Machine Learning and Tactical Asset Allocation†

Presented by Majeed Simaan[1]

[1]Assistant Professor of Finance
School of Business
Stevens Institute of Technology

May 17th, 2019
R in Finance
Chicago, IL

†An interactive vignette of this talk can be found **here**.

Disclaimer

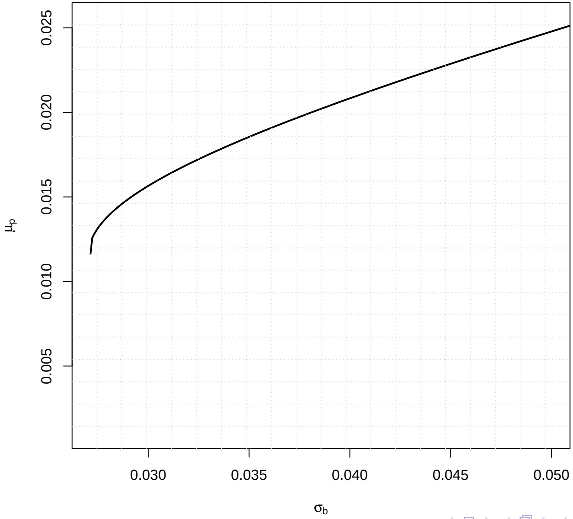
- This presentation is an illustration of an ongoing research co-authored with
  - Kris Boudt (Vrije Universiteit Brussel, University of Amsterdam and Finvex)
  - Muzafer Cela (Vrije Universiteit Brussel).
- The research is titled "In Search of Return Predictability: Evidence from Machine Learning and Tactical Allocation in R" and can be cited as Boudt, Cela, & Simaan, 2019

Introduction

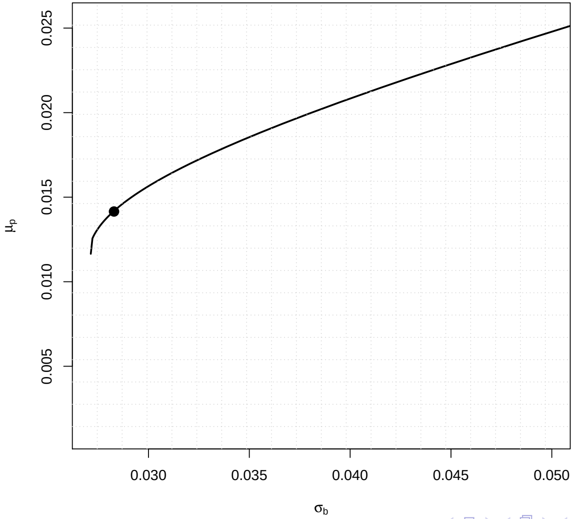
"*The first wave of quantitative innovation in finance was led by Markowitz optimization. Machine learning is the second wave, and it will touch every aspect of finance.*"

-Campbell Harvey in de Prado (2018)

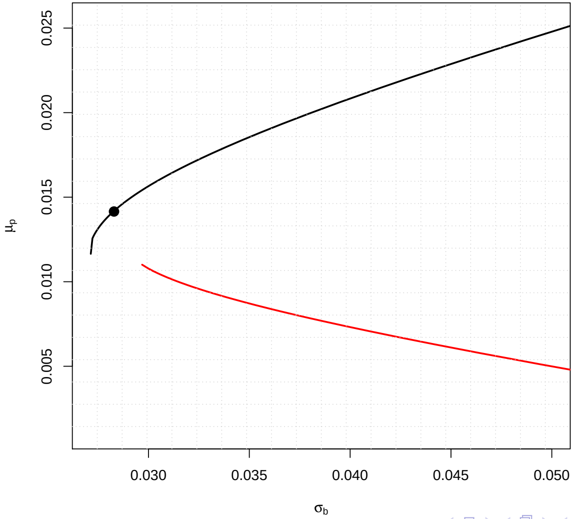
Mean-Variance Efficient Frontier (MVEF)



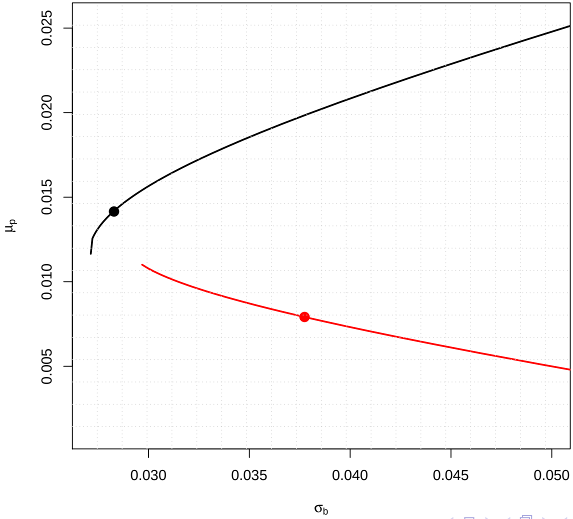
Mean-Variance Efficient Frontier (MVEF)



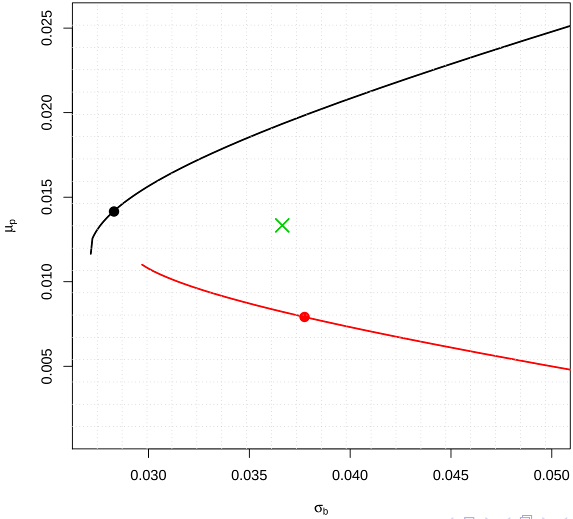
Mean-Variance Efficient Frontier (MVEF)



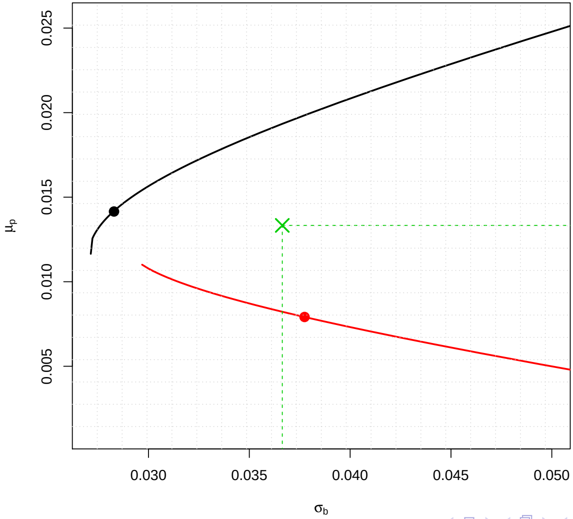
Mean-Variance Efficient Frontier (MVEF)



Mean-Variance Efficient Frontier (MVEF)



Mean-Variance Efficient Frontier (MVEF)



Challenges in Asset Allocation

- Like most models, the Mean-Variance (MV) model (Markowitz, 1952) suffers from estimation error
- Estimation error $\rightarrow$ poor out-of-sample performance, (see e.g. Michaud, 1989)
- Challenges to outperform naive allocation (equal/value weighting) (see e.g., DeMiguel, Garlappi, & Uppal, 2009)
- The conventional wisdom in asset allocation has been that better inputs leads to better outputs:
  - Implied Information (see e.g., DeMiguel, Plyakha, Uppal, & Vilkov, 2013)
  - Serial Correlation (see e.g., DeMiguel, Nogales, & Uppal, 2014)
  - Robust Statistics and Shrinkage (see e.g., Ledoit & Wolf, 2017)

Return Predictability

- Return predictability is a central issue in financial economics
- An implication of efficient market hypothesis (EMH) is that asset price follows a random walk (martingale), i.e.

$$\mathbb{E}[P_{t+1}|\Omega_t] = P_t \tag{1}$$

Return Predictability

- Return predictability is a central issue in financial economics
- An implication of efficient market hypothesis (EMH) is that asset price follows a random walk (martingale), i.e.

$$\mathbb{E}[P_{t+1}|\Omega_t] = P_t \tag{1}$$

- Nonetheless, evidence of stock return predictability is supported by numerous academic research papers
  - see e.g., Campbell & Shiller, 1988; Fama & French, 1988; Cochrane, 2007; Jiang et al., 2018

Return Predictability

- Return predictability is a central issue in financial economics
- An implication of efficient market hypothesis (EMH) is that asset price follows a random walk (martingale), i.e.

$$\mathbb{E}[P_{t+1}|\Omega_t] = P_t \tag{1}$$

- Nonetheless, evidence of stock return predictability is supported by numerous academic research papers
  - see e.g., Campbell & Shiller, 1988; Fama & French, 1988; Cochrane, 2007; Jiang et al., 2018
- Compared to the statistical evidence of return predictability, there is less literature on the economic evidence for portfolio application
  - see e.g., DeMiguel et al., 2014; Hull & Qiao, 2017; Hull et al., 2017

This Talk...

- Application of machine learning (ML) with regard to asset allocation
- Optimization will be more concerned with the signal extraction rather than solving for MV portfolios
  - focus on the input rather than the output
- Propose a cost-efficient strategy that outperforms the benchmark in terms of risk-adjusted return
- Contribution relies on open source and public data
  - reproducibility in finance
  - "poor" investor's strategy

Getting the Data

- We refer to the R `quantmod` package to download data from Yahoo Finance.

- Mainly, we look at:
  1. The SPY ETF which tracks the S&P 500 index
  2. The VIX index
  3. The GLD gold ETF,
  4. The 7-10 years treasury bond ETF,
  5. The XLF the financial sector ETF

- Merging altogether, the data dates between 2004-12-27 and 2019-05-10.
  - the GLD started trading in late 2004

Getting the Data II

Feature Space

- To construct the main feature space, we focus on
  1. The change in adjusted prices (returns)
  2. Deviation from 25 days moving averages (MA)
  3. Trading volume
- In total, the feature space, denoted X_t, consists of 14 variables

Getting the Data II

Feature Space

- To construct the main feature space, we focus on
  1. The change in adjusted prices (returns)
  2. Deviation from 25 days moving averages (MA)
  3. Trading volume
- In total, the feature space, denoted X_t, consists of 14 variables

Response Variable

- The response variable of interest, denoted by Y_{t+1}, is the next day change in the SPY price

$$Y_{t+1} = \begin{cases} +1 & r^e_{t+1} \geq -0.01 \\ -1 & r^e_{t+1} < -0.01 \end{cases} \quad (2)$$

with r^e_{t+1} denoting the return of the SPY at time $t+1$

Machine Learning Application

- The key idea behind machine learning is to find a mapping function f_t that maps X_t to Y_{t+1}, i.e.

$$f_t : X_t \to Y_{t+1} \tag{3}$$

- In practice, f_t is estimated using a information set Ω_t (data sample) available until time t
- Thus, our objective is to find an optimal function $\hat{f}_t$ in a
  - data-driven manner
  - recurring basis
- Finally, the investment decision making is facilitated based on an extracted signal from $\hat{f}_t(X_t)$

- Since Y_{t+1} is a binary variable, the mapping function should return values between 0 and 1 (probability)
- Under a binomial model, let $\hat{\pi}_{t+1}$ denote the forecasted probability that $Y_{t+1} = +1$, such that

$$\hat{\pi}_{t+1} = \mathbb{P}_t(Y_{t+1} = +1 \mid X_t) = \frac{\exp(X_t' \hat{\beta}_t^+)}{\exp(X_t' \hat{\beta}_t^+) + \exp(X_t' \hat{\beta}_t^-)} \tag{4}$$

where $\hat{\beta}_t^+$ ($\hat{\beta}_t^-$) is the estimated vector of weights that maps the feature space into $+1$ (respectively -1)

- The mapping function from (4) is determined by the weights $\hat{\beta}_t^+$ and $\hat{\beta}_t^-$
- Therefore, to implement, we need an algorithm that finds an optimal estimate of each

Machine Learning Application III

- We refer to the R `glmnet` package by Friedman, Hastie, & Tibshirani, 2010 for implementation
- At the end of each week, we estimate $\hat{\beta}_t^+$ and $\hat{\beta}_t^-$ using
  - maximum likelihood estimation (MLE)
  - net elastic penalty, $\alpha = 1/2$
  - 10-folds cross validation
  - sample size of 50 weeks (around 250 days)

Machine Learning Application III

- We refer to the R `glmnet` package by Friedman et al., 2010 for implementation
- At the end of each week, we estimate $\hat{\beta}_t^+$ and $\hat{\beta}_t^-$ using
  - maximum likelihood estimation (MLE)
  - net elastic penalty, $\alpha = 1/2$
  - 10-folds cross validation
  - sample size of 50 weeks (around 250 days)
- Given the estimated weights, we forecast the probability that the market will go up/down over the course of the following week
- The procedure is repeated until the last week of the sample (May 10th, 2019)
  - in total, there are 750 weeks

Machine Learning Extracted Signal

- The ML procedure returns a time series of the forecasted probability
- To filter noise from the signal, we use a 25-days MA
- The following figure demonstrates the probability that SPY will drop more than -1% in a single day

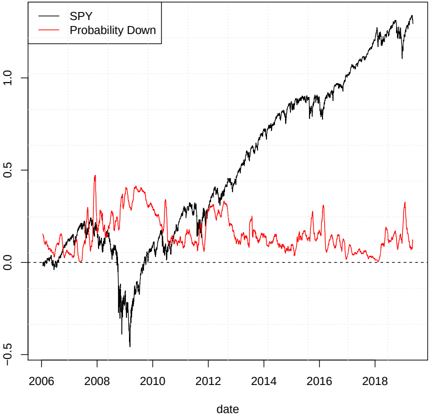


Figure: Full Sample

Machine Learning Extracted Signal

- The ML procedure returns a time series of the forecasted probability
- To filter noise from the signal, we use a 25-days MA
- The following figure demonstrates the probability that SPY will drop more than -1% in a single day

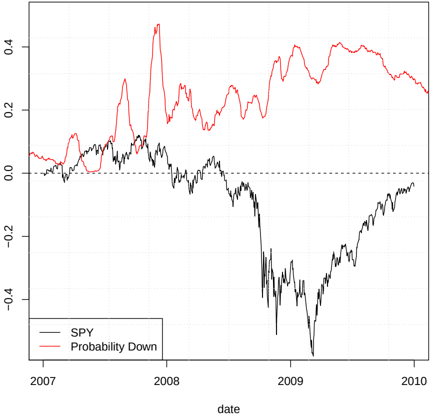


Figure: 2007-09 Sample Period

Machine Learning Extracted Signal

- The ML procedure returns a time series of the forecasted probability
- To filter noise from the signal, we use a 25-days MA
- The following figure demonstrates the probability that SPY will drop more than -1% in a single day

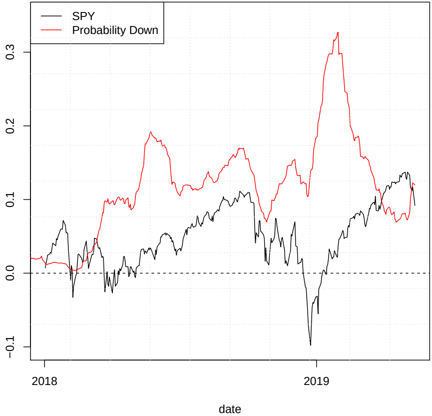


Figure: 2018-19 Sample Period

Tactical Asset Allocation Strategy

- How can we implement a tactical allocation strategy using the extracted signal?

Tactical Asset Allocation Strategy

- How can we implement a tactical allocation strategy using the extracted signal?

- Standing at time t, the probability $\hat{\pi}_{t+1}$ denotes the **level of confidence that the SPY would not drop below** -1% the next day
- Depending on the investor's level of risk tolerance, he/she will invest in the SPY versus the IEF if $\hat{\pi}_{t+1}$ is high enough

Tactical Asset Allocation Strategy

- How can we implement a tactical allocation strategy using the extracted signal?

- Standing at time t, the probability $\hat{\pi}_{t+1}$ denotes the **level of confidence that the SPY would not drop below** -1% the next day
- Depending on the investor's level of risk tolerance, he/she will invest in the SPY versus the IEF if $\hat{\pi}_{t+1}$ is high enough

- Put formally, if ω_t denotes the weight allocated to the SPY, then it follows that

$$\omega_t = \begin{cases} 1 & \hat{\pi}_{t+1} \geq a \\ 0 & \text{else} \end{cases} \tag{5}$$

  for a given level of confidence a predetermined by the investor
- On the other hand, the weight allocated to the IEF is $1 - \omega_t$

Tactical Asset Allocation Strategy II

- This implies that return of the strategy at the following period $t+1$ is given by

$$r_{s,t+1} = \omega_t r^e_{t+1} + (1-\omega_t) r^b_{t+1} \tag{6}$$

or, alternatively:

$$r_{s,t+1} = I_{[\hat{\pi}_{t+1} > a]} r^e_{t+1} + I_{[\hat{\pi}_{t+1} \leq a]} r^b_{t+1} \tag{7}$$

with

  - r^e_{t+1} and r^b_{t+1} denoting the return on the SPY and the IEF, respectively.
  - $I_{[A]}$ is an index variable returning 1 if event A takes place and zero otherwise

Tactical Asset Allocation Strategy III

- Equation (7) indicates that the return of the strategy at $t+1$ is controlled by two main components in the former period:
  1. The forecasted smoothed probability, $\hat{\pi}_{t+1}$, (level of confidence)
  2. The predetermined minimum level of confidence a

- The former acts as a forecast of the SPY return
- The latter is a control parameter determining the level of risk tolerance of the investor
  - The larger (smaller) the a, the more (less) conservative the investor is
  - In this illustration, we set $a \in \{95\%, 90\%, 85\%\}$

Backtesting

- As a benchmark, we consider 60-40 strategy that invests 60% in SPY and 40% in IEF
    - In the context of DeMiguel et al., 2009, this is a *naive* strategy
    - Additionally, we compare the strategy with each ETF alone

Backtesting

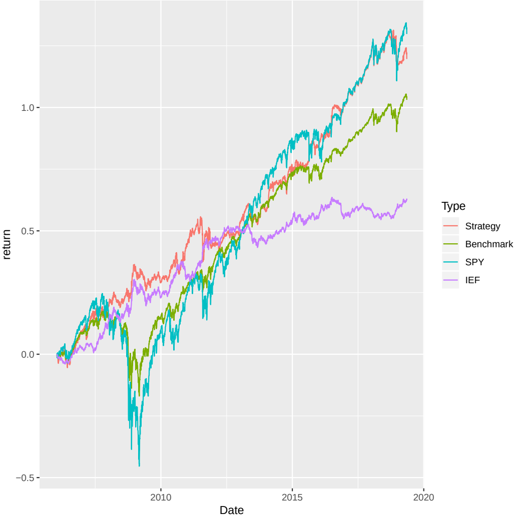
- As a benchmark, we consider 60-40 strategy that invests 60% in SPY and 40% in IEF
  - In the context of DeMiguel et al., 2009, this is a *naive* strategy
  - Additionally, we compare the strategy with each ETF alone
- Given the 50 weeks for initial estimation, the testing period begins at 2006-01-19
  - 2006-01-18 is the first available smoothed probability forecast
  - the last testing day dates back to the Friday of last week (2019-05-10)

Backtesting

- As a benchmark, we consider 60-40 strategy that invests 60% in SPY and 40% in IEF
  - In the context of DeMiguel et al., 2009, this is a *naive* strategy
  - Additionally, we compare the strategy with each ETF alone
- Given the 50 weeks for initial estimation, the testing period begins at 2006-01-19
  - 2006-01-18 is the first available smoothed probability forecast
  - the last testing day dates back to the Friday of last week (2019-05-10)
- For a given a, the backtesting procedure returns a time series of the strategy return, $r_{s,t}$, for every t in the testing period

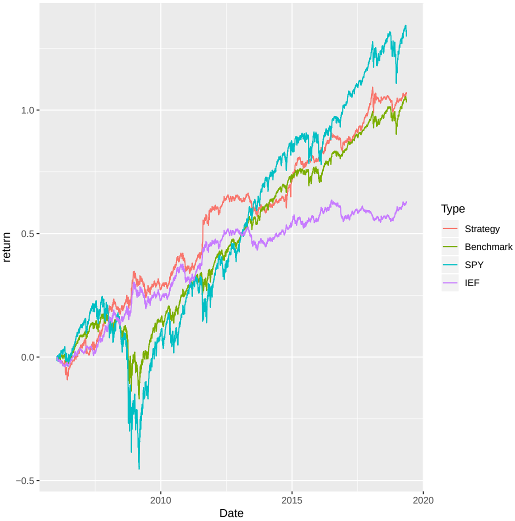
Results and Discussion

- Cumulative return of the strategy when $a = 90\%$
- In total, the strategy trades 69 times over the whole sample
    - less than 6 trades a year, on average
- Holds the SPY 47% of the time
- An interactive version of the plot can be found **here**



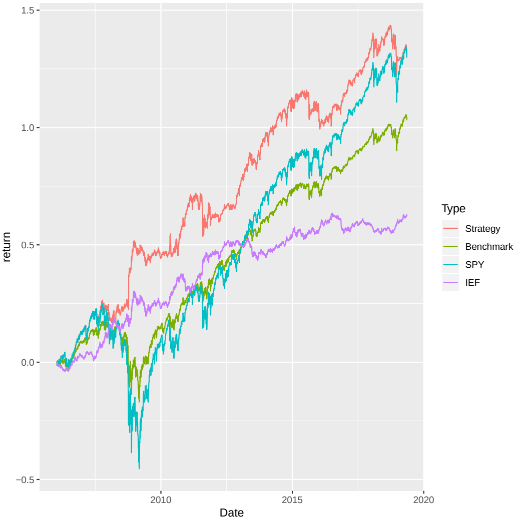
Results and Discussion II

- Cumulative return of the strategy when $a = 95\%$
- In total, the strategy trades 40 times over the whole sample
  - approximately 3 trades a year
- Holds the SPY 19% of the time
- An interactive version of the plot can be found **here**



Results and Discussion III

- Cumulative return of the strategy when $a = 85\%$
- In total, the strategy trades 44 times over the whole sample
    - approximately 3 trades a year
- Holds the SPY 68% of the time
- An interactive version of the plot can be found **here**



Results and Discussion IV

Table: Relative risk-adjusted return with respect to Benchmark/SPY

	$a = 95\%$	$a = 90\%$	$a = 85\%$
Compared to Benchmark			
Annualized Alpha	0.07	0.06	0.05
Beta	0.02	0.31	0.52
Compared to SPY			
Annualized Alpha	0.09	0.08	0.07
Beta	-0.06	0.13	0.33

- Annualized Alpha is the Jensen's alpha estimated against a given benchmark
- Beta is the beta of the strategy with respect to a given benchmark
- Statistics were computed using the `PerformanceAnalytics` package

Results and Discussion V

- Finally, consider the risk-adjusted return in terms of
  - Sharpe-ratio

$$Sharpe = \sqrt{252} \times \frac{\mathbb{E}[R_p]}{\sqrt{\mathbb{V}[R_p]}} \tag{8}$$

with the R_p is the daily return of a portfolio/asset p

  - Sortino-ratio

$$Sortino = \sqrt{252} \times \frac{\mathbb{E}[R_p]}{\sqrt{\mathbb{V}[R_p \mid R_p < 0]}} \tag{9}$$

Table: Absolute risk-adjusted return

	$a = 95\%$	$a = 90\%$	$a = 85\%$	Benchmark	SPY
Sortino	1.35	1.03	0.98	0.92	0.62
Sharpe	0.99	0.83	0.75	0.73	0.51

Concluding Remarks

- This talk demonstrates how to
  - implement machine learning using public data and open source software
  - utilize a simple cost-efficient trading strategy
- The strategy is mainly data-driven
- Nonetheless, its performance depends on a couple of specifications:
  - the level of confidence a
  - the price change in the market
- The suggested approach can be deployed to screen stocks or ETFs
  - e.g., sector rotation strategy

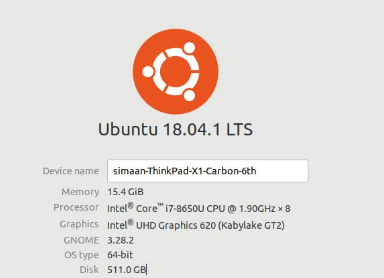
stay in touch...

Email:	`msimaan@stevens.edu`
Linkedin:	`https://www.linkedin.com/in/majeed-simaan-85383045`
RPubs:	`https://rpubs.com/simaan84`

Thank You!

Appendix - Computing Power

- The `glmnet` is efficiently utilized for cross-validation using parallel processing
- The loop in the main code can be replaced with the `mclapply` command from the `parallel` library
- The ML algorithm takes less than 10 minutes to run on a linux OS with the following specs:



Appendix - References I

[]Boudt, K., Cela, M., & Simaan, M. 2019. In search of return predictability: Evidence from machine learning and tactical allocation in r.

[]Campbell, J. Y., & Shiller, R. J. 1988. The dividend-price ratio and expectations of future dividends and discount factors. *The Review of Financial Studies*, *1*(3), 195–228.

[]Cochrane, J. H. 2007. The dog that did not bark: A defense of return predictability. *The Review of Financial Studies*, *21*(4), 1533–1575.

[]DeMiguel, V., Garlappi, L., & Uppal, R. 2009. Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy? *Review of Financial Studies*, *22*(5), 1915–1953.

[]DeMiguel, V., Nogales, F. J., & Uppal, R. 2014. Stock return serial dependence and out-of-sample portfolio performance. *The Review of Financial Studies*, *27*(4), 1031–1073.

Appendix - References II

[]DeMiguel, V., Plyakha, Y., Uppal, R., & Vilkov, G. 2013. Improving portfolio selection using option-implied volatility and skewness. *Journal of Financial and Quantitative Analysis*, *48*(06), 1813–1845.

[]Fama, E. F., & French, K. R. 1988. Dividend yields and expected stock returns. *Journal of financial economics*, *22*(1), 3–25.

[]Friedman, J., Hastie, T., & Tibshirani, R. 2010. Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*, *33*(1), 1–22. Retrieved from `http://www.jstatsoft.org/v33/i01/`

[]Hull, B., & Qiao, X. 2017. A practitioner's defense of return predictability. *The Journal of Portfolio Management*, *43*(3), 60–76.

[]Hull, B., Qiao, X., & Bakosova, P. 2017. Return predictability and market-timing: A one-month model. *Forthcoming in the Journal Of Investment Management*.

[]Jiang, F., Lee, J., Martin, X., & Zhou, G. 2018. Manager sentiment and stock returns. *Journal of Financial Economics*.

Appendix - References III

[]Ledoit, O., & Wolf, M. 2017. Nonlinear shrinkage of the covariance matrix for portfolio selection: Markowitz meets goldilocks. *The Review of Financial Studies*, *30*(12), 4349–4388.

[]Markowitz, H. 1952. Portfolio selection. *The journal of finance*, *7*(1), 77–91.

[]Michaud, R. O. 1989. The markowitz optimization enigma: is' optimized'optimal? *Financial Analysts Journal*, *45*(1), 31–42.