From Man vs. Machine to Man + Machine: The Art and AI of Stock Analyses

Sean Cao, Wei Jiang, Junbo Wang, Baozhong Yang

Discussion by Svetlana Bryzgalova[1]

[1]London Business School

ML and AI in Finance

Stockholm, 2022

Man, machine, or both?

An ongoing debate in both industry and academia: Will Algorithms/AI replace human jobs?

Key contributions:

- Focus on an important area: stock prediction by analysts and ML algos
- Information is numerical and text-based: corporate disclosures, industry trends, and macroeconomic indicators
- On average, ML is doing better but humans have an edge in complex firms with many intangible assets
- Humans catch up, especially as new tools become available
- AI power and human wisdom are complementary.

A very careful, thoughtful, and well-written paper.

Empirical design: data

Analyst: 12-month target price forecasts builds on the Thomson Reuters I/B/E/S analyst database using data from 1996 to 2018.

Information used in ML forecasts:

- Firm characteristics: history of prices and returns, realized earnings over 3 years, momentum, value, investment, profitability, intangibles, and trading frictions anomalies
- Industry characteristics: competition, product market fluidity, industry affiliation, size, and earnings
- Macro data: Fama and French (1989), Chen, Roll, and Ross (1986)
- positive and negative sentiment words from the firm-issued SEC filings following Loughran and McDonald (2011) dictionary, sentiment variables following Cao, Kim, Wang, and Xiao (2020).

NB: this is not an exhaustive list of data compared to many recent papers, yet it captures many key indicators.

Empirical design: methods

Analyst

Empirical design: methods

Analyst

Machine

- lasso, elastic-net, support vector machines, random forest, gradient boosting, and long short-term memory neural networks.
- Key approach: ensemble (average) of the *“state-of-the-art ML methods”*: random forest (1995/2001), gradient boosting (1999), and LSTM neural networks (1997).

Debiasing analysts: 70% of the performance gap

Human analysts are known to produce biased forecasts: Incentives + Psychology

Could we use machine learning to produce unbiased forecasts, a machine-debiased-man?

Debiasing analysts: 70% of the performance gap

Human analysts are known to produce biased forecasts: Incentives + Psychology

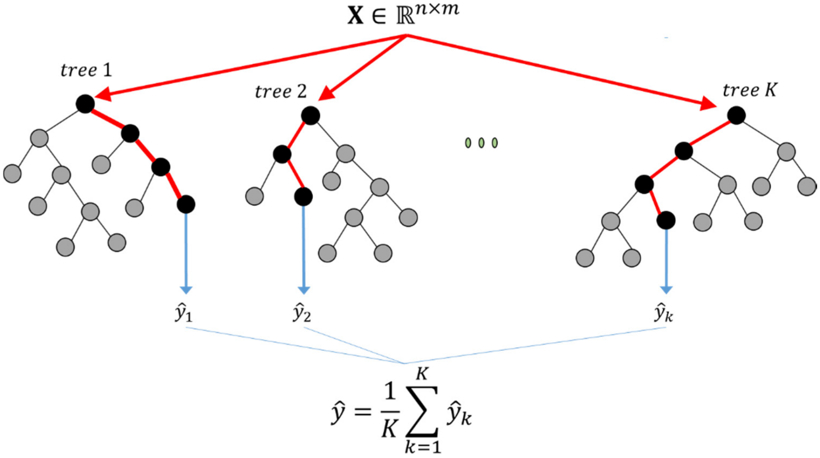
Could we use machine learning to produce unbiased forecasts, a machine-debiased-man?

I would be cautious about this statement:

- (Almost) all machine learning algorithms are by definition biased: they optimize the bias-variance trade-off
- For most methods the reason we empirically don't see a sizeable bias is because we don't measure it precisely, not because its nonexistent
- Under specific conditions, certain modifications of the methods can be (asymptotically) unbiased

Random forest: a generalization of conditional sorts

Each tree uses a random subset of X, with splits maximizing the difference between paths.



Every tree produces a biased estimate of $\hat{y}$, which vanishes when averaged over many trees.

What do we know about the forest?

- Basic bias correction: Zhang and Liu (2012)
- Asymptotic unbiasedness under some conditions: Athey an Wager (2018)
- Bias correction via boosting: Ghosal and Hooker (2015, 2021)
- Conditions for consistency: Scornet, Biau, Vert (2015)
- Rates of convergence and asymptotic normality: Peng, Coleman, Mentch (2019)
- Asymptotic distribution of predictions: Mentch and Hooker (2018)

Results of the paper will be even stronger by going beyond the MSE of prediction:

recent advances allow us to build confidence intervals for parameters and predictions, and give detailed guidance on convergence rates and required sampling schemes.

Work in progress for other ML tools, including Adversarial Neural Nets.

Similar results for lasso, ridge, etc. are already well known $\rightarrow$ use them!

Where is human advantage?

The paper goes a long way in analyzing the sources of ML competitiveness.

- machines have an advantage in processing numerical information
- machines recently caught up with image/video processing
- humans are better at processing unstructured/amorphic information, like text, intangible assets, etc.

Information in the text

Text:

– positive and negative sentiment words from the firm-issued SEC filings following Loughran and McDonald (2011) dictionary,

– sentiment variables following Cao, Kim, Wang, and Xiao (2020).

Text information is very rich and not confined to sentiment:

- both random forest and neural networks can easily handle text data
- vectorization, topics, dictionaries, etc.
- sentiment measures targeting return prediction, etc.

It would be great to see a more detailed analysis of the information in the text.

A general approach of sentiment measures

Sentiment analysis in economics/finance typically follows 3 broad steps:

- **Step 1**: Create a dictionary of relevant words for the topic
- **Step 2**: Identify the words that are positive or negative
- **Step 3**: Apply this dictionary to measure the intensity of positive and negative sentiment in a particular document: vector representation of the document, Arora et al (2017)

Typical challenges: topic-specific dictionary, automated and time-varying dictionary (ESG?), phrases instead of words, context-based, continuous sentiment score instead of binary good/bad.

Step I: Automated Dictionary

Semantic information about every word is captured by it's vector representation, based on the context/embedding in which the word/phrase is used.

- Word2vec, based on neural networks
- Applies to phrases as well as words
- Outperforms existing alternatives, Mikolov et al. (2013)

Create dictionary of most similar words

- Use distances between the vector representations of words to measure their similarities

Example:

- The word **economy** is associated with the words/phrases: *economic growth*, *inflation*, *economic*, *recession*, *economic recovery*, *consumer spending*, etc.

Step II: Automated Measure of Word Sentiment

Isolate **positive** or **negative** words to measure economic sentiment

Start with a few seed words related to sentiment:

- Positive: expansion, boom, growth, profit, optimistic, optimism, opportunity, success, successful, profitable, prosperity, profitability, bullish
- Negative: recession, bankrupt, shrinking, unemployment, loss, bankruptcy, cutback, layoff, redundancy, pessimism, contraction, unsuccessful, failure, insolvent, insolvency, bearish

Following Hamilton et al. (2016) – the gold standard to generate sentiment scores for economics and finance – we produce a continuous measure of sentiment for each word/phrase in the dictionary:

- top 10 most positive words/phrases: *successful, success, profit, opportunity, profitable, cess, proven, expansion, interview, rewarding*
- top 10 negative words/phrases: *failure, unsuccessful, insolvent, earnest, redundancy, bankruptcy, bankrupt, insolvent debtor, contraction*

Key Advantages

- Almost fully automated construction of topic specific dictionaries
  - can be easily applied to almost any other context/topic
  - no need to manually create the whole dictionary
- Reflects the context in which words/phrases are used
- Automatically deals with negation, a common challenge for word counts
  - good vs not good
- Automatic generation of a continuous measure of word sentiment score
  - good vs excellent
- Captures not just key words, but their synonyms and related language

Quarterly measure of economic sentiment

Based on the historical collection of digitized newspapers, 1736-2020 in the USA

- 13,000 local newspapers across thousands of communities
- 193 million newspaper pages, which is around 2 billion newspaper articles (compared to 800,000 for the WSJ corpus)

Figure 1: National Economic Sentiment: Quarterly data, 1850-2020
van Binsbergen, Bryzgalova, Mukhopadhyay, Sharma (2022)

Quarterly measure of economic and non-economic sentiment

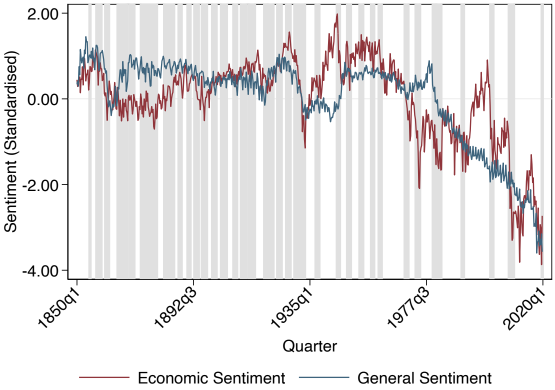


Figure 2: National Economic and Non-Economic Sentiment: Quarterly data, 1850-2020

Non-econ sentiment is different and is trending downwards from 1970.

Pure national economic sentiment

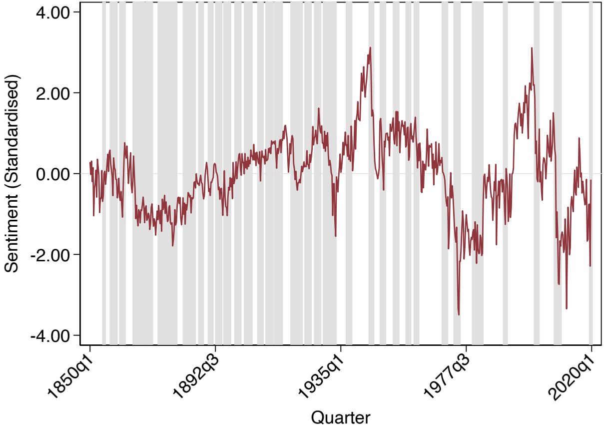


Figure 3: National Economic Sentiment (Orthogonalised): Quarterly data, 1850-2020

Orthogonalising economic w.r.t non-econ sentiment produces a stationary series

This procedure could be used to clean other existing measure of common trends

How to combine man and a machine?

1. Debiasing human forecasts with ML

How to combine man and a machine?

1. Debiasing human forecasts with ML
2. Add human forecasts among the inputs to ML

How to combine man and a machine?

1. Debiasing human forecasts with ML
2. Add human forecasts among the inputs to ML
3. Weight the forecasts

How to combine man and a machine?

1. Debiasing human forecasts with ML
2. Add human forecasts among the inputs to ML
3. Weight the forecasts
4. Use human inputs to modify the ML algorithms

How to combine man and a machine?

1. Debiasing human forecasts with ML
2. Add human forecasts among the inputs to ML
3. Weight the forecasts
4. Use human inputs to modify the ML algorithms

- Return prediction algorithm is not designed to generate high SR/alpha (volatility, risk adjustment, nonlinearity)
- Economic theory:
  - no arbitrage constraint
  - objective function: economic loss instead of just pure prediction/mse
  - shape and variation restrictions on variable importance
  - encoding narratives, causal, time series, and cross-sectional across predictions

On the importance of economic theory

Grammig et al (2021): option-implied expected returns **beat** standard ML predictions at shorter horizons (under a year).

Conclusion

Absolutely enjoyed reading the paper:

- would love to see more on the precise measure of information from different sources
- would be more careful on labeling/interpretation
- great research agenda: how to combine the benefits of humans and ML tools?

Highly recommend reading!