The use of data mining technology in financial forecasting of accounting profits: An applied study

Jinan Abdul Abbas Baqer Al Dulaimi, Assistant prof. of accounting, Al-Furat Al-Awsat Technical University, Babylon Technical Institute - Accounting Department, Hilla - Iraq. Email: jenan.aldulamy@atu.edu.iq, Orcid: https://orcid.org/0000-0003-3810-1192

Dr. Khadija Jomaa Muter, Assistant professor and Department of, Accounting College of Administration and Economics, Mustansiriyah University, Baghdad, Iraq, Email: khdija_i_m79@uomustansiriyah.edu.iq, Orcid: https://orcid.org/0000-0003-2081-9198

Nada Naif Younis, Al-Mustansiriyah University, Department of Control and Internal Audit, Baghdad, Iraq. dr.nada_naif@uomustansiriyah.edu.iq

ABSTRACT

Data mining tools have emerged as one of the most prominent modern methods in recent years. The significance of this research originates from the difficulties faced by businesses. Current methods can improve companies' capacity for financial forecasting, thereby reducing risk and bolstering financial stability. Using the regression equation to analyze time series data is one of the data mining instruments. It aids in predicting financial failure and equips businesses with the information necessary to make crucial decisions. Data mining and financial forecasting were first described to accomplish the research objectives. Then, the accounting profits of Al-Mansour company for pharmaceutical industries, medical supplies, cosmetics, and sanitary materials were predicted using data mining techniques. The results demonstrated a decline in sales ratios and a total and net loss. It is a negative indicator and undesirable situation in the industrial sector that can be explained by the lack of interest in the industrial sector, the absence of government support from the state, and the market's accessibility to foreign products. The study offers important insights into the application of data mining techniques to the forecasting of financial outcomes.

KEYWORDS: Data mining, Accounting profits, Financial forecasting, Iraq

1. INTRODUCTION

Data mining is a crucial process for knowledge discovery in databases, to discover valuable and previously unknown patterns from vast data repositories (Gupta & Chandra, 2020). It includes a variety of functionalities, techniques, and algorithms designed to derive intriguing patterns from these enormous data collections. Data is transformed from accumulated and incomprehensible information into valuable, exploitable information. To develop algorithms that are scalable and adaptive to increasing quantities of data in the search for meaningful and cognitive patterns, the data mining stage has garnered significant interest from the research community over the past decade (Madni, Anwar, & Shah, 2017). The quantity of software and algorithmic packages has increased exponentially over the past decade.

Figure 1 depicts the knowledge discovery in databases (KDD) process, which relies heavily on data mining. Time series data and regression analysis are the most common formats for representing temporal information (Efendi, Arbaiy, & Deris, 2018).

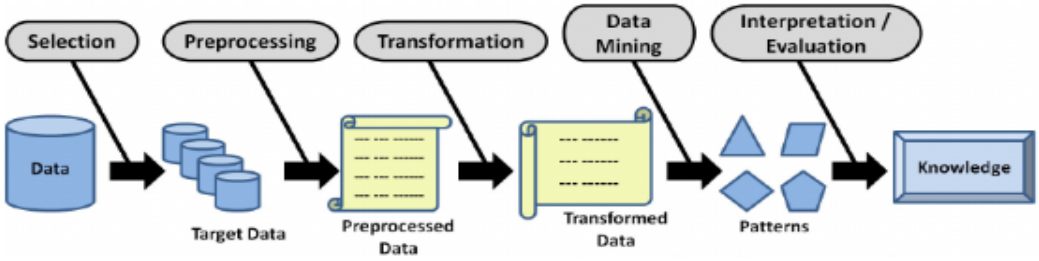

Figure 1: Data Mining for Knowledge Discovery (Adapted from Gullo (2015))

The frequency of data accumulation in financial markets is an essential factor in financial data forecasting. Moradi and Mokhatab Rafiei (2019) state that modeling dynamic financial data is becoming an increasingly crucial study area. Data mining is considered one of the techniques for analyzing and filtering vast quantities of stored data. It seeks to obtain relevant, trustworthy, and high-quality data and information for forecasting.

Time series modeling has been a prominent research field for decades, with applications in numerous domains, including business, the stock market, the weather, and electricity demand (Mills, 2019). Accurately forecasting financial information is valuable to accountants because it can guide investors in minimizing risks and investment losses (Birt et al., 2020). Thus, data mining enables the construction of future predictions and the investigation of various behaviors and trends. It leads to making the correct and appropriate decisions at the proper time.

Research Importance

The significance of this research derives from the inescapable applications of modern technologies and their capacity to increase economic units' efficiency in future profit forecasting. Moreover, companies face critical financial challenges, such as the inability to meet obligations, the risk of bankruptcy, and liquidation (Correa-Mejía & Lopera-Castaño, 2020); consequently, the present study aims to provide substantial knowledge regarding financial forecasting using data mining techniques.

Modern technologies such as data mining and one of its instruments are exemplified in the analysis of time series and the application of the regression equation. It aids in the prediction of financial catastrophes. Therefore, data mining can provide companies with sufficient benefits for enhancing precision, decision-making, and factors affecting their financial performance. Lastly, while data mining techniques have been studied in the context of Western nations (Le & Viviani, 2018; Ruzgar, 2019), evidence in the context of Iraq is limited.

Research problem

(Birt et al., 2020) Financial forecasting plays a crucial role in the decision-making process of businesses, particularly in predicting accounting profits. However, traditional forecasting methods frequently fail to account for financial data's dynamic and complexity, resulting in inaccurate forecasts. The development of information technology, software, and computers, as well as the emergence of the Internet, resulted in continuous improvements, which increased the capacity to compile financial and non-financial data, operate them, and communicate information to beneficiaries in a very brief period. Using data mining techniques to improve the precision and dependability of financial forecasting is a promising strategy.

The volume and complexity of financial data, including variables such as revenue, expenses, market trends, customer behavior, and macroeconomic indicators, is one of the greatest obstacles in financial forecasting (Hariri, Fredericks, & Bowers, 2019). This vast dataset makes it difficult for conventional forecasting methods to identify meaningful patterns and relationships, resulting in inaccurate predictions. In a similar vein, financial data frequently exhibit dynamic properties, which further complicates financial forecasting.

To address these obstacles, extensive research and the application of data mining techniques designed to comprehend financial forecasting of accounting profits are required. By utilizing the power of data mining, businesses can gain valuable insights from their financial data, uncover concealed patterns, and make informed decisions to increase profitability. This study seeks to bridge the gap by conducting applied research that investigates the use of data mining technology in financial forecasting in the context of Iraq and by providing organizations with actionable solutions and recommendations.

Research Questions

This research seeks to answer the following questions:
1. What are the different data mining techniques, processes, and challenges?
2. What is the role of financial forecasting in finance and accounting?
3. What is Al-Mansour company's economic forecast for profits using data mining techniques?

Research Objectives

This research seeks to achieve the following objectives:
4. To explore data mining techniques by discussing their process, usage, and challenges.
5. To uncover the concept of financial forecasting, including the methods used in finance and accounting.
6. To determine the use of data mining techniques in forecasting accounting profits of Al-Mansour company.

2. LITERATURE REVIEW

Anouze and Bou-Hamad (2019) analyzed various statistical and data mining techniques to evaluate bank performance. According to the research, modern data mining techniques, such as random forests and clustering, provide more accurate information for assessing bank performance. Hadi, Al-khalisy, and Abd Hamza (2019) evaluated datasets from 23 Iraqi institutions using data mining techniques for financial forecasting and failure prediction. The research indicated that the backpropagation algorithm is a valuable predictor system (Mellen, 2022).

Ruzgar (2019) examined the daily stock market price change classification for six Canadian institutions utilizing seven Weka-implemented data mining algorithms. The study was based on daily stock price fluctuations over 37 years, and each data mining algorithm was separately implemented to classify each bank's daily stock price fluctuations. The results indicated that, among the seven methods, the J48 algorithm performed the best in classifying stock price variations, individually and globally. The research demonstrated that the J48 algorithm is a promising alternative to conventional techniques for financial forecasting. The results show the potential of data mining algorithms to accurately predict and categorize stock price changes,

thereby providing investors with valuable insights for evaluating market trends and making informed investment decisions.

Le and Viviani (2018) conducted a study in which traditional and data mining techniques were used to predict bank failure. Using a sample of 3,000 banks in the United States, of which 1,438 were failed banks, the study found that artificial neural networks and k-nearest neighbor methods predicted bank failure with high accuracy.

Reviewing and analyzing previous studies shows that data mining techniques offer significant benefits; consequently, empirical evidence is required for Iraq. The present study aims to realistically and practically predict the accounting profits of Al-Mansour company for pharmaceutical industries, medical supplies, and cosmetics using the sample data. In addition, this study differed from previous research regarding the period and the statistical methods, tools, and programs employed.

3. CONCEPT OF DATA MINING

The emergence of data mining

Due to advancements in science, commerce, and information technology, the quantity of digital data has increased. Since its introduction in the late 1980s, data mining has become one of the most effective methods for analyzing massive quantities of data (Madni et al., 2017). Data mining transforms these enormous datasets from merely accumulated and incomprehensible information to knowledge that can be exploited and benefited from Sumiran (2018). In the mid-1990s, contributions from several disciplines, including statistics, databases, artificial intelligence, and machine learning, led to the emergence of the term data mining.

What is data mining?

In recent years, there have been numerous definitions of data mining owing to the interest of numerous researchers. Therefore, researchers have defined data mining differently. Previous scholars have defined data mining as "the process of discovering or extracting intriguing patterns, associations, variations, anomalies, and significant structures from vast quantities of data." This data is typically stored in file systems, databases, warehouses, and other information repositories" (Gupta & Chandra, 2020).

Data mining assists in obtaining information and resolving issues through the execution of various duties. Since this technology can carry out one or more of the abovementioned duties (Nemeth & Michalconok, 2017). Data mining identifies unexpected relationships by analyzing and extracting knowledge from vast quantities of data made available by theoretical and scientific methods. In the same context, data mining is viewed as scouring for data to remove valuable information from a large database or sub-data warehouse (Chandra & Gupta, 2018).

The conclusion is that data mining is a type of intelligent technology that extensively utilizes computers and statistics. It is based on analyzing and scrutinizing vast amounts of data to discover relationships and patterns that have not been found using traditional methods to assist with future prediction decisions.

Data mining techniques

The commonly used techniques for data mining, according to Madni et al. (2017) and Chee, Baharudin, and Karkonasasi (2016), are summarized below:

Table 1: Different Methods for Data Mining

Method	Description	Model Type
Linear Regression	Presumes a linear relationship and is relatively easy to explain and build.	Regression
Neural Networks	They are non-linear models that work on prediction and are learned through continuous training. They are similar in structure to biological neural networks.	Regression or classification
Decision Tree	They represent groups of decisions because these decisions are divided into a set of data. Decision trees have regression and classification and can detect the automatic interaction of a chi-square.	Regression or Classification
Neighbour's Approach	They represent groups of decisions because these decisions are divided into a set of data. Decision trees have regression and classification and can detect the automatic interaction of a chi-square.	Classification

Challenges Faced in Data Mining

The mining process faces several challenges, and the following difficulties are summarized as highlighted in previous literature (Gan et al., 2017; Gupta & Chandra, 2020; Madni et al., 2017):

a) ***Scalability***: It measures the extent to which existing algorithms can process a huge amount of data.
b) ***Dimensionality***: Data contains thousands of different properties, so traditional analysis methods cannot process them properly.
c) ***Heterogeneous and Complex Data***: Traditional analysis methods usually deal with data with uniform features.
d) ***Data Ownership and Distribution***: Sometimes, the need for data that is not stored in a fixed place and does not belong to a single unit requires developing techniques to distribute that data securely.
e) ***Data Quality:*** Data may be incomplete or inconsistent, negatively affecting data mining algorithms.
f) **Privacy preservation** during the data mining process**.**
g) **Stream data** that are generated sequentially and continuously.
h) **Expertise and knowledge:** Effective data mining requires technical expertise and knowledge.

How does data mining work?

(Durand & Hattingh, 2020; Gupta & Chandra, 2020). Therefore, these systems enable firms to make informed decisions, resulting in accurate predictions and advantageous outcomes. This is consistent with the objective of data mining, which is to facilitate decision-making processes via prediction and scenario formulation. Decision support systems that combine data with personal insights employ mathematical models for prediction and simulation. Forecasting models, driven by quantitative methods and computer technology, provide decision-makers with timely forecasts and vital information for mitigating risks and avoiding potential crises (Sarker, 2021). By leveraging these technologies, decision support teams can effectively measure, monitor, and control various factors, resulting in early warnings and the ability to take

proactive measures to mitigate negative impacts and ensure efficient operations (Ingle & Deshmukh, 2021). In conclusion, the data mining process includes problem identification and definition, data collection, cleaning, integration, transformation, application of mining techniques, pattern discovery, evaluation, knowledge presentation, and utilization, all of which contribute to the objective of extracting valuable insights and facilitating informed decision-making (Dogan & Birant, 2021).

4. FINANCIAL FORECASTING

According to Sezer, Gudelek, and Ozbayoglu (2020), financial forecasting is a crucial aspect of business planning and decision-making that entails estimating and predicting future economic outcomes based on historical data and other relevant factors. It aims to sketch an initial picture of the firm's or company's future financial conditions. Similarly, it provides a space to identify future financial requirements and plan accordingly (Mills, 2019). Financial forecasting enables users to recognize the opportunity regarding the company's future financial position and profitability and make necessary adjustments. Companies can benefit from financial forecasting by projecting future revenues and expenses, enabling them to modify their objectives, create effective strategies, and allocate resources effectively. This also allows them to anticipate obstacles, ensuring businesses can endure the financial crisis and achieve sustainable growth (Veganzones & Severin, 2021). In the same vein, economic forecasts are the financial information for the future founded on the anticipated future events by management. This information consists of the financial statements and their items; financial statements are viewed as a means for the outputs of the accounting system based on the economic unit, and these lists must be prepared to comply with the requirements of full disclosure (Palepu et al., 2020). Therefore, financial forecasting relies on the analysis of historical data to look ahead and foresee the future of both near and far-off events using modern techniques.

Factors affecting the process of financial forecasting

Several factors influence the financial forecasting procedure. Multiple factors, including information technology capability, firm characteristics, economic conditions, political stability, and the quality of financial reporting, have been identified by prior research as critical factors that can influence financial forecasting (Rahman, Zhang, & Dong, 2019). Analysts rely on the firms' financial reports to forecast future financial outcomes; consequently, the quality of financial reporting plays a role in accurate and trustworthy forecasting. Similarly, it has been demonstrated that economic and political stability influence future financial forecasts because of their effect on the economy (Dai & Zhang, 2019). Lastly, the development of IT enhances the capacity to employ contemporary methods. Therefore, firms implementing the most recent techniques can improve their financial forecasting process (Rahman et al., 2019).

Financial forecasting methods

Niu et al. (2020), Vilela et al. (2019), and Mills (2019) discussed some of these techniques and methods, such as using time series data and regression analysis for the financial forecasting process. Observations may be annual, monthly, weekly, or daily for time series data. The relationship between dependent and independent variables is analyzed during regression analysis. Regression analysis and other econometric techniques are widely used for financial forecasting (Shobana & Umamaheswari, 2021).

5. METHODS

Data

Profitability data

Accounting profit is the difference between revenues from financial operations during a particular period (sales, outputs, or services provided and the cost of acquiring those outputs or revenues) (Alarussi & Alhaderi, 2018). According to Tamulevičienė (2016), profit is the net increase in owners' equity from the economic unit's operations.

The importance of profits

Due to the firm's, proprietors', creditors', and potential investors' economic repercussions, profit measurement and reporting are essential financial accounting objectives. Within enterprises, dividend distribution and retention decisions are guided by profitability. Depending on factors including liquidity, investment policy, and expansion, the number of periodic profits determines whether they can be distributed to owners or retained for other purposes. Operating and non-operating results contribute to net profits, making historical profit values useful for forecasting future values. Profits also serve as an indicator of management efficiency and efficacy, revealing how well resources are managed. This metric is essential for evaluating the effectiveness of management in attaining the organization's goals. Consequently, the present study employs financial statements to collect data on profitability. Financial statements are essential for determining an economic unit's ability to repay loans and interest obligations.

Forecasting model

This model can be used to analyze actual data in data warehouses and determine some correlations in the data of economic units within a set of data (Mills, 2019). This model can predict and contain new data. Two distinct phases comprise utilizing forecasting models based on the supervised learning strategy. This method is utilized in this investigation. Using the forecasting model, the following constitutes the forecasting procedure:

Training builds a model using a large sample of historical data called the training group.

a) A test involves trying to exit the model on new data called the test set.

Data Sources

This investigation utilized secondary sources of information. The researcher collected and analyzed data and information from Arabic and foreign books, theses, periodicals, and journal articles to accomplish the first two research objectives.

Second, the researcher planned to apply data mining techniques to Al-Mansour's financial dataset. The researcher relied on the annual financial reports issued by the research sample for this purpose. The years 2011 through 2019 were obtained.

Research sample

Al-Mansour Company for Pharmaceutical Industries, Medical Supplies, Cosmetics, and Sterile Materials served as a sample for the research: Al-Mansour Company for Pharmaceutical Industries was established following the Companies Law No. 36 of 1983 on 4/19/1989 with a capital of (eight million Dinar) and a founding certificate No. 3346. The company's headquarters and factory are located in the Baghdad neighborhood of Abi Gharib. The company intends to contribute to sustaining and advancing the industrial movement in Iraq and strengthening and advancing the national economy in pharmaceuticals.

6. DATA ANALYSIS AND RESULTS

For the data analysis, the researcher utilized the data mining steps as shown in the figure below:

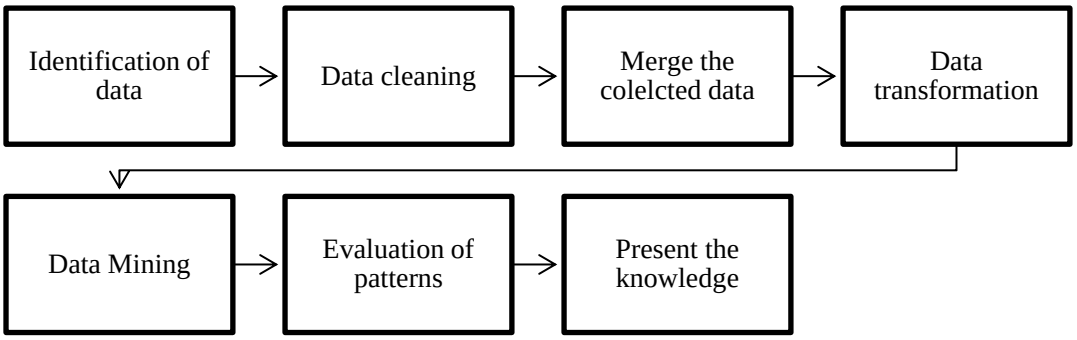

Figure 2: Stages of Data Mining Process

1. Data discovery: The data collection stage includes detecting, identifying, and describing the available data. At this stage, companies listed in the Iraq market for securities have been identified.

2. Cleaning Data: At this stage, the noise data is removed; inconsistent, conflicting data are also deleted. Al-Mansour Pharmaceutical Industries, Medical Supplies, cosmetics, and sterile materials were assigned to the research sample. The researcher selected the company as it is profitable and plays a significant role in providing the local market with drug and medical supplies to support and encourage national industries.

3. Data Integration: At this stage, similar and relevant data are collected from multiple sources and merged.

4. Data selection: At this stage, the appropriate data is determined and retrieved from the data set by evaluating the profits of the research sample according to financial ratios. Hence, profitability is a primary goal necessary for the survival and continuity of the unit's work and a goal that the shareholders aspire to. Profitability is the relationship between the profits achieved by the unit and the investments that contributed to achieving these profits. Therefore, we find a great effort is directed toward using available resources to achieve the best possible return for shareholders. Profitability ratios are used to assess the ability of an economic unit to generate profits from its operational activities compared to the burdens it incurred during a specific period. Furthermore, profitability ratios indicate a higher value compared to the previous periods. The

unit achieves an increase in sales and thus an increase in profits, and by reference to the annual reports of the research sample and the adoption of a time series for the period 2011-2019.

Table 2. The profitability ratio between (2011-2019) amounts to millions of dinars

2020	2019	2018	2017	2016	2015	2014	2013	2012	2011	
1- The ratio of the total profit to sales										
(116)	(190)	(175)	(186)	156	743	57	425	688	330	The overall profit
849	4449	2663	2255	1957	2764	1841	2778	3513	1854	÷ Net sales
-13.7%	-4.2%	-6%	-8%	8%	27%	4%	15%	20%	18%	The total profit rate *100%
2. The percentage of operating profit to sales										
(149)	(27)	(291)	180	(4)	554	430	230	375	133	Net operating profit
849	4449	2663	2255	1957	2764	1841	2778	3513	1854	÷ Net sales
-17.6%	-0.6%	-11%	8%	-0.2%	20%	23%	8%	11%	7%	The percentage of operating profit to sales * 100
3. Net profit ratio to sales										
(179)	(20)	85	193	(22)	508	423	292	373	100	Net Profit
849	4449	2663	2255	1957	2764	1841	2778	3513	1854	÷ Net Sales
-21.1%	-0.4%	3%	9%	-1%	18%	23%	11%	11%	5%	Net profit ratio to sales * 100
4. The rate of return on investment										
(179)	(960)	85	193	(22)	508	423	292	373	100	Net profit
4458	694	9504	8050	8143	8006	8161	7840	5104	3893	÷ Total assets
-4.0%	-2.8%	0.8%	2%	-0.2%	6%	5%	4%	7%	3%	The rate of return on investment * 100
5. The percentage of return to property rights										
(179)	(960)	85	193	(22)	508	423	292	373	100	Net profit
5754	5511	7666	7577	7527	7549	7272	7394	3894	3563	÷ property rights
-3.1%	-0.3%	1%	3%	-0.2%	7%	6%	4%	10%	3%	The percentage of return to property rights * 100

5. Data transformation: At this stage, the data are transformed into customized models suitable for search and retrieval by completing summary or aggregation operations. The profits of Al-Mansour Pharmaceutical Industries Company were evaluated through the use of profit rates consisting of the following proportions:

6. Mining Data: Using smart methods that apply to extract data patterns, removing as many useful models as possible. Predictive excavation has relied on Its data in advance; this model includes the most famous data excavation tool technologies, including (classification, regression, and time series analysis). The researcher adopted the period chains (2011-2020). Moreover, the regression equation was used to predict the financial status of the search sample for 2021-2030. Table 3 shows:

Table 3. The expectation of the financial proportions from 2021-2030 using the regression equation

2030	2029	2028	2027	2026	2025	2024	2023	2022	2021	Proportions
-97.84	-110.01	-122.18	-134.35	-146.52	-158.69	-170.86	-183.03	-85.67	-73.50	The ratio of the overall profit to sales
-102.71	-114.88	-127.05	-139.22	-151.39	-163.56	-175.73	-187.90	-90.54	-78.37	The percentage of operating profit to sales
-31.44	-35.79	-40.14	-44.49	-48.84	-53.19	-57.55	-61.90	-27.08	-22.73	Net profit ratio to sales
-15.80	-17.84	-19.89	-21.93	-23.98	-26.02	-28.07	-30.11	-13.75	-11.71	The rate of return on investment
-5.46	-6.53	-7.59	-8.65	-9.72	-10.78	-11.85	-12.91	-4.40	-3.33	The percentage of return to property rights

7. Pattern evaluation: At this stage, the research determined the important patterns that represent the knowledge base for using some important measures. The analysis in Table 2 shows that the research sample is making losses, which requires taking quick steps to help it avoid losses and financial failure and then bankruptcy.

8. Presentation of knowledge: The last stage of knowledge discovery in databases and the stage that the beneficiary sees. This important stage uses a visual technique to assist the beneficiary in comprehending and interpreting data extraction. To understand the data in Table 3, they were displayed graphically in Figure 1.

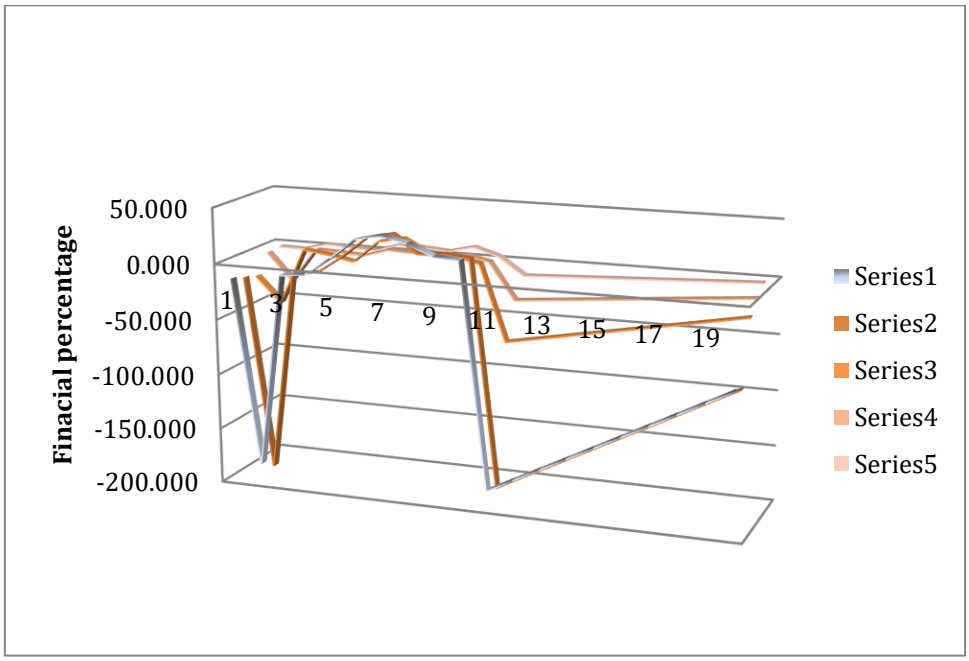

Figure 3 Linear regression formula

The findings showed a decline and decrease in sales, achieving a total and net loss. It is not a good indicator and an undesirable situation in the industrial sector. Such a decline can be associated with less interest in the industrial sector, less authority support, and market openness to foreign products. This requires paying more efforts to promote local development to support the state's general budget by issuing laws and legislation that support the regional product and the availability of government support.

7. CONCLUSION AND RECOMMENDATIONS

Conclusion

In today's digital age and era of advanced technology, incorporating new techniques in finance and accounting is essential. Data mining is an effective technique that significantly relies on statistical methods. This study aimed to provide an in-depth understanding of data mining by focusing on the concept, techniques, and procedures. Second, it provided information regarding financial forecasting and the utilization of data mining techniques for measuring and forecasting profits. It was stated that economic forecasting is essential for analyzing past events to make informed future decisions. For this purpose, the researcher gathered financial information for Al-Mansour from 2011 to 2019. The researcher utilized data mining procedures and regression analysis methods to identify patterns in the extracted data. The analysis revealed that negative financial forecasts for the company necessitate immediate action to prevent losses. Organizations can use data mining techniques and modern technologies to overcome the challenges encountered by businesses in various industries to make informed decisions.

Theoretical Contributions

This research contributes significant new knowledge to the existing data mining and financial forecasting literature. It offers a thorough comprehension of data mining techniques and economic forecasting. In addition, it added empirical evidence by employing statistical methods to Al-Mansour company's financial data. The study contributes considerably to data mining research in Iraq.

Practical Recommendations

1. There is a growing demand for data mining and its associated methods and tools to improve the culture of knowledge-based economic entities and provide all the required data stores and data warehouses to form new knowledge. It would then aid in the development of financial and accounting operations, the forecasting of future accounting profits, the reduction of fraud, and the enhancement of profitability.
2. Continual improvement and investment in essential information technology are recommended to achieve the desired objectives of increasing the efficiency of accounting operations. In addition to utilizing the most suitable accounting software, businesses must focus on providing appropriate training programs and continuous training of human cadres to equip them with knowledge of utilizing modern technologies, such as data mining, to benefit from them.
3. Paying attention to data mining by incorporating it into the primary and advanced curriculum for Accounting Department students. The subject of computer applications aids in providing students with knowledge, skills, and experience in using computers.

REFERENCES

Alarussi, A. S., & Alhaderi, S. M. (2018). Factors affecting profitability in Malaysia. *Journal of Economic Studies, 45*(3), 442-458. https://doi.org/10.1108/JES-05-2017-0124

Anouze, A. L. M., & Bou-Hamad, I. (2019). Data envelopment analysis and data mining to efficiency estimation and evaluation. *International Journal of Islamic and Middle Eastern Finance and Management, 12*(2), 169-190. https://doi.org/10.1108/IMEFM-11-2017-0302

Birt, J., Chalmers, K., Maloney, S., Brooks, A., Oliver, J., & Bond, D. (2020). *Accounting: Business reporting for decision making*. John Wiley & Sons.

Chandra, P., & Gupta, M. K. (2018). Comprehensive survey on data warehousing research. *International Journal of Information Technology, 10*(2), 217-224. https://doi.org/10.1007/s41870-017-0067-y

Chee, T. C., Baharudin, A. S., & Karkonasasi, K. (2016). Data mining framework for test time optimization in industrial electronics manufacturing enterprise. *International Journal of Applied Engineering Research, 11*(4), 2717-2722. https://www.ripublication.com/ijaer16/ijaerv11n4_94.pdf

Correa-Mejía, D. A., & Lopera-Castaño, M. (2020). Financial ratios as a powerful instrument to predict insolvency; a study using boosting algorithms in Colombian firms. *Estudios gerenciales, 36*(155), 229-238. https://doi.org/10.18046/j.estger.2020.155.3588

Dai, L., & Zhang, B. (2019). Political uncertainty and finance: a survey. *Asia-Pacific Journal of Financial Studies, 48*(3), 307-333. https://doi.org/10.1111/ajfs.12257

Dogan, A., & Birant, D. (2021). Machine learning and data mining in manufacturing. *Expert Systems with Applications, 166*, 114060. https://doi.org/10.1016/j.eswa.2020.114060

Durand, T., & Hattingh, M. (2020). Data mining and artificial intelligence techniques used to extract big data patterns. In *2020 2nd International Multidisciplinary Information Technology and Engineering Conference (IMITEC)* (pp. 1-8). IEEE. https://doi.org/10.1109/IMITEC50163.2020.9334069

Efendi, R., Arbaiy, N., & Deris, M. M. (2018). A new procedure in stock market forecasting based on fuzzy random auto-regression time series model. *Information Sciences*, *441*, 113-132. https://doi.org/10.1016/j.ins.2018.02.016

Gan, W., Lin, J. C. W., Chao, H. C., & Zhan, J. (2017). Data mining in distributed environment: a survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *7*(6), e1216. https://doi.org/10.1002/widm.1216

Gullo, F. (2015). From patterns in data to knowledge discovery: What data mining can do. *Physics Procedia*, *62*, 18-22. https://doi.org/10.1016/j.phpro.2015.02.005

Gupta, M. K., & Chandra, P. (2020). A comprehensive survey of data mining. *International Journal of Information Technology*, *12*(4), 1243-1257. https://doi.org/10.1007/s41870-020-00427-7

Hadi, R. M., Al-khalisy, S. H. J., & Abd Hamza, N. (2019). Prediction Model for Financial Distress Using Proposed Data Mining Approach. *Journal of Al-Qadisiyah for computer science and mathematics*, *11*(2), 37-44. https://doi.org/10.29304/jqcm.2019.11.2.570

Hariri, R. H., Fredericks, E. M., & Bowers, K. M. (2019). Uncertainty in big data analytics: survey, opportunities, and challenges. *Journal of Big Data*, *6*(1), 1-16. https://doi.org/10.1186/s40537-019-0206-3

Ingle, V., & Deshmukh, S. (2021). Ensemble deep learning framework for stock market data prediction (EDLF-DP). *Global Transitions Proceedings*, *2*(1), 47-66. https://doi.org/10.1016/j.gltp.2021.01.008

Le, H. H., & Viviani, J.-L. (2018). Predicting bank failure: An improvement by implementing a machine-learning approach to classical financial ratios. *Research in International Business and Finance*, *44*, 16-25. https://doi.org/10.1016/j.ribaf.2017.07.104

Madni, H. A., Anwar, Z., & Shah, M. A. (2017). Data mining techniques and applications—A decade review. In *2017 23rd international conference on automation and computing (ICAC)* (pp. 1-7). IEEE. https://doi.org/10.23919/IConAC.2017.8082090

Mellen, S. W. (2022). Following the science until she didn't – Rochelle Walensky and revised Omicron Covid-19 guidelines. *International Journal of Instructional Cases*, *6*. http://www.ijicases.com/search/following-the-science_case

Mills, T. C. (2019). *Applied time series analysis: A practical guide to modeling and forecasting*. Academic Press. https://shop.elsevier.com/books/applied-time-series-analysis/mills/978-0-12-813117-6

Moradi, S., & Mokhatab Rafiei, F. (2019). A dynamic credit risk assessment model with data mining techniques: evidence from Iranian banks. *Financial Innovation*, *5*(1), 1-27. https://doi.org/10.1186/s40854-019-0121-9

Nemeth, M., & Michalconok, G. (2017). The initial analysis of failures emerging in production process for further data mining analysis. In *2017 21st International Conference on Process Control (PC)* (pp. 210-215). IEEE. https://doi.org/10.1109/PC.2017.7976215

Niu, T., Wang, J., Lu, H., Yang, W., & Du, P. (2020). Developing a deep learning framework with two-stage feature selection for multivariate financial time series forecasting. *Expert Systems with Applications*, *148*, 113237. https://doi.org/10.1016/j.eswa.2020.113237

Palepu, K. G., Healy, P. M., Wright, S., Bradbury, M., & Coulton, J. (2020). *Business analysis and valuation: Using financial statements*. Cengage Learning Australia. https://au.cengage.com/c/isbn/9780170425186

Rahman, J. M., Zhang, J., & Dong, S. (2019). Factors affecting the accuracy of analysts' forecasts: A review of the literature. *Academy of Accounting and Financial Studies Journal*, *23*(3), 1-18. https://hdl.handle.net/20.500.12540/139

Ruzgar, N. S. (2019). Comparative Analysis of Classification Algorithms on Stock Market Price Changes. *WSEAS Transactions on Information Science and Applications*, *16*, 174-184. https://www.researchgate.net/profile/Nursel-Ruzgar/publication/346049826

Sarker, I. H. (2021). Data science and analytics: an overview from data-driven smart computing, decision-making and applications perspective. *SN Computer Science*, *2*(5), 377. https://doi.org/10.1007/s42979-021-00765-8

Sezer, O. B., Gudelek, M. U., & Ozbayoglu, A. M. (2020). Financial time series forecasting with deep learning: A systematic literature review: 2005–2019. *Applied soft computing*, *90*, 106181. https://doi.org/10.1016/j.asoc.2020.106181

Shobana, G., & Umamaheswari, K. (2021). Forecasting by machine learning techniques and econometrics: a review. In *2021 6th international conference on inventive computation technologies (ICICT)* (pp. 1010-1016). IEEE. https://doi.org/10.1109/ICICT50816.2021.9358514

Sumiran, K. (2018). An overview of data mining techniques and their application in industrial engineering. *Asian Journal of Applied Science and Technology*, *2*(2), 947-953. http://ajast.net/data/uploads/5029.pdf

Tamulevičienė, D. (2016). Methodology of complex analysis of companies' profitability. *Entrepreneurship and sustainability issues*, *4*(1), 53-63. https://doi.org/10.9770/jesi.2016.4.1(5)

Veganzones, D., & Severin, E. (2021). Corporate failure prediction models in the twenty-first century: a review. *European Business Review*, *33*(2), 204-226. https://doi.org/10.1108/EBR-12-2018-0209

Vilela, L. F. S., Leme, R. C., Pinheiro, C. A. M., & Carpinteiro, O. A. S. (2019). Forecasting financial series using clustering methods and support vector regression. *Artificial Intelligence Review*, *52*, 743-773. https://doi.org/10.1007/s10462-018-9663-x