers. Everyone who wants to understand the future of finance should read this book.

— Prof. **Frank Fabozzi**, EDHEC Business School. Editor of The Journal of Portfolio Management.

Disclaimer

- The views expressed in this document are the authors' and do not necessarily reflect those of the organizations he is affiliated with.
- No investment decision or particular course of action is recommended by this presentation.
-



Overfitting: Causes and Solutions

Prof. Marcos López de Prado
Cornell University – School of Engineering
ORIE 5256

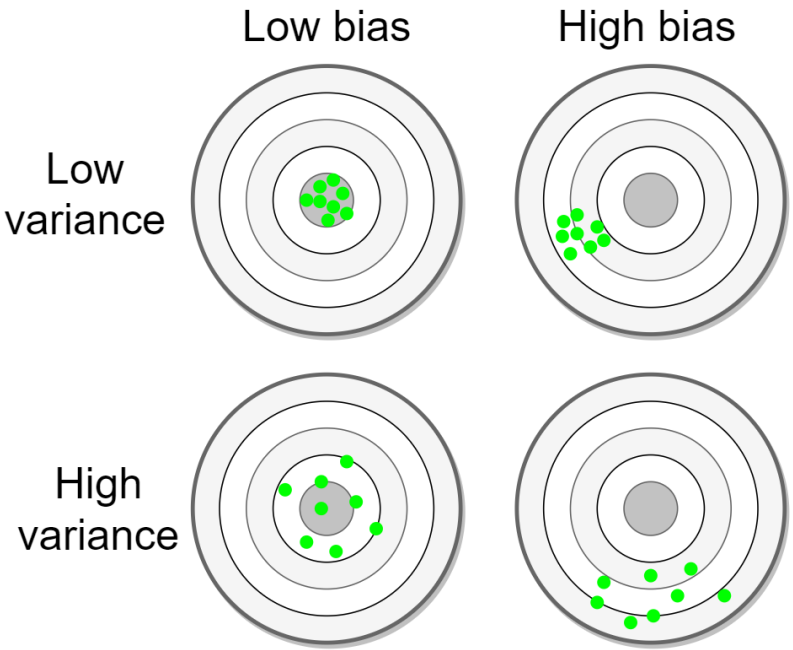
Key Points

- Scientific disciplines have successfully applied Machine Learning (ML) methods for decades
- In recent years, investment managers have begun to replace or complement classical statistical methods (e.g., Econometrics) with computer-based statistical methods (e.g., ML)
  - Well-known ML firms include RenTec, Two Sigma, DE Shaw, TGS, Capital Fund Management, etc.
- Classical methods are prone to overfitting due to their
  - reliance on train-set error estimates
  - assumption that only one trial has taken place
- **When used incorrectly, the risk of ML overfitting is higher than with classical methods**
- **This presentation reviews the causes and solutions wrt overfitting**

What is Overfitting?

Error Decomposition

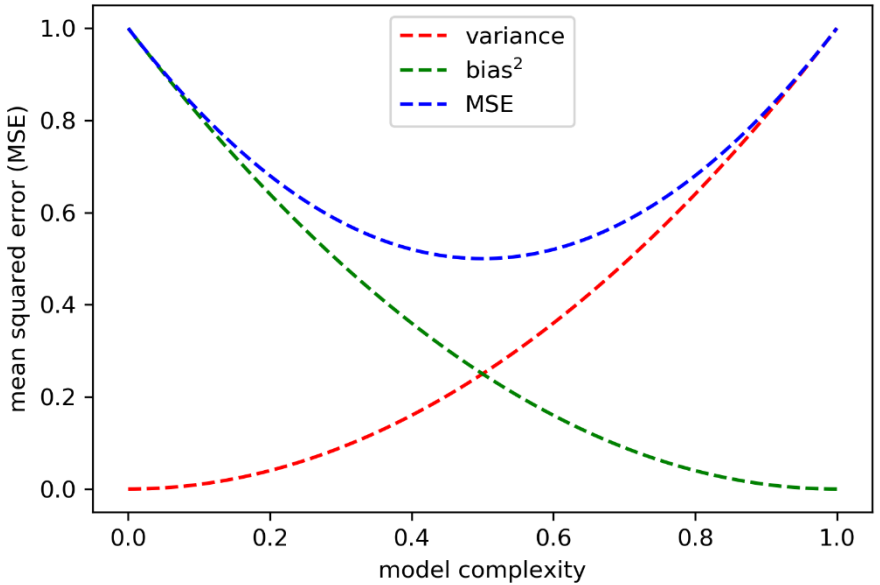
- Consider a function $f[x]$ that predicts outcomes y, such that errors $\varepsilon = y - f[x]$ are unpredictable, with $\varepsilon \sim N[0, \sigma_\varepsilon^2]$
- A statistical model proposes a function $\hat{f}[x]$ that approximates $f[x]$
- The mean squared error (MSE), $\mathrm{E}\left[\left(y - \hat{f}[x]\right)^2\right]$, is the sum of:
  - Bias squared: $\left(\mathrm{E}\left[f[x] - \hat{f}[x]\right]\right)^2$
  - Variance: $\mathrm{V}\left[\hat{f}[x]\right]$
  - Noise: $\mathrm{V}[\varepsilon]$



Combination of bias and variance in an estimator.

Bias-Variance Trade-Off

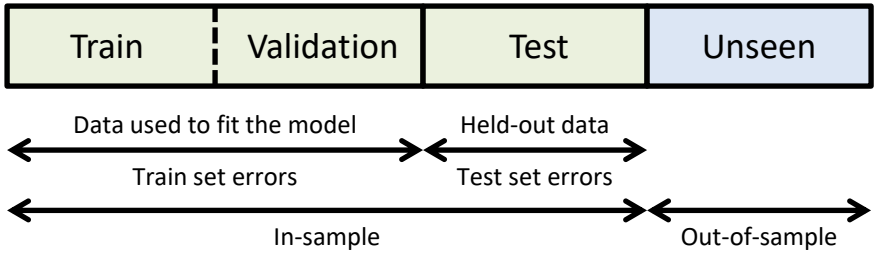
- Bias occurs when $\hat{f}[x]$ underfits the data
  - The model confounds signal for noise
- Variance occurs when $\hat{f}[x]$ overfits the data
  - The model confounds noise for signal
- In general, bias can only be reduced at the expense of increasing variance
- **Overfitting** causes model variance, because a model overfit on one set does not *generalize* well outside that set
  - The model tries to forecast noise



A good statistical model minimizes the mean squared error (MSE) by finding the optimal balance between bias and variance.

Two Kinds of Errors

- We can split a dataset into two subsets
  - **Train set**: Used to select features and fit model parameters
    - This may include a **validation set**, used to find the optimal hyper-parameters
  - **Test set**: Hold-out data, not used for fitting the model
- We can estimate two in-sample errors:
  - **Train set errors**: Errors estimated on the train set (the same data used to fit the model)
  - **Test set errors**: Errors estimated on the test set
- Overfitting can occur when we try to minimize one or both of these errors



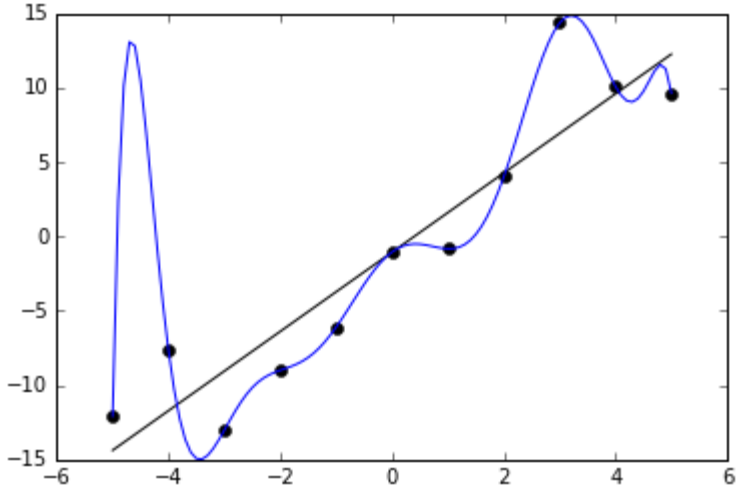
Overfitting can occur on the train set and/or the test set.

The amount of overfitting can be estimated through the **generalization error**: the model's error on data not used to choose the model.

The Two Kinds of Overfitting

Train Set Overfitting

- Train set overfitting occurs when
  - a model is chosen to minimize train set errors
  - at the expense of higher variance on test set errors
- Train set overfitting is related to model complexity
  - This overcomplexity attempts to fit signal, but it ends up fitting noise
- Train set overfitting can be easily diagnosed by estimating the generalization error on the test set, via
  - Resampling methods (e.g., cross-validation)
  - Monte Carlo
- Solutions: Simplify the model; get more data

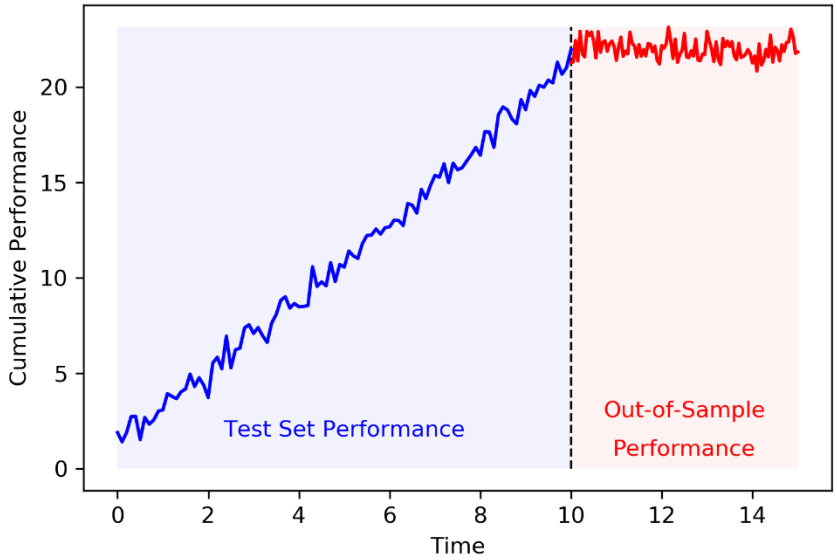


Source: Wikimedia Commons

A typical example of overfitting: The (complex) polynomial function provides a perfect fit because it explains all the noise, however it will generalize worse than a (simpler) line.

Test Set (or Backtest) Overfitting

- Test set overfitting occurs when
  - a model is chosen to minimize test set errors
  - at the expense of higher out-of-sample variance
- Test set overfitting is related to selection bias under multiple testing (SBuMT)
- Test set overfitting can be diagnosed by
  - estimating the generalization error on unseen data (out-of-sample)
  - controlling for the number and variance of the independent trials involved in the model selection
- Solutions:
  - start all over with a new (unseen) dataset
  - adjust the probability of a false positive

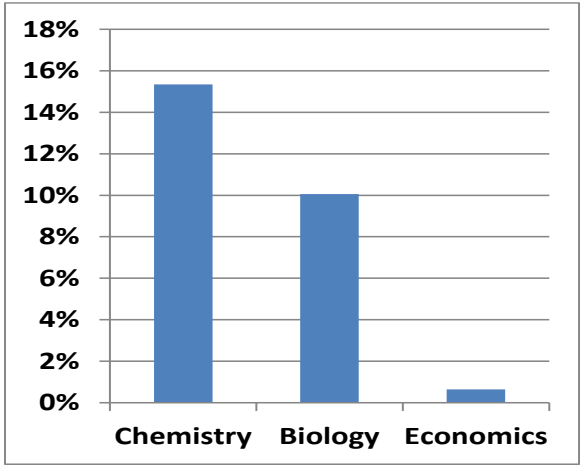


A strategy overfit on the test set will fail to perform on unseen data (out-of-sample). Note that this kind of overfitting is **entirely unrelated to model complexity**.

Classical Statistical Methods

What are Classical Statistical Methods?

- Classical statistical methods follow the research program initiated by Ronald Fisher's
  - Statistical Methods for Research Workers (1925)
  - The Design of Experiments (1935)
- This program is founded on
  - Correlation, method of moments
  - Goodness of fit, maximum likelihood estimation
  - Statistical significance, tests of hypothesis, *p*-values, ANOVA
  - Strong assumptions, needed for asymptotic properties
- This program
  - was developed before the computer age
  - was adopted by the Econometric Society (est. 1930)
  - is the backbone of the most popular Econometrics textbooks
  - has become the canon accepted/required by Financial journals

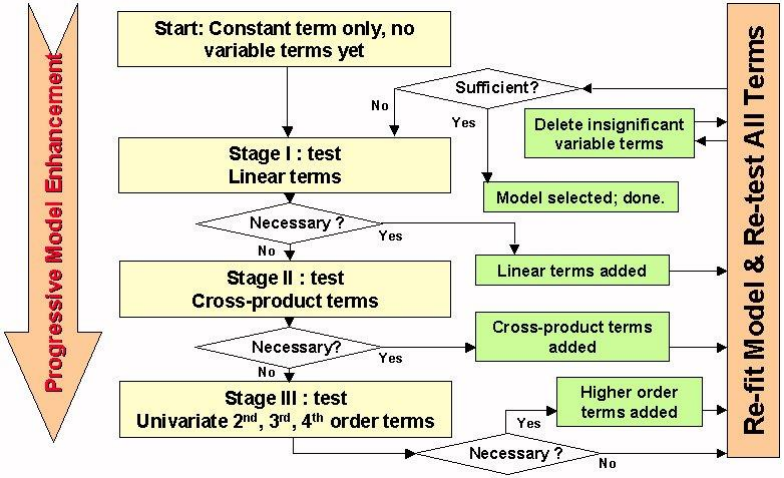


Source: The Web of Science

Fewer than 1% of journal articles in economics mention ML-related terms, such us: classifier, clustering, neural network, machine learning.

Train Set Overfitting

- Classical statistical models try to deal with train set overfitting via regularization
  - penalizing complexity (e.g., degrees of freedom)
  - reducing complexity (e.g., stepwise regression)
- However, classical models
  - do not split the data between train, validation and test sets
  - do not estimate generalization errors
- The train set is also the validation set, and the test set
  - As a result, **classical regularization fails to prevent train & test set overfitting**

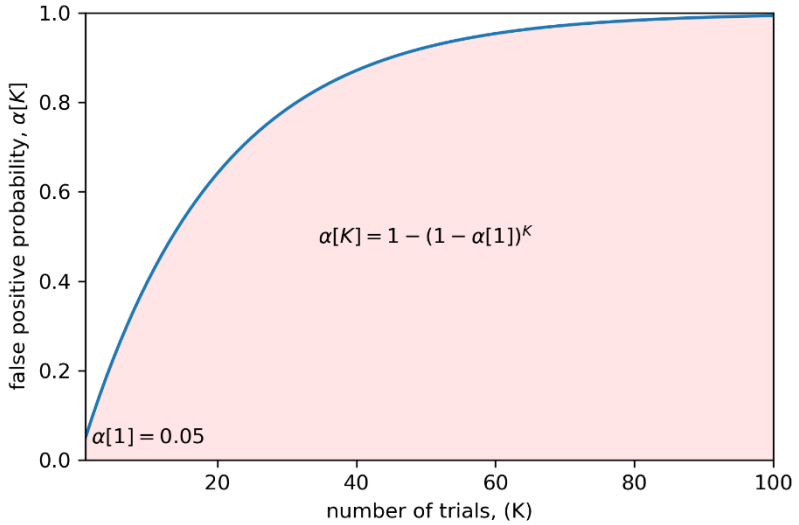


Source: Wikimedia Commons

Stepwise regression is often used in econometrics software and papers to reduce the complexity of a model and thus limit train set overfitting. Unfortunately, this makes it all but certain that the econometric model will suffer from test set overfitting.

Test Set Overfitting

- Classical statistical models were devised
  - before the invention of computers (e.g., Pearson-Neyman Lemma [1933])
  - to be run only once
- Classical statistics rarely controls for SBuMT
- One ubiquitous instance of test set overfitting is *p*-hacking
  - A researcher applies multiple statistical tests on the same data
  - Each test has a false positive rate of 5%
  - The combined false positive rate quickly converges to 100%

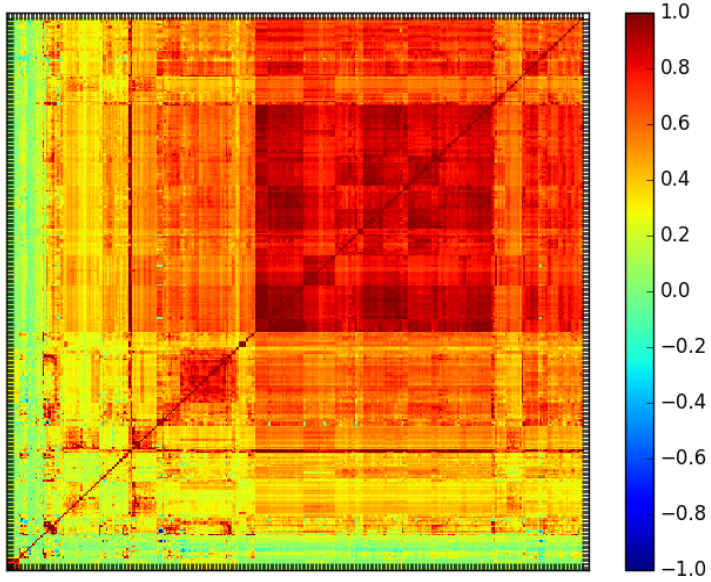


The false positive probability quickly rises after the first trial. Financial journal articles almost always present findings as if they had been the result of a single trial. Because that is rarely the case, most discoveries in finance are false.

Computer-based Statistical Methods (ML)

What is ML?

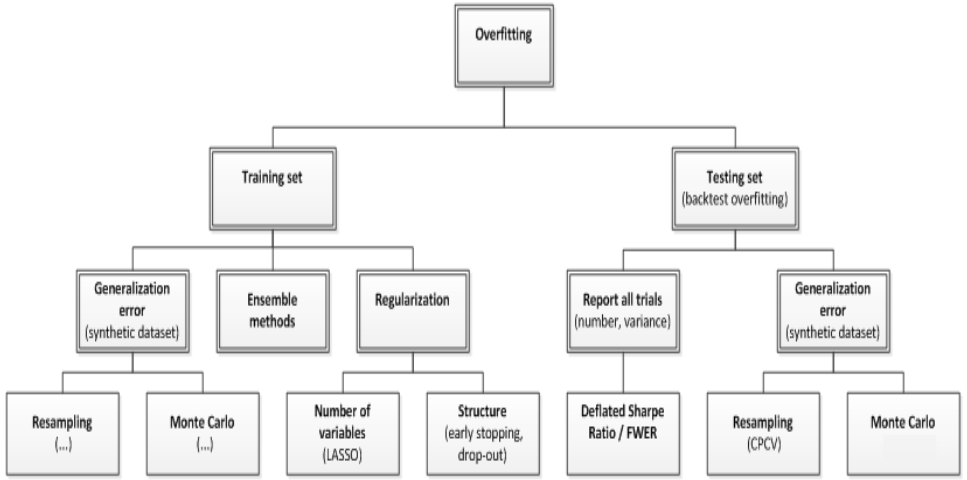
- An ML algorithm learns complex patterns in a high-dimensional space without being specifically directed
  - The ML algorithm may find a pattern that cannot be easily represented with a finite set of equations
  - Solutions often involve a large number of variables and the interactions between them
  - Unlike with other empirical tools, researchers do not impose a particular structure on the data
- ML algorithms rely on computationally-intensive methods, such as
  - estimation of the generalization error
  - ensembles, heuristics
  - experimental hypothesis testing, with minimal assumptions



Suppose that you have a 1000x1000 correlation matrix... A clustering algorithm finds that there are 3 blocks: Highly correlated, low correlated, uncorrelated.

ML Solutions to Overfitting

- There are several ML solutions for each type of overfitting
- Train set overfitting is addressed with
  - Ensembles methods
  - Regularization methods
  - Generalization errors (test set)
- Test set overfitting is addressed by
  - Reporting/controlling for all trials
  - Generalization errors (out-of-sample)
- All of these approaches require more computing power than what was available when classical methods were developed



A summarized description of various ML methods specifically designed to prevent both types of overfitting. There is no need to choose one method, and all of them can be applied simultaneously.

Train Set: Ensemble Methods

- Ensemble methods combine a set of low-correlated weak learners in order to create a learner that performs better than the individual ones
- The three main types of ensemble methods are
  - Bagging (Bootstrap aggregation)
  - Boosting
  - Stacking
- In addition, there are hybrid methods
  - E.g., random forests combine bagging with random subspaces (random sampling of features at each split, without replacement)

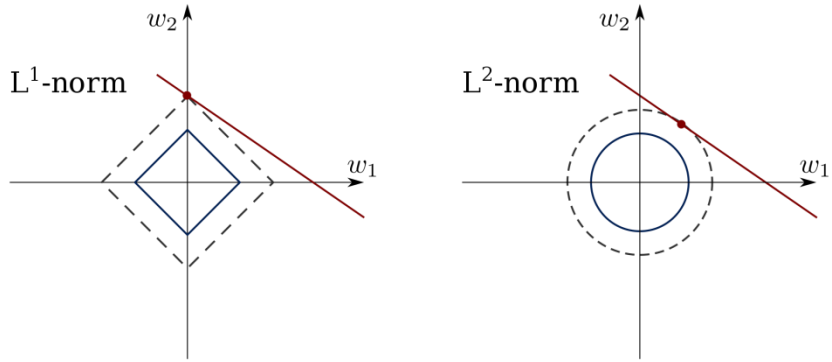
Method	Same algorithm	Parallel training	Aggregation	Primary error reduction
Bagging	Often	Yes	Deterministic	Variance
Boosting	Often	No	Deterministic	Bias
Stacking	Seldom	Yes	Meta-model (K-Fold cross-training)	Variance

Most ML algorithms can be used in ensembles. For instance, with proper parallelization, a SVC algorithm can be “bagged” to reduce train set overfitting, with minimal extra computing time.

If the weak classifiers have a minimum accuracy, bagging can also reduce bias.

Train Set: Regularization Methods

- Regularization prevents overfitting by introducing additional information to the model
- This additional information takes the form of a penalty for complexity
  - The optimization algorithm that fits the data only adds complexity if it is warranted by a certain amount of gain in explanatory power
- Three main types of regularization:
  - Tikhonov: ℓ^2 norm on the coefficients
  - LASSO: ℓ^1 norm on the coefficients
  - Elastic Net: It combines Tikhonov and LASSO

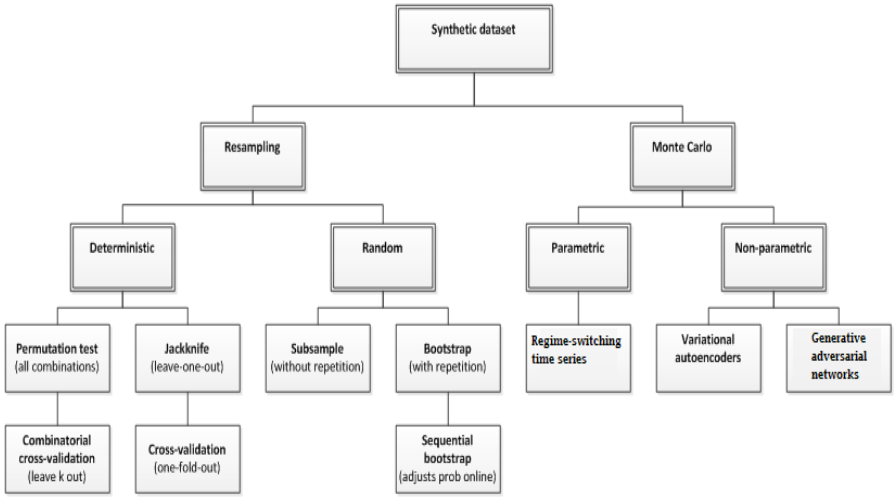


Source: Wikimedia Commons

The constraint region defined by a ℓ^1 norm is more likely to set some weights to exactly zero. In contrast, the constraint region defined by a ℓ^2 norm rarely sets any weight to zero. Elastic Nets overcome two limitations of LASSO: (a) They do not saturate when there are more variables than observations; and (b) they do not select one out of multiple multicollinear variables, discarding the rest.

Train Set: Generalization Error

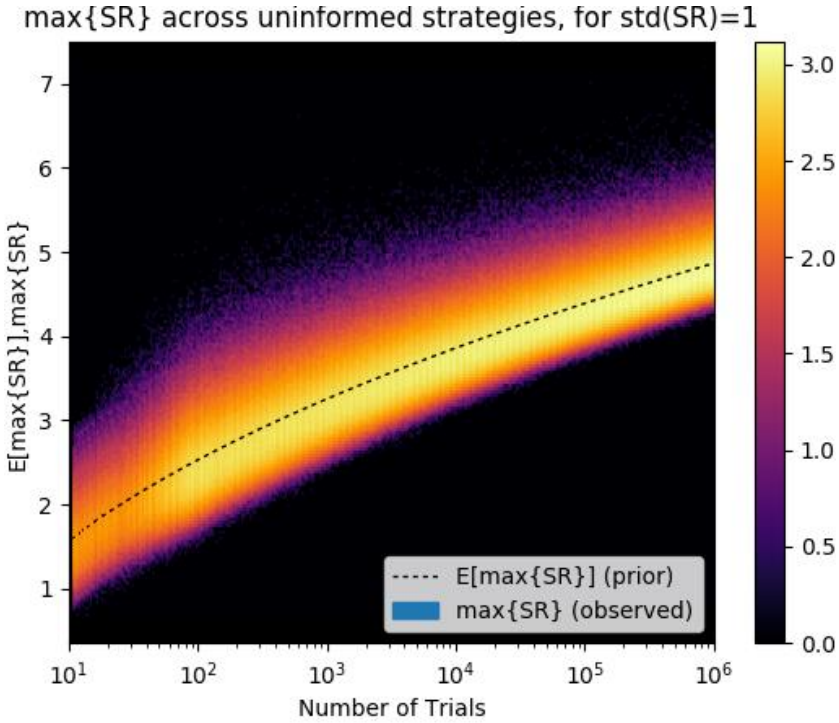
- There are two main ways of estimating the generalization error on the test set: Resampling and Monte Carlo
- **Resampling** generates synthetic datasets by sampling from the observed dataset
  - Deterministic sampling (E.g., K-fold CV)
  - Random sampling (E.g., bootstrap)
- **Monte Carlo** generates synthetic datasets by running a Monte Carlo on a data-generating process
  - Parametric (E.g., Regime-switch Markov chain)
  - Non-parametric (E.g., GAN)



A summary of ML methods that control for train set overfitting, by estimating the generalization error.

Test Set: Controlling for All Trials

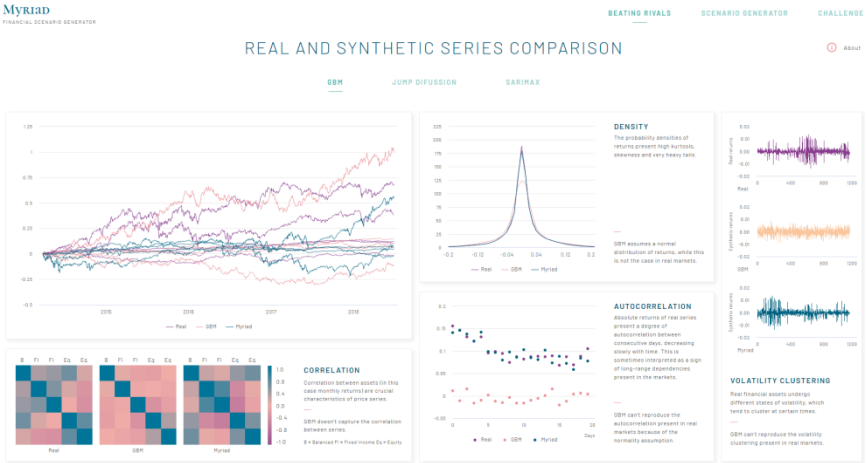
- SBuMT inflates a model's performance statistics
  - The model will perform out-of-sample worse than it did on in-sample data
- Two main approaches to control for performance inflation
  - **Parametric**: Derive the adjusted *p*-value
    - Family-wise error rate (FWER)
    - False discovery rate (FDR)
  - **Non-parametric**: Deflate the model's performance while controlling for the number and variance of the trials
    - E.g., deflated Sharpe ratio (DSR)



Non-parametric methods for SBuMT rely on fewer assumptions and tend to be more reliable.

Test Set: Generalization Error

- Once the researcher has chosen the final model, we can further estimate its generalization error on unseen data
- In order to do that, we can produce new synthetic datasets, using the same techniques described for train set generalization error
- For instance:
  - **Combinatorial cross-validation** can be used to
    - generate new test sets, different from those used by the researcher, and
    - bootstrap the entire distribution of the test set error (not only its mean), which is harder to overfit than its mean
  - **Monte Carlo methods** enable the production of arbitrarily-large new (unseen) datasets

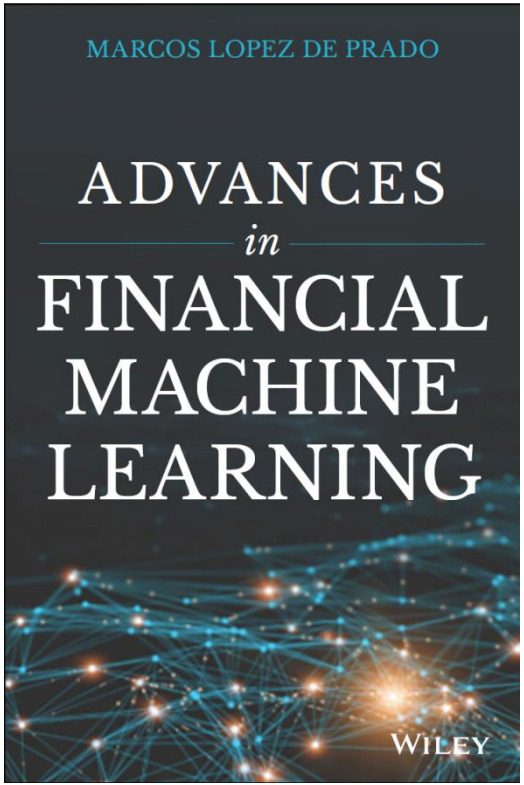


Myriad is one example of a non-parametric Monte Carlo tool that generates synthetic datasets that match the statistical properties of the observed data.

Conclusions

- When used *incorrectly*, the risk of ML overfitting is extremely high
  - Given ML's power, that risk is higher than with classical statistical methods
- However, ML counts with sophisticated methods to prevent
  - Train set overfitting
  - Test set overfitting
- Thus, the popular belief that ML overfits is false
- A more accurate statement would be that
  - **in the wrong hands, ML overfits**
  - **in the right hands, ML is more robust to overfitting than classical methods**
- **When it comes to modelling unstructured data, ML is the only choice**
  - Classical statistics should be taught as a preparation for ML courses, with a focus on overfitting prevention

For Additional Details



The first wave of quantitative innovation in finance was led by Markowitz optimization. Machine Learning is the second wave and it will touch every aspect of finance. López de Prado's Advances in Financial Machine Learning is essential for readers who want to be ahead of the technology rather than being replaced by it.

— Prof. **Campbell Harvey**, Duke University. Former President of the American Finance Association.

Financial problems require very distinct machine learning solutions. Dr. López de Prado's book is the first one to characterize what makes standard machine learning tools fail when applied to the field of finance, and the first one to provide practical solutions to unique challenges faced by asset managers. Everyone who wants to understand the future of finance should read this book.

— Prof. **Frank Fabozzi**, EDHEC Business School. Editor of The Journal of Portfolio Management.

Disclaimer

- The views expressed in this document are the authors' and do not necessarily reflect those of the organizations he is affiliated with.
- No investment decision or particular course of action is recommended by this presentation.
-

www.QuantResearch.org