CAN ARTIFICIAL INTELLIGENCE (AI) TECHNOLOGIES ESTIMATE FINANCIALS ACCURATELY? A RESEARCH ON BORSA ISTANBUL WITH CHATGPT

YAPAY ZEKA (AI) TEKNOLOJİLERİ FİNANSALLARI DOĞRU BİR ŞEKİLDE TAHMİN EDEBİLİR Mİ? CHATGPT KULLANILARAK BORSA İSTANBUL ÜZERİNDE BİR ARAŞTIRMA

İstemi ÇÖMLEKÇİ[1], Serkan ÜNAL[2], Ali ÖZER[3], Mehmet Akif ÖNCÜ[4]

ÖZ: Bu makalede ChatGPT'nin geçmiş finansal tabloları ve sektörle ilgili güncel önemli gelişmeleri yorumlayarak finansal değerleri ne ölçüde doğru tahmin edebildiği analiz edilmiştir. Araştırma kapsamında BIST100 endeksinde işlem gören 73 şirkete ait finansal veri ve Microsoft BING Chat uygulamasından temin edilen sektör spesifik gelişmelerin özetlendiği metinlerden faydalanılmıştır. İlk aşamada ChatGPT'nin 2019-2021 dönemine ait finansal değerleri yorumlayarak 2022 yılı için satış, net kâr ve temettü dağıtımı tahminleri yapması istenmiştir. İkinci aşamada ChatGPT'ye sektör spesifik gelişmelerle ilgili metin girilmiş ve tahminini yenilemesi istenmiştir. Daha sonrasında bu iki aşamada ChatGPT tarafından yapılan tahminler gerçekleşmeler ile kıyaslanmış ve ChatGPT'nin bu tahminlerde ne ölçüde başarılı olduğu anlaşılmaya çalışılmıştır. Araştırma sonuçlarına göre ChatGPT satış değerlerini yalnızca finansal tablolardan faydalandığı ilk aşamada medyan değerler olarak %18 sapma ile tahmin edebilmiş, sektörle ilgili güncel gelişmeleri kullandığı ikinci aşamada ise %18,9 sapma ile tahmin edebilmiştir. Net kâr için ilk aşamada yapılan tahminlerin medyan sapması %57,8 olurken ikinci aşamada %64,9 olmuştur. Kâr payları için medyan sapmalar ilk aşamada %54,6 iken ikinci aşamada %51,8 olmuştur. Araştırma bulgularına göre ChatGPT'nin finansal ve haber analiz yetenekleri sınırlı olmakla birlikte ciddi bir gelişme süreci yaşadığı gözlemlenmiştir.

Anahtar Kelimeler : Yapay Zeka Teknolojileri, ChatGPT, Microsoft BING, Borsa İstanbul

ABSTRACT: This article analyzes to what extent ChatGPT can accurately estimate financial values by interpreting past financial statements and current important developments in the sector. Within the scope of the research, the financial data of 73 companies traded in the BIST100 index and the texts that summarize the sector-specific developments obtained from the Microsoft BING Chat application were used. In the first stage, ChatGPT was asked to interpret the financial values for the 2019-2021 period and make estimates for sales, net profit, and dividend distribution for 2022. In the second stage, text about sector-specific developments was entered into ChatGPT and it was asked to revise its estimate. Afterward, the estimations made by ChatGPT in these two stages were compared with the realizations and it was tried to understand how successful ChatGPT was in these estimations. According to the results of the research, ChatGPT was able to estimate the sales values with 18% deviation as median values in the first stage, when it used only the financials, and with 18.9% deviation in the second stage, when it used both financials and the current developments in the sector. The median deviation of the estimates for net profit in the first phase was 57.8%, while it was 64.9% in the second phase. The median deviations for dividends were 54.6% in the first stage and 51.8% in the second stage. According to the research findings, it has been observed that ChatGPT has experienced a serious development process, although its financial and news analysis capabilities are limited.

Keywords: *Artificial Intelligence (AI) Technologies, ChatGPT, Microsoft BING, Borsa İstanbul.*

1. Doç.Dr., Düzce Üniversitesi, istemicomlekci@duzce.edu.tr, https://orcid.org/0000-0001-8922-071X
2. Doç.Dr., Ufuk Üniversitesi, serkan.unal@ufuk.edu.tr, https://orcid.org/0000-0002-7060-979X
3. Doç.Dr., Düzce Üniversitesi, aliozer@duzce.edu.tr, https://orcid.org/0000-0003-4736-3418
4. Prof.Dr., Düzce Üniversitesi, mehmetakifoncu@duzce.edu.tr, https://orcid.org/0000-0002-4557-4214

International Journal of Business, Economics and Management Perspectives - Uluslararası İşletme, Ekonomi ve Yönetim Perspektifleri Dergisi
Volume/Cilt: 7, Issue/Sayı: 1, pp/ss. 1-14
June/Haziran 2023

GENİŞLETİLMİŞ ÖZET

Çalışmanın Amacı

Yapa zekâ uygulamalarının işletmelerinin geçmiş dönem finansal tablolarını ve sektördeki güncel gelişmeleri yorumlayarak, finansal tablo kalemlerini tahmin edebilme yeteneğinin belirlenmesi bu çalışmanın amacını oluşturmaktadır.

Araştırma Soruları

ChatGPT bir işlemtenin geçmiş dönem finansal tablolarını dikkate alarak satışlarını, net karını ve temettü miktarını tahmin edebilir mi? Güncel sektör bilgileri ile yorumu revize etmesi istenirse daha tutarlı bir tahmin yapabilir mi?

Literatür Araştırması

ChatGPT'nin (Chat Generative Pre-trained Transformer) Kasım 2022'de OpenAI tarafından kullanıma sunulmasıyla insan-makine etkileşiminin önemli bir örneği ortaya konmuştur. Bu uygulamada insan-insan etkileşimine benzer iletişimin çoklu dil desteği ile yakalanması amaçlanmaktadır. İnsanlarla sohbet etmek ve sorularını cevaplamak için tasarlanan ChatGPT, hızlı metin üretebilme özelliği ile dikkat çekmektedir. Biswas (2023), ChatGPT'nin iklim tahmininde sınırlı da olsa etkili olduğunu; Lund ve Wang (2023), kütüphanecilik ve akademik araştırmalarda kullanılabileceğini, Frieder ve ark. (2023), matematiksel nesneleri arama yeteneğinin iyi olduğunu, Choi ve ark. (2023) hukuk fakültesindeki 4 dersin final sınavlarından geçer not almayı başardığını belirtmişlerdir. Dowling ve Lucey (2023), ChatGPT'nin finansal araştırmalardaki yeteneklerini incelemişler ve fikir üretme ve veri tanımlama yeteneklerine sahip olduğunu, ancak literatürü sentezleme ve uygun analizi belirleme konusunda zayıf olduğunu belirtmişlerdir.

Yöntem

Araştırma kapsamında BIST100 endeksinde işlem gören işletmelerin 2019-2021 dönemini kapsayan finansal verileri girdi olarak 2022 finansal verileri ise karşılaştırma ölçütü olarak kullanılmıştır. Araştırmada sektöre özel güncel gelişmeleri derlemek için Microsoft Bing'in Chat platformundan yararlanılmıştır. Araştırmada incelenen şirketlere ilişkin 2019-2021 dönemini kapsayan mali veriler ChatGPT 3.5 programı ile paylaşılmış ve ChatGPT'den şirketlerin 2022 yılı satış, net kâr ve temettü dağıtım tutarı tahmini istenmiştir. Bu tahminler ilk aşamada ChatGPT tarafından yapılan bu tahminler kaydedilmiştir. Bununla beraber Microsoft Bing'in Chat platformundan firmaların yer aldığı sektörlere ilişkin 2022'deki güncel gelişmeleri özetlemesi istenmiştir. Daha sonra, ChatGPT'den sektörle ilgili gelişmeleri dikkate alarak tahminlerini tekrar etmesi istenmiştir. İlk tahmin ve ikinci tahmin sonuçları karşılaştırılmıştır.

Results and Conclusions

Araştırma sonuçlarına göre ChatGPT sadece finansal tabloları kullandığı ilk aşamada satış değerini %18 sapma ile, sektör güncel gelişmeler ışığında %18,9 sapma ile tahmin edebilmiştir. Net kâr tahminlerinin ilk aşamadaki medyan sapması %57,8 olurken, ikinci aşamada %64,9 olmuştur. Temettüler için medyan sapmalar ilk aşamada %54,6 ve ikinci aşamada %51,8 oldu. ChatGPT ile yapılan çalışmada sektörel gelişmelerin ve tarihsel finansal değerlerin birlikte yorumlandığı süreç finansal tahmin sürecini iyileştirememiştir. Ancak ChatGPT'nin verilen metinleri ve finansal tabloları anlayıp yorumlayabildiği görülmüştür. Araştırmadan elde edilen bulgulara göre ChatGPT finansal tabloları yorumlayabilse de tahminlerinde yanılabiliyor. Öte yandan, sektöre özel bilgilerin yorumlanması dahil edildiğinde ChatGPT'nin tahminlerinin iyileşmediği gözlemlenmiştir. ChatGPT ile her görüşmede farklı yöntemler denenebilir ve farklı sonuçlar bulanabilir. ChatGPT'ye sorulan sorunun netliği alınan yanıtı etkileyebilir. Sorunun sorulma şekline bağlı olarak, bazı durumlarda ChatGPT, yeterli veriye sahip olmadığı gerekçesiyle yanıt vermekten kaçınabilir. ChatGPT'ye büyük miktarda veri girilmesi durumunda, ChatGPT'nin sağlanan verilerin içeriğini anlamakta güçlük çekebileceği gözlemlenmiştir. Ayrıca ChatGPT'nin sürekli olarak geliştirildiği göz önünde bulundurulduğunda gelecekte çok daha iyi sonuçların alınabileceğini belirtmek gerekir. Bu konuda ileride yapılacak çalışmalarda ChatGPT ve diğer yapay zeka araçlarının farklı veri kaynaklarını kullanabilme ve yorumlayabilme becerileri farklı boyutlarıyla test edilebilir.

1. INTRODUCTION

Information, scientific data, and their interaction with each other, which continues to increase with the effect of technology, make decision-making processes and analyzes used in decision-making more complex. In fact, solutions to problems that have become more complex with technology are sought with technological approaches. Artificial intelligence (AI), which has an important place among these technological approaches, is supported by advanced techniques in the analysis of big data. Especially since the 2000s, artificial intelligence investments and technologies based on artificial intelligence have come to the fore. The most important purpose of artificial intelligence is to create high-capacity machine intelligence by imitating human intelligence, and thus to reach information faster, to classify and analyze complex information, and to make predictions with the information obtained. With these developments, artificial intelligence has begun to be used in many different areas (Mikalef and Gupta, 2021:1-3; Davenport and Ronanki 2018:111-114).

Looking at the history of artificial intelligence, it is seen that it is based on the machine learning experiment first made by Alan Turing in 1950 (Ramesh, 2004:334). The term artificial intelligence was first brought up by McCarthy (1956) at a conference. In the 1960s, studies on artificial intelligence-based chatbots were concentrated. Thus, chatbots named Eliza in 1966 and Parry in 1972 were developed and Parry was the first to pass the Turing test. After that, although there were some initiatives until the 1990s, reasons such as the need for high budgets, difficulties in transforming theoretical knowledge into practice and the inability to achieve the desired results caused the development of artificial intelligence to not be at the desired level. Artificial intelligence initiatives that have been popular since the 1990s, especially with the A.L.I.C.E developed in 1995 and the Deep Blue computer developed by IBM in 1997, managed to attract attention once again after defeating world chess champion Garry Kasparov. Especially since the 2000s, more funds have been transferred to artificial intelligence studies. Siri, which was used by Apple in 2011, Google assistant, which was introduced by Google in 2012 and continued its development and put into use in 2016, Alexa, which was used by Amazon in 2014, and Bixby, developed by Samsung in 2017, became the most important products of artificial intelligence initiatives (Cheung et al., 2017: 229; Bory, 2019:629-630; Goel and Ganatra 2021:739).

Chatbots, the foundations of which were laid in the 1960s, have become popular again with the development of artificial intelligence since the 2000s. Every day, we come across applications with more advanced and new features in both audio format and text format. The aim is to create a feeling of human presence in human-machine interaction. Studies such as Carvalho and Ivanov (2023), Dowling and Lucey (2023) stated that chatbots are not sufficient to replace human experience, but they have important capabilities. Chatbots are used not only for entertainment purposes but also for business purposes in many fields such as education, health, e-commerce, finance, entertainment and tourism, and

International Journal of Business, Economics and Management Perspectives - Uluslararası İşletme, Ekonomi ve Yönetim Perspektifleri Dergisi
Volume/Cilt: 7, Issue/Sayı: 1, pp/ss. 1-14
June/Haziran 2023

their various abilities are utilized (Shawar and Atwell, 2007: 43-44). For this reason, many companies are making plans for chatbots, and each company allocates funds for its own chatbots. Especially financial institutions seem to be more active in this regard. Artificial intelligence-based chatbots are becoming increasingly important in terms of digital transformation and technological competition in the financial sector, and they already have an important place in areas such as banking, financial services and insurance. Artificial intelligence is used by many investors and financial institutions in risk assessment, financial analysis, investment decisions and financial forecasts. In 2017, Bank of America, a chatbot named Erica, where they can get help with banking transactions, and AmEx, the customer service chatbot of American Express in the same year, are among the firsts in this field (Hwang and Kim, 2021:2-3).

Although there are different classifications in terms of types of chatbots, we can divide them into different types according to their technological infrastructure and abilities. Menu-based chatbots are the simplest and least used artificial intelligence in terms of technology. These applications mostly serve in a certain field such as answers to frequently asked questions, and are preferred by companies in the field of customer relations, although they do not have the ability to answer different questions. (Li et al., 2020: 8). Keyword-based chatbots are more capable than menu-based chatbots. They have the ability to interpret and respond within the framework of certain keywords. Finally, there are content-based chatbot applications. These applications are the types in which artificial intelligence and machine learning are used. When these chatbots are examined in terms of artificial intelligence, it is possible to divide them into two types as rule-based chatbots based on predefined algorithms and self-learning chatbots. In particular, chatbots of the second type can answer questions at a certain level, thanks to advanced algorithms including machine learning. In addition, different algorithms have managed to attract attention with the possibility of natural language processing (NLP) with artificial intelligence. In addition, this feature of Chatbots has increased the interaction between human and machine (Wu et al., 2016:1-2; Adamopoulou and Moussiades 2020:378-379).

Along with these developments, one of the most remarkable developments has been the chatbots, which include artificial intelligence-based deep machine learning algorithms over multiple data and natural language processing (NLP) is possible. An important example of human-machine interaction was demonstrated with the introduction of ChatGPT (Chat Generative Pre-trained Transformer) by OpenAI in November 2022. In this application, it is aimed to capture communication similar to human-human interaction with multi-language support. Although ChatGPT, which is designed to chat with people and answer their questions, attracts attention with its ability to produce fast text, it also has limitations such as limited knowledge of events after 2021 and limited analysis ability (Zielinski et al., 2023:76). However, it is expected that over time, their skills in this field will improve and the problems will decrease. With the ability to customize chats to suit their specific needs and the ability to

communicate in a variety of languages, ChatGPT is an essential tool for businesses and individuals (Singh, 2023). In its current form, ChatGPT has been examined from various perspectives in many areas and has managed to attract the attention of many sectors.

Biswas (2023) stated that ChatGPT is effective, although limited, in climate prediction; Lund and Wang (2023) stated that it can be used in librarianship and academic research, but ethical problems should be considered; Frieder et al. (2023) found that its ability to search for mathematical objects is good but weak in advanced mathematical analysis; Choi et al. (2023) asked ChatGPT to answer the final exams of 4 courses in Minnesota law school and as a result ChatGPT managed to get a low but passing grade. Dowling and Lucey (2023) examined the capabilities of ChatGPT in financial research, and stated that it had capabilities in generating ideas and defining data, but was weak in synthesizing literature and determining appropriate analysis. As in these studies, many studies have focused on ChatGPT's capabilities in different areas. One of the curious issues is ChatGPT's ability to analyze and forecast using existing financial data.

Financial markets are at the forefront of the areas where AI tools such as ChatGPT can be used. In the financial markets, a fast and large amount of data flows live and this data flow can change the pricing of tens of thousands of financial instruments instantly. Therefore, an artificial intelligence instrument that can analyze and correlate data quickly without human help will be able to carry the size of competition in financial markets to a very advanced level. Therefore, for this purpose, ChatGPT's ability to read and interpret financial statements and its ability to associate industry-specific information were analyzed. It is thought that this study will contribute to the literature in this rapidly developing field.

In this study, it is aimed to measure the financial estimation skills of ChatGPT, using financial data. Various financial data of companies traded in Borsa Istanbul were given to ChatGPT and it was requested to make financial forecasts in the light of these data. Then, the BING application was asked to provide information about the sector in which the company is operating, and it was requested to revise the financial estimate of ChatGPT in the light of the sectoral information obtained from here. The financial estimation ability of ChatGPT was revealed by comparing the obtained estimations with the actual figures.

2. METHODOLOGY

2.1. Data

In the study, companies traded in the BIST100 index were used to determine the financial estimation ability of ChatGPT. The two important data inputs of the study were the financial data of the companies and the sector-specific developments in 2022. Small-scale companies were not included in

the study due to the high probability of operating in some boutique areas and the limited sector-specific news flow. For this reason, the analysis of the companies in the BIST100 index was considered sufficient. Holding companies, finance companies, and investment trusts were excluded from the scope of the research, due to the low level of significance of sales values due to their field of activity. The number of companies evaluated within the scope of the research is 73. Financial data of these companies covering the period of 2019-2021 were used as input. In addition, financial data for 2022 were also used to compare the performance of financial estimations prepared by ChatGPT. In the research, Microsoft Bing's Chat platform was used to compile industry-specific current developments. By running the chat platform in creative mode, the sectors of the analyzed companies were questioned separately, and it was requested to summarize the developments in the sector in 2022.

2.2. Method

Financial data covering the period 2019-2021 regarding the companies analyzed in the research were shared with ChatGPT and ChatGPT was asked to estimate the sales, net profit, and dividend distribution amount of the companies for 2022. These estimations made by ChatGPT 3.5 in the first stage were recorded. Meanwhile, Microsoft Bing's Chat platform was asked to summarize the current developments in 2022 regarding the sectors in which the companies are located. After receiving the industry-specific development summary from Bing in text form, ChatGPT was asked to repeat its estimations considering industry-related developments. The estimation process with ChatGPT is summarized in Figure 1. This process was repeated for all companies within the scope of the research.

Figure 1. Estimation Process Applied During the Research
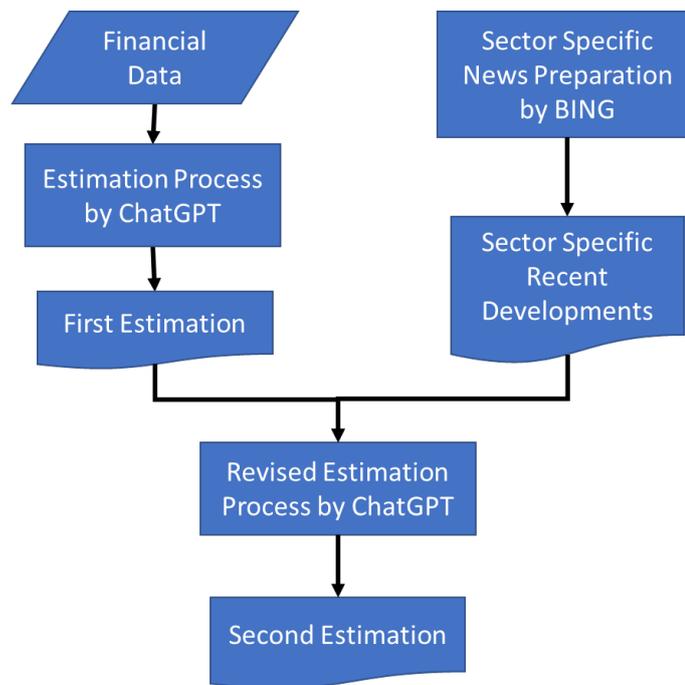
In the research, ChatGPT's ability to read and interpret financial statements, read industry-specific texts, and combine this data to effectively estimate the future was measured. In the first stage, the estimates were made using only the financial data, and in the second stage, the text showing sector-specific developments was added to the estimation process. In this way, ChatGPT's ability to combine and analyze different data together was tested.

While developing the methodology, various question texts were asked and it was observed how ChatGPT answered. Observations made during the development of the methodology are summarized below.

1- When ChatGPT was asked about its estimations at the first stage, it stated that data other than financial statements is required to respond. Apart from the financial statements, the latest situation in the industry, analyst expectations, economic growth, company expectations, and similar data were also requested by ChatGPT. In order to get an answer, ChatGPT has been told that it only needs to make estimations based on the given data.

2- It has been observed that ChatGPT focuses on directly related items while making estimations. For example, if it is making a sales estimate, it has taken into account the increasing trend of sales in previous years and the sales amount in the last year and explained in writing how it made its estimate. Although ChatGPT takes other items into account to a certain extent, when its answers are evaluated, it is observed that its ability to relate different financial items is limited.

3- When the balance sheet, income expense statement, and cash flow table were loaded at the same time, it could not understand the existence of the income and expense statement and requested the income and expense statement separately. In order to avoid this obstacle, important items in the financial statements have been simplified and uploaded to ChatGPT.

4- When the same question is asked to ChatGPT using the same data set, it has been observed that the answers and approach of ChatGPT may differ.

3.FINDINGS

In the first stage, financial data covering the period of 2019-2021 regarding the companies within the scope of the research were uploaded to ChatGPT, and values were obtained from ChatGPT for the companies' sales, net profit, and dividend distributions in 2022. Afterward, the estimations of ChatGPT were compared with the realizations, and the percentage of the deviations between realizations and the estimations was determined. In the second stage, a summary text with industry-specific developments compiled by BING was given to ChatGPT and it was requested to revise the estimation with this new information. In the second stage, the estimations produced by ChatGPT and the realizations were compared again and the deviation values were found as a percentage. The findings

International Journal of Business, Economics and Management Perspectives - Uluslararası İşletme, Ekonomi ve Yönetim Perspektifleri Dergisi
Volume/Cilt: 7, Issue/Sayı: 1, pp/ss. 1-14
June/Haziran 2023

obtained in 73 companies within the scope of the research are summarized in Table 1. Accordingly, ChatGPT was able to estimate sales with 18%, net profit with 58%, and dividend amount with 55% deviation as median values at the first stage. It is seen that there are higher deviations in the second stage in the estimations of sales compared to the first stage in terms of both mean, median values and standard deviation values. When the deviations in the net profit value are analyzed, it is seen that there is a serious correction on average in the second estimation compared to the first estimation. But this can be caused by outliers. As a matter of fact, according to the median values, the second-stage estimations are not more successful than the first-stage estimations. On the other hand, it is seen that dividend distribution amounts are estimated more effectively in the second stage. However, in general, it is seen that ChatGPT could not improve its estimations by analyzing sector-specific developments and it made estimations with similar error rates in both stages.

Table 1. Deviation Between ChatGPT's Estimates vs. Realizations (Based on 73 Companies in BIST100 Index)

Panel A - Revenues					
	Average	Median	Max	Min	Std.
1. Estimate	25.9%	18.0%	163.8%	1.2%	26.3%
2. Estimate	26.5%	18.9%	197.4%	0.2%	31.1%
Difference	0.6%	0.9%	33.6%	-1.0%	4.8%
Panel B - Net Income					
	Average	Median	Max	Min	Std.
1. Estimate	1555.8%	57.8%	105060.1%	2.2%	12199.4%
2. Estimate	273.6%	64.9%	9284.0%	0.2%	1084.1%
Difference	-1282.2%	7.1%	-95776.1%	-2.0%	-11115.2%
Panel C - Dividend					
	Average	Median	Max	Min	Std.
1. Estimate	61.3%	54.6%	156.4%	0.7%	37.1%
2. Estimate	59.4%	51.8%	187.7%	0.4%	40.2%
Difference	-1.9%	-2.8%	31.3%	-0.4%	3.1%

Note: Deviation values are found by dividing financial realizations by ChatGPT's estimate with this formula: "Financial Realization / ChatGPT's Estimate – 1"

From a sectoral perspective, some companies may be more affected by macroeconomic developments. On the other hand, it may be easier to access information about some companies and to measure their performance. In order to take these effects into account, companies are classified on the basis of their sector and divided into two separate groups, industrial companies and service companies. 43 companies were classified in the industrial companies group and 24 companies were classified in the service companies group. Since there were 73 companies within the scope of the research, it was not possible to make a detailed classification in order to preserve the statistical significance of the results. The results obtained for industrial companies are summarized in Table 2. The second stage estimates for sales did not improve when analyzed in terms of mean and median values. Estimates for net profit deteriorated as mean value but improved as median value. The estimates of the dividend amounts, on the other hand, have improved somewhat in terms of both the mean values and the standard deviation

values. However, when examined in general, it is seen that there is no significant improvement in the estimates made by ChatGPT in the second stage, similar to the analysis made for the entire sample in Table 1.

Table 2. Deviation Between ChatGPT's Estimates vs. Realizations (Based on 43 Companies in BIST Industrial Index)

Panel A - Revenues					
	Average	Median	Max	Min	Std.
1. Estimate	20.1%	13.2%	95.7%	1.2%	17.8%
2. Estimate	20.6%	16.0%	112.2%	0.2%	20.0%
Difference	0.5%	2.8%	16.5%	-1.0%	2.2%
Panel B - Net Income					
	Average	Median	Max	Min	Std.
1. Estimate	99.7%	45.8%	606.4%	2.2%	126.3%
2. Estimate	139.6%	33.0%	1136.6%	0.2%	215.7%
Difference	39.9%	-12.8%	530.2%	-2.0%	89.4%
Panel C - Dividend					
	Average	Median	Max	Min	Std.
1. Estimate	70.1%	68.5%	156.4%	4.7%	36.8%
2. Estimate	64.9%	60.1%	119.8%	9.2%	32.7%
Difference	-5.2%	-8.3%	-36.6%	4.5%	-4.0%

Note: Deviation values are found by dividing financial realizations by ChatGPT's estimate with this formula: "Financial Realization / ChatGPT's Estimate – 1"

The deviation values between the estimates made by ChatGPT for the service sector and the realizations are summarized in Table 3. Looking at the service sector in particular, it is seen that ChatGPT's ability to analyze the news flow is limited within the scope of the research. In general, it is seen that there is no significant difference between the second estimates and the first estimates in terms of the deviation of the estimates and the realizations.

Table 3. Deviation Between ChatGPT's Estimates vs. Realizations (Based on 24 Companies in BIST Services Index)

Panel A - Revenues					
	Average	Median	Max	Min	Std.
1. Estimate	35.3%	22.1%	163.8%	4.4%	36.0%
2. Estimate	37.5%	24.0%	197.4%	0.6%	44.6%
Difference	2.2%	2.0%	33.6%	-3.8%	8.6%
Panel B - Net Income					
	Average	Median	Max	Min	Std.
1. Estimate	4540.3%	70.7%	105060.1%	16.2%	20961.3%
2. Estimate	571.4%	96.2%	9284.0%	18.0%	1832.3%
Difference	-3968.9%	25.6%	-95776.1%	1.8%	-19128.9%
Panel C - Dividend					
	Average	Median	Max	Min	Std.
1. Estimate	46.1%	42.8%	100.0%	0.7%	32.4%
2. Estimate	58.8%	35.8%	187.7%	0.7%	52.9%
Difference	12.7%	-7.1%	87.7%	0.0%	20.5%

Note: Deviation values are found by dividing financial realizations by ChatGPT's estimate with this formula: "Financial Realization / ChatGPT's Estimate – 1"

International Journal of Business, Economics and Management Perspectives - Uluslararası İşletme, Ekonomi ve Yönetim Perspektifleri Dergisi
Volume/Cilt: 7, Issue/Sayı: 1, pp/ss. 1-14
*June/Haziran 2023*

In order to understand to what extent ChatGPT's estimates made in the first stage and the estimates made in the second stage are consistent, t-tests were applied. Two separate series were prepared, consisting of deviations from the realizations of the estimates made by ChatGPT in the first and second stages. Afterward, t-tests were applied to understand whether there was a difference between these series. The test procedure was repeated separately for sales, net profit, and dividends. The results obtained are summarized in Table 4. Accordingly, there is no statistically significant difference between the deviation of ChatGPT's estimates made by using only the financial statements in the first stage and the deviation of the estimates made by including industry-specific information in the second stage. This shows that ChatGPT cannot make better estimates by interpreting industry-specific information during research.

Table 4. t-tests Showing Difference Between Deviations in ChatGPT's Estimates

	Revenues		Net Income		Dividend	
	Deviation in Est. 1	*Deviation in Est. 2*	*Deviation in Est. 1*	*Deviation in Est. 2*	*Deviation in Est. 1*	*Deviation in Est. 2*
Mean	25.9%	26.5%	1555.8%	273.6%	61.3%	59.4%
Variance	7.0%	9.8%	1508913%	11917.0%	13.5%	16.5%
Observations	73	73	73	73	41	41
Pearson Correlation	0.912		0.982		0.474	
Hyp. Mean Dif.	0.000		0.000		0.000	
df	72		72		40	
t Stat	-0.411		0.977		-0.283	
P(T<=t) one-tail	0.341		0.166		0.389	
t Critical one-tail	1.666		1.666		1.684	
P(T<=t) two-tail	0.682		0.332		0.779	
t Critical two-tail	1.993		1.993		2.021	

The chart in Figure 2 was used to understand the correlation levels of the deviations of ChatGPT's estimates made by only interpreting the financial statements in the first stage and the estimates made by taking into account the industry-specific information in the second stage. In this chart, the first and second-stage deviation values of ChatGPT's sales estimates are shown. Accordingly, the estimates of ChatGPT show serious similarities in both stages. It was observed that the majority of the estimations that deviated negatively in the first stage also deviated in the negative direction in the second stage. On the other hand, in the estimations deviating in the positive direction, it was observed that the second stage was more successful than the first stage and the deviations were corrected to a certain extent. Compared to the estimations in the first stage, the number of realizations that deviate in the negative direction is 46, while the number of realizations that deviate in the positive direction is 27. When the deviations of the realizations with the estimates in the second stage are considered, it is seen that the number of realizations that deviated in the positive direction decreased to 21, while the number of realizations that deviated in the negative direction increased to 52.

Figure 2. Deviations in the First and Second Estimates of ChatGPT Related to BIST100 Companies Revenues
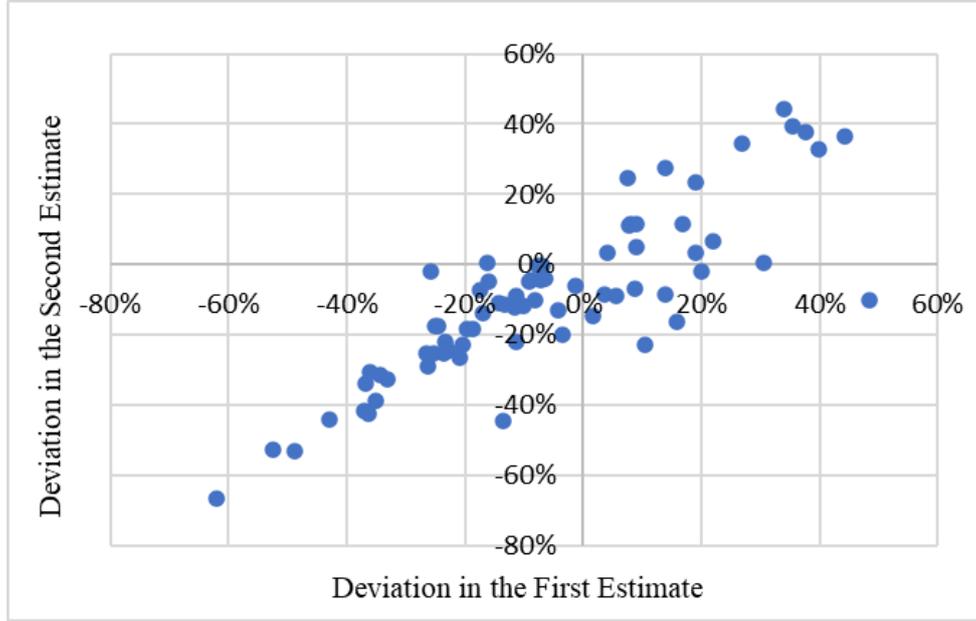
Note: Deviation values are found by dividing financial realizations by ChatGPT's estimate with this formula: "Financial Realization / ChatGPT's Estimate – 1"

Two important criteria come to the fore when ChatGPT corrects its estimations by interpreting industry-specific information. The first of these is the direction of the news. If the news about the company is negative, ChatGPT is expected to correct its estimation in the first stage downwards in the second stage. The second is the size of the correction. Even if the correction is made in the right direction, if it is overdone, this time there will be a deviation in the other direction. Figure 3 is used to understand the relationship between the deviation in the first stage and the direction and magnitude of the correction in the second stage. Corrections for the areas where the X axis is negative, and the Y axis is positive and the areas where the X axis is positive and the Y axis is negative reduce the margin of error. Corrections in other regions further increased the error rate. In the figure, it is seen that the realizations that deviated positively in the first-stage estimates were corrected more successfully in the second-stage estimates. Although some estimations made in the second stage are more successful, some estimations have a higher margin of error. To summarize, it seems that the direction of the corrections using sector-specific information has not been successful.

International Journal of Business, Economics and Management Perspectives - Uluslararası İşletme, Ekonomi ve Yönetim Perspektifleri Dergisi
Volume/Cilt: 7, Issue/Sayı: 1, pp/ss. 1-14
June/Haziran 2023

Figure 3. Deviations in the ChatGPT's First Estimate and Correction in the Second Estimate Related to BIST100 Companies Revenues

Note: Deviation values are found by dividing financial realizations by ChatGPT's estimate with this formula: "Financial Realization / ChatGPT's 1. Estimate – 1". Corrections are found by this formula: "ChatGPT's 2. Estimate / ChatGPT's 1. Estimate -1".

4. CONCLUSION

Within the scope of this research, it has been analyzed to what extent ChatGPT can accurately estimate financial values by interpreting past financial statements and current important developments in the sector. The financial data of 73 companies traded in the BIST100 index and the texts that summarize the sector-specific developments obtained from the Microsoft BING Chat application were used. In the first stage, ChatGPT was asked to interpret the financial values for the 2019-2021 period and make estimates for sales, net profit and dividend distribution for 2022. In the second stage, text about sector-specific developments was entered into ChatGPT and it was asked to revise its estimates. Afterwards, the estimations made by ChatGPT in these two stages were compared with the realizations and it was tried to understand how successful ChatGPT was in these estimations.

According to the findings obtained from the research, although ChatGPT can interpret the financials, it can be wrong in its estimations. On the other hand, it was observed that ChatGPT's estimates did not improve when the interpretation of industry-specific information was included. ChatGPT can try different methods in each conversation and find different results. The clarity of the question asked to ChatGPT can affect the answer received. Depending on the way the question is asked, in some cases ChatGPT may avoid answering on the grounds that it does not have enough data. It has

been observed that if a large amount of data is entered into ChatGPT, ChatGPT may have difficulty understanding the content of the data provided.

According to the results of the research, ChatGPT was able to estimate the sales values with 18% deviation as median values in the first stage, when it used only the financial statements, and with 18.9% deviation in the second stage, where it also used the current developments in the sector. The median deviation of the estimates for net profit in the first stage was 57.8%, while it was 64.9% in the second stage. The median deviations for dividends were 54.6% in the first stage and 51.8% in the second stage.

In summary, in the study conducted with ChatGPT, the process in which sector-specific developments and historical financial values were interpreted together could not improve the financial estimation process. However, it has been observed that ChatGPT can understand and interpret given texts and tables. In addition, considering that ChatGPT is constantly being developed, it should be noted that much better results can be obtained in the future. In future studies on this subject, the ability of ChatGPT and other AI tools to use different data resources and interpretation of them can be tested with different dimensions.

KAYNAKÇA

Adamopoulou, E. and Moussiades, L. (2020). An overview of chatbot technology. in Artificial Intelligence Applications and Innovations: *16th IFIP WG 12.5 International Conference, AIAI 2020*, Neos Marmaras, Greece, June 5–7, 2020, Proceedings, Part II. 16. 373-383.

Biswas, S. S. (2023). Potential use of ChatGPT in global warming. *Annals of Biomedical Engineering*, 1-2.

Bory, P. (2019). Deep new: The shifting narratives of artificial intelligence from Deep Blue to AlphaGo. *Convergence,* 25(4), 627-642.

Carvalho, I. and Ivanov, S. (2023). ChatGPT for tourism: Applications, benefits and risks. *Tourism Review,* https://doi.org/10.1108/TR-02-2023-0088

Cheung, C.W., Tsang I.T. and Wong, K.H., (2017). Robot avatar: A virtual tourism robot for people with disabilities. *International Journal of Computer Theory and Engineering,* Singapore, (9)3, 229-234.

Choi, J. H., Hickman, K. E., Monahan, A. and Schwarcz, D. (2023). ChatGPT goes to law school. *Minnesota Legal Studies Research Paper*, No. 23-03, 1-16.

Davenport, T. H. and Ronanki, R. (2018). Artificial intelligence for the real world. *Harvard Business Review*, 96(1), 108-116.

Dowling, M. and Lucey, B. (2023). ChatGPT for (finance) research: The Bananarama conjecture. *Finance Research Letters*, 53, 103662. 1-6.

Frieder, S., Pinchetti, L., Griffiths, R. R., Salvatori, T., Lukasiewicz, T., Petersen, P. C., Chevalier, A. and Berner, J. (2023). Mathematical capabilities of ChatGPT. (February 1, 2023), 1-29, https://doi.org/10.48550/arXiv.2301.13867.

International Journal of Business, Economics and Management Perspectives - Uluslararası İşletme, Ekonomi ve Yönetim Perspektifleri Dergisi
Volume/Cilt: 7, Issue/Sayı: 1, pp/ss. 1-14
June/Haziran 2023

Goel, P. and Ganatra, A. (2021). A survey on chatbot: Futuristic conversational agent for user interaction. *2021 3rd International Conference on Signal Processing and Communication (ICPSC).* 736-740.

Hwang, S. and Kim, J. (2021). Toward a chatbot for financial sustainability. *Sustainability*, *13*(6), 3173. 1-18.

Li, C., Pan, R., Xin, H. and Deng, Z. (2020). Research on artificial intelligence customer service on consumer attitude and its impact during online shopping. *In Journal of Physics: Conference Series*, 1-15.

Lund, B. D. and Wang, T. (2023). Chatting about ChatGPT: How may AI and GPT impact academia and libraries? (January 22, 2023). *Library Hi Tech News*, http://dx.doi.org/10.2139/ssrn.4333415

McCarthy, J. (1956). Measures of the value of information. *Proceedings of the National Academy of Sciences*, *42*(9), 654-655.

Mikalef, P. and Gupta, M. (2021). Artificial intelligence capability: Conceptualization, measurement calibration, and empirical study on its impact on organizational creativity and firm performance. *Information & Management, 58*(3), 103434. 1-20.

Ramesh, A. N., Kambhampati, C., Monson, J. R. and Drew, P. J. (2004). Artificial intelligence in medicine. *Annals of the Royal College of Surgeons of England*, *86*(5), 334-338.

Shawar, B. A. and Atwell, E. (2007). Chatbots: are they really useful?. *Journal for Language Technology and Computational Linguistics*, *22*(1), 29-49.

Singh, D. (2023). ChatGPT: A new approach to revolutionise organisations, *International Journal of New Media Studies (IJNMS), 10*(1), 57-63.

Wu, Y., Wu, W., Xing, C., Zhou, M. and Li, Z. (2016). Sequential matching network: A new architecture for multi-turn response selection in retrieval-based chatbots. *arXiv preprint* arXiv:1612.01627. 1-10.

Zielinski, C., Winker, M., Aggarwal, R., Ferris, L., Heinemann, M., Florencio-Lapen˜a, J., Pai, S., Ing, E. and Citrome, L. (2023). Chatbots, ChatGPT, and scholarly manuscripts-WAME recommendations on ChatGPT and chatbots in relation to scholarly publications. *Afro-Egyptian Journal of Infectious and Endemic Diseases*, *13*(1), 75-79.