

Development of Decision Support System on Forecasting Workload Analysis using Time-series Algorithm

Joe G. Lagarteja¹

¹Isabela State University

ARTICLE INFO

ABSTRACT

Received: 29 Dec 2024

Revised: 15 Feb 2025

Accepted: 24 Feb 2025

The growing dynamics of faculty workload planning at Isabela State University (ISU) must be addressed, this study provides the development and implementation of a Decision Support System (DSS) with time-series forecasting to enable better academic resource planning. With ISU's dynamic fluctuations in student enrollment and diversified programs, manual methods have failed to cope with dynamic workload allocations among faculty members of ISU. This study aims to bridge that gap by integrating historical information with forecasting analytics through an ARIMA-LSTM hybrid model using different metrics such as Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). The DSS analyzes trends in class enrollments, faculty assignments, and enrollment records to predict upcoming faculty workloads. RMSE, MAE, and MAPE metrics confirm the high accuracy of model forecasts with MAPE as low as 1.25%. The DSS was predictive accuracy-tested and usability-tested, offering a more data-driven insights for university administrators to make more proactive and informed decisions within ISU system.

Keywords: Decision Support System, ARIMA, LSTM, Prediction, Forecasting, Faculty Workload

INTRODUCTION

Forecasting using time-series has been extensively utilized in various industries, from finance to logistics, for predicting future performance based on past and historical trends. In the university's setting, this technique can be useful in pattern identification of faculty workloads, teaching loads, and student enrollments. Model like Autoregressive Integrated Moving Average (ARIMA) has been very promising in forecasting academic workloads, providing a sound basis for prediction of complex systems like ISU's academic system (Secaran & Sathiyamoorthy, 2020; Xu, 2023). Despite of this, the usage of this model in a real-world and comprehensive DSS designed to ISU's practical operational requirements had not been completely investigated before.

At present, Isabela State University (ISU), one of the Philippines' public universities, faculty workload management has become increasingly complex with fluctuating enrollments, a broad range of programs, and varying academic calendars. These traditional approaches often failed to keep pace with the changing needs as it mainly depend on manual planning. In order to meet these challenges, this study was purposely tasked of developing and implementing a Decision Support System (DSS) that incorporates the use of time-series forecasting for creating data-driven information. Also, another purpose is to enable university leaders to make more informed, more strategic decisions regarding resource allocation and academic planning.

The DSS is combined with a wide array of historical workload information including class sizes, faculty assignments, enrollment patterns, and applies time-series algorithms to show current trends and predict future outcomes. Equipped with these predictive views, administrators and staffs are able to plan faculty loads more effectively, efficiently, make better resource allocations, and more refined academic schedules.

Development of the DSS included the employment of forecasting model as well as designing a user-friendly interface for ISU's administrators and staffs. The DSS was tested on both the forecasting accuracy and in practical application for real-life decision-making activity. Through data transformation, raw data was converted into insightful forecasts which has given ultimate power to ISU leaders to utilize resources more efficiently and connect planning efforts to the long-term to attain university objectives.

The study was essential not only for ISU but for other universities with the same kind of challenges. It highlights the ways in which data analytics, and specifically time-series forecasting, can be employed into academic administration to enable more intelligent, and more real-time and responsive decision-making (Soriano et al., 2018; Büyükkşahin & Ertekin, 2018).

OBJECTIVES

Generally, the objective of this study is to develop and implement a Decision Support System (DSS) capable of predicting workload patterns at Isabela State University based on a time-series algorithm. Specifically, the study aims to:

1. Employ the ARIMA-LSTM Hybrid Model to analyze and forecast trends in the faculty workload data;
2. Evaluate how well the ARIMA-LSTM using different metrics such as RMSE (Root Mean Square Error), MAE (Mean Absolute Error), and MAPE (Mean Absolute Percentage Error).

METHODS

The study included essential procedures such as data gathering, analysis of data, model identification, and assessment to forecast workload analysis based on a time-series algorithm

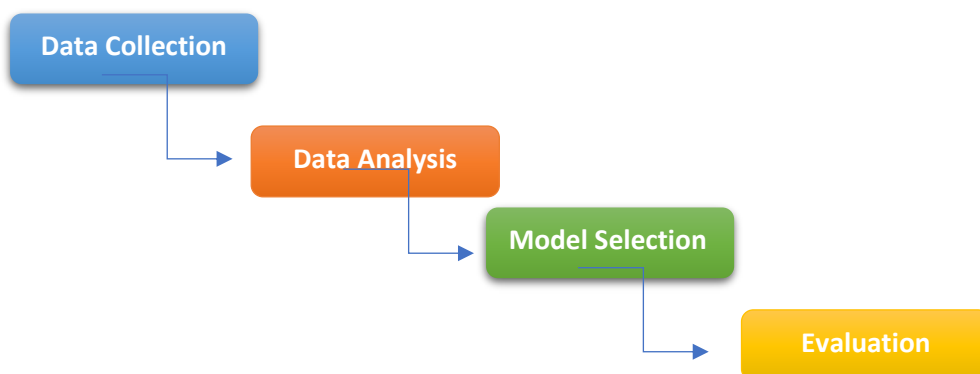


Figure 1. Summary of Methods Employed

In the process, data were collected and cleaned. The next step was to analyze the data patterns in order to guide the researcher for its implementation in ARIMA-LSTM hybrid model.

Also, researcher began with exploratory data analysis (EDA), focusing on main variables such as faculty teaching hours, student enrollment, and course loads for a period of three years, 2020 to 2023.

The researcher utilized correlation matrices to investigate associations whether or not increased enrollment was followed by greater teaching loads and tested the data for stationarity, a critical prerequisite for forecasting model. The researcher examined whether the average and variance of the data remained stable over time using the Augmented Dickey-Fuller (ADF) test. If the data series is not determined stationary, techniques such as differencing or log transformations were applied to stabilize it and render it more amenable and suited to modeling.

Additionally, the dataset was then split into training and test sets, with 80% of the data employed to train and refine the models, and the other 20% for testing the performance of the models in being able to predict future workloads.

In order to create a reliable and efficient decision support system for workload forecasting, this study was based mostly on past data obtained from Isabela State University's (ISU) administrative records. The data set comprised several indicators of academic workload, namely:

1. Teaching hours of faculty per semester, by department, subject, and rank
2. Student enrollment numbers, broken down by academic year, semester, program, and course section.
3. Course availability, such as the number of sections by course, lecture/laboratory mix, and credit hours allocated.

These criteria were pulled from ISU's academic planning, registrar, and human resources departments across a three-year period to include adequate historical breadth for comprehensive time-series modeling. The data included Academic Year 2020–2023 and both regular and special terms (e.g., summer or midyear offerings).

After being gathered, the data was subjected to rigorous preprocessing to validate quality, integrity, and readiness for time-series analysis such as data cleaning, normalization, time indexing and aggregation, feature engineering, and data integration

The resulting dataset not only provided the foundation for building the time-series forecasting model but also provided a rich source of context from which to interpret institutional workload trends over time. To achieve a good forecast, the data quality must be ensured at this stage.

Model Selection

The ARIMA-LSTM hybrid model was selected and used in forecasting the faculty workload at ISU. In order to determine how well the data can handle both linear and non-linear, the hybrid model was also tested. This included a process by capturing the general trends in the data using ARIMA and the LSTM is in charge of residue that the ARIMA could not match. RMSE, MAE, and MAPE were used to measure the accuracy to determine how well the ARIMA-LSTM performed.

Root Mean Square Error (RMSE)

The researcher used RMSE to measure how far off the predictions were from the actual results based on average. RMSE is especially good at spotting actual and potential outliers because it gives more weight to larger errors. Here's the formula used to calculate RMSE.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

Where:

y_i = Actual value for day i

\hat{y}_i = Predicted value for day i

n = Number of test samples

Mean Absolute Error (MAE)

MAE is a measure of errors between paired observations expressing the same phenomenon. MAE gives a straightforward view of how much the predictions are off, on average. Here's the formula used to calculate MAE.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

Mean Absolute Percentage Error (MAPE)

MAPE is a metric to measure the forecasting accuracy of a model. It is the mean of the absolute percent errors between the actual and forecasted values. Here’s the formula used to calculate MAPE.

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100$$

RESULTS

This section shows the results of the study. It discussed how well the ARIMA-LSTM hybrid model could predict the total workload of ISU faculty. The researcher utilized three common metric tools such as RMSE, MAE, and MAPE to measure the accuracy of the prediction. These metric tools compared the predicted faculty workloads with the actual ones for different semesters, as shown in the table below.

The Dataset		
Year-Semester	Total Workload Units	Total Sections
2020 1 st	1200	500
2020 2 nd	1000	450
2020 Summer	600	250
2021 1 st	1300	550
2021 2 nd	1100	480
2021 Summer	650	270
2022 1 st	1400	600
2022 2 nd	1200	510
2022 Summer	700	300
2023 1 st	1500	650
2023 2 nd	1300	520
2023 Summer	750	320

Table 1. Dataset

In this study, the historical dataset that was gathered and divided into two sets provided for the training and for the testing, 80% and 20% respectively. The hybrid model was critically tested against the actual workload historical data from the years 2020 to 2023.

RMSE, MAE, and MAPE were considered crucial to understanding how well the prediction model worked. The metrics aided in task to measure errors in a different way, giving valuable insight into the strengths and weaknesses of the model.

Year-Semester	Total Units(Actual)	Workload Total Units (Predicted)	Workload Error (Actual – Predicted)	Absolute Error	Absolute Percentage Error (%)
2020 1 st	1200	1180	20	20	1.67

Year-Semester	Total Units(Actual)	Workload Total Units (Predicted)	Workload Error (Actual – Predicted)	Absolute Error	Absolute Percentage Error (%)
2020 2 nd	1000	1020	-20	20	2.00
2020 Summer	600	590	10	10	1.67
2021 1 st	1300	1285	15	15	1.15
2021 2 nd	1100	1120	-20	20	1.82
2021 Summer	650	640	10	10	1.54
2022 1 st	1400	1380	20	20	1.43
2022 2 nd	1200	1190	10	10	0.83
2022 Summer	700	710	-10	10	1.43
2023 1 st	1500	1480	20	20	1.33
2023 2 nd	1300	1320	-20	20	1.54
2023 Summer	750	740	10	10	1.33

Table 2. RMSE, MAE, MAPE Result

RMSE (Root Mean Squared Error)

One of the most widely used metrics for assessing how well regression models predict outcomes is RMSE. For every data point, it computes the square of the error—the difference between the expected and actual values—averages these squared differences, and then takes the square root of the outcome. RMSE is especially helpful in this approach because of a number of important features. RMSE is sensitive to outliers because it assigns disproportionately high weight to larger errors because errors are squared before being averaged.

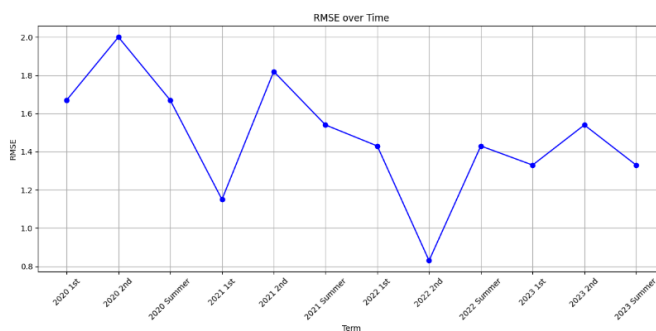


Figure 2. Result of RMSE Overtime

In this study, RMSE was calculated by first finding the squared errors (e.g. $(1200-1180)^2 = 400$, $(1000-1020)^2 = 400$) for each of the data points. The sum of these squared errors (3625) was divided by the number of data points 12 to compute the mean squared error 302.08. Finally, taking the square root of this value gives the RMSE of approximately 17.4. This means that, on average, the model's predictions deviate by 17.4 units from the actual values, but the presence of large errors would have increased this value further, reflecting the RMSE's sensitivity to outliers.

MAE (Mean Absolute Error)

Unlike RMSE, MAE does not square the errors; instead, it calculates the average of the absolute differences between the actual and predicted values. Unlike RMSE, which disproportionately penalizes larger errors, this method treats all errors equally.

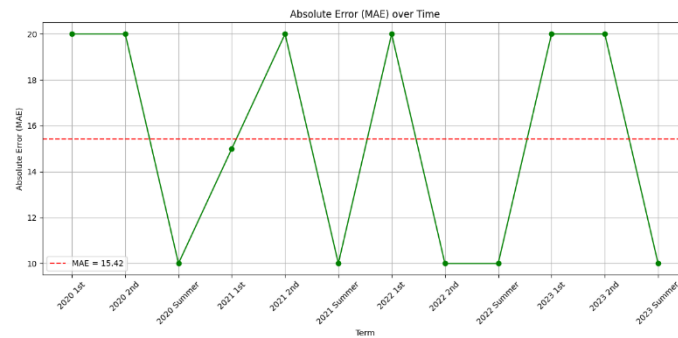


Figure 3. Result of MAE Overtime

Consequently, MAE offers a clear and simple way to quantify the average difference between predictions and actual values. An MAE of roughly 14.58 was obtained by dividing the total sum of errors 175 by the number of data points 12, which was calculated by adding up the absolute errors for each semester (for example, $|1200-1180| = 20$, $|1000-1020| = 20$). As a result, the model's predictions are generally 14.58 units off.

MAPE (Mean Absolute Percentage Error)

Prediction errors are expressed as percentages of the actual values using the MAPE metric. Since the error is normalized by the difference between the actual and actual values, it is particularly helpful for comparing errors across datasets or scales.

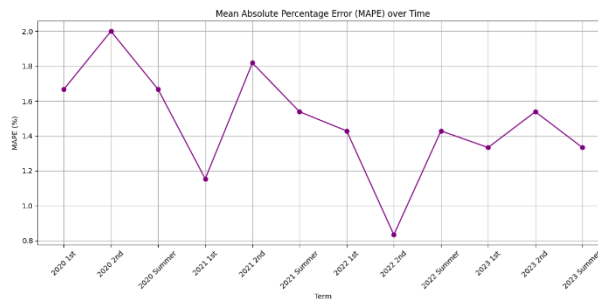


Figure 4. Result of RMSE Overtime

The percentage error for the first semester was determined to be $|(1200-1180) / 1200| \times 100 = 1.67\%$. A value of roughly 1.25% was obtained by averaging the total percentage errors for each data point. This clearly indicated that, on average, the model's predictions deviate from the actual values by roughly 1.25%.

Comparison of RMSE, MAE, and MAPE

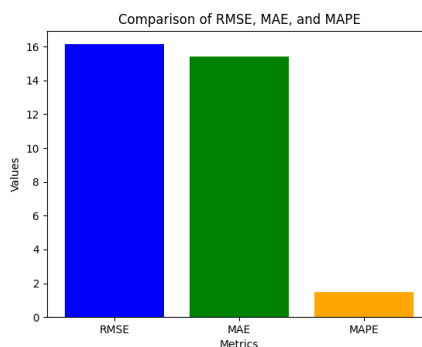


Figure 5. Comparison of RMSE, MAE, MAPE

The three metrics namely, RMSE, MAE, and MAPE offered a unique viewpoint on model accuracy. The result portrayed that RMSE should be applied when significant discrepancies are a problem because it is sensitive to large errors. RMSE metric could be useful in an instance when reducing the impact and effect of significant prediction errors. In contrast, result in MAE is a more neutral metric that treats all errors equally, making it a good choice when the objective is to guarantee a constant degree of error across all predictions without unduly penalizing significant discrepancies. Additionally, the MAPE has shown a comparative measure of error essentially important and helpful when comparing models on across datasets. In terms of percentage basis, the model's prediction is generally very close to actual values as demonstrated by the results, which clearly shows that MAPE of 1.25% has the lowest relative error. The average error of the model in actual units is represented by RMSE of 17.4 and MAE of 14.58, with RMSE exhibiting a marginally higher sensitivity to larger discrepancies as a result.

CONCLUSION

In this study, a Decision Support System (DSS) was successfully created presented to forecast faculty academic workload in Isabela State University (ISU). Using the hybrid of ARIMA and LSTM with time-series forecasting, the system provided precise predictions that are useful in managing faculty members' workloads. The three performance metrics employed in this study provided that MAPE appeared to be most reliable, as it returned the best forecast for ISU's actual academic workload requirements.

The DSS, implemented on this hybrid ARIMA-LSTM model, provided useful, data-driven advice to university officials, enhancing planning accuracy and effectiveness in academic scheduling. Through a combination of both linear and non-linear forecasting processes, the model was able to capture the subtle patterns in faculty workload, symbolizing the high degree of uncertainty in higher education. This procedure enhanced the ability of the system to forecast or predict, contributing more effective decisions in faculty workload resource allocation.

The Decision Support System (DSS) has exhibited a high performance with a low MAPE of 1.25%, indicating high accuracy in prediction. The system depicted its usability as a good forecasting tool for university administrators based on values derived using metrics like RMSE, MAE, and MAPE.

Additionally, this study focused on the identification of increasing the need to incorporate advanced analytics in academic administration. The DSS did not only resolved some of the current issues at ISU but also provided a scalable model that other institutions can replicate and further adopt. Through converting historical data into actionable insights, the system empowered educational leaders to become more strategic and proactive in running their institutions. Advances in the future, including integration of real-time data and adaptive algorithms, could make the system even more responsive and beneficial to dynamic academic environments.

REFERENCES

- [1] Secaran, Y., & Sathiyamoorthy, E. (2020). A Comparative study about Workload prediction from one time forecast with cyclic forecasts using ARIMA model for cloud environment. *EAI Endorsed Transactions on Energy Web*, 7(29). <https://eudl.eu/doi/10.4108/eai.13-7-2018.163977>
- [2] Xu, X. (2023). Forecasting Student Employment Trends in Colleges and Universities Education Using Time Series Analysis Algorithms. *Journal of Electrical Systems*, 19(1). <https://journal.esrgroups.org/jes/article/view/3131>
- [3] Büyüksahin, Ü. Ç., & Ertekin, Ş. (2018). Improving forecasting accuracy of time series data using a new ARIMA-ANN hybrid method and empirical mode decomposition. *arXiv*. <https://arxiv.org/abs/1812.11526>
- [4] Bianchi, F. M., Maiorino, E., Kampffmeyer, M. C., Rizzi, A., & Jenssen, R. (2017). An overview and comparative analysis of Recurrent Neural Networks for Short Term Load Forecasting. *arXiv*. <https://arxiv.org/abs/1705.04378>
- [5] Lim, B., & Zohren, S. (2020). Time Series Forecasting With Deep Learning: A Survey. *arXiv*. <https://arxiv.org/abs/2004.13408>
- [6] Soriano, L. T., Capucao, J. N. B., & Palaoag, T. D. (2018). Decision Support Model for Managing Academic Workload in State Universities and Colleges. *International Journal of Information and Education Technology*, 8(7), 491-495. <https://www.ijiet.org/show-102-1254-1.html>

- [7] Pan, X. (2017). Forecasting the workload with a hybrid model to reduce the inefficiency cost. *University of Kentucky Theses and Dissertations*. https://uknowledge.uky.edu/me_etds/91/
- [8] Wang, J., Xing, Y., Zhang, S., Zhou, Y., Li, L., & Nian, T. (2021). A Hybrid Total Logistics Forecasting Method Combined with ARIMA and BPNN. *Journal of Physics: Conference Series*, 1769(1), 012051. <https://iopscience.iop.org/article/10.1088/1742-6596/1769/1/012051>
- [9] Tarmanini, C., Sarma, N., Gezein, C., & Ozgonenel, O. (2023). Short term load forecasting based on ARIMA and ANN approaches. *Energy Reports*, 9(Supplement 3), 550-557. <https://doi.org/10.1016/j.egy.2023.01.060>
- [10] Siami-Namini, S., & Siami Namin, A. (2018). Forecasting Economics and Financial Time Series: ARIMA vs. LSTM. *arXiv*. <https://arxiv.org/abs/1803.06386>
- [11] Suriya, S., & Arvinth, M. S. (2024). Prediction of Workloads in Cloud using ARIMA-ANN. *Journal of IoT in Social, Mobile, Analytics, and Cloud*, 6(4), 327-342. <https://irojournals.com/iroismac/article/view>