Proceedings of the 3rd International Conference on Business and Policy Studies
DOI: 10.54254/2754-1169/77/20241678

Enterprise Digital Retail Business Data Analysis and Forecasting Based on Time Series Analysis

Dan Liu[1,a,*]

[1]*Business School, China University of Political Science and Law, Beijing,100088, China*
a. vanessaliu0427@hotmail.com
[*]*corresponding author*

Abstract: Digital enterprise management refers to the use of digital technology and information systems to manage various businesses and processes of an enterprise in order to improve its efficiency and competitiveness. Digital enterprise management is crucial to the development and operation of enterprises, which can help them realize more efficient production and management, improve their efficiency and competitiveness, as well as improve their innovation and flexibility, so that they can be more adaptive to market changes and customer needs. This paper focuses on the significance and methods of time series analysis of a retail enterprise's sales. By analyzing the historical sales data, we can predict the future sales trend and change rule, formulate corresponding sales strategies and plans, and improve sales efficiency and profit. Time series analysis is a commonly used data analysis method, in which AR model, MA model and ARMA model are three commonly used models, which have different principles and applications.AR model is a model based on the historical data of the time series itself for forecasting, which assumes that the future values are only related to the past values.MA model is a model based on the historical data of the error term of the time series for forecasting, which assumes that the future values are only related to the past values. The ARMA model is a combination of the AR model and the MA model, which can take into account both the autocorrelation of the time series and the correlation of the error term. By comparing the performance of the three models, the most suitable model can be selected for forecasting. For retail companies, sales are a very important indicator. Through time series analysis, it is possible to understand the seasonal and cyclical pattern of sales, predict the future sales trend and change pattern, adjust the product structure, inventory management and marketing strategy in time, and improve sales efficiency and profit. In addition, it can also understand the development trend and potential of the enterprise by comparing the sales of different time periods, which can provide support and reference for the strategic decision-making of the enterprise.

Keywords: Digital enterprise management, Digital enterprise management, ARMA.

1. Introduction

Digital enterprise management refers to the use of digital technology and information systems to manage various businesses and processes of an enterprise in order to improve its efficiency and competitiveness [1]. Digital enterprise management is crucial to the development and operation of enterprises, which can help enterprises realize more efficient production and management, improve

DOI: 10.54254/2754-1169/77/20241678

the efficiency and competitiveness of enterprises, and also improve the innovation ability and flexibility of enterprises, so that they can be more adaptive to market changes and customer needs [2,3].

Machine learning is a kind of artificial intelligence technology, which can automatically discover the laws and patterns in data and make predictions and decisions by analyzing and learning from the data [4]. In digital enterprise management, machine learning can help companies better understand the market and customer needs, forecast sales and demand, optimize production and supply chain management, and improve customer satisfaction and loyalty [5].

In digital enterprise management, sales forecasting is a very important application scenario. By analyzing and forecasting historical sales data, enterprises can better understand the market and customer demand, develop more reasonable sales strategies and plans, and improve sales efficiency and effectiveness [6]. At the same time, sales forecasting can also help enterprises better manage inventory and supply chain to avoid excess or out-of-stock situations and improve the efficiency and effectiveness of production and supply. Amazon uses machine learning algorithms to analyze and learn from massive amounts of data to predict the sales volume and demand of different products at different times, so as to optimize inventory and supply chain management and improve sales efficiency and effectiveness [7]; Walmart uses machine learning algorithms to analyze and learn from historical sales data and weather data, etc., to predict the sales volume and demand of different products at different times and locations, so as to optimize inventory and supply chain management to improve sales efficiency and effectiveness [8]; Baidu uses machine learning algorithms to analyze and learn from users' search and click behavior to predict the click rate and conversion rate of different advertisements, thus optimizing the advertisement placement strategy and effect, and improving advertisement revenue and customer satisfaction [9,10].

In conclusion, digital enterprise management is crucial for the development and operation of enterprises, and machine learning, as an advanced artificial intelligence technology, can help enterprises better understand the market and customer needs, predict sales and demand, optimize production and supply chain management, and improve customer satisfaction and loyalty, etc. In digital enterprise management, sales forecasting is a very important application scenario, which can help enterprises better understand the market and customer demand, develop more reasonable sales strategies and plans, and improve sales efficiency and effectiveness.

2. Introduction to data sets and statistical analysis of data

This paper uses a sales dataset. This dataset contains sales data spanning from January 2013 to December 2017 for a U.S. retailer. The dataset includes the daily sales volume of each product in each store, as well as some basic information about the store and the product, such as the city, type and name of the store, and the type and name of the product. In total, there are 50 different stores and more than 100,000 different products.

The task of this dataset is to predict the sales volume of each product in each store for the next 28 days. This task can be either a regression problem or a time series problem. Predicting future sales is very important for companies to help them optimize inventory management, develop promotional strategies, forecast revenues, etc., so this dataset has great practical value.

The challenge of this dataset is that the sales volume of each product in each store may be affected by different factors in different time periods, such as holidays, promotions, weather, and so on. Therefore, these factors need to be feature engineered to extract the factors that have an impact on sales volume and add them to the model for prediction.

This dataset can be used to train time series models such as ARIMA, Prophet, LSTM, etc. or machine learning models such as Random Forest, XGBoost, etc. The study of this dataset can help organizations to better understand sales data and improve the accuracy of sales forecasts.

DOI: 10.54254/2754-1169/77/20241678

The sales of this retailer are statistically analyzed according to the time series and a line graph is plotted as shown in Figure 1.

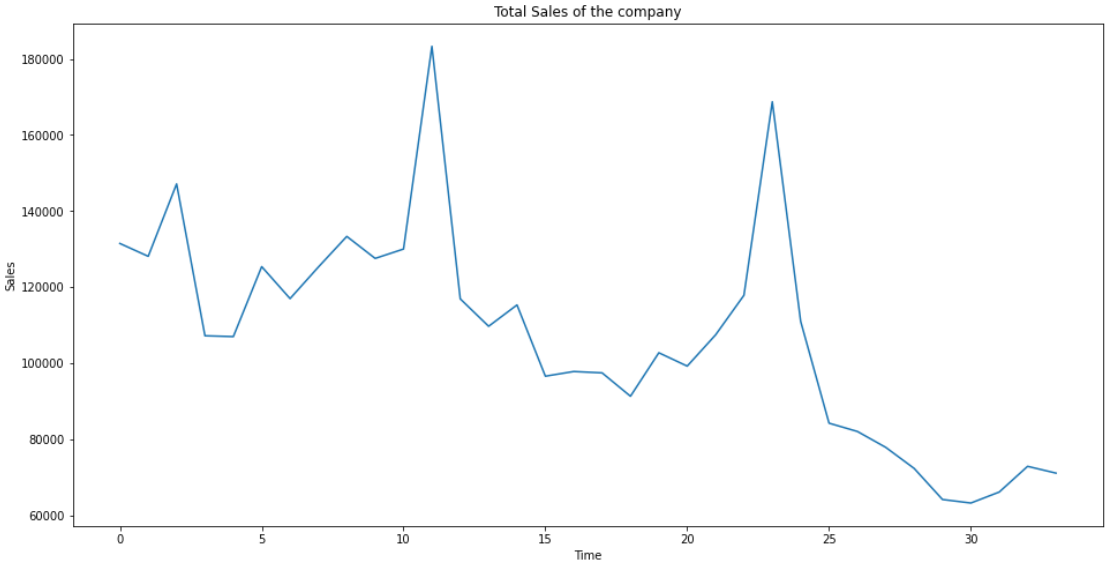

Figure 1: Line graph.
(Photo credit: Original)

3. Time series analysis

Autoregressive (AR), Moving Average (MA) and Autoregressive Moving Average (ARMA) models are commonly used in time series analysis. They can be used to forecast future time series values and also to study the characteristics and patterns of time series.

3.1. Autoregressive model

Autoregressive modeling is a time series analysis method and one of the most commonly used methods in time series forecasting. It predicts future data based on past data. The core idea of the autoregressive model is that the data of the current moment is associated with the data of the previous moment, so the data of the current moment can be predicted by the data of the previous moment.

The autoregressive model was first proposed by British statisticians Box and Jenkins in the 1970s. The basic idea of the autoregressive model is to view the time series as an autoregressive process, i.e., the data at the current moment is obtained from the data at the previous moment plus an error term. The error term is a random variable that represents the difference between the data at the current moment and the predicted value.

The autoregressive model has two important parameters, the lag order and the error term variance. The lag order refers to how many lagged terms are included in the model and the error term variance indicates the magnitude of the variance of the error term in the model. The determination of these two parameters usually requires model fitting and model testing.

Autoregressive modeling is very predictive, especially in long-term forecasting. It can predict the trend of future data from historical data, and for some highly cyclical time series data, the autoregressive model can better capture the pattern of its cyclical changes.

Autoregressive model also has some advantages, such as fewer model parameters, easy to understand and interpret, and less demanding on data. However, autoregressive models also have some disadvantages, such as only being able to deal with univariate time series, requiring high smoothness of data, and being easily affected by outliers. In practical applications, autoregressive

DOI: 10.54254/2754-1169/77/20241678

models usually need to be combined with other models to make predictions. For example, an autoregressive model can be combined with a moving average model to form an ARMA model to improve forecasting accuracy. In addition, the ARMA model can be combined with other models, such as ARIMA model and VAR model.

In conclusion, autoregressive model is a very important time series analysis method with wide application prospects. In practical application, it is necessary to choose the appropriate model according to the specific situation and carry out model fitting and model testing to improve the prediction accuracy.

3.2. Moving average model

Moving average modeling is a time series analysis method used to forecast future data. It predicts future data based on past data and is similar to autoregressive modeling. The core idea of the moving average model is that the data at the current moment is correlated with the data at previous moments, so that the data at the current moment can be predicted from the data at previous moments.

The moving average model was first proposed by British statisticians Box and Jenkins in the 1970s. The basic idea of the moving average model is to view the time series as a moving average process, i.e., the data at the current moment is obtained by adding an error term to the average of the data at previous moments. The error term is a random variable that represents the difference between the data at the current moment and the predicted value.

The moving average model has two important parameters, the lag order and the moving average window size. The lag order refers to how many lagged terms are included in the model, and the moving average window size indicates the size of the time window for moving averages in the model. The determination of these two parameters usually requires model fitting and model testing.

The moving average model is very predictive, especially in short-term forecasting. It can predict the trend of future data through historical data, and for some more cyclical time series data, the moving average model can better capture the pattern of its cyclical changes.

The moving average model also has some advantages, such as fewer model parameters, easy to understand and interpret, and less demanding on data. However, the moving average model also has some disadvantages, such as only being able to deal with univariate time series, requiring higher smoothness of data, and being easily affected by outliers.

3.3. Autoregressive moving average model

Autoregressive moving average (ARMA) modeling is a time series analysis method that combines the advantages of autoregressive (AR) and moving average (MA) models for forecasting future data.The core idea of ARMA modeling is that the data at the current moment is associated with the data at the previous moments and the error terms at the previous moments, and thus the data and error terms at the previous moments can be used to forecast the current moment's data.

The ARMA model is very predictive, especially in long-term forecasting. It can predict the trend of future data through historical data, and for some more cyclical time series data, the ARMA model can better capture the pattern of its cyclical changes.The ARMA model also has some advantages, such as fewer model parameters, easy to understand and interpret, and less demanding on data. However, ARMA model also has some drawbacks, such as it can only deal with univariate time series, requires high smoothness of data, and is easily affected by outliers.

In practical applications, we usually need to perform model selection and parameter estimation for time series. Model selection can be carried out by observing the autocorrelation and partial autocorrelation plots of the time series, the autocorrelation plots reflect the correlation of the time

DOI: 10.54254/2754-1169/77/20241678

series, and the partial autocorrelation plots reflect the partial correlation of the time series. Parameter estimation can be performed using maximum likelihood estimation or Bayesian estimation.

When using AR, MA and ARMA models for forecasting, we need to pay attention to the range of applicability of the model and forecasting accuracy. The applicability range of the model is affected by the characteristics and patterns of the time series, and the forecasting accuracy is affected by the order of the model and the accuracy of the parameter estimation. Therefore, when using AR, MA and ARMA models for forecasting, we need to carry out an in-depth study of the characteristics and laws of the time series, as well as optimization of the model selection and parameter estimation, in order to improve the forecasting accuracy and reliability.

In conclusion, AR, MA and ARMA models are commonly used in time series analysis, and they can be used to forecast future time series values as well as to study the characteristics and laws of time series. In practical applications, we need to conduct in-depth research on the characteristics and laws of time series, as well as optimization of model selection and parameter estimation, in order to improve forecasting accuracy and reliability.

4. Result

The AR, MA and ARMA models were used to analyze the time series of the sales of the retailer, respectively, to compare the predicted and actual values, and to draw Q-Q plots, as shown in Figures 2, 3 and 4.

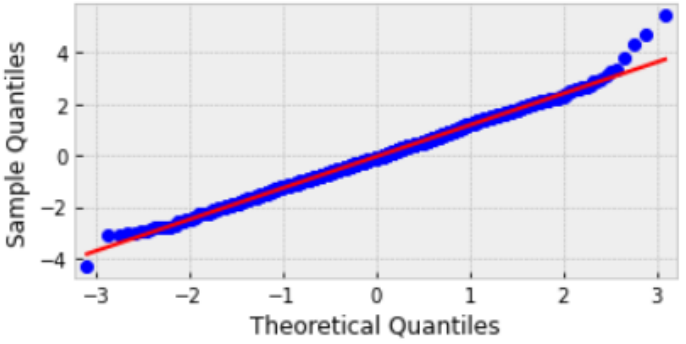

Figure 2: AR.
(Photo credit: Original)

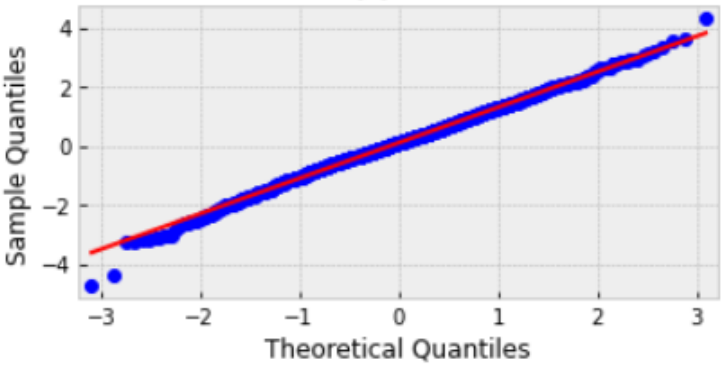

Figure 3: MA.
(Photo credit: Original)

DOI: 10.54254/2754-1169/77/20241678

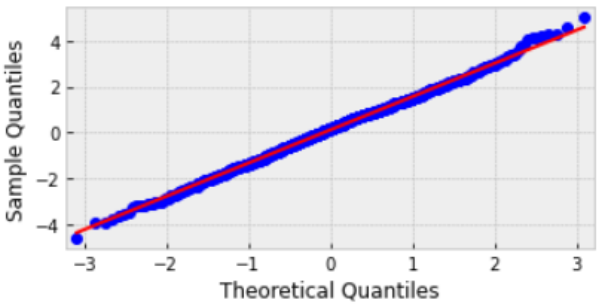

Figure 4: ARMA.
(Photo credit: Original)

As can be seen from the above figure, the predicted and actual values of the AR, MA and ARMA models are roughly on a straight line, indicating that the three time series models have well tapped into the pattern of time series changes. In addition, there are a few points that are not predicted accurately at the beginning and end of the AR model, and there are also a few points that are not predicted accurately at the beginning of the MA model, and the one that performs best is the ARMA model, which basically predicts accurately at all points.

The image of the actual values of the ARMA model is plotted against the predicted points as shown in Fig. 5. In addition, the fitted curve of sales volume is plotted based on the time series model as shown in Figure 6.

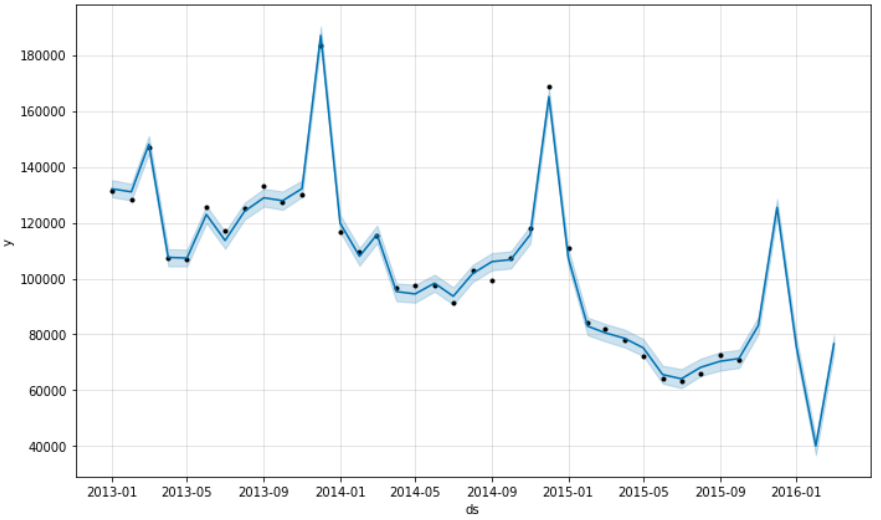

Figure 5: ARMA scatter discount plot.
(Photo credit: Original)

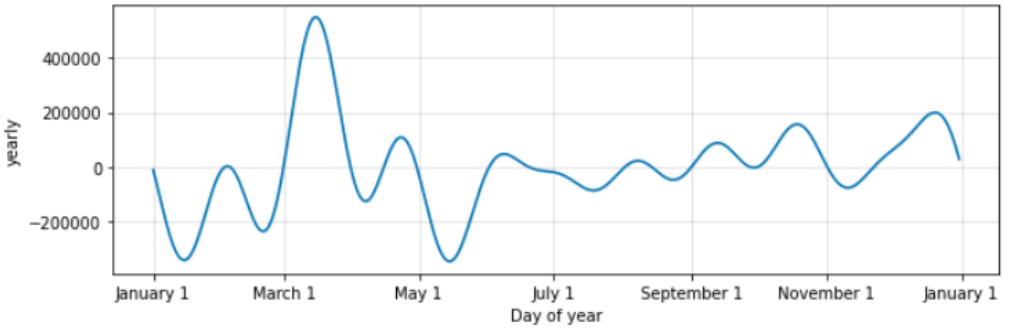

Figure 6: Timing analysis diagram.
(Photo credit: Original)

DOI: 10.54254/2754-1169/77/20241678

As can be seen from the above graph, the Prophet's trend and seasonality is similar to the trend and seasonality we previously obtained using traditional methods.

5. Conclusion

Time series analysis is a commonly used data analysis method, which can predict future trends and changing patterns by analyzing historical data. For retail enterprises, sales is a very important indicator, through time series analysis can better understand the sales trends and patterns, develop corresponding sales strategies and plans, and improve sales efficiency and profits.

AR model, MA model and ARMA model are three models commonly used in time series analysis, and their principles and applications are different.

The MA model is a model for forecasting based on historical data of time series error terms, which assumes that the future values are only related to the past errors.The predicted values of the MA model are mainly dependent on the past errors, so if there is a periodicity or trend in the time series, the accuracy of the MA model prediction will be affected as well.The MA model has the advantage of eliminating the stochastic fluctuations of the time series very well, but the for time series data with longer periods, its prediction may be less satisfactory.

ARMA model is a combination of AR model and MA model, which can consider the autocorrelation of the time series and the correlation of the error term at the same time.The predicted value of ARMA model can more accurately reflect the change rule of the time series, and therefore the performance effect is the best.The advantage of ARMA model is that it can deal with the cyclical and trending time series data very well, but for the nonlinear and nonsmooth time series data, its prediction effect may be less satisfactory. series data, its prediction effect may be less satisfactory.

By analyzing the time series of sales of retail enterprises, some valuable information and conclusions can be obtained. Firstly, the seasonal and cyclical patterns of sales can be understood and corresponding sales strategies and plans can be formulated. Secondly, it can predict the future sales trend and change rule, adjust the product structure, inventory management and marketing strategy in time, and improve the sales efficiency and profit. In addition, it can also understand the development trend and potential of the enterprise by comparing the sales of different time periods, providing support and reference for the strategic decision-making of the enterprise.

References

[1] *Alberto J M ,R. P P ,Félix V .Evolution of digital transformation in SMEs management through a bibliometric analysis[J].Technological Forecasting Social Change,2024,199123014-.*

[2] *Tim H ,Reinhard M .More capable, more innovative? An empirical inquiry into the effects of dynamic managerial capabilities on digital firms' innovativeness[J].European Journal of Innovation Management,2022,25(6):892-915.*

[3] *Ajay J ,Rohit S ,Ashish D , et al.Sustainable recovery for digital entrepreneurs with shared resources: enablers, challenges and solutions[J].Journal of Asia Business Studies,2022,16(3):515-537.*

[4] *S P R V ,Poruri R .Power of Apis to Build Digital Enterprises[J].Journal of Global Economy, Business and Finance,2021,3(9):*

[5] *UK Wealth Management Research Portfolio Tribe Digital Entrepreneurs[J].M2 Presswire,2013,*

[6] *Business News; BMC Database Management Solutions Support DB2[Registered trade-mark] 12 - A Leap Forward in Innovation and Support for the Digital Enterprise[J].Information Technology Business,2017,*

[7] *Thanachawengsakul N .A Conceptual Framework for the Development of a MOOCs-Based Knowledge Repository to Enhance Digital Entrepreneurs' Competencies[J].International Journal of Information and Education Technology,2020,10(5):*

[8] *PARABLU: An award winning provider of secure data management software, engineers new-age cloud data protection solutions for the digital enterprise.[J].M2 Presswire,2022,*

[9] *Daniel L ,G. D A . Topological data analysis and machine learning[J]. Advances in Physics: X,2023,8(1).*

[10] *Chaolin L ,Can P ,Hong Y , et al.Social capital and innovation performance of digital firms: Serial mediation effect of cross-border knowledge search and absorptive capacity[J].Journal of Innovation Knowledge,2022,7(2):*