MJRET

OPEN ACCESS

Multidisciplinary Journal of Research in Engineering and Technology

www.mjret.in

eISSN:2348-6953

Study of Different Prediction and Forecasting Techniques for Students' Performance Prediction

Gayatri Pawashe, Sangeetha Sahadevan, Shivangi Mishra,
Vaishnavi Kalgutkar, Prof Rajesh D. Bharati
Department of Computer Engineering, Dr. D.Y. Patil Institute of Technology,
Pimpri, India

Abstract: *The performance prediction of a student in different subjects is important so that the professors as well as the students can gain an insight into their future performance and thus improve before failing. Lately, the number of dropouts has increased in the context of advanced education, affecting the image of schools and their occupation. Different machine learning algorithms and models can be used for the prediction and forecasting purpose. The existing system explores the capacity to send and collect knowledge from undergraduates but, in addition, it is impossible to find useful information. The proposed system will predict the performance of the student. This paper presents the study of different available algorithms, which will be suited for the proposed system. The paper concludes with a new system which can overcome the shortcomings of the present system that only stored the student information and various algorithms suited for it.*

Keywords: *ANN Algorithm, ARIMA, Forecasting, Linear Regression Algorithm, Machine Learning, Prediction, SARIMA*

1. INTRODUCTION

It is a tedious task for people to analyze the trend of the enormous amount of data to shape simplified useful information. Forecasting is the technique that takes the data and predicts the future value of the data looking at its unique trends. For example, predicting average annual company turnover based on the last year's data. Predictive analysis factors in the variety inputs and predicts the future behaviour and not just a number.

The highprecision prediction in the students ' success is useful because it helps to distinguish students with little academic achievement early in academia.University student success is related to academic performance and enrollment. The measures to help educate the weak academics are:

1. Generation of data base predictive variables.
2. Identification of various traits or variables that affect the academic career performance of stud ents.
3. Create a predictive model based on statistical variables.
4. Validation of the model developed for universities with student success.

Different machine learning algorithms are used for the purpose of prediction and forecasting. This paper contains the study of different existing systems for predicting students' performance such as Naive Bayes, Regression, ARIMA, ANN, etc.

2. LITERATURE SURVEY

Ms. Timsy Devasia, Ms. Vinushree TP and Mr. Vinayak Hedge [12] found that the percentage of students dropping out of schools could be decreased if the current system that juts the students' data is turned into a system they proposed. They suggested and smart program, which was a web-based application, which allows use of the Naive Bayes algorithm to collect valuable knowledge and predict potential student outcomes. We performed the experiment with data from 700 students, with 19 attributes, and found that the most successful was the Naive Bayes algorithm.

Based on the C4.5 and Naive Bayes classifiers, Mr. Masna Wati and Mr. Wahyu Indravan [13] compared the performance of data mining algorithms to predict students learning outcomes. We found the accuracy of both was above 60% but the precision of c4.5 was just 58.82%. Most previous researchers concentrated on predicting the outcome of success using data mining in academic database and evaluated the algorithms of data mining in each case they took. In this study, they studied more predictive models. In RapidMiner Studio, they used the training data set to predict how well a model would work.

Vaccaro et al [7] the paper proposes hybrid architecture for electricity price forecasting. The proposed architecture combines the reward of the easy- to-use and comparatively simple to- tune Auto regressive Integrated Moving Average (ARIMA) model and the approximation power of local learning tech- niques. The architecture is robust and more accurate than the individual forecasting methodologies on which it is based, since it combines a reliable built-in linear model (ARIMA) with an adaptive dynamic corrector (Lazy Learning algorithm). The corrector model is sequentially updated, in or- der to adjust the whole architecture to varying market conditions. Detailed simulation studies show the effectiveness of the proposed hybrid learning methods for forecasting the volatile Hourly Ontario Energy Prices (HOEPs) of the Ontario, Canada, and electricity market.

Guoqiang Liu et al [5] in the software program reliability boom section, the character of the failure records in an experience, decided by means of the software testing method. A hybrid version is proposed for medium and lengthy-term software program failure time forecasting on this paper. The hybrid version consists of two techniques, Singular Spectrum Analysis (SSA) and ARIMA on this version, the time series of software failure time are firstly decomposed into numerous sub-series corresponding to some tendentious and oscillation (periodic or quasiperiodic) components and noise by using the us- age of SSA after which every sub-series is expected, respectively, through the best ARIMA version, and lastly a correction system is conducted for the sum of the prediction results to ensure the residual to be a natural random collection.

Ling Wang et al [9] presents Based totally on the evaluation of the measured put on records and the wear characteristics in terms the wheels of Guangzhou Metro Line1, the accumulative put on prediction approach of metro wheels based totally on the ARIMA (p, d, q) model is proposed in this paper. According to the time series modelling method of the ARIMA (p, d, q) model, the stationarity evaluation and transformation of the metro wheel wear records are described at first. Then, with the application of the AIC criterion and the most likelihood Estimation method, the model order is determined and the version parameters are derived for the ARIMA (p, d, q) model in the end, the flange thickness and the diameter of the metro wheels will be anticipated by means of this advanced

ARIMA(p, d, q) model. The effects show that the proposed prediction method is straightforward and effective for the quick-term prediction of the metro wheel.

W. Jacobs et al [1] presents This study aims to predict the standards of the time series of milk insist in a dairy industry by combining forecasting of Box-Jenkins and artificial neural network replica and contrast the results to the individual models, exemplifying the joint forecast for the production planning. Eight prediction combining technique were used and, behind the use of statistical methods, the consequences obtain by fitting the Box- Jenkins and artificial neural network template were contrast with the results obtained in the future combinations. The results showed that the mixture of seasonal Box-Jenkins and de-seasonalise artificial neural network model by the inverse denote square method, provide a performance in the forecast for six months in front 66.5 higher than the individual models, where the combination of forecasts provided a root mean square error of 1.43 and mean absolute percentage error of 2.16. The forecast for 12 months ahead, the presentation of the mixture was 56.5 higher compare to entity models, with root mean square error of 2.86 and mean absolute percentage error of 3.70. In both cases, the combination of predictions showed superior results.

Kalid Yunus et al [8] presents this paper, a modified auto regressive integrated moving average (ARIMA) modeling technique, which could capture time correlation and possibility distribution of determined wind-pace time collection records, is offered. The technique introduces frequency decomposition (split- ting the wind-speed information into high frequency (HF) low frequency (LF) components), shifting, and limiting further to differencing and energy trans- formation that are used within the trendy ARIMA modeling system.

3. METHODOLOGY

Various algorithms for prediction are reviewed in this paper out of which some of the best algorithms are discussed in this section.

Artificial Neural Network:

The term "Artificial neural network" refers to a biologically inspired sub-field of artificial intelligence modeled after the brain. An Artificial neural network is usually a computational network based on biological neural networks that construct the structure of the human brain. Similar to a human brain has neurons interconnected to each other; artificial neural networks also have neurons that are linked to each other in various layers of the networks. These neurons are known as nodes. An Artificial Neural Network in the field of Artificial intelligence where it attempts to mimic the network of neurons makes up a human brain so that computers will have an option to understand things and make decisions in a human-like manner. The artificial neural network is designed by programming computers to behave simply like interconnected brain cells. There are around 1000 billion neurons in the human brain. Each neuron has an association point somewhere in the range of 1,000 and 100,000. In the human brain, data is stored in such a manner as to be distributed, and we can extract more than one piece of this data when necessary from our memory simultaneously. We can say that the human brain is made up of incredibly amazing parallel processors.

Sarima:

Seasonal Autoregressive Integrated Moving Average, SARIMA or Seasonal ARIMA, is an extension of ARIMA that explicitly supports univariate time series data with a seasonal component.

It adds three new hyperparameters to specify the autoregression (AR), differencing (I) and moving average (MA) for the seasonal component of the series, as well as an additional parameter for the period of the seasonality.

The algorithm is easy to use and does not require expert knowledge. It is useful for the high-level automation of forecasting jobs that are characterized by the following:

- Forecasts are calculated for the near future.
- Products typically have long lifecycles.
- Considerable sales history is available.
- There are no noticeable dependencies between the historical sales and other variables OR
- Data on other demand influencing variables isn't available.

SARIMA Process

In case of seasonal auto-ARIMA (auto-SARIMA), all process steps are performed on seasonal cycles instead of distinct time periods. The length of the seasonal cycle can be provided by you or identified by the algorithm.

4. FUTURE SCOPE AND APPLICATION

Future Work

This system can be used by educational institutions to predict their result for the following semesters as well as pay more attention towards weak students. Students can use the proposed system to track their academics. The number of college dropouts could be decreased

Application

Schools and colleges will use this new program for predicting students' performance which will help in reducing college dropouts.Also, it will help the faculty members to emphasis on students' weaker areas and provide the extra guidance in those areas.

5. CONCLUSION

In this paper various machine learning algorithms for students' performance prediction are assessed. Different machine learning algorithms can be used for predicting student results. Algorithms like naïve bayes, regression, artificial neural network, sarima. Sarima algorithm overcomes current system's disadvantages. Seasonal Autoregressive Integrated Moving Average, SARIMA or Seasonal ARIMA, is an extension of ARIMA that explicitly supports univariate time series data with a seasonal component.

ACKNOWLEDGEMENT

The authors would like to thank the management of Dr. D. Y. Patil Institute of Technology, Pimpri for their expertise and assistance throughout all aspects of our study and for their help in writing the manuscript.

REFERENCES

[1]. *W. Jacobs, A. M. Souza and R. R. Zanini Combination of Box-Jenkins and MLP/RNA Models for Forecasting Combining Forecasting of Box- Jenkins, IEEE LATIN AMERICA TRANSACTIONS, 14(4), 2016.*

[2]. *S. I. Vagropoulos, G. I. Chouliaras, E. G. Kardakos, C. K. Simoglou and A. G. Bakirtzis, "Comparison of SARIMAX, SARIMA, Modi- fied SARIMA and ANN-based Models for short-term PV generation forecasting", IEEE International Energy Conference (ENERGYCON), Leuven 1-6, 2016.*

[3]. *Rodrigo N. Calheiros ; Dept. of Comput. Inf. Syst., Univ. of Mel- bourne, Melbourne, VIC, Australia; Enayat Masoumi; Rajiv Ranjan; Rajkumar Buyya, Workload Prediction Using ARIMA Model and Its Impact on Cloud Applications, QoS IEEE Transactions on Cloud Com- puting, 3(4), 2015.*

[4]. *Takaomi HIRATA, Takashi KUREMOTO, Masanao OBAYASHI, Shingo MABU, Time Series Prediction Using DBN and ARIMA, International Conference on Computer Application Technologies, IEEE, 2015*

[5]. *G. Liu, D. Zhang and T. Zhang, "Software Reliability Forecasting: Sin- gular Spectrum Analysis and ARIMA Hybrid Model", Theoretical As- pects of Software Engineering (TASE), 2015 International Symposium on Nanjing, 111-118, 2015.*

[6]. *Sornpon Wichaidit, Surin Kittitornkun, Predicting SET50 Stock Prices Using CARIMA, International Computer Science and Engineering Con- ference (ICSEC), 2015.*

[7]. *A. Vaccaro, T. H. M. EL-Fouly, C. A. Caizares and K. Bhattacharya, "Local Learning-ARIMA Adaptive Hybrid Architecture for Hourly Elec- tricity Price Forecasting", PowerTech, 2015 IEEE Eindhoven, 1-6, 2015.*

[8]. *Kalid Yunus, Torbjorn Thiringer, and Peiyuan Chen, ARIMA-Based Frequency-Decomposed Modeling of Wind Speed Time Series, IEEE Transaction on Power Systems, 31(4), 2546-2556, 2016.*

[9]. *S. Li, H. Wang, Y. Tian, Y. Shen and A. Aitouche, "Wind Speed Fore- casting Based on Fuzzy-Neural Network Combination Method", The 27th Chinese Control and Decision Conference (2015 CCDC), Qing- dao, 4811-4816, 2015.*

[10]. *E. Guresen, G. Kayakutlu, T. U. Daim, Using Artificial Neural Network Models in Stock Market Prediction, Expert Systems with Applications, 38, 10389-10397, 2011.*

[11]. *T. M. Choi, Y. Yu, K. F. Au, A Hybrid SARIMA Wavelet Transform Method for Sales Forecasting, Decision Support Systems, 51, 130-140, 2011.*

[12]. *Ms.Tismy Devasia1 ,Ms.Vinushree T P2, Mr.Vinayak Hegde3, "Prediction of Students Performance using Educational Data Mining ".*

[13]. *Masna Wati1, Wahyu Indrawan2, Joan Angelina Widians3, Novianti Puspitasari, "Data Mining For Predicting Students' Learning Result".*