

# Data Mining in Financial Analytics to Forecast Loan Behavior: An Integrated Approach Combining Time Series and Machine Learning

Assala Laifa<sup>1</sup>, Fatiha Bouhrine<sup>2</sup>

## Abstract

*This study aims to apply time series as a predictive data mining technique to analyze future loan trends based on historical customer behavior. Using the Power BI business intelligence tool, the research presents a logical analysis of the financial status of the selected public bank. It also seeks to support the enhancement of decision-making quality and direction. The study concluded that state-subsidized loans are expected to rise significantly. In contrast, loans requiring personal contributions are projected to remain relatively stable. Partially subsidized loans, however, are likely to fluctuate in response to economic conditions and financial policies.*

**Keywords:** *Time series, forecasting, future loans, customer behavior.*

Received: 12.05.2025

Accepted: 09.08.2025

## Introduction

The transformation in data processing methods has led to a growing interest in advanced information technologies that support automated data analysis. Among the most prominent of these technologies is data mining, which has gained wide recognition across many fields, particularly in finance and banking.

Data mining integrates knowledge from various disciplines such as statistics, mathematics, information technology, artificial intelligence, and computer science. This integration supports the development of tools capable of solving complex problems in diverse sectors.

This shift has made it increasingly difficult to forecast future trends due to the volume and variety of data produced by modern information systems. Consequently, there is a growing need for advanced analytical tools to extract hidden knowledge. Such insights can enhance the efficiency and effectiveness of banks by enabling a deeper understanding of market trends and predicting changes in customer preferences based on their historical behavior.

This understanding can improve strategic planning, helping banks manage risks related to specific loan types or services. From this perspective, it becomes essential to examine how predictive data mining techniques contribute to managing banking operations, particularly loan management, amidst the rising challenge of data proliferation. These techniques help in making timely and accurate decisions at minimal cost.

Although visual analysis tools are not new—having been used to convey ideas since early human history in forms such as cave paintings—their role has expanded significantly with the emergence of smart systems and modern information technology. These tools now support the analysis of complex, multi-dimensional data and help users quickly understand patterns, trends, and relationships.

While visual analytics is widely used in engineering, science, and education, its application in economic inquiry has greatly enhanced decision-making in financial contexts. This is especially true in banking, where loan analysis plays a crucial role in assessing risk and repayment capacity. Graphical representations of loan

---

<sup>1</sup> University of Constantine 2 – Abdelhamid Mehri, Algeria. Email: [assala.laifa@univ-constantine2.dz](mailto:assala.laifa@univ-constantine2.dz)

<sup>2</sup> University of Constantine 2 – Abdelhamid Mehri, Algeria. Email: [fatiha.bouhrine@univ-constantine2.dz](mailto:fatiha.bouhrine@univ-constantine2.dz)

data—such as loan amounts, repayment periods, and client profiles—make it easier to identify patterns and relationships. This leads to better financial decisions.

### Research Problem

In light of the above, the main research question can be stated as: To what extent can predictive data mining techniques contribute to analyzing the future direction of granted loans?

### Research Objectives

This study aims to assess the impact of using time series models to forecast future loans granted by a public Algerian bank between 2023 and 2027. The analysis relies on a loan database covering the five years from 2018 to 2022. The goal is to provide the bank with a clear outlook to support rational and informed decision-making.

### Theoretical Background

Forecasting is one of the core functions of data mining. It helps in building models that predict the occurrence of future events or determine the likely direction of a certain behavior based on its historical data. There are various predictive data mining techniques that users may choose from, depending on the objective of the analysis, the nature of the problem, and the type of data being processed. Some of these techniques include:

#### Time Series Analysis

A time series, as its name implies, is a sequence of events, values, or observations recorded at regular time intervals. It represents an ordered series of  $n$  variables, expressed by the following equation:

$$T = (t_1, t_2, t_3, \dots, t_n), t_i \in \mathbb{R}$$

This means that the values in the series are real numbers. (Philippe Esling, 2012, p. 2)

The aim of time series analysis is to identify the changes that have occurred in a given phenomenon over time, analyze their causes and consequences, and understand their trends. These trends are then used as inputs in forecasting the future behavior of that phenomenon. (Saif Al-Din Othman & Jaafar Al-Shafi'i, 2014, p. 6)

In addition, time series analysis provides insights that support better strategic decision-making. These benefits make time series models highly relevant today. Two main types of time series models can be distinguished:

#### Additive Model

This model is based on the assumption that a time series is made up of several components added together. It is expressed as:

$$Z(t) = M(t) + S(t) + C(t) + E(t)$$

Where:

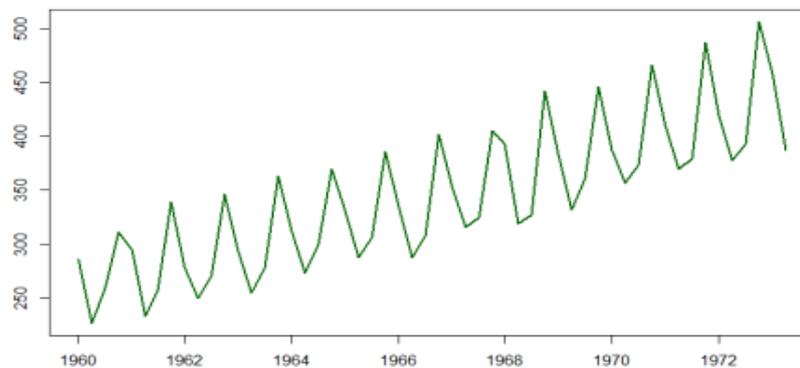
- $Z(t)$ : the observed time series
- $M(t)$ : the overall trend

- $S(t)$ : seasonal variations
- $C(t)$ : cyclical variations
- $E(t)$ : random (irregular) variations

(Introduction to Time Series Analysis and its Applications, 01/08/2022)

In this model, the graph of the time series shows how seasonal, cyclical, and random factors appear as deviations from the main trend.

**Figure 01: Additive Model of Time Series**



**Source:** Introduction to Time Series Analysis and its Applications, 01/08/2022. Available at: [https://www.researchgate.net/publication/362389180\\_INTRODUCTION\\_TO\\_TIME\\_SERIES\\_ANALYSIS\\_AND\\_ITS\\_APPLICATIONS](https://www.researchgate.net/publication/362389180_INTRODUCTION_TO_TIME_SERIES_ANALYSIS_AND_ITS_APPLICATIONS), Accessed on 04/04/2023 at 18:32 PM.

From Figure 01, we observe that the additive time series model appears as a relatively simple linear curve. It includes several deviations, which represent the differences between the actual and expected values. These deviations correspond to the seasonal, cyclical, and random components mentioned earlier.

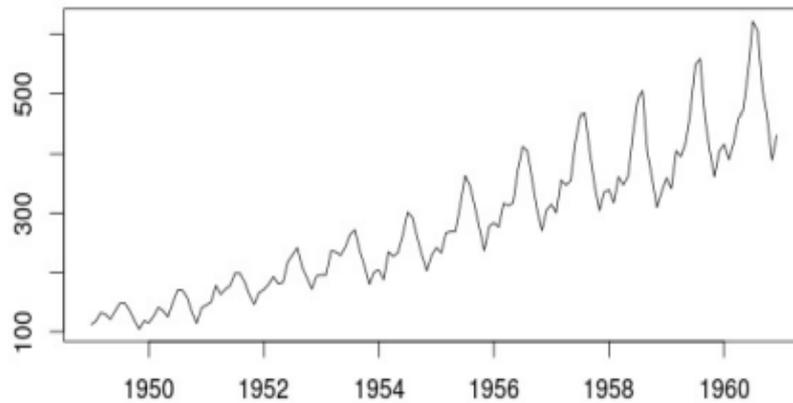
### Multiplicative Model

This model assumes that the components of the time series interact through multiplication, and is expressed as follows:

$$Z(t) = M(t) \cdot S(t) \cdot C(t) \cdot E(t)$$

(Introduction to Time Series Analysis and its Applications, 01/08/2022)

This model is illustrated in the following figure. It assumes that seasonal and cyclical variations are expressed as ratios. Among time series models, the multiplicative model is the most widely used.

**Figure 02: Multiplicative Model of Time Series**

**Source:** Introduction to Time Series Analysis and Its Applications, 01/08/2022. Available at: [https://www.researchgate.net/publication/362389180\\_INTRODUCTION\\_TO\\_TIME\\_SERIES\\_ANALYSIS\\_AND\\_ITS\\_APPLICATIONS](https://www.researchgate.net/publication/362389180_INTRODUCTION_TO_TIME_SERIES_ANALYSIS_AND_ITS_APPLICATIONS). Accessed on 04/04/2023 at 18:32 PM.

As shown in Figure 02, the multiplicative model is a linear model used in time series analysis, especially when the variables are unstable over extended periods. It multiplies the seasonal, cyclical, and random components to produce a more accurate series.

There are several time series algorithms used for forecasting. One of the simplest is the Moving Average algorithm, which calculates the average of past values and uses it to estimate future values. Another is the ARMA algorithm, which combines autoregression and moving average components. A more advanced version is ARIMA, one of the most widely applied algorithms in time series forecasting. (Saadia Abdelkarim, 2012, pp. 378–379)

Another method is the Exponential Smoothing (ETS) algorithm. It is popular because of its simplicity and ease of use for most users.

Time series techniques, and predictive data mining more broadly, have received considerable international attention—particularly in the banking sector. These tools are essential for improving decision-making, analyzing customer behavior, and identifying transaction patterns. The value of applying predictive (and also descriptive) models in banking can be seen in several key areas:

- **Operational management:** Data mining supports better access to reliable and accurate data, helping improve business efficiency and daily performance.
- **Performance analysis:** It enables the creation of scientific, evidence-based reports for both external audiences—such as markets or regulatory bodies—and internal stakeholders, including executive leadership and decision-makers.
- **Data integrity:** Improved consistency and reliability of data builds confidence in decisions and enhances the credibility of management reports, especially in the context of audits.
- **Marketing and sales:** Forecasting customer behavior allows the bank to better understand client needs, improve service offerings, and ultimately build loyalty and increase profitability. (Yuvika Priyadarshini, 2017, p. 06)

When applying data mining in banking—regardless of the objective—it is important to follow a structured process to ensure the success of the implementation. This process is illustrated in the following figure:

**Figure 03: Stages of Applying Data Mining in the Banking Sector**



**Source:** Prepared by the author based on: Chitra K., Subashini B., Data Mining Techniques and its Applications in the Banking Sector, International Journal of Emerging Technology and Advanced Engineering, Vol. 3, No. 8, 2013, p. 2020.

From Figure 03, it becomes clear that the first step in applying data mining is to understand the bank's financial, economic, and social context. This step is essential to identify the risks and opportunities that the bank may face. It also helps clarify the goal of data mining, which guides the data selection process.

For instance, if the goal is to identify non-performing loans and forecast default risk, then not all bank activities should be included. Only relevant data should be used, and the mining algorithms should be directed accordingly. The type of mining method—supervised or unsupervised—helps determine what needs to be done at each stage of the data mining life cycle.

### Using Data Mining Techniques to Improve Banking Performance

Data mining plays a key role in providing insights that support accurate decision-making. It improves banking performance by increasing the efficiency of operations. This is done by identifying patterns and predictions from large volumes of data that traditional methods cannot process quickly or accurately. Some of the main applications in banking performance improvement include:

## Customer Relationship Management (CRM)

Data mining supports all stages of the customer relationship management process. This process includes four key areas:

- Identifying customers, segmenting them, and analyzing target profiles;
- Acquiring new customers;
- Measuring customer satisfaction;
- Developing customer relationships. (Adrian Ionut, p. 162)

For existing customers, the focus is on:

- Increasing customer value;
- Improving customer retention. (Preethi M., Vijayalakshmi M., 2017, p. 03)

CRM uses data mining in the following ways:

- Grouping customers with similar traits or preferences to better understand their needs. (Preethi M., Vijayalakshmi M., 2017, p. 03)
- Predicting which customers are most likely to pay late or default, which helps in assessing credit and liquidity risk.
- Identifying reasons why customers remain loyal to a specific bank or why they leave for competitors.
- Explaining customer behavior to improve loyalty, provide better support to current clients, and design new services that match external customer behavior trends. (Application of Data Mining Techniques in the Banking Sector, 12/2022)

To ensure the success of CRM through data mining, three key steps should be followed:

- Measuring the customer retention rate;
- Identifying internal and external factors that affect customer behavior;
- Designing appropriate actions and addressing the root causes that influence customer decisions. (Dilmuradjon Zakirov, 2015, p. 06)

## Risk Management

Since data mining requires a clear understanding of the data type and the goal, this section focuses on the different types of risks banks face. The financial sector, especially banking, is a core part of the economy. This makes it vulnerable to many types of risks.

Data mining supports risk management in areas such as:

- Distinguishing between borrowers who repay on time and those who do not;
- Forecasting bad loans, especially when borrowers are in default;

- Analyzing whether a customer may delay or miss payments;
- Evaluating customer reliability and behavior during credit card transactions;
- Predicting future loan behavior. (K. Chitra, B. Subashini, 2013, p. 226)

## **Fraud Detection**

Fraud detection is one of the most common uses of data mining in banking. Fraudulent activities are a major concern for all financial institutions. (Imran Kazi, Bassar Qazi, 2012, p. 741)

Fraud detection involves identifying unusual or suspicious transactions. This is done by classifying transactions, spotting outliers, modeling past user behavior, and comparing it to current behavior. If significant deviations are detected, this may indicate fraud. (Sreekumar Pulakkazhy, 2013, p. 1257)

There are two main approaches developed by the banking and financial industry:

- The bank relies strictly on internal data to detect fraud patterns.
- The bank uses external databases from third-party providers. These are cross-referenced with internal data using data mining techniques to uncover fraud. (Imran Kazi, Bassar Qazi, 2012, p. 741)

Other techniques used in fraud detection through data mining include:

- Generating alerts when measurable deviations are found in customer transaction patterns;
- Using clustering techniques to classify transactions and detect outliers;
- Modeling the probability density of past credit card behavior and comparing it with current usage patterns;
- Detecting fraud in financial records that are hard to analyze through standard audit procedures. These may include asset understatements or overstatements of sales, profits, or assets.

To support this, classification techniques like regression and neural networks are used. These help describe and interpret fraud indicators using non-fraudulent data as a reference. (Sreekumar Pulakkazhy, 2013, p. 1257)

## **Among the Most Common Fraud Detection Applications: Anti-Money Laundering**

One of the most prominent areas of fraud detection is money laundering. This refers to the process of converting illegal financial flows and income into seemingly legitimate assets. If not detected and prevented, money laundering can have a serious impact on banks, both financially and reputationally.

Modern banks use data mining to identify and limit complex money laundering patterns. This includes:

- Identifying patterns of suspicious transactions;
- Detecting customer- or account-related risk models;
- Grouping accounts based on discovered similarities through clustering (Adrian Ionut, pp. 164–165)

## Banking Operations Management

Data mining also contributes to the management of banking operations in several key areas:

### Investment Banking Services

Banks often offer investment services to encourage clients to invest in specific assets or instruments. Data mining helps improve the performance of these services by:

- Predicting asset prices, such as stock values, which can lead to significant investment returns;
- Selecting the best investment options based on the client's profile and the projected return on investment. (Sreekumar Pulakkazhy, 2013, p. 1257)

### Supporting Accounting Decisions

Data mining supports accounting in various ways, including:

- Providing accurate, error-free data that is timely and not duplicated. This helps improve the quality of decisions made by accountants;
- Managing costs by analyzing cost data, determining optimal expenditure, and forecasting cost trends;
- Forecasting cash flows and the bank's financial performance. (Waheed Mahmoud & Ghazi, 2021, p. 343)

It is also important to highlight the role of data mining in analyzing financial statements. This provides a clearer picture of the bank's financial standing. It helps in evaluating its capacity to grant loans and accept deposits in the future, based on analyzing financial ratios.

### Applied Case Study

To support the theoretical framework discussed earlier, this applied case study focuses on predicting future loan behavior based on clients' historical borrowing patterns in a selected Algerian public bank during the period 2023–2027.

The analysis relies on historical loan data from 2018 to 2022. It aims to link loan risk management with customer behavior forecasting and relationship management. Predictive data mining techniques were employed using Power BI, which offers advanced tools for data analysis and visualization.

The study followed a structured scientific methodology consisting of several sequential phases. These included collecting, cleaning, processing, and analyzing historical loan data. The loans were categorized into three types:

- **Fully subsidized loans**, where the interest rate equals zero;
- **Partially subsidized loans**, where the client pays part of the interest rate and the state covers the rest;
- **Loans with a personal contribution**, where the client covers all or most of the cost.

Following this classification, predictive models were built and applied to estimate future trends for each loan type. The study benefited from the analytical features of Power BI, particularly the built-in predictive

algorithms based on Time Series models, including ARIMA (Autoregressive Integrated Moving Average), as well as machine learning forecasting tools.

The study used the additive model discussed earlier in the theoretical part. The platform also enabled interactive and visual representation of data, which simplified the analysis of trends and patterns in loan behavior.

This applied section aims to answer core questions related to the future of the bank's lending activity, such as:

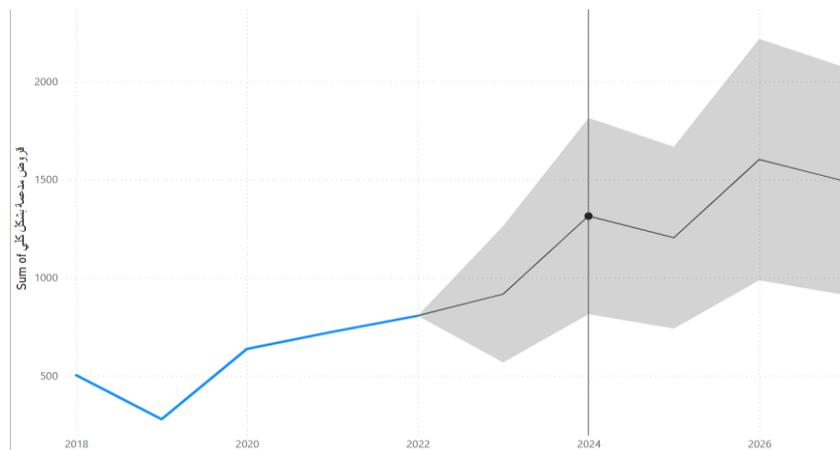
- What types of loans are likely to grow in the coming years?

The goal is to enhance the bank's competitiveness and ensure the sustainability of its financial performance in a shifting banking environment.

The following figures present the projected estimates for each type of loan.

### Fully Subsidized Loans

Figure 04: Forecasted Estimates for Fully Subsidized Loans



### Source: Power BI Results

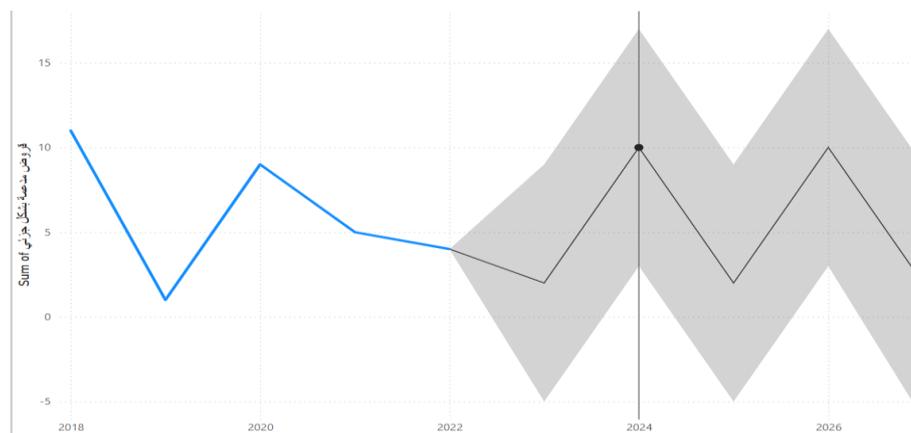
The results show a steady increase in **fully subsidized loans** (i.e., loans with zero interest rates) over the next five years. This upward trend can be explained by several key factors:

- The Algerian government may be adopting **stimulus policies** to support the economy by expanding the volume of subsidized loans. This is especially relevant in light of current economic challenges and the national strategy to diversify the economy beyond oil dependency.
- The state appears to be directing its support toward **specific sectors** such as agriculture, manufacturing, small and medium enterprises, or housing. This is likely being done by offering **interest-free loans** to encourage investment in these areas.
- The **forecasting model in Power BI** may have been based on a rising trend in historical data for subsidized loans from 2018 to 2022. As a result, the model extrapolated that trend into the future.

To support the accuracy of these forecasts, it may be useful to analyze the relationship between subsidized loan volumes and other economic indicators. These may include GDP growth rates, planned government spending, and public budget trends during the forecast period.

### Partially Subsidized Loans

Figure 05: Forecasted Estimates for Partially Subsidized Loans



### Source: Power BI Results

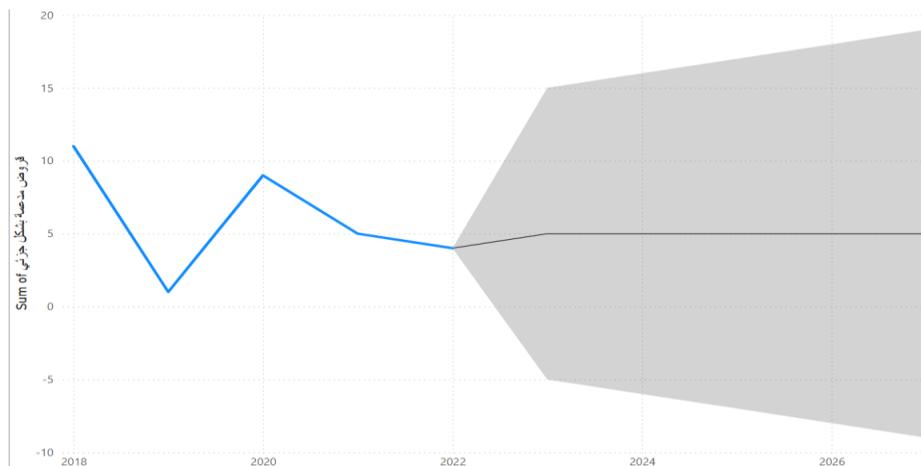
The observed fluctuations in the forecast curve for partially subsidized loans may be attributed to a range of overlapping economic and administrative factors. These may include:

- Changes in interest rates, which can significantly influence both individual and corporate borrowing decisions. Lower interest rates may encourage borrowing, while higher rates may discourage it.
- Variations in loan demand resulting from shifting individual or business needs. These may stem from broader economic conditions or sector-specific developments, such as increased loan applications to finance specific projects, or reduced demand due to austerity measures.

Thus, the volatility seen in the partially subsidized loan forecasts may reflect a complex interaction of these elements.

## Loans with Personal Contribution

**Figure 06: Forecasted Estimates for Loans Requiring Personal Contribution**



### Source: Power BI Results

A slight increase is observed at first in the forecast curve for loans with personal contribution, followed by a projected period of stability. These types of loans are often linked to investment loans due to the higher loan amounts, which usually require the borrower to contribute a portion of the funding. This expected stability may be explained by several factors closely tied to economic conditions, the behavior of financial institutions, and government policy.

One key reason behind this stability may be the existence of clearly defined, long-term government support policies for investment loans. For example, if the government continues to offer incentives or subsidies to investors, this can lead to consistent levels of loan disbursement that require personal contributions.

Accordingly, the projected estimates regarding customer preferences—particularly for fully subsidized loans—highlight how predictive data mining tools, powered by artificial intelligence, can support better customer relationship management. The anticipated growth in demand may reflect a governmental response to social needs such as improving living standards and reducing unemployment. This often leads to an increase in subsidized loans to stimulate economic activity.

However, this trend may also present risks for the bank. These risks could emerge if there are changes in government subsidy policies or imbalances in the bank's income due to the low interest generated from subsidized loans, compared to those with higher rates. Addressing this risk is essential to ensure the bank's long-term sustainability.

## Conclusion

This study concludes that forecasting using time series analysis, as one of the key predictive data mining techniques, provides valuable insight when applied to historical customer behavior and preferences. The analysis shows that continued reliance on government-backed financing policies encourages a sustained customer inclination toward subsidized loans.

The findings suggest that past behavior strongly influences future preferences, especially when previous experiences were positive. In this case, customers appear to view fully subsidized loans as satisfactory, as confirmed by the prediction results.

It is important to note, however, that these conclusions are based on the assumption that the existing government support policies will remain in place over the long term.

Therefore, the results reinforce the importance of relying on historical customer data as a primary input for understanding future behavior. The time series model used in this study proves effective not only in forecasting but also in supporting risk management, especially in relation to customer preferences and their growing focus on fully subsidized loans. These trends, if not managed properly, may increase exposure to loan-related risks.

## References

- Abdelkarim, S. (2012). The use of time series analysis to forecast the number of cancer patients in Anbar Governorate. *Anbar University Journal of Economic and Administrative Sciences*, 4(8), 378–379.
- Abd Elwahid, G., & Mahmoud, W. (2021). Requirements for applying data mining in auditing: Identifying audit risks as a model. *Journal of Research in Financial and Accounting Sciences*, University of M'sila, 6(1), 343.
- Chitra, K., & Subashini, B. (2013). Data mining techniques and its applications in the banking sector. *International Journal of Emerging Technology and Advanced Engineering*, 3(8), 2020–2025.
- Esling, P., & Agon, C. (2012). Time series data mining. *ACM Computing Surveys*, 45(1), 1–34. <https://doi.org/10.1145/2379776.2379788>
- Ionut, A. (2018). Data mining concepts and applications in the banking sector. *Annals of the “Constantin Brâncuși” University of Târgu Jiu, Economy Series*, (1), 161–167.
- Kazi, I., & Qazi, B. (2012). Use of data mining in banking. *International Journal of Engineering Research and Applications (IJERA)*, 2(2), 741–745.
- Osman, S., & Al-Shafie, J. (2014). Data mining and decision-making: An applied model for Khashm Al-Qirba reservoir. *White Nile Journal for Studies and Research*, 3, Sudan.
- Priyadarshini, Y. (2017). The effectiveness of data mining techniques in banking. *Computer Applications: An International Journal*, 4(1), 6–12.
- Preethi, M., & Vijayalakshmi, M. (2017). Data mining in the banking sector. *International Journal of Advanced Networking and Applications (IJANA)*, 8(5), 3–7.
- Pulakkazhy, S., & Balan, R. V. S. (2013). Data mining in banking and its applications: A review. *Journal of Computer Science*, 9(10), 1252–1257.
- Zakirov, D., & Momtselidze, N. (2015). Application of data mining in the banking sector. *Journal of Technical Science and Technologies*, 4(1), 1–6.
- Online Sources:
- Introduction to Time Series Analysis and Its Applications. (2022, August 1). Retrieved April 4, 2023, from [https://www.researchgate.net/publication/362389180\\_INTRODUCTION\\_TO\\_TIME\\_SERIES\\_ANALYSIS\\_AND\\_ITS\\_APPLICATIONS](https://www.researchgate.net/publication/362389180_INTRODUCTION_TO_TIME_SERIES_ANALYSIS_AND_ITS_APPLICATIONS)
- Application of Data Mining Techniques in the Banking Sector. (2022, December). Retrieved July 22, 2023, from <https://www.researchgate.net/publication/366324432>