International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2019 IJSRCSEIT | Volume 5 | Issue 2 | ISSN : 2456-3307
DOI : https://doi.org/10.32628/CSEIT1952164

An Effective College Prediction System using Time Series Analysis

Ravi Parmar[1], Dhaval Tandale[1], Avdhut Kanago[1], Rahul Repal[1], Prof. Geeta S. Navale[2]

[1]Department of Computer Engineering, Sinhgad Institute of Technology and Science, Narhe, Pune, Maharashtra, India
[2]Head of Department of Computer Engineering, Sinhgad Institute of Technology and Science, Narhe, Pune, Maharashtra, India

ABSTRACT

In recent times time series analysis has gained more importance with increasing applications. The time series data is related with a time stamp for each data. One of the possible applications is the prediction of cut-off of a college using time series analysis over is previous cut-offs. It is very important for a student to secure the best possible college for his graduation degree. For his further undergraduate studies, the student needs to apply for a list of colleges. It is very crucial which colleges the student applies for and what are the chances of him getting admission into that college. The future cut-off of a college can be predicted using methods such as time-series analysis which will aid the students to decide which colleges to apply to.

Keywords : Data Analytics, Time-Series, Data Prediction, ARIMA Model, Time Series Forecasting

I. INTRODUCTION

Often students are confused about which colleges to apply. Also, most students especially who are from some other city or area do not know about the colleges and are unaware about the details of various colleges Thus, the students need some guidance while they select the list of colleges to apply for. While applying for a college there are many factors which need to be taken into consideration. It is a complicated process. Factors such as the cut-off marks of the college, locality of the college, etc. must be taken into consideration.

The trend in the cut-off marks of the college shows the incline or decline in the demand of that college which also shows the performance of the college. Thus, by studying the previous trends in the cut-off marks of a college and by establishing a graph of the cut-offs over a time series we can analyse it and furthermore predict the future cut-off of that college.

The following paper is a survey of various techniques and models that can be used for analysis and forecasting of such academic data.

II. LITERATURE REVIEW

Sr No.	Paper Title	Paper theme/Idea	Advantages
1	Research on time series data prediction based on clustering algorithm - A case study of Yuebao	This paper compares different models for time series analysis and the results generated by using ARMA model on Yuebao case study	The model of comprehensive clustering and forecasting Provide better forecasting results. ARMA model can avoid the complicated task of estimating multiple parameters in AR and MA models.

Sr No.	Paper Title	Paper theme/Idea	Advantages
2	Mining Time Series Data with Apriori Tid Algorithm	A modified version of Apriori Tid is proposed which is a Association Rule Mining (ARM) algorithm, is implemented over time series data to find out frequent item sets.	The output of Apriori Tid and modified Apriori Tid algorithm is same, however difference lies in case of execution time. Modified Apriori Tid algorithm outperforms Apriori Tid.
3	Decision Tree Classification and Forecasting of Pricing Time Series Data	Different methods and algorithms for solving these problems have been discussed, implemented and tested	The time series analysis shows that it is possible to use the information within each class to do predictions, and a simple vector autoregressive model used to perform it shows good predictive results
4	Performance comparison and future estimation of time series data using predictive data mining techniques	The Time Series mining techniques like Linear Regression and Autoregressive integrated moving averages (ARIMA) model are compared	Both the models almost give the same results. Due to low Root mean square error (RMSE)Linear Regression model performs better

III. METHODS AND MATERIAL

Model selection

Time Series is nothing but graph of values with respect to time. For modelling the time series analysis number of questions are to be focused on. Such as is the time series stationary? Disorders? Based on the type of data various models can be used for prediction of the future values in a time series.

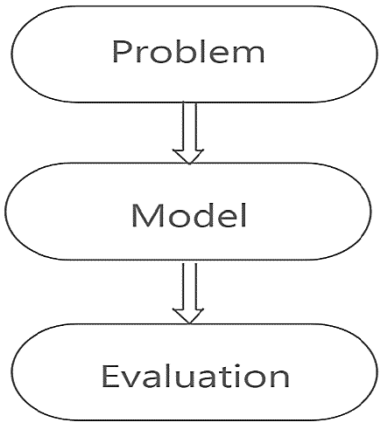

Fig 1 : Procedure for time series modelling

We have to pick up the model based on structure of the problem and validity of assumptions. Each selected model will have to be evaluated by testing it on data and its parameters must be future modified to improve accuracy.

Another important aspect in time series is that some time series appear cyclic features at a fixed time interval. Namely, time series exhibits a feature of seasonal varying [5]. We need to consider an appropriate model to deal with such data.

The following components of time series need to be studied before selection of model - 'levels' that is the baseline value for the series if it were a straight line, 'trend' which is the pattern of linear increasing or decreasing behavior in the time series and 'seasonality' which is the phenomenon of repeating patterns over a particular time.

Models used

Various models can be used for prediction of future values in a time series. The most commonly used models are autoregressive (AR) model, purely moving average (MA) model and autoregressive moving

average (ARMA) model. AR model and MA model can be regarded as a special case of ARMA model [5]. ARMA model is the most commonly used model that fits a smooth sequence. The problem with these models is that they work accurately only for stationary data. For data like cut-off of a college, which vary every year and may show a trend, these models are not applicable.

ARIMA model ARIMA(p,d,q) is the model used for non-stationary data. The seasonal components are handled in this model. The essence of the ARIMA model is the difference operator with a combination of the ARMA model. This shows that any non-stationary sequence will be stable if by appropriate order difference, then we can do ARMA model fitting after the difference [5]. We generate a new series in ARIMA model by taking the difference from the previous time point to the current tie point. This will create a stationary series. The limitations for this model is that it can only predict accurate data over a short time span and it does not take into consideration other factors which may affect the future value. For example, a change in college management may lead to drastic change in the cut-off of that college due to increase or decrease in the demand of that college, but ARIMA model cannot take that into consideration.

IV. CONCLUSION

A comparative analysis on existing models shows us various parameters to consider while selecting a model for time series forecasting, the common models used and their limitations. The papers studied can be combined for an creating an effective system for prediction of future cut-off of a particular college by studying the time series of its previous cut-offs.

V. REFERENCES

[1]. Xu Lu, and Tianzhong Zhao, “Research on time series data prediction based on clustering algorithm - A case study of Yuebao”, 2017, AIP Conference Proceedings 1864, 020152 (2017); doi: 10.1063/1.4992969.

[2]. EMIL LUNDKVIST, “Decision Tree Classification and Forecasting of Pricing Time Series Data”, July 2014.

[3]. Hiren Kumar Deva Sarma, Swapnil Mishra, “Mining Time Series Data with Apriori Tid Algorithm”, 2016 International Conference on Information Technology.

[4]. Tanwar, H, Kakkar M, “Performance comparison and future estimation of time series data using predictive data mining techniques”, 2017 International Conference on Data Management, Analytics and Innovation (ICDMAI). doi:10.1109/icdmai.2017.8073477

[5]. Xiaoguo Wang , and Yuejing Liu, “ARIMA Time Series Application to Employment Forecasting”, 2009 , Proceedings of 2009 4th International Conference on Computer Science & Education, 978-1-4244-3521-0/09/$25.00©2009IEEE

Cite this article as :

Ravi Parmar, Dhaval Tandale, Avdhut Kanago, Rahul Repal, Prof. Geeta S. Navale, "An Effective College Prediction System using Time Series Analysis", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 5 Issue 2, pp. 1186-1188, March-April 2019. Available at doi : https://doi.org/10.32628/CSEIT1952164 Journal URL : http://ijsrcseit.com/CSEIT1952164