

GPT - The future of business data analysis and prediction - A case study of Hanvon Technology.

Kaisheng Yang, Shengrong Zhang

Nanjing University Business School(NJUBS);Nanjing University Software Institute(NJUSI), Nanjing University, Nanjing 210046

Mentor: Zhenyu Chen

Department of Computer Science and Technology, Nanjing University, Nanjing 210046

Whether the paper is published: [No]

Abstract

The purpose of this research is to validate the Generative Pre-trained Transformer's (GPT)[1] effectiveness and accuracy when used for commercial data collection and analysis based on large language models. (LLMs). The current study makes an effort to determine whether the GPT has a similar mindset to human analysts as language models by combining quantitative and qualitative analysis techniques. It then evaluates its effectiveness promotion and substitution effect to human analysts when conducting business data collection and analysis. The investigation shows that: 1) GPT can reduce the amount of human work required for data collection by an average of 50%, even though some of the data it collects still needs to be verified for accuracy; and 2) when using a specific method for data analysis, GPT not only reduces the level of expertise required for analysts, but also, on average, shortens their work by about 70%. The majority of the general data collection and processing tasks can be handled by this efficiency improvement, which will have a big impact on the future of the labour market and the knowledge base for business analysis. Furthermore, despite not operating as well as it could, GPT's resemblance to human thought and its broader range of thought reveal its tremendous potential for making future business predictions.

Key words:

Business data analysis; Artificial Intelligence (AI); Generative Pre-trained Transformer; Pearson correlation coefficient

Originality

The use of LLMs for commercial data collecting and prediction is the main topic of this essay. Despite LLMs' widespread use in disciplines like language production and natural language processing, their use in the collecting of commercial data is still in its infancy. The research focus of this work has a lot of practical application.

The GPT model's application is the main topic of the paper. Despite the fact that there are many different kinds of LLMs models available, GPT is now one of the most well-liked models due to its high accuracy and effectiveness. In order to thoroughly examine the use of LLMs in commercial data collecting and prediction, this work chooses the GPT model as the research object.

It is suggested to use a novel technique for commercial data collecting prediction. This study suggests a novel GPT-based methodology. Our technique can better adapt to particular commercial data prediction jobs and increase prediction accuracy and efficiency by tweaking the GPT model.

We investigate any potential GPT data collecting and prediction restrictions. There are still considerable restrictions in commercial data collecting and prediction, despite the fact that GPT has made significant progress in areas like language production and natural language processing. In order to improve the use of GPT for commercial data collecting and prediction, this study addresses these limitations in detail and suggests some alternatives.

1 Literature Review

A dominant force in artificial intelligence study in recent years has been the development of large language models (LLMs), such as the Generative Pre-trained Transformer (GPT)¹, which exhibit sophisticated language-based functionalities, including but not limited to code generation and writing. Multiple factors are responsible for this success. For instance, GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models (Tyna Eloundou et al., 2023)[2] shows that in practical applications, people with bachelor's, master's, and professional degrees tend to experience greater efficiency gains in specific labour tasks when using GPTs compared to those without formal education. LaMDA: Language Models for Dialog Applications (Thoppilan et al., 2022)[3] also shows LaMDA's versatility in handling a variety of tasks. In general, jobs that demand specialised abilities are less replaceable.

However, we discovered that prior study on financial quantitative analysts had largely ignored the potential efficiency gains and substitutability of their particular job processes and instead concentrated more on the level of exposure of their work (Agrawal et al., 2022) [4]. Large-scale unstructured data analysis using artificial intelligence (AI) could result in significant improvements in accuracy and productivity for the economy. Thus, understanding GPTs' technical and particular application techniques in business analysis, as well as how to incorporate them into the fields of business forecasting and data analysis, is therefore of significant academic and practical value.

Our research goal is to introduce GPT into the field of business data analysis. To accomplish this, we combined our study objective with ideas from the books and research field already in existence. Like (Felten et al., 2023)[5], we will concentrate our study on the effects and particular application strategies of GPTs rather than on more general machine learning or technological advancements. In choosing our study subject, we made an effort to identify businesses that had undergone significant events that significantly impacted their operational performance during the COVID-19 pandemic-related economic downturn. As a result, "Hanvon Technology," a business that released ChatGPT in 2022 to help with content creation, has been selected as our study subject.

We suggest gathering time indicators for both human and AI users with different levels of AI proficiency in data collection and processing in order to build our main dataset for efficiency and accuracy comparisons. To show GPT's excellent performance in data collection and processing as well as its huge economic potential going forward, we also collect time indicators for data analysts and further analyse the data using Pearson correlation coefficients. Additionally, we compare the depth, breadth, and similarity of AI and human analysis, demonstrating the enormous promise of GPT in business analysis and forecasting, to evaluate the accuracy of business forecasting.

Despite GPT's stellar performance in data gathering, there are still improper operation methods in use, which can result in unreliable feedback and the possibility of receiving false information. The "ready-to-use" nature of GPT may make it somewhat unreliable for these jobs. (Abid et al., 2021[6]; Schramowski et al., 2022[7]; Goldstein et al., 2023[8]; OpenAI, 2023a[1]). However, it unquestionably offers a practical and effective method for experts to select from in terms of data handling. Based on their professional expertise, users can then evaluate and select particular solutions and use them to carry out particular data processing duties. In

¹ In the remaining sections of this paper, we use GPT to refer to large language models, usually using ChatGPT and NewBing as examples from our study.

other words, even though GPT does not foresee the future, it is unquestionably a potent instrument. Users only need to be adept with the steering wheel with its assistance to travel to remote locations.

2 Methods and Data Collection

2.1 Data Collection

We primarily rely on sources like "cninfo,"² "National Government Service Platform," and "Sina News" for our research and data verification in order to gather pertinent information about the company's operational situation, including financial statements and related news. Balance sheets, revenue statements, and cash flow statements are all examples of financial statements. The balance sheet illustrates the make-up and condition of a company's assets by examining the amount and division of its available financial resources. In order to evaluate the risks the business faces, it can also analyse the total debt. The results and operating effectiveness of a business can be seen in the income statement. The cash flow statement displays a company's cash flow and evaluates its possibility that it will generate future net cash flows, be able to pay debt, and be able to spend money efficiently. it and our study objective.

We primarily concentrate on two common language models, ChatGPT and NewBing, as our major research objects in terms of AI's capacity for data collection and its application in data-based business analysis. From two aspects, we assess their prowess in gathering and analysing data.

To measure the efficiency improvement AI brings to humans, we also use the Pearson correlation coefficient between the AI's proficiency level measured by human analysts and the efficiency improvement brought by the AI under the premise of consistency and accuracy.

Pearson product-moment correlation coefficient:

The Pearson product-moment correlation coefficient (PMCC)[9], which is used in statistics, is a technique for evaluating how comparable two vectors are. The outcome range is from -1 to +1, where 0 denotes no correlation, a value of -1 denotes a negative correlation, and a value of +1 denotes a positive correlation.

The strength of the linear correlation between two variables is frequently assessed in academic study using the Pearson correlation coefficient. The covariance between two variables is equal to their covariance split by the product of their standard deviations, and is known as the Pearson correlation coefficient. The Pearson correlation coefficient is frequently denoted by the Greek character rho, lowercase. The sample correlation coefficient (also known as the sample Pearson coefficient), which is typically denoted by the lowercase symbol r in English, can be calculated by estimating the sample covariance and standard deviation:

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}$$

- n is the number of observations in the data
- x_i and y_i are the individual values of x and y , respectively, for $i = 1, 2, \dots, n$
- \bar{x} and \bar{y} are the sample means of x and y , respectively

The degree to which recorded sample data agrees with a given hypothesis model is measured by the p-value, a statistic used in hypothesis testing. The p-value can be used to determine whether the sample data for a Pearson correlation coefficient adequately supports the presence of a significant linear relationship between two variables. The null hypothesis can be rejected and it can be determined that there is a significant linear relationship between the two variables if the p-value is less than the significance threshold, which is typically 0.05.

The following is the formula for calculating the p-value:

² www.cninfo.com.cn

1. Find the degrees of freedom using the formula $df = n - 2$, where n is the sample number.
2. Using the t-distribution table or computer software, calculate the critical number t based on the degrees of freedom and significance level.
3. Using the formula $t = r * \frac{\sqrt{df}}{\sqrt{1-r^2}}$, where r is the Pearson correlation coefficient, determine the t-value.
4. $P = P(|T| > |t|)$, where T is the t-distribution and $|T|$ represents the exact value of T , can be used to calculate the p-value.

2.1.1 Timeframe of Data Collection

In terms of time usage, we compared the total amount of accurate data obtained with the sum amount of accurate data obtained by data analysts and AI users of various levels of experience. As shown in Figure 3.1.1³, we used this comparison to ascertain the effectiveness of data collection for various study subjects.

Following are the various degrees of AI user proficiency that we defined:

- Non-AI user/inexperienced AI user: Users who are unaware of or have a limited grasp of how AI is used.
- Briefly exposed AI user: Users with minimal exposure to AI are those who have heard of or learned about its use.
- Experienced AI user: Experienced AI users are those who have used AI to handle straightforward issues in the past.
- Highly experienced AI user: those who have used AI to handle challenging problems and are adept at organising work for the best results.
- Senior AI user: a user with in-depth knowledge of how to use AI, who can input and output data correctly and effectively, and who can validate the outcomes of using AI.

time-consumed	Non-AI user using ChatGPT	Senior AI user using ChatGPT	Non-AI user using NewBing	Senior AI user using NewBing	Human
0	0	0	0	0	0
10	5	15	5	16	5
20	5	30	10	19	5
30	7	45	23	21	6
40	11	60	40	22	7
50	13	70	55	23	9
60	16	78	67	28	13
70	20	83	76	33	17
80	30	88	83	40	23
90	45	91	88	50	29
100	60	93	91	65	40
110	70	95	93	73	48
120	76	96	95	78	57
130	83	97	96	83	65
140	88	98	97	85	71
150	92	99	98	87	79

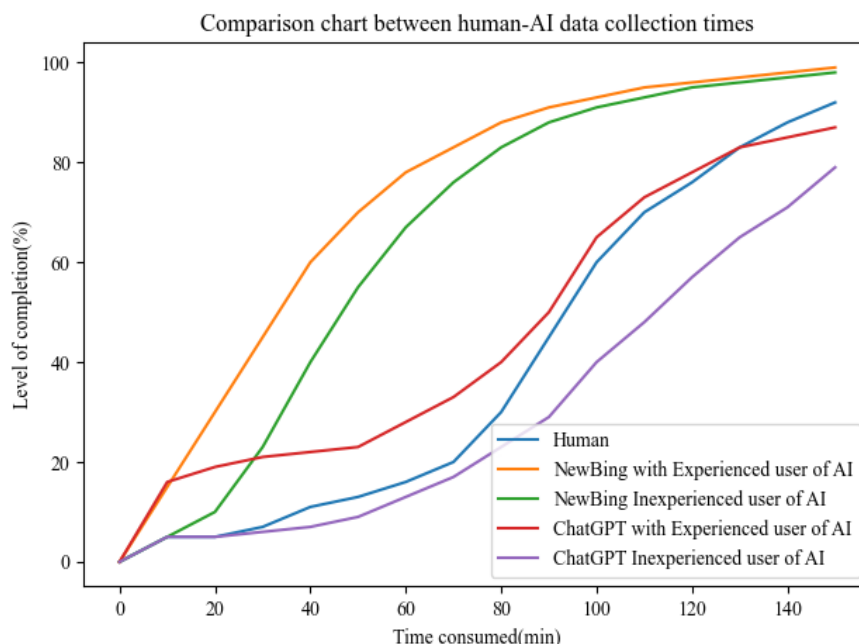
Table 1: Data collection timeline - Taking NewBing as an example

Correlation measurement	PMCC	p-value
20%	-0.82	0.04
50%	-0.78	0.06
90%	-0.68	0.13

2.1.2 Data Collection Accuracy

In terms of accuracy, we started by manually verifying the data that human researchers had collected. Then, we compared the performance of the AI itself through numerous experiments, as well as the similarity of the

³ Despite the lengthy gathering period, the marginal impact of commercial data is negligible due to its complexity and extensive correlation. Because of this, we choose 95% as a representative example of our collection of all commercial data; however, this only reflects the degree of data quantity, not the real situation.



data gleaned from the same dataset by AI and human analysts. We carried out weighted quantitative assessments based on the significance and clarity of the data. ⁴

By examining the relationship between the variance of AI's multiple analyses and the correlation between the scores and the data, we were able to measure the accuracy of AI's scoring in repeated trials under various circumstances.

2.2 Data Analysis

In order to gather AI data, we divide a company's data types into internal and external environments and create quantitative scoring criteria that correspond to those categories. These criteria include human evaluations based on the accuracy of the data given by AI as well as the AI's scoring of the data. By comparing the speed and accuracy of AI and human analysis, this article seeks to analyse the impact of AI's replacement of humans in the field of AI data gathering.

AI data analysis is split into two categories: data processing and critical analysis. The judgmental type refers to logical forecasts of an enterprise's future development trends that are made by AI or humans based on current financial data, operational circumstances, and exterior environments. Using the same data and fixed analysis techniques, the data processing type entails determining whether humans and AI can arrive at the same conclusion.

2.2.1 Data Processing in Data Analysis

In order to study the efficiency and precision of AI in data processing, we primarily used the DuPont analysis technique ⁵ as an example and also developed some customised methods. We assessed the difference

⁴ As explained in the methodology's limitations in section 2.3, the scoring factors used here may be a little subjective.

⁵ A financial analysis technique called the DuPont Analysis is used to describe changes in a company's profitability. By breaking down a company's profitability into three essential components—net profit margin, total asset turnover, and financial leverage—it finds the primary factors that have an impact on a company's profitability. The DuPont Analysis offers a deeper comprehension of a company's financial situation and how various factors influence its profitability. The DuPont Analysis helps investors better comprehend a company's earnings pattern, asset utilisation efficiency, and debt ratio by breaking down a company's profitability into three essential components.

in time consumption by simultaneously measuring the degree of completion of a particular technique in terms of time consumption, consistent with the processing method used during data collection. It is important to note that because a specific method is set at this point, the process can be measured precisely with steps, so there is no issue with errors occurring during data gathering. In a similar vein, when comparing the processing times of humans and artificial intelligence, we continue to use the time required to acquire accurate data as the actual time consumed, and we use time as the independent variable and completion percentage as the dependent variable to plot the chart. Both kinds of GPT have demonstrated incredible benefits in terms of data processing accuracy. If input data is given, they will unavoidably produce accurate and reliable outcomes. However, since the data volume for the DuPont analysis technique used as an example is constrained, we have set a relatively high threshold for data volume within the measurable range. The obtained results are still accurate, but due to limitations on the number of dialogues when working with enormous amounts of data, we cannot be certain that GPT will generate answers with significant errors and high misleadingness. In section 2.3, this method limitation of ours will be explained in more depth.

2.2.2 Data analysis judgement

In data analysis, the use of judgmental analysis measures one's capacity for information abstraction, gathering, and generalisation in addition to their aptitude for data processing and analysis. In other words, the capacity of AI to engage in rational thought processes, such as conceptualization, judgment, and inference, will determine whether it can accurately mimic human brain patterns and exhibit "thinking" distinctions.

However, subjectivity is a necessary component of judgment, so we can only "average" when testing with a human test group to establish quantitative judgement standards. Here we request a quantitative evaluation implementation technique from an evaluation team made up of financial managers, tax system managers, and university-bound economics students.

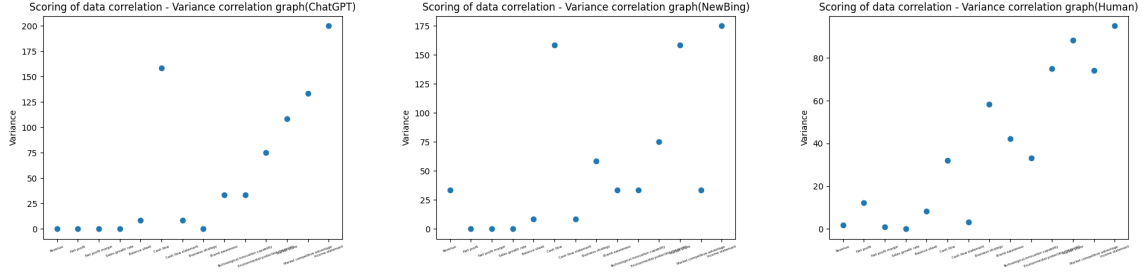
The evaluation committee will specifically assess the company based on its 2021 financial report and pertinent operating data from prior to 2021. The evaluation team will then revise the evaluation findings by giving news about the company in 2022, analysing the event's effects on the company's operating conditions, and making comparisons. There is, however, still a current, inevitable limitation. People frequently alter the outcomes of the second judgement to suit expectations unintentionally as a result of the consistency effect in human psychology. Section 2.3 of our approach contains a detailed explanation of this restriction.

First, we tested how well AI and humans both evaluated the quality of the data in order to determine whether AI has certain assessment skills.

The Evaluated Project	ChatGPT Scoring	ChatGPT Variance	NewBing Scoring	NewBing Variance	Human Scoring
Revenue	90.00	0.00	65	0.00	80.00
Net profit	70.00	0.00	40	0.00	40.00
Net profit margin	70.00	0.00	40	133.33	50.00
Cash flow	71.67	158.33	36.67	33.33	70.00
Return on Equity (ROE)	81.67	8.33	50.00	100.00	60.00
Return on Assets (ROA)	71.67	8.33	53.33	158.33	65.00
Equity multiplier	70.00	0.00	55	175	60.00
Market share	81.67	108.33	61.67	158.33	70.00
Sales growth rate	70	0.00	78.33	2.35	80.00
Brand awareness	73.33	33.33	73.33	33.33	85.00
Market competitive advantage	76.67	133.33	86.67	33.33	85.00
Technological innovation capability	83.33	33.33	93.33	8.33	90.00
Business strategy	90.00	0.00	83.33	8.33	85.00
Balance sheet	71.67	8.33	75.00	25.00	75.00
Income statement	60	200	55.00	175.00	50.00
Cash flow statement	71.67	8.33	70.00	0.00	70.00
Environmental protection measures	88.33	75.00	81.76	8.33	85.00

Table 2: Table for AI-human scoring of a company's operational status

Using the data relevance assessment score as an independent variable and the input data as a fixed and precise condition, we also evaluated the variance of AI's repeated tests. Additionally, based on the connection between variance and data relevance, we used Python as a trustworthy analytical tool to acquire the related Pearson coefficient.

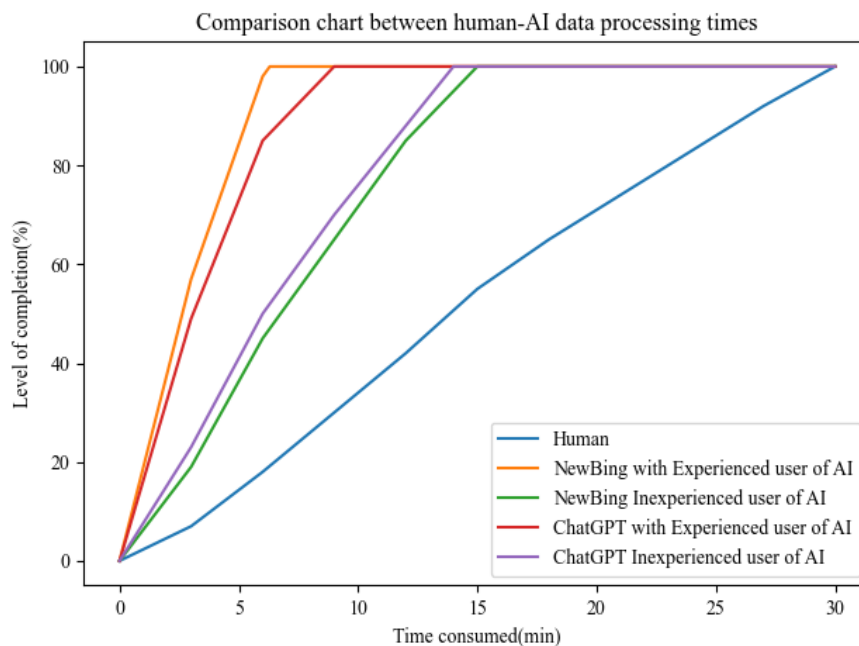


Correlation measurement	PMCC	p-value
Human	0.90	1.17e-5
ChatGPT	0.70	4.4e-3
NewBing	0.57	0.03

Table 3: Measurement of the strength of the relationship between data and evaluated projects and assessment of the stability of the correlation

There is no question that humans and AI have comparable analytical patterns, and that both have a strong linear relationship between the relevance of the data and the evaluation standards. AI is capable of making judgements that are more stable than those made by humans when the correlation between the data and evaluation criteria is strong. In contrast, when there is little correlation between the data and the assessment criteria, people make more reliable judgements.

We examine the possible effects of each news item on different aspects of a company when using news content for business forecasting. We do this both by humans and AI, and we assess the AI's conceptualization and analytical reasoning abilities. We included direct news that could have an impact on a company's financial



success in the test, such as "Hanvon Technology announces the launch of ChatGPT to support creative writing." We should be aware that forecasting the future unavoidably involves some level of uncertainty. Instead of assessing the accuracy of the AI's future analysis, our goal is to compare the AI's path and angle of "thinking" with that of humans. Our study's objective is to examine how artificial intelligence (AI) improves and replaces human analysis of data. Our study and analysis still start with humans, and comparing humans and AI is what makes this research significant.

We chose five news stories to analyse for the potential effects on a company's financial success using AI and data analysts:

- Report1: Report1: The United States has decided to continue the 301 tariff on Chinese goods.
- Report2: Report2: The Ministry of Science and Technology, along with five other departments, has issued the "Guiding Opinions on Accelerating Scenario Innovation to Promote High-level Application of Artificial Intelligence and High-quality Economic Development".
- Report3: Hanwang Technology's direct competitor, the Hong Kong listed company SenseTime-W (00020), experienced a sharp drop in stock price at today's opening, with a decline of up to 47.79
- Report4: The United States and South Korea have successively introduced new regulations on AI governance, strictly limiting, and even prohibiting the development and application of related artificial intelligence technologies in certain scenarios.
- Report5: Hanwang Technology has announced the introduction of ChatGPT to assist in content creation.

Effective Issues	Report 1 AI / Human	Report 2 AI / Human	Report 3 AI / Human	Report 4 AI / Human	Report 5 AI / Human
Revenue	-2 / -1	1 / 2	0 / 2	-1 / -1	1 / 1
Net profit margin	-1 / -1	2 / 1	0 / 0	-1 / -1	1 / 3
Cash flow	-1 / -1	1 / 1	0 / 1	-1 / -1	1 / 1
Market share	-1 / -1	1 / 3	0 / 1	-1 / -2	1 / 3
Sales growth rate	-1 / 0	1 / 0	1 / 3	-1 / -1	1 / 2
Brand awareness	0 / 0	1 / 0	0 / 0	0 / 0	1 / 1
Market competitive advantage	-1 / 0	1 / 0	1 / 1	-1 / -1	1 / 1
Business strategy	-1 / 0	1 / 2	-1 / -1	-1 / 0	0 / 1
Technological innovation capability	0 / 0	2 / 3	0 / -1	-1 / 0	1 / 1
Environmental protection measures	0 / 0	1 / 0	0 / 0	0 / 0	-1 / 1
Aspect amount	7 / 4	10 / 6	3 / 7	8 / 6	9 / 10

Table 4: Table for Event Impact Assessment

2.3 Limitations

2.3.1 Data Collection Restrictions

The findings reached in this paper may have some flaws due to time restrictions and challenges with using AI during the data gathering process. When a company is freshly listed, for example, its online information may be incredibly deficient, and on-site visits are the only way to get financial information. Furthermore, it is still unknown how well the scoring framework works, how fully operational performance can be broken down into data and scores, and whether or not some types of events or data are purposefully ignored in order to avoid conducting a thorough evaluation.

There is also the problem of the study group. Although the evaluation team has a professional bias towards business, which helps them to accurately assess the performance of the company, this group is not diverse in terms of their occupations, which could result in theoretical, singular, and biased assessments of the company. We also recognise that the evaluation team should have adequate political insight and foresight in addition to knowledge or experience in business data analysis in order to produce results for each judgement that are more

accurate. We cannot promise that our evaluation is adequately accurate and diverse in these hard-to-quantify evaluation areas.

The problem of LLM's physical memory constraints is another concern. Even though we have chosen a large data threshold, i.e. the minimum data amount, in the data processing experiment using the DuPont analysis method as an example, we still cannot guarantee whether there will be data loss due to the rollback memory limit in some situations requiring extremely large amounts of data because LLM's dialogue rollback has a certain upper limit at the physical level (4). That is, AI presently lacks the capability to divide tasks into subtasks and may experience issues like logical errors when dealing with large amounts of data. Our findings, however, are unquestionably highly credible in terms of general data analysis.

2.3.2 Substantive Psychiatric Impacts

The subjective nature of grading is a problem that cannot be avoided when gathering data. Despite our choosing of an assessment team, there may still be some one-sidedness because of the cognitive phenomenon of comparative cognition in psychology⁶, which means that the evaluators' personal experiences will unavoidably influence the scores when quantitatively scoring the company's operational performance.

Additionally, the coherence and commitment principles may have an impact on the data iteration in section 2.2.2.⁷. We might inescapably run into some evaluation error when assessing the company's operational success in 2022. It is important to keep in mind, though, that if we switched evaluation teams, the thorough research performed by various evaluation teams might be even more prone to random errors. The direction of our qualitative conclusions based on the data's trend have a high degree of credibility despite the possibility that our quantitative data may not be correct.

Future research is necessary to confirm these findings, and it should focus on the different limitations that are presented here.

3 Conclusion

Our research shows that AI has greatly increased business analysis productivity and to some extent replaces skilled human analysts. However, our evaluation of the potential for GPT to replace human analysis of business data is limited because it did not completely take into account all factors, multitasking, and the intricate and extensive business data correlations that could greatly affect analysis findings. Long-term, multidimensional, and cross-disciplinary productivity growth may also result from the advancement of GPT and its auxiliary uses. By emphasising the significant advancement that mastering this universal technology can bring to business analysis through the LLM, this paper also aims to highlight the substantial transformation that this general-purpose technology (GPT) may bring to society.

3.1 AI's Capability for Data Collection

3.1.1 AI could replace humans in data collection

According to the findings of our study, the efficiency of human data collection can be greatly increased (by over 50%) by GPT-based AI. Although there may be instances where ambiguous data needs manual verification, LLM is still the preferred method for data gathering.

⁶ comparative cognition is the comparative study of the mechanisms and origins of cognition in various species, and is sometimes seen as more general than, or similar to, comparative psychology. en.wikipedia.org/wiki/Comparative_cognition.

⁷ According to the "Influence" book, the principle of commitment and consistency refers to people's propensity to identify with the promises they make, even when they are aware that they are unwise. Since the commitments made have unconsciously influenced their thoughts, this results in a wish for consistency in their actions.

Furthermore, it can be deduced that data collection combined with GPT will have a significant impact on current methods in the future and may also change the allocation of the labour market because of the strong linear correlation between human proficiency in GPT and productivity improvement within a certain range of data volume [10]. Knowledge of GPT could become crucial in this area.

3.1.2 Techniques for gathering AI data

Although GPT has not yet developed specialised methods for precise commercial data collection, this paper provides a way of using AI for data collection given its significant efficiency in the production process.

Our investigation into ChatGPT revealed that it has a very constrained capacity for data searching in a single input. The accuracy of the information given by ChatGPT may suffer significantly if there is an excessive demand for information in a single conversation, and even the majority of the information may need to be manually verified and modified. On the other hand, there are two proper ways to operate, one of which is to direct ChatGPT to find outcomes through a series of quick searches. One is provided as follows:

Author: Are you familiar with Hanvon Technology in China?

ChatGPT: Yes, Hanvon Technology is a well-known provider of intelligent solutions in China, specializing in fields such as intelligent text recognition, speech recognition, and artificial intelligence. Its products and technologies are widely used in government, finance, logistics, medical and other fields. Hanvon Technology was founded in Nanjing, China in 1998 and has since developed into one of China's largest providers of intelligent solutions.

Author: From which websites can I obtain information about its financial reports?

ChatGPT: You can obtain information about Hanvon Technology's financial reports from the following websites: Shanghai Stock Exchange official website: <http://www.sse.com.cn/>. Hanvon Technology is a company listed on the Shanghai Stock Exchange, and you can find the company's announcements, financial reports, stock prices and other information on this website. Eastmoney.com: <http://www.eastmoney.com/>. This is one of China's better-known financial news websites, and you can find Hanvon Technology's financial data, company news, industry analysis and other information on this website. Tonghuashun Finance: <http://finance.thsi.cn/>. This is also a professional financial website, which provides basic company data, company announcements, and industry research information. Please note that the above websites are for reference only, and specific information should be based on official information released by the company.

Author: Can you provide the company's net profit for 2021?

ChatGPT: As an AI model, I cannot directly obtain the latest financial data. However, according to the financial data disclosed in Hanvon Technology's 2021 annual report, the company achieved a net profit of RMB 472 million in 2021. It should be noted that financial data may change due to accounting standards, audits, and other factors. Specific data should be based on the official financial reports released by the company.

However, we do not advise using ChatGPT as a reliable data provider. Currently, it has a negligible effect on how effectively direct data gathering works and even poses a risk of errors. In one experiment, we discovered that the accurate information we were looking for was not provided, but rather the official website of another business with a like spelling.

Author: From which websites can I obtain information about its financial reports?

ChatGPT: You can obtain information related to Hanwang Technology's financial reports from the company's official website's investor relations page: <http://ir.hwamei.com/>

As a GPT trained on search engines, NewBing, a popular large-scale language model, has demonstrated

outstanding data search performance. It also has some issues with data linkage, though. We advise beginning with data structure when using NewBing as a data gathering tool, choosing the necessary data structure and analytical strategy beforehand, and expressing the data requirements plainly and directly. Please send me Hanwang Technology's financial information for 2021, including cash flow, net profit, and net profit percentage.

It should be mentioned that when formulating data structures, using multiple GPT models can significantly increase efficiency and guarantee accuracy. For ChatGPT, as an illustration: Please list the five key financial metrics for gauging a business' operational success. For NewBing: I need precise financial information for Hanwang Technology in 2021 for a particular indicator. It is important to remember that some of the data in this paper were obtained using this technique as well.

It is crucial to note that GPT 4.0 has shown remarkable learning capacity and rapid growing capacity in the most recent testing, almost reaching the level of divine calculation. As a result, the conclusions reached in this study about the application of artificial intelligence to data collecting and analysis may have some temporal relevance, and the quantitative information contained in the conclusions and answers may vary significantly in the future. As a result, this research concentrates more on qualitatively evaluating the likelihood that AI will eventually replace humans and emphasizing the enormous potential of GPT in subsequent work.

3.2 The Impact of AI in Data Processing and Business Forecasting

3.2.1 Data Processing Ability of AI

According to our study's findings, AI is superior to humans in some "zero-based" data analysis tasks because it can quickly and precisely search for relevant information. Additionally, LLM can significantly improve data processing efficiency for humans with limited business data analysis knowledge and experience (by 70%) due to its extensive and in-depth grasp of the majority of business data methods.

Additionally, the data used for particular business data analysis is typically accurate and precise; as a result, AI has demonstrated significant potential for increasing efficiency in our analysis examples.

For example, when analysing the DuPont method, AI only produced 100 percent accurate findings when given the right information and the name of the method. Despite the possibility of imprecision issues brought on by large amounts of data (see 2.3), data analysts do not even need the necessary skilled expertise. They require the capacity to match the method name to the measurement content. In other words, process-oriented knowledge structures may be questioned and replaced by object-oriented ones as business data analysis progressively. In other words, we only need to focus on learning how to drive in order to analyse business statistics in the future.

However, it is clear that AI falls short when it comes to economic forecasting. Although AI appears to approach problems from a broader viewpoint than humans do, its predictions differ significantly from human predictions in both aspects and degrees. In other words, AI only lists the options because it lacks the awareness and tools to recognise the "focus".

3.2.2 Methodology for AI Data Processing

Similar to GPT's success in data gathering, GPT has been able to significantly increase data analysis efficiency despite not having created the complementary applications necessary for commercial analysis. As a result, a methodology for applying AI to data analysis is also presented in this article.

A general methodology is given here because the two GPTs have comparable data processing input processes when using particular methods.

Example: Do you have any experience with DuPont analysis? Please compute the company's ROE using the DuPont analysis technique given [data input].

It should be mentioned that, as fully detailed in section 2.3, errors might occur when processing large amounts of data due to GPT's physical memory limitation. Similar to this, GPT's performance is still precise

and accurate when defining methods that we are unfamiliar with, even though we employ a similar strategy by defining methods with high complexity thresholds. However, given these methods' high levels of complexity, we cannot ensure their applicability. However, the techniques described here have a lot of application in general data analysis.

4 Discussion

4.1 It is anticipated that GPT will be widely used in business data analysis.

We discussed the use of GPT in the particular area of commercial data analysis in the literature review, as well as any possible changes in accuracy or efficiency that it might bring. This implies that it must also have matching complementary applications and expandable user interfaces. After the release of ChatGPT, the effectiveness of UI enhancements on LLM became clear because ChatGPT far outperformed the use of GPT2 and GPT3 API interfaces.[17] Therefore, it is expected that as software and digital tools are developed to complement and innovate GPT, it will have a significant effect on a variety of economic activities, including commercial analysis.

However, there are still practical obstacles preventing the broad adoption of these models. The degree of confidence and acceptance people have in them, as well as their related knowledge bases and usage patterns, are important determinants of their usefulness. For instance, data analysts who have used AI show a preference for using it for data collection and processing and selectively accepting or rejecting analysis results based on the analysis process in the commercial analysis discussed in this article. On the other hand, data researchers who are unfamiliar with AI tend to be sceptical and unaccepting of its analysis and processing.

4.2 The effect of GPT on the market for business analysis

4.2.1 The entry bar is lowered by GPT

The entry requirements for business data analysts can be greatly lowered thanks to GPT, a representative of natural language processing technology, as was stated in the previous section. Business analysts need to have a certain level of expertise in economics, statistics, and computer science as well as related skills in data gathering and processing in order to perform conventional business analysis. However, as AI capabilities advance over time and become capable of handling more challenging tasks and test cases, it is possible that they may have a specific and significant effect on business data analysis, particularly on its entry standards.

We predict that the standards for expert knowledge and data processing skills, which are two crucial competencies in conventional business data analysis, will be significantly lowered in the market for business analysis in the future. In contrast, the ability to assess data accuracy and dependability as well as the creation and application of AI analysis tools may end up being the deciding elements in establishing the entry requirements for the future business data analysis market.

4.2.2 The GPT modifies the information structure.

Beyond that, the general knowledge system of commercial data analysis may further transition from learning to application as a result of the changes in industry admission standards brought about by GPT. The knowledge structure needed for applied commercial data analysis will eventually change from the current approach, which emphasises process and methodology and necessitates corresponding knowledge of mathematical and scientific disciplines, to a logic- and object-oriented approach, as stated in section 3.2.1. To put it another way, the ability to directly apply knowledge to achieve desired results will take precedence over the manipulation

of specific knowledge in future business data analysis. The core components of the future knowledge system of business analysis will revolve around effectively mastering potent tools like GPT and drawing logical inferences, conclusions, and generalisations from them.

5 Acknowledgment

We thank Lili Zhou from China Mobile and Xin Yang from the Tax Bureau for their extensive involvement in the study as data researchers as well as their unbiased and trustworthy evaluation. We also acknowledge Ziyao Yi (Nanjing University) for encouraging us to investigate statistical techniques and Desu Liu (Nanjing University) for giving technical assistance.

We are grateful to Yulong Song (Nanjing University) for helping us develop our writing and analytical abilities in English. In our English language change process, Zhuochen Liu (Beijing University of Posts and Telecommunications) performed a significant part. We appreciate Jie Zhang's (Nanjing University) advice and recommendations regarding the use of LaTeX.

We would like to extend a special thank you to Yi Liu (Lanzhou University) for improving our work and pointing out its flaws. We also thank the OpenAI team, in particular ChatGPT, for their assistance with the translation process by giving precise descriptions of technical terms and assisting with other editing duties.

References

- [1] OpenAI (2022). Introducing chatgpt. OpenAI (2023a). Gpt-4 system card. Technical report, OpenAI. OpenAI (2023b). Gpt-4 technical report. Technical report, OpenAI.
- [2] Tyna Eloundou¹, Sam Manning^{1,2}, Pamela Mishkin^{□1}, and Daniel Rock³ ¹OpenAI ²OpenResearch ³University of Pennsylvania (2023) GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models. arXiv:2303.10130
- [3] Thoppilan, R., De Freitas, D., Hall, J., Shazeer, N., Kulshreshtha, A., Cheng, H.-T., Jin, A., Bos, T., Baker, L., Du, Y., et al. (2022). Llama: Language models for dialog applications. arXiv preprint arXiv:2201.08239.
- [4] Agrawal, A., Gans, J., Goldfarb, A. 2022. "ChatGPT and How AI Disrupts Industries" Harvard Business Review. Available: <https://hbr.org/2022/12/chatgpt-and-how-ai-disrupts-industries>
- [5] Felten, E., Raj, M., and Seamans, R. (2023). How will language modelers like chatgpt affect occupations and industries? arXiv preprint arXiv:2303.01157.
- [6] Abid, A., Farooqi, M., and Zou, J. (2021). Persistent anti-muslim bias in large language models. In Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society, AIES '21, page 298–306, New York, NY, USA. Association for Computing Machinery.
- [7] Schramowski, P., Turan, C., Andersen, N., Rothkopf, C. A., and Kersting, K. (2022). Large pre-trained language models contain human-like biases of what is right and wrong to do. *Nature Machine Intelligence*, 4(3):258–268
- [8] Goldstein, J. A., Sastry, G., Musser, M., DiResta, R., Gentzel, M., and Sedova, K. (2023). Generative language models and automated influence operations: Emerging threats and potential mitigations
- [9] 杜秀英. 基于 Pearson 相关分析的期刊引用关系研究 [J]. 科技文献信息管理, 2012(2):18-2
- [10] Felten, E. W., Raj, M., and Seamans, R. (2018). A method to link advances in artificial intelligence to occupational abilities. *AEA Papers and Proceedings*, 108:54–57.
- [11] Yuntao Bai*, Andy Jones, Kamal Ndousse,(2022) Training a Helpful and Harmless Assistant with Reinforcement Learning from Human Feedback. arXiv:2204.05862v1 [cs.CL] 12 Apr 2022
- [12] Goldfarb, A., Taska, B., and Teodoridis, F. (2023). Could machine learning be a general purpose technology? a comparison of emerging technologies using data from online job postings. *Research Policy*, 52(1):104653.
- [13] Grace, K., Salvatier, J., Dafoe, A., Zhang, B., and Evans, O. (2018). When will ai exceed human performance? evidence from ai experts. *Journal of Artificial Intelligence Research*, 62:729–754.

- [14] Tolan, S., Pesole, A., Martínez-Plumed, F., Fernández-Macías, E., Hernández-Orallo, J., and Gómez, E. (2021). Measuring the occupational impact of ai: tasks, cognitive abilities and ai benchmarks. *Journal of Artificial Intelligence Research*, 71:191–236.
- [15] Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., Sutskever, I., et al. (2019). Language models are unsupervised multitask learners. *OpenAI blog*, 1(8):9. 34 ResumeBuilder.com (2023). 1 in 4 companies have already replaced workers with chatgpt.
- [16] Rock, D. (2019). Engineering value: The returns to technological talent and investments in artificial intelligence. Available at SSRN 3427412.
- [17] Chow, A. R. (2023). Why ChatGPT Is the Fastest Growing Web Platform Ever | Time
- [18] Webb, M. (2020). The impact of artificial intelligence on the labor market. Working paper, Stanford University
- [19] Weidinger, L., Uesato, J., Rauh, M., Griffin, C., Huang, P.-S., Mellor, J., Glaese, A., Cheng, M., Balle, B., Kasirzadeh, A., Biles, C., Brown, S., Kenton, Z., Hawkins, W., Stepleton, T., Birhane, A., Hendricks, L. A., Rimell, L., Isaac, W., Haas, J., Legassick, S., Irving, G., and Gabriel, I. (2022). Taxonomy of risks posed by language models. In 2022 ACM Conference on Fairness, Accountability, and Transparency, FAccT ' 22, page 214–229, New York, NY, USA. Association for Computing Machinery
- [20] Brynjolfsson, E., Mitchell, T., Rock, D. (2018). What can machines learn, and what does it mean for occupations and the economy? *AEA Papers and Proceedings*, 108, 43–47. <https://doi.org/10.1257/pandp.20181019>
- [21] Genz, S., Gregory, T., Janser, M., Lehmer, F., Matthes, B. (2021). How do workers adjust when firms adopt new technologies?. ZEW-Centre for European Economic Research Discussion Paper, (21-073).