AI Intensity and Stock Returns: Do Companies with more AI job posting have higher returns?

Alisa Yang

28751949

This paper introduces a new dataset on Artificial Intelligence (AI) skill intensity of companies in the Dow Jones and S&P 500 obtained from online job boards. This AI-intensity is then used to relate to ex-post returns for the years 2015-2020. I find that most of the excess returns of the portfolio are explained by Fama and French 3-factor model. I discuss the drawbacks and benefit from using job boards as a novel information source.

1 Introduction

Artificial Intelligence (AI) has been rising in popularity in recent years, a simple way to visualize this is by using Google Ngram[1] to search "Artificial Intelligence" (Google, 2021). We can see that the popularity of this word has gone up and down throughout the years (See Appendix A Figure 1), but what is AI exactly? Simply put, AI is a field in computer science set "on designing machines that can mimic human behavior" (Khanam, Tanweer and Khalid, 2020). There have been some famous AI programs such as AlphaGo. It played and even won Go matches against world champions[2] (Mozur, 2017), but these intelligent programs are not just limited to table games. AI has permeated into many aspects of our daily lives. Some other examples closer home are things like self-driving cars, virtual personal assistants (e.g. Siri, Cortana, etc.) and even Google's own search algorithms (Khanam, Tanweer and Khalid, 2020).

AI has become ever so prevalent, so it is not surprising that there has been an increased interest on the topic even in finance literature. There have been some recent papers that show very promising results of using AI and ML on empirical asset pricing. Gu, Kelly, and Xiu (2020) use different types of machine learning (e.g. neural networks, decision trees, etc.) to form portfolios. They find that there are "large economic gains to investors using machine learning forecasts, in some cases doubling the performance of leading regression-based strategies from the literature."

We also have other authors that look at AI from a different but equally interesting angle. They try to quantify the impact of AI in the labour market, such as Grennan and Roni (2019) that describe the role of AI in the high-skilled labour market. They find that "AI serves as (a) direct substitute for analysts' work but also as a complement." This is the angle I will be taking on this paper, so rather than using AI to create portfolios, I will create portfolios based on how much AI companies use.

Current literature in AI and its effect on the labour market is still scarce but there have been some interesting papers in the recent years. Babina et al. (2020) "provide(s) a comprehensive picture of the use of AI technologies and their impact among US firms over the last decade"

[1] Google N-gram is software that shows how often a word has been used in the Google book archives each year.
[2] "Google's AlphaGo Defeats Chinese Go Master in Win for A.I." this headline appeared in The New York Times in 2017.

while Alekseeva et al. (2020), find that there is a "dramatic increase in the demand for AI skills over 2010-2019 in the U.S. economy across most industries and occupations." What stood out the most from these papers is how they figured out which companies were using AI technology, and as I will show later, it is neither an easy nor cheap task.

Seamans and Raj (2018), have a working paper that points out the lack of firm level data on AI and the difficulties of creating one. Since it is a very new technology, there are not established and publicly accessible databases on AI. This is the reason why most researchers in this field get creative and look for more unconventional sources. For example, Damioli, Roy and Vertesy (2021) use a database on companies that have filed patents for AI inventions. While others like Babina et al. (2020) and Alekseeva et al. (2020), find that "demand for labor with AI skills can proxy for the use of AI technology". Both authors use information from Burning Glass Technology (BGT), which is a platform that hosts information about online job posting since 2010; this includes resumes, employee skill sets, job vacancies, etc., to predict how much AI technology each company uses. The rationale is simple: companies that use AI technology must have human capital that knows how to create and maintain it.

It is a straightforward idea that yields very interesting results. Babina et al. (2020) find that there is a "positive feedback loop between AI investments and firm size: AI investments concentrate among the largest firms, and as a firm invests in AI, it grows larger, gaining sales, employment, and market share", but since AI is such a new technology, a lot of firms do not know how to use it more efficiently, that is why there is no evidence for short-run productivity gains. Alekseeva et al. (2020) finds that there has been a tenfold increase in its usage since 2010 to 2019. Where most of the demand for AI is concentrated in "computer, engineering and science occupations" but also in other less obvious industries, such as "farming, forestry and finishing". Like Babina et al, Alekseeva et al. finds that "larger firms, firms that invest in R&D and firms that have more cash holdings" are the ones that hire more AI-skilled workers.

This idea of taking advantage of publicly available information to gain insight into the hidden workings of a company is not new. Sheng (2021) does something very similar, where he "uses employees' online forecast as a proxy for their information and examines the investment value of their information." He uses the publicly available review information from Glassdoor, a large job review website, to create scores such as *AbnOutlook*, which measure how good or bad the employee sees the company. He uses these scores to create long-short portfolios and runs multiple factor model regressions on them. He finds that "employee outlooks contain value-relevant information and predict future stock returns", though the "the return predictability of employee outlook decays over five months."

I use this central idea of demand for AI as a proxy for AI-intensity and take advantage of publicly accessible job boards to create a novel database of job posting related to AI. I obtained my information from Indeed, which is one of the biggest online job boards in the US with 250 million unique visitors per month (Indeed, 2021). For each company in the S&P 500 and the Dow Jones, I collected the number of job vacancies related to AI and the total number of jobs. Like Sheng (2021), I will be using this AI-score to create a long-short portfolio and run a three factor Fama-French model to find alphas and betas of the returns of these portfolios from 2015-2020. By creating my own dataset, I also explore the benefits and limitations of web scraping public available job boards and portals.

This paper fits into the emerging literature of AI in the labour market. Specifically, Alekseeva et al., (2019), Grennan and Roni, (2019), and Babina et al. (2020). I created a new cross-sectional dataset on the AI intensity/scores of all companies in the Dow Jones and S&P 500 taken from a publicly accessible job board. This paper also adds to the more classic finance literature on information and empirical asset pricing. Sheng (2021), Cooper et al (2010), and Savor and Wilson (2013) are just some examples that I look at to see how novel information can predict abnormal excess returns.

I end up finding that cumulative returns are higher for the High AI (ML) portfolios compared to the Low AI (ML) portfolios, but after running the Fama and French 3 factor model, there is insignificant alphas for ten out of twelve portfolios I test, most of the variation is explained by one or most of the three risk factors in the Fama French model. There is a positive 0.23% and 0.24% monthly alpha for the High AI and ML S&500 portfolios, respectively, but this result is only significant at a 95% confidence level.

2 Data

AI is a very new field in the context of Finance which means there are not many established databases for AI at a company level, especially publicly accessible ones (Seamans and Raj ,2018), which prompted me to create my own. The first step was to determine which companies use AI and which do not. Babina et al, (2020) uses the job titles, resumes, and employers' profile as a proxy for AI intensity of different companies. They scrape information from Burning Glass Technology (BGT) to create an AI score for each company. As a student, I do not have the resources to access private databases like BGT, so my next best option is to use free job boards like Indeed to see job vacancies for companies. It is a similar approach Sheng (2021) used, where he scrapped company reviews from Glassdoor to predict stock returns.

Taking the same approach as Alekseeva et al. (2020), I give each company an AI-score which is the "ration of AI job postings over the total number of job postings". I built two scores, one of AI and another of Machine Learning (ML). I included Machine Learning since it is related to AI and Babina et al. (2020) also includes it when building their AI score. They include Natural Language Processor (NLP) and Computer Vision (CV) as well, but because of the time constraint of the course, I do not add them to mine.

ML is a subset of AI and a lot of companies have postings about ML and little to none about NLP and CV. I also focus on ML since "Artificial intelligence is the human intelligence that is exhibited by machines. Machine learning is an approach to achieve artificial intelligence" (Khanam, Tanweer and Khalid, 2020). ML is the means to creating Artificial Intelligence, so I expect companies to look for ML skills, rather than just AI skills, and as I will show later, there are more ML jobs than AI, which prove my suspicions correct.

The information about job vacancies is obtained from Indeed. I chose this platform because I will be looking at Dow Jones and S&P 500 companies which are all from the US and Indeed is one of the biggest job boards in the US. It also has a simple and stable user interface that made it easy to scrape information from. For each company, I go to their corresponding Indeed company page, and in there, I find the total number of jobs and then the jobs related to AI and ML. This added step of going to each company's page, prevents me from including jobs from companies that have a similar name.

The biggest issue with this data set is that it is a cross sectional and not panel. I collect AI related jobs for that company at one point in time. Due to the nature of job boards, jobs expire and disappear from the website, so I cannot go back in time and note down all jobs related to AI for a company. It would have been better if I could collect data for a set period. For this paper, I was only able to collect jobs posted on March 1-10th of 2021. We are also going to use present AI and ML scores to look at past returns (2015-2020), so all the results will have an ex-post selection bias.

Another thing to keep in mind is that any company that might have a big AI department but is not actively looking for employees, will not appear in my dataset. This is something Alekseeva et al. (2020) also brings up in their paper. “What we can observe are changes in potential hiring and we consider it a proxy for which type of workforce is likely to grow or shrink in the firm in the near future.” Babina et al. (2020) also finds a strong correlation between job openings and actual employment. Which is why I also believe that job opening will be a good measure for AI intensity.

Job boards are also not perfect, they rely on hiring managers to post jobs with them, or web scrapers. It can also happen that companies do not want to post jobs in Indeed for a variety of reasons (e.g. companies do not want too many applicants). All the Dow Jones companies have at least one job posting while 14% of the companies in S&P 500 had zero job postings in Indeed.

Given all the mentioned drawbacks, a better way to approach this project in the future would be to also look into people's resumes and see if they have any AI related jobs in the company they work at (or used to work at). This information is partially available in LinkedIn, but due to privacy and time constraint challenges, it might not be feasible. It would make the dataset more robust, if I could get the job posting directly through company websites. This would reduce the gap in information from companies, but this requires a bigger time commitment.

Having said this, I will be diving into the dataset that I was able to compile.

Panel A. Summary Statistics for Dow Jones		
	AI Score	ML Score
Mean	2.00%	5.75%
Std. Deviation	2.71%	6.08%
Median	1.12%	4.08%
75th	2.99%	7.55%
Correlation with AI Score	1	1.17***
Correlation with ML Score	0.23***	1
Industries with Highest AI/ML Score	E & OEE & C [9.13%], Business Services [3.43%], Non-depository Credit Institutions [3.34%]	Food and Kindred Products [19.64%], E & OEE & C [15.47%], Non-depository Credit Institutions [10.24%]

Panel B. Summary Statistics of S&P 500		
	AI Score	ML Score
Mean	0.93%	3.06%
Std. Deviation	5.09%	8.39%
Median	0.00%	0.57%
75th	0.50%	2.70%
Correlation with AI Score	1	0.66***
Correlation with ML Score	0.24***	1
Industries with Highest AI/ML Score	E & OEE & C [6.85%], Business Services [2.65%], Depository Institutions [1.72%]	E & OEE& C [10.14%], Nondepository Credit Institutions [9.27%], Transportation by Air [8.71%]

E & OEE & C stands for 'Electronic & Other Electrical Equipment & Components'

* $p<0.1$, ** $p<0.05$, *** $p<0.01$

Table 1. *Summary Statistic of AI and ML scores for Dow Jones and S&P500 Companies.*

We can see in Table 1. that the average AI and ML scores are higher for companies in the Dow Jones compared to S&P 500. This is mainly because Dow Jones has only 30 large companies while S&P500 has 505 with many companies with many AI and ML scores of zero, which pushes down the average. Proof of this is that the median AI score is zero for S&P500, means that more than half the companies in the S&P 500 have zero AI related jobs. We can also see that standard deviation is higher for S&P 500, so within the companies with some AI and ML related jobs, there is a great variance.

Alekseeva et al. (2020) mentions that " Professional, Scientific, and Technical Services sector has AI Share just below 2% and the next groups of industries are Finance and Insurance, Administrative and Support Services, Agriculture, Forestry, Fishing and Hunting, and Manufacturing – all with AI Share around 1%." It can be seen in *Table 1*. that average AI score for companies in the Dow Jones is 2%. When we separate the Dow Jones into industries according to their two digits SIC (Standard Industrial Classification) code, the Service sector has the highest average with an AI Score of 3.43%, followed by the Manufacturing Sector with a 2.7%. These scores are a little high compared to Alekseeva et al. (2020). If we instead look at our S&P 500 companies, we see closer values to those of Alekseeva et al.(2020), since the Service Industry has a 1.89% and the Manufacturing Industry with a 1.31% and Finance, Insurance, And Real Estate with a 0.59% of AI share. We also did find some discrepancies with the authors since there is 0% share of AI related jobs, in Agriculture, Forestry, And Fishing industry as well as the Construction Industry.

In Table 1, we can see more granular industries than the ones mentioned above, with Electronic and Other Electrical Equipment and Component (E&OEE&C), always making it in the top three in both AI and ML score for both Dow Jones and S&P 500. E&OEE&C falls within the Manufacturing Industry, as is the Food and Kindred Products. Depository (e.g. banks) and Non-Depository Credit Institutions fall within Finance, Insurance, And Real Estate, while Business Services fall within the Service sector. (NAICS, 2021)

The very high ML score for the Food and Kindred Products sector in the Dow Jones is solely due to The Coca Cola Company, which is the only company in this sector, having a ML score of 19.64%. The other sector that has a large ML score is Transportation by Air, that includes major airlines and delivery companies.

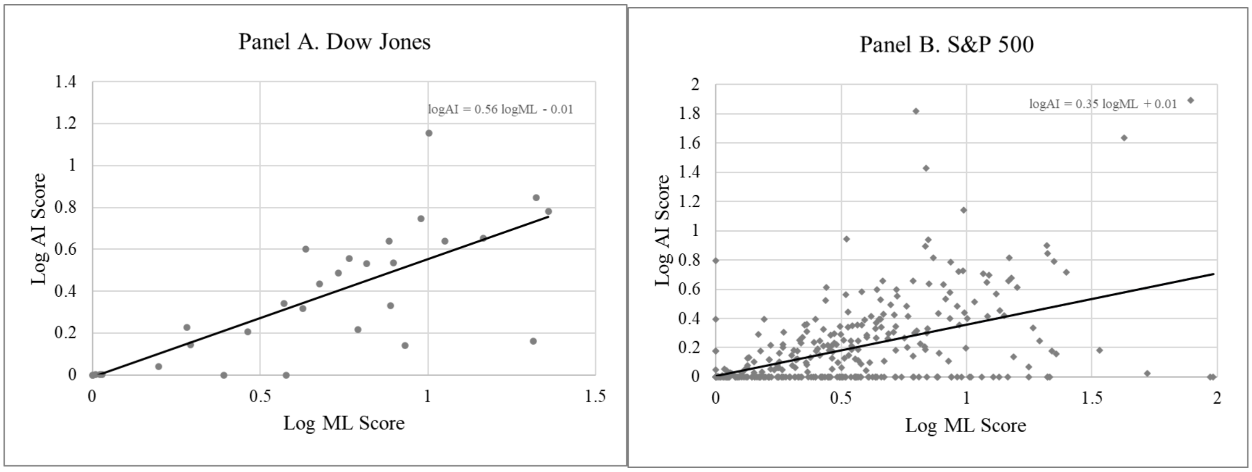

Figure 1. Correlation between ML and AI score. Since there were many zero scores, we applied a log (x+1) transformation.

In *Figure 1.*, we can see a positive correlation between the ML and AI Score (statistically significant at a 99% confidence level). In general, companies tend to have a higher ML score compared to their AI score. Though this is not always the case as seen from the multiple data points lying flat on the x-axis. This represents the companies that have a zero AI score, yet they have positive ML scores.

An extreme example of this is Verisk Analytics, which has 110 job postings and 105 are related to machine learning, while there are zero jobs related to AI. Looking more closely, we see the reason for this is that part of the company's description mentions using machine learning as an analytical tool and since the majority of the job postings have a company description, Indeed shows all of the job posting, even if in reality, they are not related to machine learning. These outliers are very small in number (only two of the companies have very high ML score and zero AI score) and they are only found in the ML dataset. These data points were not dropped from the dataset since the portfolios are not formed by weight of AI/ML, so the magnitude of the ML score is not what is important. We just need to distinguish the ones that have relatively high or low AI/ML intensity.

The positive relationship between AI and ML is due to the fact that many jobs require both these skills. A company can put a job description that requires both AI and ML skills, so the job will show up twice in the dataset, once as an AI job and once as a ML job. There are ways to overcome this double counting but requires more time. What is interesting is that there is a higher proportion of ML compared to AI jobs, which means companies describe more jobs as machine learning than artificial intelligence. As mentioned earlier, this is because ML is a more concrete skill required to develop AI.

3 Empirical Methods

To create the high/low portfolios, we first create a dummy variable that becomes one if the AI score is higher than the medium and zero if it is lower. This is the same process for ML score in both the Dow Jones and S&P500. Since the medium AI score in the S&P500 is zerp, I will assign any company that has a non-zero AI score into the high AI portfolio. I am aware that by just assigning one or zero to each company, we are losing information and the portfolios can only be equally weighted. The reason for this is that, as mentioned in the previous section, AI (ML) scores are noisy and they only reflect information from one point in time.

The information on returns comes from the Center for Research in Security Prices (CRSP), where I download the monthly holding returns for each of the Dow Jones and S&P500 companies. The date spans from January 2015 to December 2020. High portfolios are formed by companies with AI (ML) scores higher than the median, and while Low portfolios are created with the remaining of the companies.

AI score is created from job posts from the present and the performance analysis is done on returns from 5 years in the past. This is an issue since companies that are doing well (high return), have more resources to invest in novel technology such as AI, so that is why present High AI companies did so well in the past. This is related to Babina et al. (2020) findings, where there is a positive feedback loop where bigger companies invest more in AI and the new technology helps them grow bigger. So, it might be the case that companies with higher return tend to invest more in novel technology, this is part of the ex-post selection bias that was mentioned earlier in the paper. Therefore, we cannot say that the AI or ML scores are explaining the excess returns, rather it’s just to see how these present AI/ML scores relate to these past returns.

I will be using a Fama-French 3-Factor Model to conduct the analysis. Their model builds upon the older work of Sharpe, (1964) and Lintner (1965) on the now famous Capital Asset Pricing Model (CAPM). As Fama and French (1993) discuss regarding their model: "(The intercept) is the average abnormal return needed to judge whether a manager can beat the market, that's, whether he can use special information to generate average returns greater than those on passive combinations". This is exactly what we want, to know if this dataset on AI/ML is generating special information. This factor model is a good fit for our analysis since it controls for small market cap, high book-to-market ratio and excess market return, that have been historically known to be associated with higher earnings (Fama and French, 1993). Especially since we are using companies from the Dow Jones and S&P 500, which are indexes of big cap companies in the US. Fama and French (2004) also point out that: "the empirical record of the model (CAPM) is poor—poor enough to invalidate the way it is used in applications.", so the factor model is preferred over the CAPM.

To be able to run this regression, we would need data on these factors and luckily, we can download them on a monthly frequency from the Kenneth French Data Library. The first factor is an excess return from the market, meaning it is the returns from the market minus a risk free return (one-month Treasury Bill rate), the SMB factors stands for 'Small minus Big', which is the "average return on the three small portfolios minus the average returns on the three big portfolios." and HML stands for High Minus Low, which addresses "average return on the two value portfolios minus the average return on the two growth portfolios." (French, 2020)

Fama and French (1993), show that "size and BE/ME (book-to-market ratio) are related to systematic patterns in relative profitability and growth that could well be the source of common risk factors in returns" and by including them in the performance analysis, we can truly see if High AI/ML scores are a good indicators to generate excess returns.

4 Results

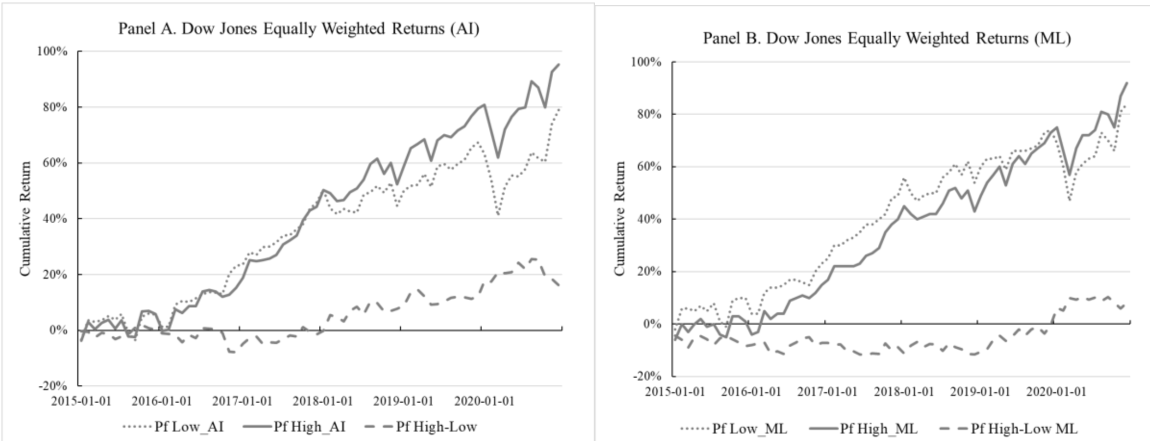

Figure 2. Cumulative Equally Weighted Monthly Return from January 2015 - December 2020 for companies in the Dow Jones (N=30).

From *Figure 2*, we see that at the beginning, the portfolio with only High AI companies trails very closely with the portfolio with Low AI companies; this changes at the start of 2018 and by the end of 2020, the High AI portfolio is 16 percentage points higher than the Low AI portfolio. For the ML based portfolio, the one with High ML stays below the Low ML portfolio for quite a long time, from the start of 2015 until the end of 2019. By 2020, it seems to be doing better than the Low ML portfolio and at the end of 2020, it is 8 percentage points higher than the Low ML portfolio. We can see the same story play out in the long-short portfolios, which is to be expected since these portfolios are formed from taking the difference between the High/Low portfolios.

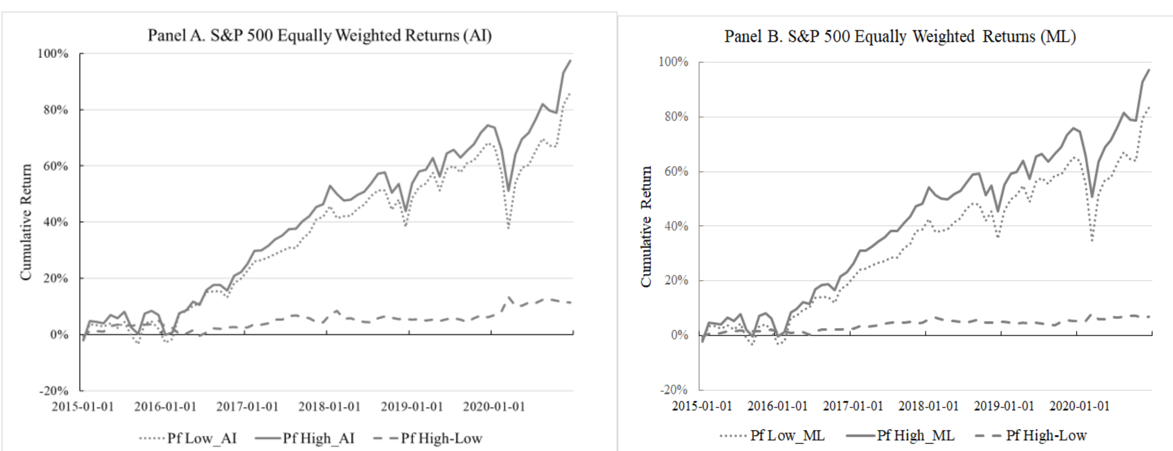

Figure 3. Cumulative Equally Weighted Monthly Return from January 2015 - December 2020 for companies in the S&P 500 (N=505).

From *Figure 3*, we see that the High AI portfolio is always above the Low AI portfolio, but they follow each other very closely, and there is never more than 11 percentage points of difference between these two portfolios, which is the same difference they have at the end of 2020. The ML based portfolio tells a similar story, where the High ML portfolio is always higher than the Low ML one, the long-short portfolio cumulative returns is maintained above zero during all the time periods. From a first glance, it seems that High AI/ML portfolios are doing better than their counterpart, and the High-Low portfolios are also doing moderately well since they have positive cumulative returns, albeit conservative positive returns.

Panel A. Portfolio Sumary Statistic - Dow Jones			
	Mean	Std. Deviation	Sharpe Ratio
High_AI	1.25%	4.45%	0.28
Low_AI	1.02%	4.38%	0.23
High-Low_AI	0.23%	2.32%	0.10
High_ML	1.20%	4.25%	0.28
Low_ML	1.09%	4.51%	0.24
High-Low_ML	0.11%	2.14%	0.05
Panel B. Portfolio Sumary Statistic - S&P 500			
	Mean	Std. Deviation	Sharpe Ratio
High_AI	1.28%	4.69%	0.27
Low_AI	1.12%	5.01%	0.22
High-Low_AI	0.16%	1.19%	0.13
High_ML	1.28%	4.73%	0.27
Low_ML	1.08%	5.06%	0.21
High-Low_ML	0.19%	1.26%	0.15

Table 2. Summary statistics for twelve portfolios returns, Sharpe ratio (monthly) is the portfolio's average *excess* return over portfolio's *excess* return standard deviation.

From *Table 2.* we can see that the mean and standard deviation between the High and Low portfolios are very similar between AI/ML and Dow Jones/S&P500. The mean returns are always higher for High AI/ML portfolios. We can also see that the High - Low portfolio both lower returns and standard deviation, which makes sense since by taking a long/short strategy we are able to mitigate the natural movements of the market. From the Sharpe ratio (Sharpe, 1994), we can see that our portfolio returns are subjected to high volatility, especially our High-Low portfolios. This is a bad sign for our portfolios, since the returns aren't making up for the high risk we are incurring as seen from the high standard deviation.

Panel A. Famma French three factor regression of Dow Jones Portfolios

%	Equally Weighted AI Porfolios			Equally Weighted ML Porfolios		
	(1) High	(2) Low	(3) High-Low	(4) High	(5) Low	(6) High-Low
Alpha	0.111	0.271	-0.159	0.221	0.166	0.0543
	(0.191)	(0.179)	(0.251)	(0.178)	(0.203)	(0.269)
Market - Risk Free	0.978***	0.889***	0.0888	0.926***	0.940***	-0.0145
	(0.0428)	(0.0401)	(0.0561)	(0.0397)	(0.0455)	(0.0602)
Small Mins Big	-0.277***	-0.154*	-0.123	-0.268***	-0.167*	-0.101
	(0.0764)	(0.0716)	(0.100)	(0.0709)	(0.0812)	(0.108)
High Minus Low	-0.121**	0.262***	-0.383***	0.0137	0.117	-0.104
	(0.0599)	(0.0562)	(0.0786)	(0.0556)	(0.0637)	(0.0844)
N	72	72	72	72	72	72

Panel B. Famma French three factor regression of S&P 500 Portfolios

%	Equally Weighted AI Porfolios			Equally Weighted ML Porfolios		
	(1) High	(2) Low	(3) High-Low	(4) High	(5) Low	(6) High-Low
Alpha	0.227**	0.157	0.0700	0.235**	0.13	0.105
	(0.0991)	(0.139)	(0.137)	(0.0971)	(0.156)	(0.150)
Market - Risk Free	0.996***	1.012***	-0.0164	0.998***	1.014***	-0.0153
	(0.0222)	(0.0310)	(0.0307)	(0.0217)	(0.0349)	(0.0336)
Small Mins Big	0.0542	0.121*	-0.0664	0.0747	0.116	-0.0414
	(0.0396)	(0.0554)	(0.0548)	(0.0388)	(0.0623)	(0.0599)
High Minus Low	0.0423	0.184***	-0.142**	0.0603	0.201***	-0.141***
	(0.0310)	(0.0434)	(0.0430)	(0.0304)	(0.0489)	(0.0470)
N	72	72	72	72	72	72

*p<0.1, **p<0.05, ***p<0.01

Table 3. Alphas and Betas from Fama and French 3-factor model, on twelve different portfolios; all units are in percentages

We ran a Fama and French 3-factor regression on our returns to get a clearer picture of the results. In Panel A, we see that the High AI portfolio has an insignificant alpha, it has a positive coefficient 0.978 in the RM-RF which stands for excess market returns. This suggests that the High Portfolio returns is mainly explained by the market movement. We also have a significant negative coefficient for the SMB and HML factors, which means that our portfolio tends to be big cap and growth companies. Low AI portfolio has a similar story for the RM-RF and SMB factors, but it has significant positive coefficient, which means it behaves like a portfolio of value stocks. When we take on a long-short strategy, the only factor that is statistically significant is that negative HML coefficient, which means that our portfolio behaves like a growth stock portfolio.

Looking at Panel B, we see a very similar story for the High ML portfolio, its variation is explained mostly by the market excess return, it tends to have big cap companies but now we have an insignificant HML coefficient. For the Low ML portfolio, it behaves very similarly to the High ML. All three ML based portfolios have insignificant alphas, which means the excess returns are being explained by the 3-factor model. The Dow Jones has thirty large companies making up the index, so it makes sense that almost all portfolios have a negative correlation to the SMB risk factor. From previous literature, specifically Babina et al. (2020) we know that bigger firms are the ones investing in AI since investment in AI helps firms grow and explore more markets. This could also explain the negative coefficient in the SMB risk factor.

Now let's look at the S&P 500 Portfolio, in Panel B, we get that there is a significant positive alpha of 0.23%, at a 95% confidence level, which is quite small, economically speaking. Most of the excess return behaves according to the market excess return and are not correlated to the other two risk factors of size and book-to-market ratio. The Low AI portfolio does not have a significant alpha, but it does have a positive and significant coefficient for all the three factors. The positive excess returns are explained by the portfolio behaving like the market, have smaller cap and value companies. The High-Low have insignificant coefficients for most risk factors, except for the HML which is negative, meaning it behaves like growth stocks portfolios.

The ML based portfolios in Panel B are also behaving very similarly. The High ML portfolio has a positive monthly alpha 0.24% at a 95% confidence level, and only the excess market return seems to be explaining portfolio excess return. Low ML portfolio is strongly linked with the market and is positively correlated with the HML risk factor, and the High-Low portfolio is negatively correlated to the HML risk factor.

It is not surprising that in all four portfolios (High/Low and AI/ML), there is a significant positive relationship with the market excess returns, since we are picking stocks around 50% of stocks from the Dow Jones and S&P500 which are indexes, specifically used as a benchmark. We can also see that for all High-Low portfolios, there is not a statistically significant coefficient for the market excess returns. This is due to the portfolio being a long-short portfolio and these "generally have lower market expose" (Fung & Hsieh, 2011). It is also not surprising that we have an insignificant alpha for the long-short portfolio, since Fung & Hsieh, (2011) find that "in terms of risk, the standard four-factor model accounts for over 80% of the variation in the returns of these (long-short) hedge funds." Meaning that only 20% of 3038 long-short equity hedge funds they sampled (1994-2008) had a positive alpha. We only control for three factors (we do not include momentum), and even then, we do not have any positive alphas.

Since our High AI and ML portfolios have a statistically significant alphas, I wanted to subject them to a robustness test, where I remove the stock return of companies that are related to information technology. This sector includes very popular companies like Google, Amazon, Facebook, Microsoft, etc. We dropped forty companies with SIC equal to 737 which stands for "Computer Programming, Data Processing, and other Computer Related Services" NAICS, (2021). By running the 3-factor model again on these modified portfolios, we can see if our previous results were driven by these bigger IT sector companies. This is relevant since one might argue that the AI and ML scores are just picking stocks form the IT sectors. Maybe figuring out AI and ML scores is just another even more convoluted way to create an IT based portfolio.

%	S&P 500 AI	S&P 500 ML
Alpha	0.185*	0.206*
	(0.1091)	(0.1072)
Market - Risk Free	0.986***	0.996***
	(0.0222)	(0.0240)
Small Mins Big	0.0846*	0.107**
	(0.0436)	(0.0428)
High Minus Low	0.097***	0.113***
	(0.0342)	(0.0336)
N	72	72

*p<0.1, **p<0.05, ***p<0.01

Table 4. Alphas and Betas from Fama and French 3-factor model on High AI/ML portfolio Minus IT sector company stocks.

We see that after removing the forty IT companies, our monthly alpha drops from 0.227% to 0.185% for our high AI portfolio and for the ML portfolio it drops from 0.235% to 0.206%. The results are now only statistically significant at a 90% confidence level. We also have a positive coefficient in the HML factor which is statistically significant at a 1% level, but they are very small only 0.1% and 0.11%. That suggests that those forty IT companies were driving our alphas up, but since they don't disappear completely, there must be some other companies giving us returns not predicted by the other factors. From our summary statistics, we know there is a great proportion of business/financial sector companies still in the high AI and ML portfolio, but without a deeper breakdown of returns by industry, we won't be able to tell what is driving our returns.

Our portfolios are not a good trading strategy, since the majority have negligible alpha and most of the returns are explained by well known risk factors. From the Sharpe ratio, we also see that on average, our portfolios are exposed to high volatility compared to their conservative returns, that do not do better than market indexes. After removing the IT sector, our alpha drops and is only statistically significant at a 90% level. If we were to control for productivity of the firm, we would probably have negligible alphas since we know from previous literatures, (Babina et al, 2020; Damioli, Roy and Vertesy, 2021), that AI is positively correlated with productivity.

This of course does not mean that AI and ML scores are bad sources of private information. As mentioned before, the dataset used to for these High/Low/High-Low portfolios were imperfect. Since 1) we ran this analysis in the past; the AI and ML scores for each company were made in 2021, but we conducted the 3-factor model regression with returns from five years in the past, and 2) it was a cross sectional data, we could not re-weight the portfolios each month/quarter. I also believe that as more companies start incorporating AI into their daily processes, information on how, what and who are using these technologies will be crucial.

5 Conclusion

I created an AI intensity database from an online job board, and used job opening related to AI/ML to serve as a proxy for AI/ML intensity for each company. I find that high AI and ML companies from S&P 500 do have higher returns with a monthly alpha of 0.23% and 0.24% respectively, at a 95% confidence level. The Low and High-Low AI/ML portfolio give negligible alpha, and excess returns are mostly explained by the risk factors in the Fama and French model. This result has an ex-post selection bias since the AI/ML scores are from a cross sectional dataset created in 2021 while the returns are from 2015-2020. In its current state, the dataset isn't a good way to create high performing portfolios. The returns are very conservative considering the high volatility they are subjected to and they don't really give an edge over just a market index-based portfolio. Future work should focus on creating a panel dataset of AI job vacancies and use it to build a weighted portfolio. This could be done using public job boards which one will have to follow for a couple of years, or one could pay for market analysis services like Burning Glass Technology or Cognims, to access archived AI job vacancies. It is fascinating how much information there is around us, and with a little help from computers, we can compile it and turn it into useful data. I believe that this age of information, where information grows exponentially, will bring along new exciting possibilities to this seeming over saturated field of asset pricing.

References

[1] Alekseeva, Liudmila and Azar, Jose and Gine, Mireia and Samila, Sampsa and Taska, Bledi, 2019. The Demand for AI Skills in the Labor Market. *CEPR Discussion Paper No. DP14320.* https://ssrn.com/abstract=3526045

[2] Babina, Tania and Fedyk, Anastassia and He, Alex Xi and Hodson, James 2020. Artificial Intelligence, Firm Growth, and Industry Concentration *American Finance Conference (2021).* http://dx.doi.org/10.2139/ssrn.3651052

[3] Cooper, Michael J., et al. (2010) Corporate Political Contributions and Stock Returns *The Journal of Finance* 65(2) 687-724. https://doi.org/10.1111/j.1540-6261.2009.01548.x

[4] Damioli, G., Van Roy, V. & Vertesy, D. (2021). The impact of artificial intelligence on labor productivity. *Eurasian Bus Rev* 11, 1-25. https://doi.org/10.1007/s40821-020-00172-8

[5] Fama, Eugene F., French, Kenneth R., (1993) Common Risk factors in the returns on stock and bonds. *Journal of Finance* 33, 3-56. https://rady.ucsd.edu/faculty/directory/valkanov/pub/classes/mfe/docs/fama_french_jfe_1993.pdf

[6] Fama, Eugene, F., and Kenneth R. French. (2004). The Capital Asset Pricing Model: Theory and Evidence. *Journal of Economic Perspectives*, 18 (3), 25-46. https://www.aeaweb.org/articles?id=10.1257/0895330042162430

[7] French, Kenneth R. 2020. Description of Fama/French Factors *Kenneth R. French Data Library*. https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/f-f_factors.html

[8] Fung, William and Hsieh, David A. (2011). The risk in hedge fund strategies: Theory and evidence from long/short equity hedge funds. *Journal of Empirical Finance*, 18(4), 547-569. https://doi.org/10.1016/j.jempfin.2011.04.001

[9] Grennan, Jillian and Michaely, Roni. (2021). Artificial Intelligence and the Future of Work: Evidence from Analysts *American Finance Conference*. https://conference.nber.org/conf_papers/f130049.pdf

[10] Google. (2021). Google Book Ngram Viewer. https://books.google.com/ngrams

[11] Gu, Shihao and Kelly, Bryan and Xiu, Dacheng. (2020) Empirical Asset Pricing via Machine Learning, *The Review of Financial Studies*, 33 (5), 2223-2273. https://doi.org/10.1093/rfs/hhaa009

[12] Indeed. (2021) Indeed About. *Indeed*. https://www.indeed.com/about

[13] Lintner, J. (1965). The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets. *The Review of Economics and Statistics*, 47(1), 13-37. https://www.jstor.org/stable/1924119?seq=1

[14] Mozur, Paul (2017) Google's AlphaGo Defeats Chinese Go Master in Win for A.I. *The New York Times*. https://www.nytimes.com/2017/05/23/business/google-deepmind-alphago-go-champion-defeat.html

[15] NAICS Association. (2021). NAICS & SIC Code list. *NAICS. https://www.naics.com/search/*

[16] Khanam, Sana and Tanweer, Safdar and Khalid, Syed. (2020). Artificial Intelligence Surpassing Human Intelligence: Factual or Hoax, *The Computer Journal*. https://doi-org.ezproxy.library.ubc.ca/10.1093/comjnl/bxz156

[17] Savor, P., & Wilson, M. (2013). How Much Do Investors Care About Macroeconomic Risk? Evidence from Scheduled Economic Announcements. *Journal of Financial and Quantitative Analysis*, 48(2), 343-375. 10.1017/S002210901300015X

[18] Seamans, Robert and Raj, Manav. (2018). AI, Labor, Productivity and the Need for Firm-Level Data. *Working Paper. National Bureau of Economic Research.* https://www.nber.org/papers/w24239

[19] Sharpe, William F. (1964), Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk. *The Journal of Finance*, 19, 425-442. https://doi.org/10.1111/j.1540-6261.1964.tb02865.x

[20] Sharpe, William F. (1994), The Sharpe Ratio. *The Journal of Portfolio Management*, 21 (1), 49-58. https://doi.org/10.3905/jpm.1994.409501

[21] Sheng, Jinfei, 2021 Asset Pricing in the Information Age: Employee Expectations and Stock Returns. *Working Paper, University of California, Irvine*. http://dx.doi.org/10.2139/ssrn.3321275

Appendix A

Figure 1. Google Ngram of the word Artificial Intelligence. It had its peak in 1989, then again in 2002, and has been rising in 2019.