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

Sean Cao
J. Mack Robinson College of Business
Georgia State University
scao@gsu.edu

Wei Jiang
Columbia Business School, NBER, and ECGI
wj2006@columbia.edu

Junbo Wang
E.J. Ourso College of Business
Louisanna State University
junbowang@lsu.edu

Baozhong Yang
J. Mack Robinson College of Business
Georgia State University
bzyang@gsu.edu

First Draft: March 2021

This Draft: May 2021

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

ABSTRACT

An AI analyst we build to digest corporate financial information, qualitative disclosure and macroeconomic indicators is able to beat the majority of human analysts in stock price forecasts and generate excess returns compared to following human analyst. In the contest of "man vs machine," the relative advantage of the AI Analyst is stronger when the firm is complex, and when information is high-dimensional, transparent and voluminous. Human analysts remain competitive when critical information requires institutional knowledge (such as the nature of intangible assets). The edge of the AI over human analysts declines over time when analysts gain access to alternative data and to in-house AI resources. Combining AI's computational power and the human art of understanding soft information produce the highest potential in generating accurate forecasts. Our paper portraits a future of "machine plus human" (instead of human displacement) in high-skill professions.

Key Words: Artificial Intelligence; Machine Learning; FinTech; Stock Analyst; Alternative Data; Disruptive Innovation

1. Introduction

Since its inception and as it rises, artificial intelligence (AI) constantly makes human beings to rethink their own roles. While AI is meant to be intelligence augmentation for humans, concerns abound that AI could replace human tasks and increasingly skilled ones, and thus displace jobs by those currently performed by the better-paid and better-educated workers (Muro, Maxim, and Whiton, 2019). Such a concern and the associated debates have motivated a quickly growing literature. Recent work by Webb (2020), Acemoglu, Autor, Hazell, and Restrepo (2020), Babina, Fedyk, He, and Hodson (2020), and Jiang, Tang, Xiao, and Yao (2021) have all conducted large-sample analyses on the extent of job exposure and vulnerability to AI-related technology as well as the consequences on employment and productivity.

The existing literature has been mostly focusing on characterizing the type of jobs that are vulnerable to disruption by, as well as those that could be created due to, AI evolution. In other words, the sentiment of the existent studies mostly involves a theme of "man-versus-machine," i.e., to characterize the contest between human and AI, to explore ways human adapts, and to predict the resulting job redeployment. In such settings, human beings are often rendered passive or reactive – dealing with disruptions and looking for new opportunities defined by the AI landscape. There has been relatively little research devoted to prescribing how skilled human workers could tap into a higher potential with enhancement from AI technology, presumably the primary goal for human beings to design and develop AI in the first place. This study aims to connect the contest of "man-versus-machine" ("man v. machine" hereafter) to a potential equilibrium of "man-plus-machine" ("man + machine" hereafter).

Our study could be motivated by the experience of chess grand master Garry Kasparov. The story that IBM's Deep Blue beat the then reigning grand master Kasparov in 1997 was well-known. Multiple contests repeated in a similar setting afterwards killed any remaining suspense for the outcome of man v. machine in chess playing. What is far less known is that

humans, despite having lost interest in man-versus-machine chess contests, have not lost interest in either the game or the machine. In fact, the encounter with the Deep Blue was a catalyst for people like Kasparov to pioneer the concept of man + machine matches, in which a chess player equipped with AI assistance (a "centaur" player) competes against AI. Up to today the centaur keeps an upper hand against machines; and even more encouragingly, there have been more and better human chess players with the advent of affordable AI-engineered chess programs.[1] If AI can help more humans become better chess players, it stands to reason that it can help more of us become better at many skilled jobs, from pilots, medical doctors, to investment advisors. In this study, we zoom into the profession of stock analysis, whose data availability allows us to calibrate both man vs. machine and man + machine. Stock analysts are among the most important information intermediaries in the market place (e.g., Brav and Lehavy, 2003; Jegadeesh, Kim, Krische, and Lee, 2005; Crane and Crotty, 2020). Their job, which require both institutional knowledge and data analytics, has not been spared by AI as more and more investors start to heed to recommendations about stock picking and portfolio formation made by AI-powered tools.[2]

To trace out the path from "man v. machine" to "man + machine," we decided to build our own AI model for year-end stock predictions so that we have a consistent and time-adapted benchmark for AI performance which we understand and are able to explain.[3] Target prices and earnings are the two primary subjects of analyst forecasts, we choose the former as the latter are subject to managerial discretion, which a large body of accounting literature on earnings management manifests. Our "AI analyst" is built on training a com-

---

[1]Source of information: The Inevitable, by Keven Kelly, Penguin Publishing Group, 2016. See also "Defeated Chess Champ Garry Kasparov Has Made Peace With AI," *Wired*, February 2020.

[2]Sources: "What Machine Learning Will Mean for Asset Managers," Robert C. Pozen and Jonathan Ruane, *Havard Business Review*, December 3, 2019. "How Startup Investors Can Utilize AI To Make Smarter Investments," Jia Wertz, *Forbes*, January 18, 2019.

[3]We focus on year-end prices because these are a typical focal point for investors and corporations for tax and reporting reasons. For example, "Credit Suisse Raised Its SP 500 Target – Earnings Are Too Good to Ignore," Jacob Sonenshine, *Barron's*, April 30, 2021.

bination of current machine-learning (ML) tool kits[4] using timely publicly available data and information. More specifically, we collect firm-level, industry-level, and macro-economic variables, as well as textual information from firms' disclosure (updated to right before the time of an analyst forecast) as inputs or predictors, but deliberately exclude information from analyst forecasts (past and current) themselves. We resort to machine learning models, instead of traditional economics models (such as regressions) due to the advantages to the former in managing high dimensional unstructured data, and in their flexibility in optimizing and fitting unspecified functional forms. More recent development in the area also allows us to mitigate over-fitting and to improve out-of-sample performance.

We kept training and improving the model until we were confident that our AI analyst is able to beat human analysts as a whole: The AI analyst based on the final "ensemble" model outperforms 53.7% of the target price predictions made by all IBES analysts during the sample period of 2001-2016.[5] Moreover, a monthly rebalanced long-short portfolio based on the differences in the opinions of AI and human analysts[6] is able to generate a monthly alpha of 0.84% to 0.92% using the Fama-French-Carhart four-factor model. Though building an AI analyst is not the ultimate goal of this study and though we do not claim our AI analyst to be the best of the kind, its performance already suggests that the profession of financial analysts is subject to technology disruption as our model is a lower bound of the state-of-the-art. To the extent that we have, at our disposal, an AI analyst that beats the average of its human counterparts, we are able to explore the relative advantages of, and potential synergies between, the two sides.

First, we examine the circumstances under which human analysts retain their advantage, in that a forecast made by an analyst beats the concurrent AI forecast in terms of lower

---

[4]We start with two versatile quasi-linear ML models, Elastic-Net and Support Vector Regressions, that are adept at tasks with a large number of variables. We then add on three highly nonlinear ML models, Random Forest, Gradient Boosting, and Long Short-Term Memory (LSTM) Neural Networks. Random Forest and Gradient Boosting can both capture complex and hierarchical interactions among the input variables while the LSTM model is designed to model time-series patterns in the data.

[5]In comparison, the predictions by an OLS model only outperform 19.3% of analyst forecasts.

[6]Such a portfolio would long a stock if the AI forecasts a higher target price than the median analyst forecast over a time horizon and short a stock if otherwise.

absolute forecast error relative to the ex post realization (i.e., the actual year-end stock price). We find human analysts perform better for more illiquid, smaller firms, and firm with asset-light business models (i.e., higher intangible assets), consistent with the notion that such firms are subject to higher information asymmetry and require better institutional knowledge or industry experience to decipher. Analysts affiliated with large brokerage houses also stand a higher chance of beating the machine, a combination of their abilities and the research resources available to them. Moreover, analysts are more likely to have the upper-hand when the associated industry is experiencing distress, suggesting that the AI has yet to catch up on relatively infrequent changes such as an industry recession. This is consistent with the limitation of current machine learning and AI models which lack reasoning functions and thus cannot learn effectively from infrequent events.[7] As expected, AI enjoys a clear advantage in its capacity to process information, and is more likely to out-smart analysts when the volume of public information is larger.

Just like the "centaur" chess player which Kasparov pioneered, the superior performance of an AI analyst does not rule out the value of human inputs. If human and machine have relative advantage in information processing and decision making, then human analysts may still contribute critically to a "centaur" analyst, i.e., an analyst who makes forecasts that combine their own knowledge and the outputs/recommendations from AI models. After we add analyst forecasts to the information set of the machine learning models underlying our AI analyst, the resulting "man + machine" model outperforms 57.3% of the forecasts made by analysts, and outperforms the AI-only model in all years. Thus, AI analyst does not displace human analysts yet; and in fact an investor or analyst who combines AI's computational power and the human art of understanding soft information can attain the best performance.

We are thus interested in knowing when the incremental value of human to a man + machine model is the highest, as manifested in the relative performance of the man + machine model versus the pure AI model. Similar to previous findings, we find inputs from analysts

---

[7]Source: "What AI still can't do," Brian Bergstein, *MIT Technology Review*, February 19, 2020

are more valuable when covering firms that are more illiquid and firms with more tangible assets. Moreover, analyst inputs have more incremental value in long-horizon forecasts, and during the time period an industry is experiencing difficulty. Importantly, the incremental value of human does not decrease as the volume of information (hence demand for processing capacity) increases, though this constitutes a human disadvantage when alone. Similarly, analysts from small brokerage houses make similar level of contribution to the man + machine model compared to their counterparts from larger banks, suggesting that AI could potentially help level the disparity in institutional resources.

Finally, we resort to an event study to sharpen the inference of the impact of integrating man and machine in stock analyses. In recent years, the infrastructure of "big data" has created a new class of information about companies that is collected and published outside of the firms, and such information provides unique and timely clues into investment opportunities. An important and popular type of alternative data captures "consumer footprints," often times in the literal sense such as satellite images on retail parking lots. Such data, which have to be processed by machine learning models, have been shown to contain incremental information for stock prices (Zhu, 2019; Katona, Painter, Patatoukas, and Zeng, 2020). We build on data from Katona, Painter, Patatoukas, and Zeng (2020) on the staggered introduction of several important alternative data bases, and conduct a difference-in-differences test of analysts' performance versus our own AI model before and after the availability of the alternative data. The underlying premise is that analysts who cover firms that are served by the alternative data could be in the situation of man + machine, as they have the opportunity to use the additional, AI-processed, information. Indeed, we find that post alternative data, analysts covering affected firms improve their performance relative to the AI-only forecast model we build. Furthermore, such improvement concentrates in the subset of analysts who are affiliated with brokerage firms with strong AI capabilities, measured by AI-related hiring

using the Burning Glass U.S. job posting data[8] and the classification algorithm developed in Babina, Fedyk, He, and Hodson (2020).

Overall, results support the hypothesis that analyst capabilities could be augmented by AI, and moreover, analysts' work possesses incremental value such that they, with the assistance of AI, can still beat a machine model without human inputs, analogous to the centaur chess players' out-performing machines as we discussed earlier. If there is some external validity from chess and stock analysis to skilled workers in general, the inference from our study is encouraging news for humans in the age of AI.

Our work is related to the rapidly growing literature on the competition and threat to human workers posed by new technology including robots and AI.[9] The literature overall finds that when low- or intermediate-skill jobs are replaced by machines, humans tend to move to high-skill jobs that are more difficult to replace. However, the most recent wave of AI innovations disrupt many of the high-skill jobs. Our study focuses on humans' relative advantage over machines and, more importantly, the potential synergies between humans and machines.[10] We envision a future in which AI and machines can assist humans with the more tedious and quantitative tasks and democratize access to information, allowing humans to be more creative and productive.[11]

A few contemporaneous papers also study the impact of big data and AI in the financial industry. Abis (2020) studies how quantitative investment strategies influence mutual fund

---

[8]Burning Glass is currently the leading data vendor in job postings in the U.S. The postings are scraped from web sites, news letters, and agency reports, and cover the period of 2007, and then 2010-2019. Acemoglu, Autor, Hazell, and Restrepo (2020) show that Burning Glass data cover 60-80% of all U.S. job vacancies.

[9]An incomplete list of recent papers include Aghion, Jones and Jones (2017), Acemoglu and Restrepo (2018), Acemoglu and Restrepo (2019), Brynjolfsson, Mitchell and Rock (2018), Webb (2020), Ray and Mookherjee (2020), Cao, Cong, and Yang (2019), Acemoglu, Autor, Hazell, and Restrepo (2020), and Jiang, Tang, Xiao, and Yao (2021).

[10]An example in a non-finance setting comes from Sports. After trying Video Assistant Referee (VAR) for a few seasons, the English Premier League decided not to let VAR over-power referee judgment. One main reason is that players will reverse-engineer the rules underlying the VAR decisions and play to their advantage, such as committing more "low-trade" (to the machine) but atrocious (to humans) fouls. This can be remedied by having a human referee having the final say. See "Why Has The Introduction Of Video Technology Gone So Badly In Soccer?" James Reade, *Forbes*, December 10, 2020.

[11]Due to the complementary nature of AI and humans, the advent of AI technologies can potentially create more jobs than they destroy. See "Artificial Intelligence To Create 58 Million New Jobs By 2022, Says Report," Amit Chowdry, *Forbes*, September 18, 2018.

performance. Abis and Veldkamp (2020) examine the change in labor shares in the financial industry driven by the new data management and AI jobs. Coleman, Merkley, and Pacelli (2020) compare the performance of robot analysts from FinTech companies with human analysts. Grennan and Michaely (2020) study how analysts perform and adjust in response to the advent of AI-processed recommendations in the markets. Rossi and Utkus (2021) compare human asset managers with robot advisors. Our paper differs from the existent literature in that we are able to explore the internal mechanism of machines we constructed ourselves, and to identify its relative advantages to humans more directly. More importantly, our set-up allows us to integrate human and machines so that we are able to investigate the incremental value of humans relative to machines in such a cooperation.[12]

Finally, we contribute to the literature of building and assessing the performance of machine learning models in financial applications, e.g., in predicting asset prices (Gu, Kelly, and Xiu, 2020), robo-advising (D'Acunto, Prabhala, and Rossi (2019)), portfolio management (Chen, Pelger, and Zhu, 2020; Cong, Tang, Wang, and Zhang, 2020), estimating values of artwork (Aubry, Kraeussl, Manso, and Spaenjers, 2020), earnings forecast (van Binsbergen, Han, and Lopez-Lira, 2020), lending decisions (Liu, 2019), innovation evaluation (Zheng, 2021) and estimating bank risk (Hanley and Hoberg, 2019). While we also employ machine learning models to build our AI analyst, we focus more on understanding the difference between humans and machines, and the complementary value humans can offer in the age of AI. Our study is one of the first to study the implications of the combination of humans and machines in the financial markets.[13] Given the increasing presence of machines and AI, understanding how they can best complement and improve humans will bring to fruition the original mission of AI development.[14]

---

[12]To make sure our AI model is consistent with those in industry practice, we communicated with investment companies who also develop AI analyst models and found that they reachedmany similar conclusions regarding the incremental value of human to AI.

[13]Armour, Parnham, and Sato (2020) study the impact of AI and the associated digital technologies on the law profession. They find that AI-enabled services will augment the capabilities of human lawyers and also generate new roles for legal experts to produce such services.

[14]This echoes the mission of the Stanford Human-Centered AI Institute, "to advance AI research, education, policy and practice to improve the human condition." See https://hai.stanford.edu/about.

2. Information Sources, Data Construction, and Machine Learning Models

2.1. Sample of forecasts

Our sample of analyst target price forecasts builds on the Thomson Financial I/B/E/S analyst database, over the period from 1996 to 2016.[15] After merging I/B/E/S with CRSP and Compustat, the final sample consists of 685,888 analyst target price forecasts on 6,118 firms, issued by 10,287 analysts from 755 brokerage firms.

We choose to analyze analysts' price forecasts instead of earning forecasts because earnings are subject to managerial discretion or even manipulation. Earnings could be endogenous to analyst forecasts due to the well-documented feedback loop due to managerial incentive to meet and beat analyst consensus (Abarbanell and Lehavy, 2003; Doyle, Jennings, and Soliman, 2013). Prices are more difficult to control by insiders and thus provide a more objective benchmark to assess performance of human analysts and machines.

2.2. Building the information set for AI analyst

Given our goal to build an AI analyst to compete with professional analysts, we need to define the information set available to such a professional whenever a price forecast is made. The unit of analysis in our main set up is a forecast on the year-end stock price for firm i by (human) analyst k on date t (in year u). We focus on year-end prices because these are a focal point for performance evaluation, tax, and corporate reporting purposes. The information set, $\mathcal{I}_t$, would, in an ideal setting, include all publicly available data and information up to $t-$. We assume that professional analysts do not have access to material non-public information, which is essentially the requirement of Regulation FD.[16] We approximate $\mathcal{I}_t$

---

[15]The I/B/E/S coverage prior to 1996 was limited, with fewer than 2,000 in total of target price forecasts.

[16]Regulation FD ("fair disclosure"), implemented in 2000, generally prohibits public companies from disclosing previously nonpublic, material information to certain parties unless the information is distributed to the public first or simultaneously.

with firm and industry information from CRSP and Compustat, textual information from firms' SEC Filings, including annual reports (10-K), quarterly reports (10-Q), ad hoc disclosure of material corporate news and developments (8-K), and other reports, as well as macroeconomic data from the Federal Reserve Economic Data at Fed at St. Louis.

To operationalize time adaptation, we adopt the following rolling-window approach. For a given forecast made by a (human) analyst on date t in year u, all forecasts in the previous five years $u-5, u-4, \ldots, u-1$ form the training sample. That is, inputs data up to the dates of those forecasts (but excluding the forecasts themselves) and the corresponding realized prices were used to train our machine learning models. Move to the estimation sample, we then feed inputs data available up to date $t-1$ in year u into the trained model to make the prediction at time t for the year-end price. Because our sample starts in 1996, our AI Analyst makes its first prediction in 2001 given the requirement of five years as the training period. Though we allow (public) information to be updated till $t-1$, most of the information inputs came from disclosed quarterly data from the previous eight quarters.[17]

2.3. Information and variables as inputs to machine learning

Firm Characteristics The firm characteristics fed into machine learning models are retrieved or processed based on information from standard databases accessed via WRDS, especially CRSP/Compustat and Thomson Reuters Ownership Databases. The first set of predictors include annual returns and end-of-year stock prices for the past five years, as well as the realized earnings within the past 3, 6, 9, 12, 24 and 36 months. We also include a number of firm characteristics known to predict cross-sectional differences of the stock prices. In particular, we include anomalies from each of the six broad categories of anomalies considered in Hou, Xue, and Zhang (2020), i.e., the momentum, value versus growth,

---

[17]Data that come in different frequencies, e.g., corporate news releases (8-K), are aggregated at the quarterly level.

investment, profitability, intangibles, and trading frictions categories.[18] Variables in this group are constructed quarterly using information available at the previous quarter-end.

Industry Variables We compose a set of industry-level variables that capture competition, industry dynamics, and other factors relevant for firm valuation based on the existent literature. The set of variables include the following: (i) The competition measure from 10K following Li, Lundholm, and Minnis (2013), which captures the degree of competition resulting from rivalry within and across industries as perceived by the management; (ii) the product market fluidity measure following Hoberg, Phillips and Prabhala (2014), which quantifies product market poaching threat posed by the movement of competitors toward the focal firm; (iii) industry affiliation with the Fama-French 48 industries (48 indicator variables); (iv) industry size, measured by the number of firms in the Fama-French 48 industry, within the past 3, 6, 9, 12, 24 and 36 months; and (v) equally-weighted industry average earnings per share realized within the past 3, 6, 9, 12, 24 and 36 months.

Macro Variables Macro-economic and stock market development are common factors to all firms' valuation and returns (e.g., Fama and French, 1989; Chen, Roll, and Ross, 1986). For this category we include the following variables: (i) Industrial Production Index; (ii) Consumer Price Index; (iii) Crude oil price (WTI); (iv) three-month treasury bill rate; (v) ten-Year treasury constant-maturity rate; and (vi) The BAA-AAA yield spread. These macro variables are obtained from the Federal Reserve Economic Data at Federal Reserve St. Louis on a monthly frequency.

Textual Sentiment Information One leading strength of machine over human being is the former's ability to digest large volume of information. One new edge that machine learning models boast over traditional statistical methods is its capacity to process unstructured textual data based on firms' SEC Filings, including annual report (10-K), quarterly report

---

[18]We list all variables serving as inputs into the machine learning models, their definitions, and sources in Table A1 in Appendix A.

(10-Q), corporate news (8-K), and other reports. The new developments allow researchers to quantify information which was considered qualitative or "soft," commonly termed "sentiments."

Two different sets of sentiment variables from textual data serve as inputs to our AI analyst. The first set of sentiment variables is based on the Loughran and McDonald (2011) sentiment, which has been widely used in the academic literature. We calculate the frequency of positive and negative sentiment words from the firm-issued SEC filings following Loughran and McDonald (2011). The second set of machine-learning-based sentiment variables follow Cao, Kim, Wang, and Xiao (2020), where a deep-learning neural network model to incorporate contextual information and syntactic relations between words that are performance-related. The second approach aims to isolate managerial sentiment related to the firm's future performance from sentiment regarding other issues (such as location and weather).

2.4. Potential factors for the relative performance of AI and human analysts

A main objective of the study is to assess the factors that contribute to the relative performance of AI vs. human, as well as the synergy between the two. We hypothesize that these factors are related to the information environment of the firm, industry, and analysts.

We first consider the following firm-level variables: the *Amihud Illiquidity* measure which is the ratio of absolute daily stock return to the daily trading volume (in dollars); *Stock Volatility* (the total return volatility); *Log Market Cap* (the logarithm of market capitalization); *# 8K Reports*, which is the number of 8K reports filed each year and represents the volume of available information about the firm; and *Intangible Assets*, defined as the average of three proxies: one minus the ratio of PP&E to assets, organization capital scaled by assets, and knowledge capital scaled by assets. The last two measures, derived from the accumulation of SG&A and R&D expenditures, are constructed following Ewens, Peters, and Wang (2020).

11

We further include a number of variables that characterize the information environment and resources for analysts: *Forecast Horizon* is the number of days from the forecast to the year end; *# Analysts in Brokerage Firm*, is the number of analysts and proxies for the size and resources of the brokerage firm; *# Previous Predictions* is the number of previous predictions by all I/B/E/S analysts in the same firm-year on the firm-year; *# Institutional Owners* is the total number of institutional owners; *Industry Distress*, which equals 1, if the industry return is negative and *Time Trend* equals the number of years elapsed from the beginning of the sample (2001). The final set of variables are related to analysts' access to alternative data and AI resources, which will be introduced in Section 5.2.

2.5. Machine Learning Models

We compare a number of machine learning models and obtain our main model as an ensemble of the three best-performing models. The machine learning models are essential tools but not the ultimate objectives of this study. For completeness we provide a brief overview in Appendix B of the models we examined, which include: Elastic-Net, Support Vector Machines, Random Forest, Gradient Boosting, and Long Short-Term Memory Neural Networks. Moreover, the overview focuses on the economic intuition of each methodology's mechanism as well as strength. For further details, we refer the reader to the most representative references in this field, e.g., Hastie, Tibshirani and Friedman (2009) and Goodfellow, Bengio, and Courville (2016).

3. Building and Performance of AI Analyst

Our ultimate goal is not to build an AI analyst per se, but instead is to analyze the relative strength as well as synergy between machine and human. Such a goal, however, sets the premise that the AI analyst at our hand needs to be strong enough to compete with, or even beat, human analysts. This section describes how we build an AI analyst that meets

the calibre.

3.1. The predictive models

For each stock i at time t, where t is the day when an analyst makes a forecast, $F_{i,j,t}^{Man}$ (where i, j, t are indices for the stock, the analyst, and the date; and the superscript Man indicates human as opposed to AI), of the year-end target price, our model makes an attempt at predicting the same target, i.e., the stock price $P_{i,T(t)}$, where $T(t)$ is the last trading day of the calendar year containing t. Included in the predictive information set is all public information (as described in Section 2 and Appendix A) up to $t-$. We summarize the prediction model as:

$$F_{i,t}^{AI} = f(X_{i,t-}) + \epsilon_{it}. \tag{1}$$

In (1), stock price is expressed in natural logarithm in order to mitigate skewness.

Next we compare the AI forecast $F_{i,t}^{AI}$ and the analyst forecast $F_{i,j,t}^{Man}$, in terms of their accuracy relative to the ex post realized price $P_{i,T(t)}$. AI beats human if $|F_{i,t}^{AI} - log(P_{i,T(t)})| < |F_{i,j,t}^{Man} - log(P_{i,T(t)})|$, and vice versa. We define $Beat$ to be an indicator variable for AI winning. Moreover, we $F_{i,t}^{AI} > F_{i,j,t}^{Man}$ to be a "buy" signal and "sell" otherwise. Figure 1 shows the relative performance of AI vs. Analyst forecasts over time.

[insert Figure 1 here]

Figure 1 shows that out of 652,659 forecasts made by analysts during 2001-2016, our AI Analyst we build is able to beat 53.4% of them. The p-value for the percentage to be drawn from a distribution with the neutral probability of half for this sample size is less than 0.01%. However, the AI advantage is volatile from year from year ranging from 37.8% in 2013 to 64.4% in 2011 with a negative drift, as shown by the trend-fitted line. We conjecture (which will be tested) that the waning advantage of the AI analyst is precisely because human analysts increasingly have access to, and are assisted by improving technologies in data collection, statistical packages, and machine learning tools.

To put the performance comparison in perspective and to illustrate progressive improvement of the AI models, Table 2 compares various models, from the simple OLS model to a variety of machine learning algorithms, in their ability to beat analyst forecasts. As the starting line, the simple OLS model only beats analysts in 19.3% of the cases, suggesting that analysts' job is indispensable with the availability of standard statistics tools such as linear regression models. The performance of quasi-linear machine learning models is better than the plain OLS, at 23.3% (Elastic Net) and 27.2% (Support Vector Machines), respectively. As explained in Appendix B, the quasi-linear models are better at identifying the most important factors among a large number of independent variables, or at dimension-reduction tasks, resulting in better out-of-the-sample performance compared to the OLS. However, the quasi-linear models are still decidedly inferior to the average analyst, suggesting that humans are capable of processing flexible (nonlinear) relations and interactions among different informative signals better than quasi-linear and non-linear models.

[Insert Table 2 here]

We next apply the LSTM neural networks model, a model that is known for capturing nonlinear and time-series dynamics (Chen, Pelger, and Zhu, 2020).[19] The LSTM model outperforms just about half, or 50.1%, of the analyst forecasts, indicating that the neural network model is indeed helpful in capturing nonlinear patterns in our sample.

Finally, we progress to two other nonlinear machine-learning models that themselves are ensemble models built on decision trees, the random forest model and gradient boost models (see Appendix B for details). These two models are known to be especially proficient at capturing complicated interactions of multiple independent variables. The random forest and gradient boost models perform better than 51.7% and 52.0% of analysts, respectively. Given that the LSTM, random forest, and gradient boost models capture important but

---

[19]The LSTM model involves random factors in the optimization process; therefore, one run of LSTM model cannot predict precisely given the randomness. To alleviate this concern, we form an ensemble LSTM model that aggregates the results of 20 runs of the LSTM model by taking the mean prediction. We also examine the ensemble methods using 10 runs and 60 runs and the performance of these models stays the same.

different nonlinear and interactive patterns and features, we form another layer of ensemble of these three models to obtain synergistic benefits of the three models. The resulting "AI analyst" outperforms 53.7% of analysts.

The progression of forecast accuracy demonstrated by Table 2 signifies the relentless upgrading of AI technologies over time, to the point that AI reaches or even beats human intelligence in one after another skilled task. Financial analysts are thus vulnerable, especially considering that our model, which already beats the average analyst, is a lower bound of what state-of-art AI technology could accomplish.

3.2. AI vs. Analysts with Persistent Performance

Analysts are a large group with heterogeneous skill levels such that forecast performance would to be persistent if skills are innate. Moreover, the market recognizes, at least partially, the relatively more skilled analysts by responding more strongly to their forecasts or recommendations (Mikhail, Walther, and Willies, 2007; Li, 2005). Thus, a higher hurdle is for our AI analyst to beat the subset of skilled analysts. We assess the relative performance withe respect to the higher hurdle with two tests. First, we sort all analysts into the top and bottom halves based on their average prediction error (normalized by stock prices) over the past one, two, three, four, and five years. We then track the percentage of their future forecasts during the one year horizon that are beat by our AI analyst. In the second test, we repeat the same procedure except selecting the analysts that are among top and bottom quantiles in each of the past one, two, three, four and five years. Table 3 reports the results.

[Insert Table 3 here]

Results in Table 3 show that the AI comfortably beats the analysts in the low skill quantiles. It is basically neck-to-neck to the more successful analysts, and is almost even with analysts (beat ratio of 49.2%) who demonstrated five years' superior performance in the recent past, which constitutes only 3.9% of all the analysts.

15

3.3. Performance of portfolio following AI Recommendations

Analysts make forecasts as a way to advise portfolio formation or turnover. The performance of a portfolio following the analyst advice is a natural metric for analyst skill. For the same reason, we can form portfolios based on the different opinions between the AI and human analysts. The performance of the resulting portfolio is testament of their relative proficiency.

In each month, we gather all predictions made by all analysts and the corresponding AI forecasts in past 30, 60, 90 and 360 days. For each pair of predictions, if AI predicts a higher (lower) price, it is considered as a buy (sell) signal. If analysts predicts a higher price. During the given time horizon, the portfolio will long the stock if there are more buy than sell signals; and short the stock otherwise. The portfolio is equal weighted. In a monthly- (or six month-) rebalanced portfolio, we hold the position for one month (six months), or till the signals reverse. The portfolio contains 400 to 1100 stocks with signals from past 30 days to 360 days, and long/short positions are roughly even.[20]

Table 4 Panel A reports the performance of the monthly rebalanced long-short portfolio, in terms of average return and alpha estimated using Fama-French three factor, Carhart four factor, Fama-French five factor and Fama-Fench six factor models. Panel B reports the same statistics for the semi-annually rebalanced portfolio.

[Insert Table 4 here]

Results in Table 4 are highly encouraging in that AI is able to generate superior returns/alpha, relative to analysts, on the order of 65 to 91 basis points monthly, significant at less than the 5% levels in all cases. When we separately examine the long and short portions of the portfolio, we discover that the superior returns are significant (all at the 1% level) only on the long-side (for which the transaction costs are lower). Such an asymmetry could be driven by the well-documented positive bias in analyst forecasts (Lim (2002)). While

---

[20]The average monthly turnover rate of the monthly-rebalanced portfolios ranges from 10% to 58%; and that for the six-month rebalanced portfolio is 6% to 12%.

AI forecasts contain no average directional bias, we confirm that the median analyst price forecast in our sample contains a 5.2% positive bias In such a scenario, signals are not as informative when analysts are more optimistic than the AI.[21]

3.4. Combined wisdom of Man and Machine

Results from the previous sections suggest that the analyst profession could be seriously disrupted by AI technology. However, the superior performance of the AI Analyst does not rule out the possibility that analyst forecasts contain information that is incremental to AI-produced forecasts. In other words, if analysts are able to possess information that is not picked up by the AI, then AI forecast is not sufficient for the analyst forecast, even though analysts lose to AI in forecasting accuracy. An investor who combines the wisdom of both should attain even better performance.

To assess the performance of combined power, we consider adding the analyst forecasts to the information set to our machine learning model. That is, the information set $\mathcal{I}_t$ now includes the analyst forecasts, $F_{i,j,t}$, made on the same firm i during the 90-day window ending on date t. In particular, we obtain the consensus and mean square error of the forecasts by analysts in the previous 90 days, and build a "Man + Machine" hybrid analyst. We find that the hybrid analyst outperforms 57.3% of all human analysts, and 56.6% of the AI-alone forecasts. Figure 2 plots the relative performance of the hybrid analyst against human analysts in the time-series. We find that the hybrid analyst beats more than 50% of analysts for almost all years. Furthermore, the average performance of the hybrid analyst vs. human analysts is now relatively stable over time with a slight decline. Thus, the hybrid model continues to provide substantial benefits even in recent years.

[Insert Figure 2 here]

---

[21]As Lim (2002) pointed out, analyst forecast bias is related to analyst and company characteristics as analysts may trade off bias to improve management access and forecast accuracy. As such, simply "debiasing" analyst forecasts lead to even worse performance.

Figure 3 plots the relative performance of the hybrid analyst (Man + Machine) vs. the plain AI (Machine). Interestingly, Man + Machine outperforms plain Machine in every single year. Such outperformance, which captures the incremental value of human analysts to machines, increases substantially with time. The combined results portrays a bright future for Man + Machine: not only the combination attains better performance than either side alone, but also the incremental value of human does not weaken with the technology.

[Insert Figure 3 here]

4. Man vs Machine: Relative Advantage of Analyst vs AI

In this section, we strive to understand when human analysts perform better than the AI and when otherwise. Such understanding will help to "unbox" the black box associated with AI or machine learning, and provides intuition and guidance on the applicability of AI for researchers and investors.

We consider a number of variables at the analyst, firm, and industry levels that are potentially relevant for the performance of human analysts and AI, which are defined in Section 2.4. We group these variables into several classes. First, we consider a number of proxies for information asymmetry or opacity, including *Amihud Illiquidity*, *Stock Volatility*, *Log Market Cap*, and *# Institutional Owners*. Second, we include variables representing the volume of information (*# 8K Reports*) and the tangibility of information (*Intangible Assets*). Third, we examine several variables that affect the information and resources available to the analyst, e.g., *Forecast Horizon*, *# Analysts in Brokerage Firm*, and *# Previous Predictions*. Finally, we consider *Industry Distress*, highlighting the uncertainties at the industry level (Opler and Titman, 1994), and finally *Time Trend*, which can help to capture temporal patterns.

For each target price forecast, we define two variables that measure the relative performance of humans vs AI. First, the indicator variable *Analyst Beats AI* equals one if the absolute forecast error of the analyst is smaller than that of the AI, and zero otherwise.

18

Second, the continuous measure *Forecast Error Difference* is the difference between the absolute prediction error (of log price as defined in Equation (1)) of the AI and that of the analyst, scaled by the maximum of these two prediction errors if the difference is non-zero. A positive and large value of *Forecast Error Difference* is in favor of analyst accuracy.

We estimate the follow regression, on the panel data of firm i, analyst j, and date t, to understand the determinants of the relative strengths of humans and AI,

$$Relative\ Performance_{i,j,t} = X'_{i,j,t}\beta + \alpha_i + \alpha_{year} + \epsilon_{i,j,t}, \tag{2}$$

where the dependent variable *Relative Performance* is either *Analyst Beats AI* (Panel A) or *Forecast Error Difference* (Panel B), $X_{i,j,t}$ include the independent variables discussed above, and α_i and α_{year} represent firm and year fixed effects, respectively. The results are reported in Table 5.

[Insert Table 5 here]

Table 5 show that, controlling for year and firm fixed effects, humans are more likely to outperform when covering illiquid, small firms with higher intangible assets, consistent with the notion that such firms are subject to higher information asymmetry and require deeper institutional knowledge to understand. On the other hand, equipped with vast processing power, AI performs better for firms with a larger volume of disclosed information, as proxied by *# 8K Reports* each year.

Human analysts perform better when the forecast horizon is shorter, indicating that they are better at taking cues from, or perhaps have access to, informal value-relevant information ahead of a shorter forecast horizon. Those working for larger brokerage hours performance better, both because of positive skill matching to more prestigious work places and because of the more abundant resources and research capacity at such places.

Outside the firms, humans perform better when the associated industry is experiencing a downturn, suggesting that the AI has more difficulty handling less frequent and more

uncertain events such as an industry recession. Finally, when year fixed effects are not included, Columns (2) and (4) show that the comparative performance of humans increases with time. This is probably due to the fact that human analysts are increasingly assisted by AI and big data technologies over time.

5. Man + Machine: Combining Strengths and Incremental Contributions

5.1. Incremental value of analyst and AI in forecasts made by Man + Machine

Acknowledging that Man + Machine is superior to either Man or Machine alone, it is still instructive to understand the respective incremental values of Man and Machine in the combo. Analogous to the previous section, we define relative performance measures of the hybrid analyst vs the AI (the human analyst) to capture the incremental value of Man (Machine). We then reestimate Equation (2) with these relative performance measures as dependent variables. Table 6 presents the results. Given that incremental value is measured better with the continuous measure of absolute forecast error difference, we focus on the results in Panel B.

[Insert Table 6 here]

Similar to the previous findings, we find inputs from analysts are more valuable when covering firms that are more illiquid and firms with more tangible assets. Moreover, analyst inputs have more incremental value in long-horizon forecasts, and during the time period an industry is experiencing difficulty.

We note that while analysts in larger brokerage firms perform better than other analysts when pitted against the AI (Table 5), the former do not add significantly more incremental value in the Man + Machine model than the latter (Columns (1) and (2) in Panel B of Table

6). Such a contrast in the results highlight that democratizing AI technology levels the playground: When we let all analysts (from large and small brokerage houses) be equipped with AI assistance in the Man + Machine model, disparity in institutional resources does not significantly affect the incremental value of human inputs. In a similar vein, the incremental value of human does not decrease as the volume of information (hence demand for processing capacity) increases, though this constitutes a human disadvantage when forecasting alone without AI inputs.

5.2. Impact of Man + Machine: An event study

In this section, we resort to an event study to sharpen the inference of the impact of integrating man and machine in stock analyses. In recent years, the infrastructure of "big data" has created a new class information about companies that is collected and published outside of the firms, which can provide unique and timely clues into market demand, profit prospects, and investment opportunities. An important and popular type of such alternative data captures "consumer footprints," often times in the literal sense such as satellite images on retail parking lots. Such data, which have to be processed by machine learning models, have been shown to contain incremental information for earnings and stock prices conditional on corporate disclosure and news coverage (Zhu, 2019; Katona, Painter, Patatoukas, and Zeng, 2020).

We build on data from Katona, Painter, Patatoukas, and Zeng (2020) on the staggered introduction of several important alternative data bases, and conduct a difference-in-differences test of analysts' performance versus our AI model before and after the availability of the alternative data on specific firms. The underlying premise is that analysts who cover firms that are served by the alternative data are potentially in the situation of Man + Machine, as they have the opportunity to use the additional, AI-processed, information. We define two variables based on the staggered introduction of alternative data coverage. The first is *Alt Data Covered*, which is one if satellite imaging data are available for the firm at any point

in our sample period (based on the list of covered firms and coverage start dates in Table 1 in Katona, Painter, Patatoukas, and Zeng, 2020), and zero otherwise; The second if *Post*, which is an indicator variable that is one if satellite data are currently available (based on coverage start dates in Table 1 in Katona, Painter, Patatoukas, and Zeng, 2020), or if the firm is not listed in that table but the date is after 2014,[22] and zero otherwise.

Alternative data tend to be large in volume and unstructured in recording. Such data are hard to process with traditional tool kits. Despite that commercial data vendors may have preprocessed the alternative data, e.g., satellite imaging data are converted into car counts for each business location, substantial additional analysis is still needed to render such data useful for stock analysis. Whether analysts covering the alternative data "treated" firms could capitalize on the novel information source depends on the AI resources in their work place. We measure AI resources analysts have access to by the variable *AI Hiring*, which is the ratio of the number of AI jobs to the total number of job postings using the Burning Glass U.S. job posting data and following the classification algorithm developed in Babina, Fedyk, He, and Hodson (2020).

We estimate the following difference-in-differences model,

$$
\begin{aligned}
\textit{Analyst Beats AI}_{i,j,t} = {} & \beta_1 \textit{Alt Data Covered}_i \times \textit{Post}_{i,t} \times \textit{AI Hiring}_{j,t} \\
& + \beta_2 \textit{AI Hiring}_{j,t} + \beta_3 \textit{Post}_{i,t} \\
& + \beta_4 \textit{Alt Data Covered}_i \times \textit{Post}_{i,t} \\
& + \beta_5 \textit{Alt Data Covered}_i \times \textit{AI Hiring}_{j,t} \\
& + \beta_6 \textit{Post}_{i,t} + \textit{Controls}_{i,j,t} + \alpha_i + \alpha_{year} + \epsilon_{i,j,t}.
\end{aligned} \tag{3}
$$

Note that *Alt Data Covered* and *Post* are indexed by firm i and date t while *AI Hiring* is indexed by the analyst j (or the brokerage firm associated with the analyst) and date t. Table 7 reports the results. The sample here is smaller than those in Tables 5 and 6 due to

---

[22]Based on anecdotal evidence from news and discussion with industry experts, 2014 is the year most alternative data become widely available.

the requirement that the *AI Hiring* be observable.

[Insert Table 7 here]

Columns (1) and (2) of Table 7 show that analysts associated with brokerage houses with greater AI capabilities generally perform better against our AI model, a piece of direct evidence that humans complemented by AIs enjoy a step up in predictive capabilities. Columns (3) and (4) show that post alternative data, analysts covering affected firms improve their performance relative to the AI model, but only significantly so when interacting with *AI Hiring.* In other words, the improvement of predictive performance post alternative data concentrates in the subset of analysts who are affiliated with brokerage firms with strong AI capabilities. Overall results suggest that augmenting humans with new technologies constitutes a promising direction for the analyst profession.

6. Concluding Remarks

In this paper we build an AI analyst to digest corporate disclosure and other and information (qualitative and quantitative), and perform forecast tasks similar to those of stock analysts. Our AI analyst is able to beat the majority of human analysts in stock forecasts. A portfolio following the difference between AI and analyst's forecasts generates a monthly excess return of more than 60 basis points. In the contest of "Man vs Machine," we find that the relative advantage of such an AI Analyst is stronger when when information is more transparent and voluminous. Human analysts remain competitive when critical information requires institutional knowledge (such as the nature of intangible assets). The edge of the AI analyst over human analysts declines over time, especially when analysts gain access to alternative data and to in-house AI resources. Combining AI and the art of human experts produce the highest potential in generating accurate forecasts in settings when the two skills are complementary, suggesting a future of "Man + Machine" in high-skilled professions.

The complementarity between humans and machines documented in this study also

provides guidance about how humans can adapt to survive and thrive in the age of machines. For example, reforming education and professional training to strengthen soft skills and creativity can help human professionals to better prepare for the incoming future.

References

Abarbanell, Jeffery and Reuven Lehavy, 2003, Biased forecasts or biased earnings? The role of reported earnings in explaining apparent bias and over/underreaction in analysts earnings forecasts, *Journal of Accounting and Economics* 36, 105–146.

Abis, Simona, 2020, Man vs. Machine: Quantitative and Discretionary Equity Management, Working paper.

Abis, Simona and Laura Veldkamp, 2020, The Changing Economics of Knowledge Production, Working paper.

Acemoglu, Daron, David Autor, Jonathon Hazell, and Pascual Restrepo, 2020, AI and Jobs: Evidence from Online Vacancies, Working paper.

Acemoglu, Daron and Pascual Restrepo, 2018, The Race Between Man and Machine: Implicationsof Technology for Growth, Factor Shares and Employment, *American Economic Review* 108, 1488–1542.

Acemoglu, Daron and Pascual Restrepo, 2019, Automation and New Tasks: How Technology Displaces and Reinstates Labor, *Journal of Economic Perspectives* 33, 3–30.

Aghion, Philippe, Benjamin F. Jones, and Charles I. Jones, 2017, Artificial Intelligence and Economic Growth, Working paper.

Agrawal, Ajay, Joshua Gans, and Avi Goldfarb, 2019, The Economics of Artificial Intelligence: An Agenda, National Bureau of Economic Research.

Amihud, Yakov, 2002, Illiquidity and Stock Returns: Cross-Section and Time Series Effects, *Journal of Financial Markets* 5, 31–56.

Armour, John, Richard Parnham, and Mari Sato, 2020, Augmented Lawyering, Working paper.

Aubry, Mathieu, Roman Kraeussl, Gustavo Manso, and Christophe Spaenjers, 2020, Biased Auctioneers, Working paper.

Babina, Tania, Anastassia Fedyk, Alex Xi He, and James Hodson, 2020, Artificial Intelligence, Firm Growth, and Industry Concentration, Working paper.

Barbee, William C., Sandip Mukherji, and Gary A. Raines, 1996, Do Sales-Price and Debt-Equity Explain Stock Returns Better than BookâMarket and Firm Size? *Financial Analysts Journal* 52, 56–60.

van Binsbergen, Jules H., Xiao Han, and Alejandro Lopez-Lira, 2020, Man vs. Machine Learning: The Term Structure of Earnings Expectations and Conditional Biases, Working paper.

Basu, Sanjoy, 1983, The relationship between earnings' yield, market value and return for NYSE common stocks: Further evidence, *Journal of Financial Economics* 12, 129–156.

Bhandari, Laxmi Chand, 1988, Debt/Equity Ratio and Expected Common Stock Returns: Empirical Evidence, *Journal of Finance* 43, 507–528.

Boudoukh, Jacob, Roni Michaely, Matthew Richardson, and Michael Roberts, 2007, On the Importance of Measuring Payout Yield: Implications for Empirical Asset Pricing, *Journal of Finance* 62, 877–915.

Bradshaw, Mark T., Scott A. Richardson, and Richard G. Sloan, 2006, The relation between corporate financing activities, analysts' forecasts and stock returns, *Journal of Accounting and Economics* 42, 53–85.

Brav, Alon and Reuven Lehavy, 2003, An Empirical Analysis of Analysts' Target Prices: Short-term Informativeness and Long-term Dynamics, *Journal of Finance* 58, 1933–1968.

Breiman, Leo, Random Forests, *Machine Learning* 45, 5–32.

Brynjolfsson, Erik, Tom Mitchell and Daniel Rock, 2018, What Can Machines Learn, and What Does It Mean for Occupations and the Economy? *AEA Papers and Proceedings* 108, 43–47.

Campbell, John Y., Jens Dietrich Hilscher, and Jan Szilagyi, 2008 In Search of Distress Risk, *Journal of Finance* 63, 2899–2939.

Cao, Sean, Lin William Cong, and Baozhong Yang, 2019, Financial Reporting and Blockchains: Audit Pricing, Misstatements, and Regulation, Working paper.

Cao, Sean, Wei Jiang, Baozhong Yang, and Alan L. Zhang, 2021, How to Talk When a Machine is Listening: Corporate Disclosure in the Age of AI, Working paper.

Cao, Sean, Yongtae Kim, Angie Wang, and Houping Xiao, 2020, Power of Deep Learning: Quantifying Language to Explain Cross-Sectional Returns, Working paper.

Chan, Louis K. C., Josef Lakonishok, and Theodore Sougiannis, 2001, The Stock Market Valuation of Research and Development Expenditures, *Journal of Finance* 56, 2431–2456.

Chen, Luyang, Markus, Pelger, and Jason Zhu, 2020, Deep Learning in Asset Pricing, Working paper.

Chen, Mark A., Qinxi Wu, and Baozhong Yang, 2019, How valuable is FinTech innovation? *Review of Financial Studies* 32, 2062–2106.

Chen, Long, Robert Novy-Marx, and Lu Zhang, 2011, An Alternative Three-Factor Model, working paper.

Chen, Nai-Fu, Richard Roll, and Stephen Ross, 1986, Economic Forces and the Stock Market, *The Journal of Business* 59, 383–403.

Cheng, Si, Ruichang Lu, and Xiaojun Zhang, 2020, What Should Investors Care About? Mutual Fund Ratings by Analysts vs. Machine Learning Technique, Working paper.

Cohen, Lauren, and Dong Lou, 2012, Complicated Firms, *Journal of Financial Economics* 104, 383–400.

Cohen, Lauren, Christopher Malloy, and Quoc Nguyen, 2020, Lazy prices, *Journal of Finance*, forthcoming.

Coleman, Braiden, Kenneth J. Merkley, and Joseph Pacelli, 2020, Man versus Machine: A Comparison of Robo-Analyst and Traditional Research Analyst Investment Recommendations, Working paper.

Cong, Lin William, Tengyuan Liang, Baozhong Yang, and Xiao Zhang, 2020, Analyzing textual information at scale, Information to Facilitate Efficient Decision Making: Big Data, Blockchain and Relevance (ed. Kashi Balachandran), World Scientific Publishers.

Cong, Lin William, Ke Tang, Jingyuan Wang, and Yang Zhang, 2020, AlphaPortfolio: DirectConstruction Through Reinforcement Learning and Interpretable AI, Working paper.

Cooper, Michael J., Huseyin Gulen, and Michael J. Schill, 2008, Asset Growth and the Cross-Section of Stock Returns, *Journal of Finance* 63, 1609–1651.

Crane, Alan and Kevin Crotty, 2020, How Skilled are Security Analysts, *Journal of Finance* 75, 1629–1675.

D'Acunto, Francesco Nagpurnanand Prabhala and Alberto G Ross, The Promises and Pitfalls of Robo-Advising, *The Review of Financial Studies* 32, 1983–2020.

Daniel, Kent and Sheridan Titman, 2006, Market Reactions to Tangible and Intangible Information, *Journal of Finance* 61, 1605–1643.

Doyle, Jeffrey T. , Jared N. Jennings and Mark T. Soliman, 2013, Do managers define non-GAAP earnings to meet or beat analyst forecasts? *Journal of Accounting and Economics* 56, 40–56.

Ewens, Michael, Ryan H. Peters, and Sean Wang, 2020, Measuring Intangible Capital with Market Prices, Working paper.

Fama, Eugene F. and Kenneth R. French, 1989, Business conditions and expected returns on stocks and bonds, *Journal of Financial Economics* 25, 23–49.

Fama, Eugene F. and Kenneth R. French, 1992, The Cross-Section of Expected Stock Returns, *Journal of Finance* 47, 427–465.

Fama, Eugene F. and Kenneth R. French, 1997, Industry costs of capital, *Journal of Financial Economics* 43, 153–193.

Fama, Eugene F. and Kenneth R. French, 2006, Profitability, investment and average returns, *Journal of Financial Economics* 82, 491–518.

Fama, Eugene F. and Kenneth R. French, 2015, A five-factor asset pricing model, *Journal of Financial Economics* 116, 1–22.

Goodfellow, Ian, Yoshua Bengio, and Aaron Courville, 2016. *Deep Learning*, MIT Press.

Grennan, Jillian and Roni Michaely, 2020, Artificial Intelligence and High-Skilled Work: Evidence from Analysts, Working paper.

Gu, Shihao, Bryan Kelly, and Dacheng Xiu, 2020, Empirical Asset Pricing via Machine Learning, *Review of Financial Studies* 33, 2223–2273.

Hanley, Kathleen Weiss, and Gerard Hoberg, 2010, The information content of IPO prospectuses, *Review of Financial Studies* 23, 2821–2864.

Hanley, Kathleen Weiss, and Gerard Hoberg, 2019, Dynamic interpretation of emerging risks in the financial sector, *Review of Financial Studies* 32, 4543–4603.

Hastie, Trevor, Robert Tibshirani, and Jerome Friedman, 2009, *Elements of Statistical Learning*, Springer Publisher.

Hirshleifer, David, Kewei Hou, Siew Hong Teoh, and Yinglei Zhang, 2004, Do investors overvalue firms with bloated balance sheets? *Journal of Accounting and Economics* 38, 297–331.

Hoberg, Gerard, Gordon Phillips and Nagpurnanand Prabhala, 2014, Product Market Threats, Payouts, and Financial Flexibility, *The Journal of Finance* 69, 293–324.

Hochreiter, Sepp and Jürgen Schmidhuber, 1997, Long Short-Term Memory, *Neural Computation* 9, 1735–1780.

Hou, Kewei, Chen Xue, and Lu Zhang, 2020, Replicating anomalies, *Review of Financial Studies* 33, 20192133.

James, Gareth, Daniela Witten, Trevor Hastie and Rob Tibshirani, 2013, An Introduction to Statistical Learning: with Applications in R, Springer Press.

Jegadeesh, Narasimhan, Joonghyuk Kim, Susan D. Krische and Charles M. C. Lee, 2005, Analyzing the Analysts: When Do Recommendations Add Value? *Journal of Finance* 59, 1083–1124.

Jegadeesh, Narasimhan and Sheridan Titman, 1993, Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency, *Journal of Finance* 48, 65–91.

Jiang, Wei, Yuehua Tang, Rachael Jiqiu Xiao, and Vincent Yao, 2021, Surviving the FinTech Disruption, Working paper.

Katona, Zsolt, Marcus Painter, Panos N. Patatoukas, and Jean Zeng, 2020, On the Capital Market Consequences of Alternative Data: Evidence from Outer Space, Working paper.

Kothari, S.P., Eric C. So, and Rodrigo S. Verdi, 2016, Analysts' Forecasts and Asset Pricing: A Survey, *Annual Review of Financial Economics* 8, 197–219.

Lakonishok, Josef, Andrei Shleifer, and Robert W Vishny, 1994, Contrarian Investment, Extrapolation, and Risk, *Journal of Finance* 49, 1541–1578.

Lamont, Owen, Christopher Polk, and Jesus Saa-Requejo, 2001, Financial Constraints and Stock Returns, *Review of Financial Studies* 14, 529–554.

LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton, 2015, Deep learning, *Nature* 521, 436–444.

Li, Xi, 2005, The persistence of relative performance in stock recommendations of sell-side financial analysts, *Journal of Accounting and Economics* 40, 129–152.

Li, Congcong, An-Ping Lin, and Hai Lu, 2020, Analyzing the Analysts: The Effects of Technical and Social Skills on Analyst Performance and Careers, Working paper.

Li, Feng, Russell Lundholm, and Michael Minnis, 2013, A Measure of Competition Based on 10-K Filings, *Journal of Accounting Research* 51, 399–436.

Li, Kai, Feng Mai, Rui Shen, and Xinyan Yan, 2020, Measuring corporate culture using machine learning, *Review of Financial Studies*, Forthcoming.

Lim, Terence, 2002, Rationality and Analyst Forecast Bias, *The Journal of Finance* 56, 369–385.

Liu, Miao, 2019, Assessing Human Information Processing in Lending Decisions: A Machine Learning Approach, Working paper.

Loughran, Tim and Bill McDonald, 2011, When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks, *Journal of Finance* 66, 35–65.

Loughran, Tim and Bill McDonald, 2016, Textual analysis in accounting and finance: A survey, *Journal of Accounting Research* 54, 1187–1230.

Loughran, Tim and Ritter, Jay R., 1991, The New Issues Puzzle, *Journal of Finance* 50, 23–51.

Loughran, Tim and Jay W. Wellman, 2011, New Evidence on the Relation between the Enterprise Multiple and Average Stock Returns, *Journal of Financial and Quantitative Analysis* 46, 1629–1650.

Lyandres, Evgeny, Le Sun and Lu Zhang, 2008, The New Issues Puzzle: Testing the Investment-Based Explanation, *Review of Financial Studies* 21, 2825–2855.

Mikhail, Micheal, Beverly Walther, and Richard Willies, 2007, When Security Analysts Talk, Who Listens? *The Accounting Review* 82, 1227–1253.

Novy-Marx, Robert, 2011, Operating Leverage, *Review of Finance* 15, 103–134.

Novy-Marx, Robert, 2013, The other side of value: The gross profitability premium, *Journal of Financial Economics* 104, 162–185.

Muro, Mark, Robert Maxim, and Jacob Whiton, 2019, Automation and Artificial Intelligence: How machines are affecting people and places, Brookings.

Ohlson, James A., 2008, Financial Ratios and the Probabilistic Prediction of Bankruptcy, *Journal of Accounting Research* 18, 109–131.

Ortiz-Molina, Hernan and Gordon Phillips, 2014, Real Asset Illiquidity and the Cost of Capital, *Journal of Financial and Quantitative Analysis* 49, 1–32.

Ray, Debraj and Dilip Mookherjee, 2020, Growth, Automation and the Long Run Share of Labor, Working paper.

Richardson, Scott A., Richard G. Sloan, Mark T. Soliman and Irem Tuna, 2005, Accrual reliability, earnings persistence and stock prices, *Journal of Accounting and Economics* 39, 437–485.

Ritter, Jay R., 1991, The Long-Run Performance of initial Public Offerings, *Journal of Finance* 46, 3–27.

Rosenberg, Barr, Kenneth Reid and Ronald Lanstein, 1985, Persuasive evidence of market inefficiency, *Journal of Portfolio Management* 11, 9–16.

Rossi, Alberto G. and Stephen P. Utkus, 2021, Who Benefits from Robo-advising? Evidence from Machine Learning, Working paper.

Sloan, Richard G., 1996, Do Stock Prices Fully Reflect Information in Accruals and Cash Flows about Future Earnings? *The Accounting Review* 71, 289–315.

Soliman, Mark T., 2008, The Use of DuPont Analysis by Market Participants, *The Accounting Review* 83, 823–853.

Tetlock, Paul C, 2007, Giving content to investor sentiment: The role of media in the stock market, *Journal of Finance* 62, 1139–1168.

Thomas, Jacob K. and Huai Zhang, 2002, Inventory Changes and Future Returns, *Review of Accounting Studies* 7, 163–187.

Titman, Sheridan, K. C. John Wei and Feixue Xie, 2004, Capital Investments and Stock Returns, *The Journal of Financial and Quantitative Analysis* 39, 677–700.

Webb, Michael, 2020, The Impact of Artificial Intelligence on the Labor Market, Working paper.

Xing, Yuhang, 2008, Interpreting the Value Effect Through the Q-Theory: An Empirical Investigation, *Review of Financial Studies* 21, 1767–1795.

Zheng, Xiang, 2021, How can Innovation Screening be Improved? A Machine Learning Analysis with Economic Consequences for Firm Performance, Working paper.

Zhu, Christina, 2019, Big Data as a Governance Mechanism, *The Review of Financial Studies* 32, 2021–2061.

Zou, Hui. and Trevor Hastie, 2005, Regularization and variable selection via the elastic net, *Journal of the Royal Statistical Society Series B.* 67, 301–320.

Figure 1: Man vs. Machine: The Performance of AI vs. Analysts

This figure plots the beat ratio, or the proportion of AI price forecasts that are more accurate than the corresponding analyst price forecasts in each year. The blue line in the middle plots the annual beat ratios, and the surrounding blue-dotted lines indicate the 95% confidence interval of the beat ratio. The red line gives the best linear approximation of the time-series trend in beat ratios.

31
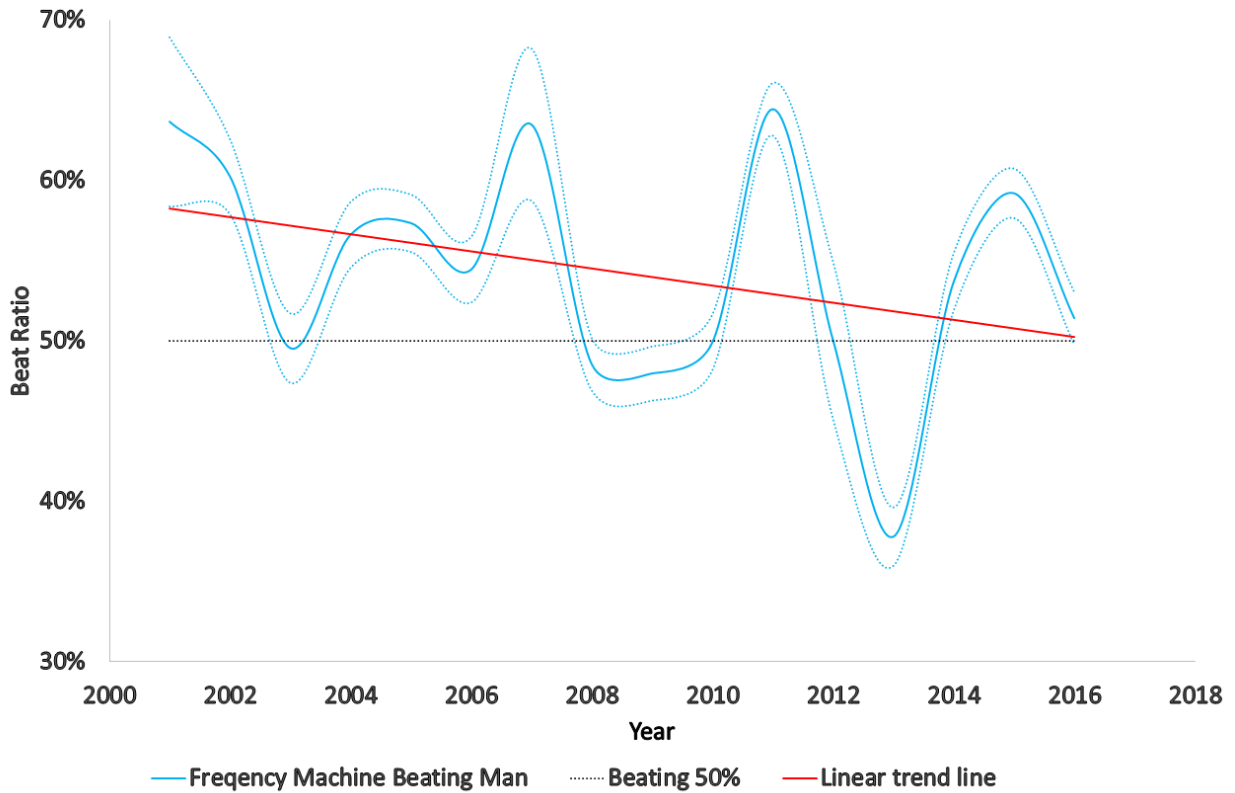
Figure 2: Man + Machine: The Performance of AI-assisted Analyst vs Analysts

This figure plots the proportion of AI-assisted Analyst recommendations that are more accurate than the Analyst recommendations alone on an annual basis. The blue line in the middle gives the annual AI-assisted Analyst beat ratios, the blue-dotted lines above and below are the 95% confidence interval of the beat ratio, and the red line gives the best linear approximation of the trend in beat ratios.

Figure 3: Man + Machine: The Performance of AI-assisted Analysts vs AI

This figure plots the proportion of AI-assisted analyst recommendations that are more accurate than the AI recommendations alone on an annual basis, or the "beat ratio". The blue line in the middle gives the annual AI-assisted analyst beat ratios, the blue-dotted lines above and below are the 95% confidence interval of the beat ratio, and the red line gives the best linear approximation of the trend in beat ratios.
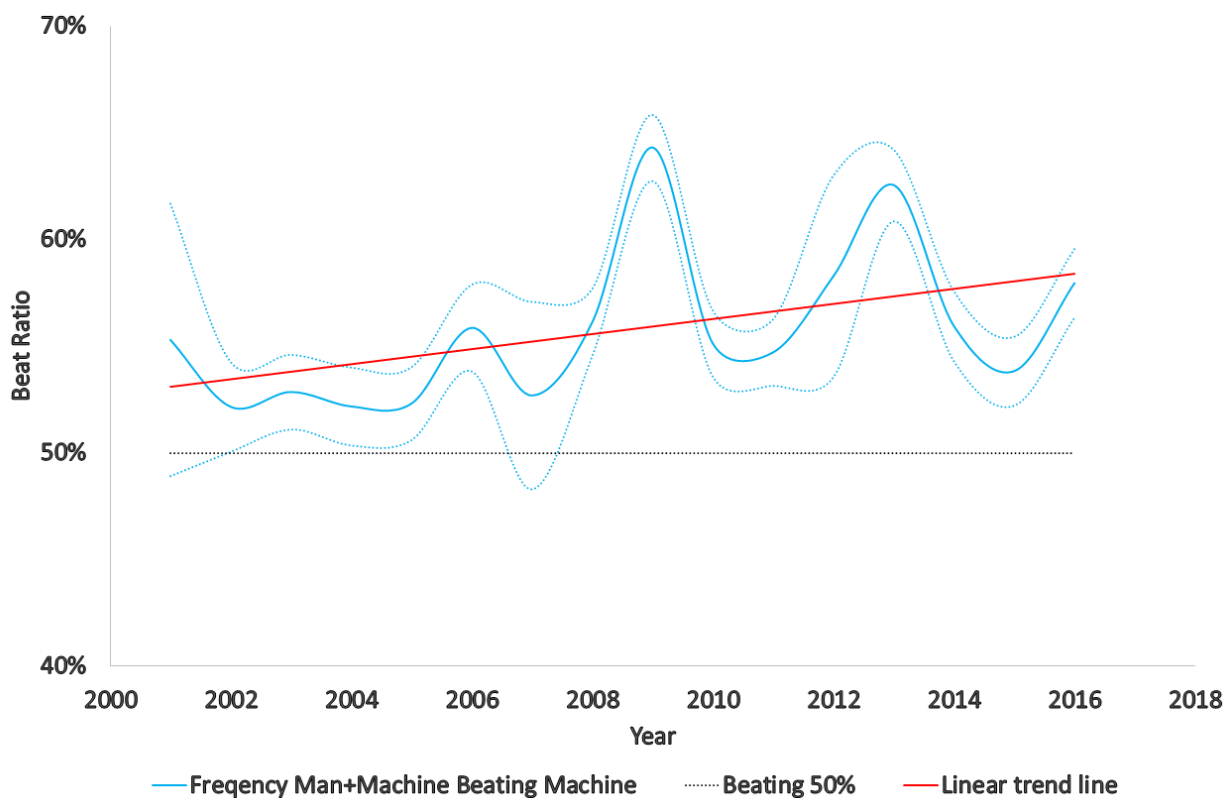
Table 1: Summary Statistics
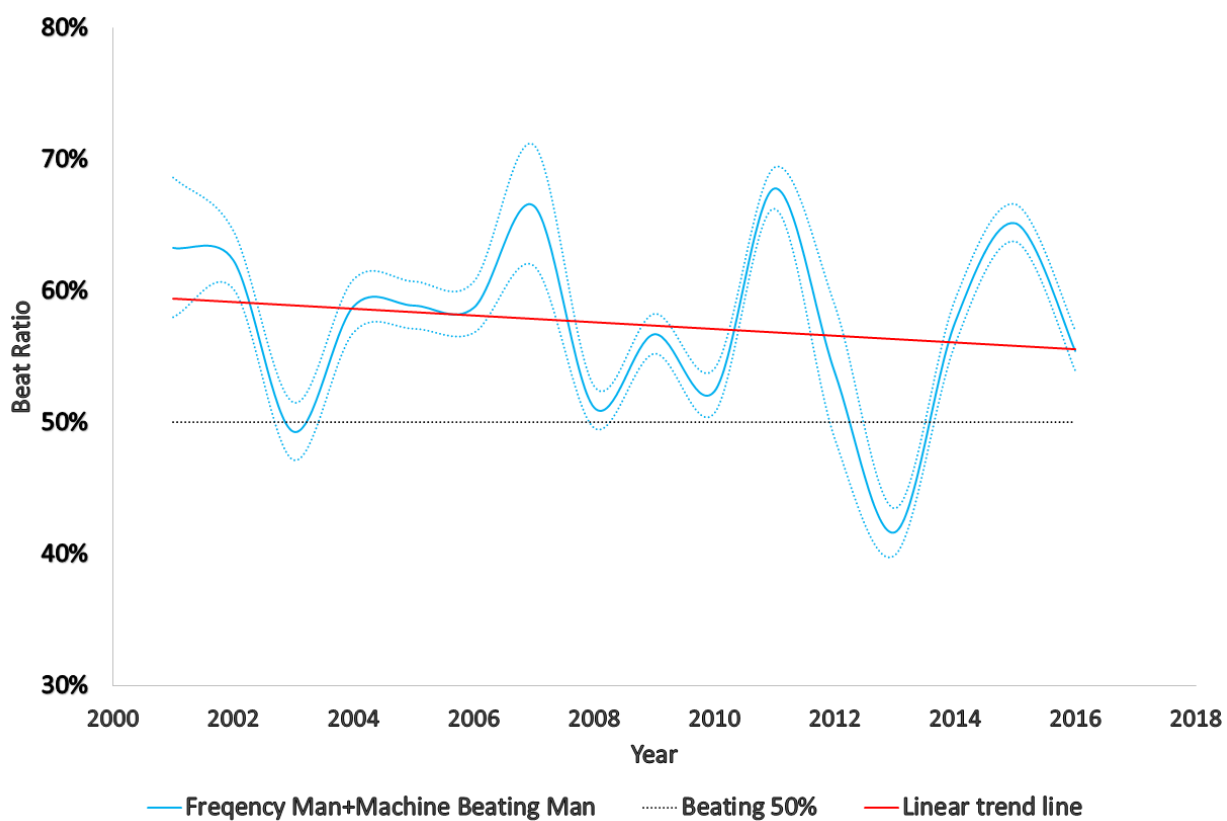
This table reports the summary statistics of key variables. The firm-level, industry-level and macroeconomic variables are defined in Section 2.4. The AI and satellite coverage variables include *AI Hiring*, *Alt Data Covered* and *Post*. *AI Hiring* is the ratio of the number of AI jobs to the total number of job postings. *Alt Data Covered* is one if the firm is covered by satellite, and zero otherwise. *Post* is equal to one in the following to cases and zero otherwise: (1) if a firm is covered by satellite and the date is after the starting cover date. (2) When a firm is not covered by satellite, we assume the post is 1 if the date is after 2014. The mean, median, standard deviation, 25 percentile, 75 percentile and number of observations are reported in the table.

Variables	Mean	Median	Std	P25	P75	N
Panel A. Firm-level, industry-level and macroeconomic variables						
Amihud Illiquidity	0.69	0.01	91.54	0.00	0.02	215,934
Stock Volatility	2.59	2.26	1.32	1.68	3.16	215,934
Log Market Cap	7.99	7.91	1.69	6.79	9.14	215,934
# 8K Reports	3.08	1.00	20.37	0.00	2.00	215,934
Intangible Assets	0.24	0.28	0.09	0.20	0.32	215,934
Forecast Horizon	190.06	188.00	106.18	88.00	287.00	215,934
# Analysts in Brokerage Firm	3.36	3.47	0.98	2.77	4.06	215,934
# Previous Predictions	31.80	17.00	41.91	7.00	40.00	215,934
# Institutional Owners	2.87	1.94	3.09	1.01	3.72	215,934
Industry Distress	0.25	0.00	0.43	0.00	1.00	215,934
Panel B. AI and satellite coverage variables						
AI Hiring	0.10	0.00	2.56	0.00	0.00	85,950
Alt Data Covered	0.04	0.00	0.20	0.00	0.00	85,950
Post	0.24	0.00	0.43	0.00	0.00	85,950

Table 2: Comparison of Machine Learning Models

We apply a variety of machine learning models to predict stock prices (log). We use the past five years as a rolling window to train our models. We compare the AI predicted stock price and the analyst predicted ones with the end-of-year actual price. If the absolute difference between AI predicted values and the actual price is smaller than that between analyst values and actual price, we define this case as AI beats analysts. This table presents the beat ratio (the number of times AI beats analysts/total number of predictions).

Method	Beat ratio	Method description
OLS Regression	19.3%	The traditional linear method that minimizes the in-sample sum of squares of prediction errors
Elastic-Net Regression	23.3%	A quasi-linear ML method that combines features of LASSO and Ridge regressions to allow dimension reduction
Support Vector Regression	27.2%	A quasi-linear ML method that focuses on marginal points with higher estimation errors to allow dimension reduction
LSTM Neural Networks	50.1%	A deep learning model that uses recurrent neural networks to incorporate sequential or time-series information in the predictors
Random Forest	51.7%	A decision-tree based ML model that combines decision trees generated with random sets of predictors to capture complex interactions
Gradient Boosting	52.0%	A decision-tree based ML model that combines a series of decision trees and sequentially approximates the predicted variable to capture complex interactions
Our Final AI Model	53.7%	An ensemble ML model that combines the LSTM Neural Networks, Random Forest and Gradient Boosting methods

Table 3: Persistence of Performance of AI Analyst

Each year, analysts are sorted by mean squared prediction errors based on past one, two, and up-to-five years. If the mean square error over last year is among the smallest/largest 50%, the analyst is the top/bottom in 1 year. If the mean square error over last two (three, four and five) years is among the smallest/largest 50%, the analyst is the top/bottom in 2 (3, 4, and 5) years. The panel A reports the AI beat ratio with respect to analysts who are among the top or bottom 50% over the last year, two years until five years. On the other hand, if the mean square errors of each of the past two (three, four and five) years are among the smallest 50%, the analyst is persistently top over 2 (3, 4, and 5) years. The panel B reports the AI beat ratio with respect to analysts who are persistently among the top or bottom 50% over the last year, two years until five years.

Panel A: Beat ratio with respect to analysts who are among the top or bottom 50%

	1 year	2 years	3 years	4 years	5 years
Analyst Top	51.02%	51.00%	50.82%	50.82%	50.92%
Analyst Bottom	55.66%	55.43%	55.60%	55.59%	55.47%

Panel B: Beat ratio with respect to analysts who are persistently among the top or bottom 50%

	1 years	2 years	3 years	4 years	5 years
Analyst Persistent Top	51.02%	50.45%	49.83%	48.91%	49.24%
Analyst Persistent Bottom	55.66%	56.41%	56.45%	55.84%	55.82%

Table 4: Portfolio Performance following Machine vs Man Recommendations

For each month/every six months, all predictions of AI and analysts are collected over past 30, 60, 90 and 360 days. When comparing AI predictions with analysts', if AI prediction is higher than analyst prediction, it is a buy signal. Otherwise, it is a sell signal. Among all signals over past 30, 60, 90 and 360 days, if there are more buy signals, the stock is longed. Otherwise, the stock is shorted. The stocks are rebalanced every month/semi-annually. The monthly percentage return of the long-short, long-leg (stocks only with a buy sign) and short-leg portfolios (stocks only with a short sign) as well as alphas generated from FF3, FFC4, FF5 and FF6 models are presented. The OLS standard error is used to construct t-stats. The t-stats are reported in parentheses. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels (two-tailed), respectively. Panel A/B reports those results for monthly/semi-annually rebalanced portfolios. The number of stocks in each portfolio is reported at the bottom of each panel.

Panel A. Portfolio returns with monthly rebalancing

		Portfolio information horizon			
Long-Short		30 days	60 days	90 days	360 days
	Ret	0.91***	0.90***	0.91***	0.85***
		(2.94)	(3.25)	(3.34)	(3.78)
	FF3	0.85***	0.84***	0.86***	0.82***
		(2.82)	(3.16)	(3.27)	(3.89)
	FFC4	0.92***	0.90***	0.91***	0.84***
		(3.05)	(3.40)	(3.50)	(4.00)
	FF5	0.69**	0.68**	0.70***	0.63***
		(2.30)	(2.56)	(2.67)	(3.03)
	FF6	0.74**	0.72***	0.74***	0.65***
		(2.48)	(2.75)	(2.86)	(3.16)
Long-Leg		30 days	60 days	90 days	360 days
	Ret	1.56***	1.51***	1.53***	1.50***
		(3.15)	(3.13)	(3.21)	(3.27)
	FF3	0.87***	0.82***	0.85***	0.83***
		(4.85)	(4.96)	(5.27)	(5.66)
	FFC4	1.03***	0.97***	1.00***	0.95***
		(6.58)	(6.74)	(7.11)	(7.19)
	FF5	0.80***	0.74***	0.77***	0.74***
		(4.50)	(4.54)	(4.84)	(5.14)
	FF6	0.89***	0.83***	0.85***	0.81***
		(5.89)	(6.02)	(6.40)	(6.46)
Short-Leg		30 days	60 days	90 days	360 days
	Ret	0.64	0.61	0.62	0.65
		(1.15)	(1.13)	(1.15)	(1.26)
	FF3	0.02	-0.02	-0.01	0.01
		(0.07)	(-0.08)	(-0.04)	(0.06)
	FFC4	0.12	0.07	0.08	0.11
		(0.42)	(0.30)	(0.34)	(0.56)
	FF5	0.11	0.07	0.07	0.11
		(0.38)	(0.26)	(0.28)	(0.58)
	FF6	0.16	0.11	0.12	0.16
		(0.56)	(0.46)	(0.48)	(0.85)

Panel B. Portfolio returns with semi-annual rebalancing

Long-Short		Portfolio information horizon			
		30 days	60 days	90 days	360 days
	Ret	0.92***	0.93***	0.93***	0.76***
		(2.96)	(3.04)	(3.06)	(3.17)
	FF3	0.89***	0.89***	0.90***	0.73***
		(3.02)	(3.08)	(3.12)	(3.26)
	FFC4	0.94***	0.94***	0.94***	0.73***
		(3.20)	(3.25)	(3.26)	(3.27)
	FF5	0.62**	0.64**	0.64**	0.52**
		(2.16)	(2.25)	(2.27)	(2.40)
	FF6	0.67**	0.68**	0.68**	0.54**
		(2.35)	(2.43)	(2.44)	(2.48)
Long-Leg		30 days	60 days	90 days	360 days
	Ret	1.58***	1.60***	1.60***	1.54***
		(3.31)	(3.42)	(3.44)	(3.39)
	FF3	0.86***	0.88***	0.88***	0.82***
		(5.01)	(5.58)	(5.76)	(5.81)
	FFC4	1.00***	1.00***	1.00***	0.92***
		(6.44)	(7.08)	(7.31)	(7.17)
	FF5	0.72***	0.74***	0.74***	0.68***
		(4.33)	(4.87)	(5.06)	(5.04)
	FF6	0.81***	0.82***	0.83***	0.75***
		(5.61)	(6.27)	(6.52)	(6.36)
Short-Leg		30 days	60 days	90 days	360 days
	Ret	0.66	0.67	0.67	0.78
		(1.15)	(1.19)	(1.19)	(1.50)
	FF3	-0.03	-0.02	-0.02	0.09
		(-0.11)	(-0.07)	(-0.08)	(0.43)
	FFC4	0.06	0.06	0.06	0.20
		(0.21)	(0.24)	(0.25)	(0.95)
	FF5	0.10	0.10	0.10	0.16
		(0.37)	(0.38)	(0.39)	(0.75)
	FF6	0.14	0.14	0.14	0.21
		(0.53)	(0.53)	(0.55)	(1.05)

Table 5: Man vs. Machine: The Relative Advantage of Analyst vs AI

Panel A presents the coefficients and t-stats of regressing the analyst-beat-AI indicator on the firm-level, industry-level and macroeconomic variables presented in Table 1. *Analyst beats AI* indicator is 1/0 if analysts beats/does not beat AI. Panel B presents the coefficients and t-stats of regressing *Forecast Error Difference Analyst vs. AI* on the same variables. *Forecast Error Difference Analyst vs. AI* is defined as the difference between absolute prediction error between AI and Analysts, divided by the maximum value of these two prediction errors. The number is positive if analyst has smaller absolute error, i.e. analyst beat AI. To calculate t-statistics, standard errors are clustered at the firm level. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels (two-tailed), respectively.

Panel A. Man vs. machine: when analyst beats AI

Variables	Dependent variable Analyst beats AI			
	(1)	(2)	(3)	(4)
Amihud Illiquidity	0.224***	0.224***	0.358***	0.346***
	(6.40)	(6.76)	(10.33)	(10.01)
Stock Volatility	-0.449	-1.548***	0.952***	0.530
	(-1.49)	(-4.34)	(2.86)	(1.29)
Log Market Cap	-0.001	-0.003	-0.039***	-0.030***
	(-0.26)	(-0.88)	(-5.26)	(-4.03)
# 8K Reports	-0.273***	-0.385***	-0.217*	-0.343***
	(-2.84)	(-3.56)	(-1.86)	(-2.59)
Intangible Assets	0.199***	0.187***	0.245**	0.209**
	(3.96)	(3.77)	(2.32)	(1.96)
Forecast Horizon	-0.340***	-0.316***	-0.354***	-0.327***
	(-11.62)	(-11.09)	(-13.93)	(-13.18)
# Analysts in Brokerage Firm	0.007***	0.008***	0.005***	0.005***
	(3.64)	(3.97)	(3.04)	(3.39)
# Previous Predictions	-0.150	-0.020	-0.125	0.020
	(-0.97)	(-0.13)	(-0.99)	(0.17)
# Institutional Owners	0.145	0.106	1.037	1.015
	(0.79)	(0.59)	(1.27)	(1.28)
Industry Distress	0.041***	0.025**	0.031***	0.026**
	(5.31)	(2.20)	(3.96)	(2.15)
Time Trend	0.005***		0.008***	
	(5.89)		(5.14)	
Constant	0.418***	0.466***	0.636***	0.492***
	(12.53)	(6.53)	(10.04)	(5.17)
Year fixed effects	No	Yes	No	Yes
Firm fixed effects	No	No	Yes	Yes
Observations	215,934	215,934	215,934	215,934
Adjusted R-squared	0.01	0.02	0.120	0.128

Panel B. Man vs. machine: Relative prediction error

| Variables | Dependent variable | | | |
| | Forecast Error Difference Analyst vs. AI | | | |
	(1)	(2)	(3)	(4)
Amihud Illiquidity	0.347***	0.347***	0.419***	0.375***
	(8.16)	(8.79)	(8.93)	(8.11)
Stock Volatility	-1.216***	-2.891***	0.630	-0.169
	(-3.26)	(-6.41)	(1.63)	(-0.36)
Log Market Cap	0.001	-0.002	-0.045***	-0.036***
	(0.29)	(-0.43)	(-5.37)	(-4.18)
# 8K Reports	-0.241*	-0.335***	-0.148	-0.277*
	(-1.85)	(-2.60)	(-1.01)	(-1.85)
Intangible Assets	0.243***	0.225***	0.305**	0.265**
	(3.68)	(3.47)	(2.32)	(2.02)
Forecast Horizon	-0.462***	-0.428***	-0.495***	-0.459***
	(-13.15)	(-12.52)	(-15.62)	(-14.84)
# Analysts in Brokerage Firm	0.009***	0.010***	0.006***	0.007***
	(3.73)	(4.21)	(3.52)	(3.95)
# Previous Predictions	-0.131	0.029	-0.185	-0.020
	(-0.66)	(0.15)	(-1.09)	(-0.12)
# Institutional Owners	0.095	0.039	0.396	0.361
	(0.35)	(0.14)	(0.43)	(0.40)
Industry Distress	0.051***	0.024*	0.036***	0.022
	(5.25)	(1.69)	(3.69)	(1.53)
Time Trend	0.005***		0.010***	
	(4.37)		(5.48)	
Constant	-0.068*	0.011	0.201***	0.060
	(-1.72)	(0.14)	(2.71)	(0.52)
Year fixed effects	No	Yes	No	Yes
Firm fixed effects	No	No	Yes	Yes
Observations	213,164	213,164	213,164	213,164
Adjusted R-squared	0.01	0.02	0.147	0.155

Table 6: Man + Machine: The Incremental Value of Analyst and AI

Panel A presents the coefficients and tT-stats of regressing *Analyst + AI Beats AI* or *Analyst + AI Beats Analyst* indicator on the firm-level, industry-level and macroeconomic variables presented in Table 1. *Analyst + AI Beats AI/Analyst + AI Beats Analyst* indicator is 1/0 if analysts+AI beats/does not beat AI/analyst. Panel B presents the coefficients and t-stats of regressing *Incremental Value AI/Analyst* on the same variables. *Incremental Value AI/Analyst* is defined as the absolute prediction error of AI/analyst minus absolute prediction error of analysts+AI, divided by the maximum value of these two prediction errors. The number is positive if analysts+AI has smaller absolute error than AI/analyst. To calculate the t-stats, standard errors are clustered at the firm level. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels (two-tailed), respectively.

Panel A. Analyst + AI beats analyst/AI.

| | Dependent variable | | | |
| | *Analyst + AI Beats AI* | | *Analyst + AI Beats Analyst* | |
Variables	(1)	(2)	(3)	(4)
Amihud Illiquidity	0.136***	0.283***	-0.254***	-0.347***
	(5.94)	(8.66)	(-5.93)	(-9.70)
Stock Volatility	-0.403	0.004	1.720***	-0.151
	(-1.33)	(0.01)	(5.02)	(-0.37)
Log Market Cap	-0.002	-0.005	0.002	0.034***
	(-0.51)	(-0.68)	(0.45)	(4.50)
# 8K Reports	-0.500**	-0.486	0.192*	0.175*
	(-2.23)	(-1.62)	(1.88)	(1.76)
Intangible Assets	0.144***	0.279**	-0.138***	-0.147
	(3.82)	(2.28)	(-2.72)	(-1.34)
Forecast Horizon	-0.181***	-0.188***	0.267***	0.278***
	(-6.15)	(-6.01)	(10.63)	(10.33)
# Analysts in Brokerage Firm	-0.002	-0.002	-0.009***	-0.006***
	(-0.97)	(-1.12)	(-4.98)	(-3.96)
# Previous Predictions	0.168	0.172	0.256**	0.238
	(1.25)	(1.10)	(2.22)	(1.63)
# Institutional Owners	0.167	1.006**	-0.037	-0.811
	(0.93)	(2.12)	(-0.21)	(-1.23)
Industry Distress	0.028**	0.031**	-0.029***	-0.030***
	(2.45)	(2.51)	(-2.61)	(-2.63)
Constant	0.629***	0.482***	0.615***	0.502***
	(5.96)	(3.61)	(7.69)	(4.60)
Year fixed effects	Yes	Yes	Yes	Yes
Firm fixed effects	No	Yes	No	Yes
Observations	212,960	212,960	215,934	215,934
Adjusted R-squared	0.01	0.102	0.02	0.118

Panel B. Incremental value of analysts/AI.

Variables	Incremental Value Analyst		Incremental Value AI	
	(1)	(2)	(3)	(4)
Amihud Illiquidity	0.071***	0.085***	-0.342***	-0.350***
	(4.63)	(3.70)	(-7.31)	(-6.44)
Stock Volatility	0.054	0.481	3.057***	0.789*
	(0.21)	(1.45)	(7.25)	(1.74)
Log Market Cap	-0.003	-0.003	-0.000	0.038***
	(-1.04)	(-0.49)	(-0.03)	(4.47)
# 8K Reports	-0.281	-0.320	0.065	-0.004
	(-1.33)	(-1.22)	(0.64)	(-0.04)
Intangible Assets	0.067**	0.243***	-0.169***	-0.109
	(2.22)	(2.68)	(-2.70)	(-0.83)
Forecast Horizon	-0.132***	-0.134***	0.361***	0.392***
	(-4.62)	(-4.34)	(12.21)	(12.08)
# Analysts in Brokerage Firm	-0.002	-0.001	-0.011***	-0.008***
	(-1.55)	(-1.15)	(-5.22)	(-4.53)
# Previous Predictions	0.257*	0.260	0.236	0.303
	(1.66)	(1.49)	(1.48)	(1.60)
# Institutional Owners	0.103	0.239	0.009	-0.153
	(0.60)	(0.54)	(0.04)	(-0.22)
Industry Distress	0.022**	0.022**	-0.016	-0.015
	(2.54)	(2.41)	(-1.23)	(-1.09)
Constant	0.057	-0.094	0.036	-0.142
	(1.05)	(-1.19)	(0.40)	(-1.12)
Year fixed effects	Yes	Yes	Yes	Yes
Firm fixed effects	No	Yes	No	Yes
Observations	212,960	212,960	215,934	215,934
Adjusted R-squared	0.01	0.062	0.02	0.140

Table 7: Man + Machine Event Study: Alternative Data Covered Firms

This table presents the coefficients and *t*-stats of regressing *Analyst Beats AI* indicator on brokerage *AI Hiring*, *Alt Data Covered*, *Post*, and the interactions among these variables. *Analyst Beats AI* indicator is 1/0 if analysts beats/does not beat AI. *AI Hiring* is the ratio of the number of AI jobs to the total number of job postings. *Alt Data Covered* is one if the firm is covered by satellite, and zero otherwise. *Post* is equal to one in two case: (1) if the firm is covered by satellite and the date is after the starting cover date. (2) if the firm is not covered by satellite, we assume the post is 1 if the date is after 2014. The control variables are the firm-level, industry-level and macroeconomic variables presented in Table 1. Standard errors are clustered at the firm level. ***, **, * denote statistical significance at the 0.01, 0.05, and 0.10 levels (two-tailed), respectively.

Variables	Dependent variable: *Analyst Beats AI*			
	(1)	(2)	(3)	(4)
Alt Data Covered × Post × AI Hiring			6.674***	4.807**
			(3.27)	(2.12)
AI Hiring	0.141**	0.192**	0.206**	0.243**
	(2.00)	(2.52)	(2.56)	(2.44)
Alt Data Covered			-0.011	
			(-0.18)	
Post			0.109*	0.004
			(1.69)	(0.05)
Alt Data Covered × Post			0.057	-0.028
			(0.74)	(-0.31)
Alt Data Covered × AI Hiring			-2.880**	-1.429
			(-2.02)	(-0.89)
Post × AI Hiring			-0.322	-0.282
			(-1.46)	(-1.20)
Constant	0.354***	0.527***	0.360***	0.526***
	(6.30)	(4.20)	(6.56)	(4.19)
Controls	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Firm fixed effects	No	Yes	No	Yes
Observations	85,950	85,950	85,950	85,950
Adjusted R-squared	0.02	0.166	0.02	0.167

Appendix A. List of Variables

Table A1: List of All Variables Used in AI Algorithms

All Variables (and the definition/source) used in the machine learning algorithms are provided.

Firm Characterisitics	Definition and/or Source
Momentum	Past 12-month return, Jegadeesh and Titman (1993)
Composite Equity Issuance	Daniel and Titman (2006)
Gross Profits-to-Assets	Novy-Marx (2013)
Asset Growth	Cooper, Gulen, and Schill (2008)
Investment-to-Assets	Titman, Wei, and Xie (2004) and Xing (2008)
Net Operating Assets	Hirshleifer, Hou, Teoh, and Zhang (2004)
Accruals	Sloan (1996)
Net Stock Issues	Ritter (1991) and Loughran and Ritter (1995)
Failure Probability	Campbell, Hilscher, and Szilagyi (2008)
O-Score	Ohlson (1980)
Return on Assets	Fama and French (2006) and Chen, Novy-Marx, and Zhang (2011)
Book-to-Market Equity	Rosenberg, Reid and Lanstein (1985)
Debt-to-Market	Bhandari (1988)
Earnings-to-Price	Basu (1983)
Cash Flow-to-Price	Lakonishok, Shleifer, and Vishny (1994)
Payout Yield	Boudoukh, Michaely, Richardson, and Roberts (2007)
Five-year Sales Growth Rank	Lakonishok, Shleifer, and Vishny (1994)
Enterprise Multiple	Loughran and Wellman (2011)
Sales-to-Price	Barbee, Mukherji, and Raines (1996)
Abnormal Corporate Investment	Titman, Wei, and Xie (2004)
Investment-to-Assets	Cooper, Gulen, and Schill (2008)
Changes in PPE and Inventory/Assets	Lyandres, Sun, and Zhang (2008)
Investment Growth	Xing (2008)
Inventory Changes	Thomas and Zhang (2002)
Operating Accruals	Sloan (1996)
Total Accruals	Richardson, Sloan, Soliman, and Tuna (2005)
Net External Finance	Bradshaw, Richardson, and Sloan (2006)
Return on Net Operating Assets	Soliman (2008)
Profit Margin	Soliman (2008)
Asset Turnover	Soliman (2008)
Operating Profits-to-Equity	Fama and French (2015)
Book Leverage	Fama and French (1992)
Advertising Expense-to-Market	Chan, Lakonishok, and Sougiannis (2001)
R&D-to-Market	Chan, Lakonishok, and Sougiannis (2001)
Operating Leverage	Novy-Marx (2011)
Financial Constraints	Kaplan-Zingales index, Lamont, Polk, and Saa-Requejo (2001)
Asset Liquidity	Scaled by book assets, Ortiz-Molina and Phillips (2014)
Asset Liquidity	Scaled by market assets, Ortiz-Molina and Phillips (2014)
IBES Actual Earning	IBES actual earning 4 quarter before scaled by adjusted price
Number of Institutional Owners	Number of 13F institutional investors that own the stock
Ownership Concentration	Herfindahl-Hirschman Index
Total Institutional Ownership	Percent of shares outstanding owned by 13F investors

Industry Variables	Definition and/or Source
Competition Measure from 10-K	Li, Lundholm, and Minnis (2013)
Fluidity	Product market Fluidity, Hoberg, Phillips and Prabhala (2014)
48 Industry Dummy	Dummy variables that indicate Fama-French 48 industries
Industry Size	Industry Size within past 3, 6, 9 ,12, 24 and 36 months
Industry Earning	Industry earning within past 3, 6, 9 12, 24 and 36 months
Macro Variables	Definition and/or Source
IP	Industrial Production Index
CPI	Consumer Price Index
Oil price	Crude Oil Price
Tbill3	3-month Treasury Bill
TBond10	10-Year Treasury Constant Maturity Rate
Credit Spread	Baa-AAA yield spread
Textual Variables	Definition and/or Source
Neg 10KQ	Percentage of negative words from 10K/10Q
NegPos 10KQ	Percentage of negative minus positive words from 10K/10Q
Neg 8k	Percentage of negative words from 8K
NegPos 8K	Percentage of negative minus positive from 8K
Neg Other	Percentage of negative words from other reports
NegPos Other	Percentage of negative minus positive from other reports
ML-based Sentiment	ML-based negative tones minus ML-based positive tones scaled by the length of SEC filings, Cao, Kim, Wang, and Xiao (2020)
ML-based Neg Sentiment	ML-based negative tones scaled by the length of SEC filings

Appendix B. Details of the Machine Learning Models

In this section, we briefly describe the basic structure and strengths of machine learning models considered in our paper. Interested readers are referred to representative references for more details, such as Hastie, Tibshirani and Friedman (2009) and Goodfellow, Bengio, and Courville (2016).

B.1. Quasi-linear Models

Quasi-linear machine learning models generalize linear regressions and classification models, and are more flexible and can accommodate a larger number of variables than the traditional linear regressions, by their built-in dimension-reduction capabilities. Quasi-linear models are typically efficient in model training because they are typically associated with fast algorithms, such as linear and quadratic programming techniques.

B.1.1. Elastic-Net

The Elastic-Net model is a generalization of the OLS linear regression model. When there is a large number of predictors, the OLS tends to have good in-sample performance (*small bias* in the terms of machine learning) and bad out-of-sample performance (*large variation* in the terms of machine learning). Furthermore, the OLS can generate significant loadings on a large number of independent variables, making the interpretation of the model difficult. One class of models, the shrinkage models, generalize the OLS by imposing a penalty on the number and size of non-zero coefficients in the estimation, effectively limiting the model to focus on a subset of the independent variables and achieving dimension reduction.

The Elastic-Net model, introduced by Zou and Hastie (2005), is a shrinkage model in which the penalty function is a linear combination of L^1 and L^2 norms of the coefficients. In particular, the Elastic-Net model minimizes the following objective function,

$$\min_{\beta} \sum_{i=1}^{N} (y_i - \beta_0 - \sum_{j=1}^{p} x_{ij}\beta_j)^2 + \lambda_1 \sum_{j=1}^{p} |\beta_j| + \lambda_2 \sum_{j=1}^{p} |\beta_j|^2. \tag{A1}$$

The Elastic-Net model is also a generalization of the well-known LASSO and Ridge regression models. When the hyperparameter $\lambda_2 = 0$ in (A1), we recover the LASSO model. When $\lambda_1 = 0$, we recover the Ridge model. In general, the LASSO model tends to select a few strong predictors while setting the coefficients of other predictors to essentially zero, but can make random choices among several strong and correlated variables. The Ridge model usually

includes more predictors and shrink the cofficients of correlated variables together. The Elastic-Net model strikes a balance between these charateristics, allowing both a selection of strong features and the averaging of correlated features.

B.1.2. Support Vector Regression

The support vector regression (SVR) is motivated by support vector machines (SVM) for classification problems. Consider a classification problem with 2 classes and n predictors, i.e., each observation belongs to one of two classes and is a point in the n-dimensional space. Given a training sample, i.e., a set of labeled n-dimensional points, a linear classifier is equivalent to a $(n-1)$-dimensional hyperplane that separates the two classes of points in the n-dimesional space. The linear SVM searches for the *maximum separating* $(n-1)$-dimensional hyperplane such that it maximizes the distance from the hyperplane to the closest data points. To deal with the case that separating hyperplanes may not exist, the SVM also tolerates misclassified points within some bounds. The SVM thus focuses on points close to the separating hyperplane, i.e., observations that are on the boundary of the two classes. In fact, the results the SVM do not depend on observations far away from the boundary. One key advantage of the SVM is that it performs well when there is a large number of features relative to the sample size, e.g., in the case of texual and image analysis. The SVM is regarded as one of the best out-of-the-box machine learning algorithms and has been widely applied in the classification of text, image, hand-writing, and proteins.

The support vector regression optimizes the following objective function.

$$\min_{\beta} \sum_{i=1}^{N} V(y_i - \beta_0 - x_i^T \beta) + \lambda \sum_{j=1}^{p} |\beta_j|^2. \tag{A2}$$

where the cost function $V(\cdot)$ is given by

$$V(z) = \begin{cases} 0, & \text{if } |z| < \epsilon, \\ |z| - \epsilon, & \text{otherwise.} \end{cases} \tag{A3}$$

The cost function is insensitive to the signs/sizes of errors if the error size is less than ϵ, i.e., it is more sensitive to points where the estimation error is larger. These marginal points, in turn, are instrumental in determining the estimated coefficients. The benefit of the support vector regression is that it allows efficient dimension reduction even with a very large number of features. However, it may not perform well if the underlying pattern is far away from

47

being linear.

B.2. Decision-Tree Based Models

The quasi-linear models considered above may not work well if there are nonlinear relationships among the predictive variables. In this section, we discuss a class of versatile nonlinear models – decision trees and derived models.

B.2.1. Decision Trees

Decision trees are modeled after human decisions. A decision tree is a series of binary decisions based on cutoffs of independent variables at each branching point. The tree thus will divide the rectangular feature space into smaller rectangular blocks. The decision tree regression then use the sample mean of the dependent variable in each block as the prediction for any point in the block.

Decision trees have the benefit of being easily interpretable because it is modeled after human decisions (similar to a step-by-step instructions) and can also be displayed graphically (as binary trees). Trees are also a flexible non-linear model that can model a variety of nonlinear patterns given the large degree of freedom in specifying the sequences of branching rules.

However, trees do not have a high level of accuracy by themselves because of the restrictive form of the binary branching process, which forces the sample to be split into rectangular regions and may not approximate the real underlying patterns (whether linear or nonlinear) well. Trees are also non-robust. In addition, a small change in the data can lead to large changes in the structure of the estimated tree because the tree structure is discrete, not continuous. Several methods, including random forest and gradient boosting, use trees as basic building blocks to form ensemble predictors and achieve superior performance.

B.2.2. Random Forest

A random forest (introduced by Breiman, 2001) proceeds in the following way. First, it involves drawing a bootstrapped sample (drawing with repetition) from the original sample. Second, on the bootstrapped sample, one builds a decision tree, selecting a splitting predictor among only a random m features of the total p predictors. Third, one repeats the above two steps to build a number of decision trees, and form the ensemble predictor by taking the

mean predictor of all the trees.

Random forests perform better than simple trees for several reasons. First, through aggregating predictions over bootstrapped samples, it reduces the variance and non-robustness of single trees. Second, the random feature selection in the second step above ensures that the estimated trees are not too correlated, avoiding relying only on a few prominent features and further reducing the variance of the model.

B.2.3. Gradient Boost

Boosting also combines a number of weak models to generate a stronger model. In boosting of trees, a number of trees are constructed sequentially, i.e., each tree is constructed using information based on the previously constructed trees. In gradient boosting, each decision tree is fit to the residuals of the model, not to the outcome. Once a new tree is obtained, it is added to the predictive function to update it, usually with a learning weight multiplied to the tree predictor to adjust the rate of learning new information. Then new residuals are obtained from the updated predictive function and the process is repeated for a number of times to obtain the final ensemble predictor. Because boosting models aggregate results of decision trees sequentially, each component tree does not need to be very precise and can be simple, i.e., having a low depth.

In a sense, gradient boosting is similar to the Newton's gradient algorithm in optimization. It approximates the true underlying function sequentially by improving on the predicted residuals/errors gradually. This allows the final predictive function to have a much richer and more flexible structure and thus much better performance than single decision trees. It also reduces the non-robustness of single trees through using an ensemble of trees. For these reasons, gradient boosting is one of the best off-the-shelf machine learning methods.

B.3. Deep Learning Model: Long-Short Term Memory Neural Networks

The neural networks models, initial motivated by the neuron structures in the brains of humans and animals, blossomed after breakthroughs in algorithms and computing power (LeCun, Bengio and Hinton, 2015). Neural networks models, also called deep learning models, have become some of the most powerful models and achieved near- or super-human capabilities in a wide variety of applications, such as natural language processing, speech recognition, computer vision, game playing, and autonomous driving.

There are many different architechtures of neural networks, such as the simplest Feed-

forward Neural Networks for straightforward classification tasks, the Convoluntional Neural Networks for image and pattern recognition, and Recurrent Neural Networks (RNN) that can process sequential data such as speech and text. Long-Short Term Memory (LSTM) Neural Networks are a special type of RNN that is the key to the many successes of RNN, including speech recognition, language modeling, and translation.

In a neural network, there are nodes (neurons) that are connected to each other. There are three types of nodes: input nodes that are used to receive data; output nodes that produce desired outcomes or predictions; and intermediate nodes that process the data from input nodes and convert them to outputs. The connections of the nodes determine the structure of the neural network and its features. RNNs are neural networks with loops, or nodes that are connected to themselves.

LSTM networks are introduced by Hochreiter and Schmidhuber (1997) to solve the problem that standard RNNs have trouble retaining "memory" of the much earlier parts of sequential input data, when processing the later parts of the data. Since sequential data may have long-term dependencies, i.e., parts far away in the sequence may be related, it is important to have "long-term memory" to handle them. LSTM networks have a sequence of nodes that are specifically designed to retain long-term information and update it continuously with new information in a flexible way. As a result, LSTM can capture both short-term and long-term relations in sequential or time-series data very well, suggesting its potential applications in financial economics given the abundance of time-series financial data.