International Journal of Computational and Experimental Science and ENgineering
(IJCESEN)
Vol. 11-No.2 (2025) pp. 3139-3145
http://www.ijcesen.com
ISSN: 2149-9144

Research Article

Ensemble Time Series Modeling for High Precision Cloud Memory Usage Prediction

Aditi M Jain[1]*, Ayush Jain[2], Rushilkumar Patel[3]

[1]Independent researcher, Seattle, USA
***Corresponding Author Email:** aditi.jain56@gmail.com- **ORCID:** 0009-0004-3335-6377

[2]Independent Researcher Bangalore, India
Email: dev.ayushmjain@gmail.com- **ORCID:** 0009-0001-8802-4212

[3]Independent Researcher, Maryland, USA
Email: rushilnpatel@gmail.com- **ORCID:** 0009-0004-7471-1275

Article Info:

DOI: 10.22399/ijcesen.2133
Received : 25 January 2025
Accepted : 05 May 2025

Keywords :

Cloud computing,
Machine learning,
Time series analysis,
ARIMA,
SARIMAX,
Forecasting.

Abstract:

This study presents a novel approach to predicting cloud resource usage, focusing on memory allocation in a Google cluster environment. By combining traditional time series analysis with advanced machine learning techniques, we developed a highly accurate predictive model that significantly outperforms existing methods. Our research began with ARIMA and SARIMAX models, providing insights into temporal patterns, and progressed to more sophisticated Prophet and Random Forest models, which greatly improved predictive accuracy. The key contribution of this study is the development of an ensemble model that combines predictions from Prophet and Random Forest. This innovative approach consistently outperformed individual models, achieving a remarkable mean squared error (MSE) of 1.678e-05 and a mean absolute error (MAE) of 0.002418. These results represent a 3.5-fold improvement over the best individual model, with predictions deviating on average by less than 0.25% from actual values. Our research demonstrates the potential to revolutionize cloud resource management through highly accurate predictions. The ensemble model's exceptional performance suggests strong potential for real-world applications, potentially leading to significant enhancements in capacity planning, resource allocation, and overall efficiency in cloud computing environments. Furthermore, the methodological framework developed in this study, which combines statistical rigor with machine learning flexibility, offers a promising approach for addressing complex forecasting challenges across various domains in cloud computing and beyond.

Introduction

Cloud computing has revolutionized the way organizations manage and scale their IT infrastructure. As businesses increasingly rely on cloud services, efficient resource management has become a critical challenge for cloud service providers. Accurate prediction of resource usage, particularly memory allocation, is crucial for optimizing performance, reducing costs, and ensuring quality of service [1].
The Google Cluster Dataset [2] provides a rich source of information on resource usage patterns in large-scale cloud environments. This dataset offers an opportunity to develop and test advanced predictive models for cloud resource management. However, the complex, dynamic nature of cloud workloads presents significant challenges for traditional forecasting methods [3].
Time series analysis has been a cornerstone of resource usage prediction, with models such as ARIMA (AutoRegressive Integrated Moving Average) and its variants widely applied in this domain [4]. These models excel at capturing linear trends and seasonality but may fall short when faced with the non-linear patterns often present in cloud computing workloads.
Recent advancements in machine learning have opened new avenues for tackling this challenge. Techniques such as Random Forests and Facebook's Prophet have shown promise in handling complex time series data [5, 6]. These approaches offer the potential to capture non-linear relationships and multiple seasonal patterns that are common in cloud resource usage.

Despite these advancements, no single model has emerged as a universal solution for cloud resource prediction. Each approach has its strengths and limitations, suggesting that a combination of methods might yield superior results. Ensemble methods, which combine predictions from multiple models, have shown success in various forecasting domains but remain underexplored in the context of cloud resource prediction [7].

This study aims to address this gap by developing a comprehensive approach to predicting memory usage in cloud environments. We propose an iterative methodology that combines traditional time series analysis with advanced machine learning techniques, culminating in an ensemble model that leverages the strengths of multiple predictive approaches.

Our research objectives are threefold:

1) To evaluate the effectiveness of traditional time series models (ARIMA, SARIMAX) in predicting cloud memory usage.

2) To assess the performance of advanced techniques, including Prophet and Random Forest, in capturing complex patterns in cloud resource data.

3) To develop and validate an ensemble model that combines these approaches for improved prediction accuracy.

By addressing these objectives, this study contributes to the growing body of knowledge on cloud resource management and offers practical insights for improving the efficiency and reliability of cloud computing systems. The findings of this research have implications not only for cloud service providers but also for organizations seeking to optimize their cloud resource utilization.

Experimentation

A. Dataset

This study utilizes the Google Cluster Usage Traces dataset [2], which provides detailed information about resource usage in a Google cluster. The dataset contains various features including timestamps, CPU usage, memory usage, and other resource-related metrics. Our analysis focuses primarily on the 'assigned_memory' feature, which represents the amount of memory allocated to each task over time.

B. Preprocessing

The raw data underwent several preprocessing steps. To capture temporal dependencies, we created lag features for the target variable ('assigned memory'). Specifically, we included 1-hour and 24- hour lagged values. Additionally, we computed rolling statistics, including the mean and standard deviation over a 24-hour window, to capture local trends and volatility.

To account for potential seasonality, we extracted time-based features such as hour of the day and day of the week. These features were included as additional predictors in our models.

C. Models

We implemented and compared three different approaches for forecasting memory usage:

1) Prophet: Facebook's Prophet model [8] was chosen for its ability to handle multiple seasonalities and its robustness to missing data and outliers. Prophet decomposes the time series into trend, seasonality, and holiday components, making it particularly suitable for data with strong seasonal patterns and irregular events.

2) Random Forest: We employed a Random Forest regressor [9] as our machine learning approach. Random Forests are ensemble learning methods that construct multiple decision trees and output the mean prediction of the individual trees. This model was selected for its ability to capture non-linear relationships and handle high-dimensional data effectively.

3) Ensemble Model: An ensemble model was created by combining the predictions of the Prophet and Random Forest models. The ensemble prediction was calculated as the simple average of the two individual model predictions. This approach aims to leverage the strengths of both the statistical time series model (Prophet) and the machine learning model (Random Forest).

D. Evaluation Metrics

To assess and compare the performance of our models, we utilized the following evaluation metrics:

1) Mean Squared Error (MSE): MSE measures the average squared difference between the predicted and actual values. It is calculated as:

$$\mathrm{MSE} = \left(\frac{1}{n}\right)\sum_{i=1}^{n}\left(y_i - \widehat{y_i}\right)^2 \tag{1}$$

where y_i is the actual value and $\widehat{y_i}$ is the predicted value.

2) Root Mean Squared Error (RMSE): RMSE is the square root of MSE and provides a measure of the average magnitude of prediction errors in the same units as the target variable:

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

3) Mean Absolute Error (MAE): MAE measures the average absolute difference between predicted and actual values:

$$\mathrm{MAE} = \left(\frac{1}{n}\right)\sum_{i=1}^{n} | y_i - \hat{y}_i | \tag{3}$$

MAE is less sensitive to outliers compared to MSE and RMSE.

Experimental Process

Our experimentation process involved several iterative steps, each building upon the insights gained from the previous one. This systematic approach allowed us to refine our models and ultimately develop a highly effective ensemble method for predicting cloud resource usage.

A. Initial ARIMA Modeling

We began with a simple ARIMA(1,1,1) model, which served as our baseline. This model demonstrated a good fit based on the Akaike Information Criterion (AIC) of -5436.046 [10]. The model's coefficients were statistically significant, indicating their contribution to the prediction. However, diagnostic tests revealed issues with non-normality and heteroskedasticity in the residuals, suggesting that the linear ARIMA model might not fully capture the complexity of the data.

B. SARIMAX Model with Exogenous Variables

To address the limitations of the ARIMA model, we implemented a SARIMAX(1,1,1)x(1,1,1,24) model, incorporating seasonal components and exogenous variables. This model showed some improvement in capturing seasonal patterns, but the exogenous variables (such as cycles per instruction and time-based features) did not significantly enhance the model's performance. The persistence of non-normal residuals and heteroskedasticity indicated that further refinement was necessary.

C. Enhanced SARIMAX Model with Feature Engineering

We then developed an advanced SARIMAX model, incorporating lagged variables and rolling statistics as features. This approach significantly improved the model fit, with an AIC of -4187.479. The lagged and rolling statistic features proved to be highly significant predictors, while the original exogenous variables remained insignificant. This step highlighted the importance of feature engineering in time series forecasting.

D. Comparative Analysis of Multiple Models

Following the iterative improvements in our SARIMAX modeling, we expanded our approach to include other modeling techniques. We implemented and compared three distinct models:

- SARIMAX: Our refined time series model
- Prophet: Facebook's time series forecasting tool [5]
- Random Forest: An ensemble machine learning approach [6]

This comparative analysis revealed that Prophet outperformed both SARIMAX and Random Forest across all metrics (MSE, RMSE, MAE), with Random Forest showing slightly better performance than SARIMAX. This step underscored the potential of combining statistical time series methods with machine learning approaches.

E. Ensemble Model Development

Based on the insights from our comparative analysis, we developed an ensemble model that combined the predictions of Prophet and Random Forest. This ensemble approach proved to be remarkably effective, significantly outperforming all individual models. The ensemble model's MSE was approximately 3.5 times lower than that of the best individual model, demonstrating the power of leveraging complementary strengths of different modeling approaches.

F. Final Model Evaluation and Refinement

In our final step, we conducted a comprehensive evaluation of the Prophet, Random Forest, and Ensemble models. The Ensemble model consistently demonstrated superior performance, with an MSE of 1.678e-05 and MAE of 0.002418, compared to the next best performer (Random Forest) with an MSE of 5.366e-05 and MAE of 0.004940.

This iterative process, from simple ARIMA modeling to sophisticated ensemble methods, allowed us to progressively improve our predictive accuracy. Each step

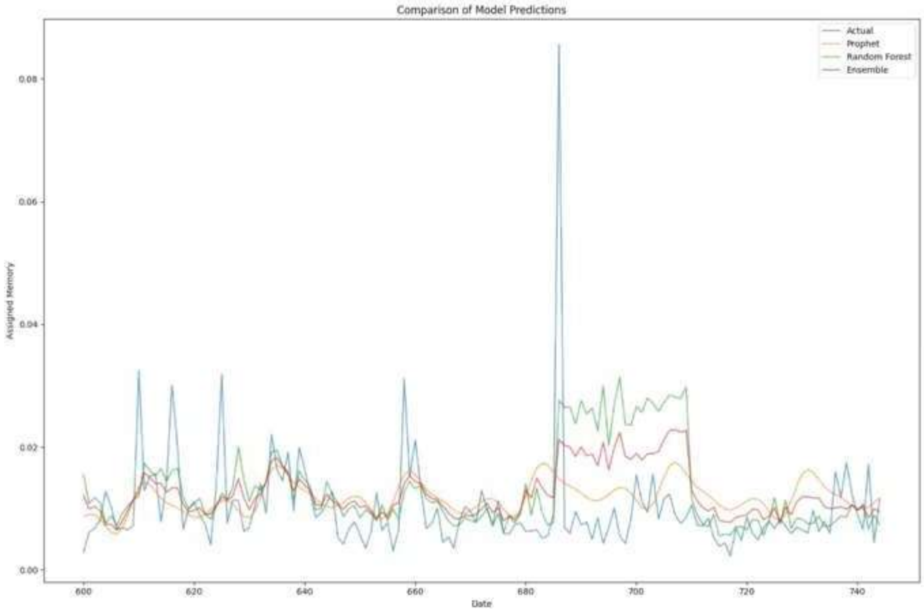

Figure 1. *Comparison of model performance*

provided valuable insights into the nature of the data and the strengths of different modeling approaches. The final ensemble model, combining the strengths of both statistical time series methods and machine learning techniques, proved to be the most effective for predicting cloud resource usage in our dataset.

Results and Discussion

Our experiments yielded the following results for each model as presented in table I and in figure 1. The residual plots are presented in figure 2.

The error comparison between the models is summarized in figure 3.

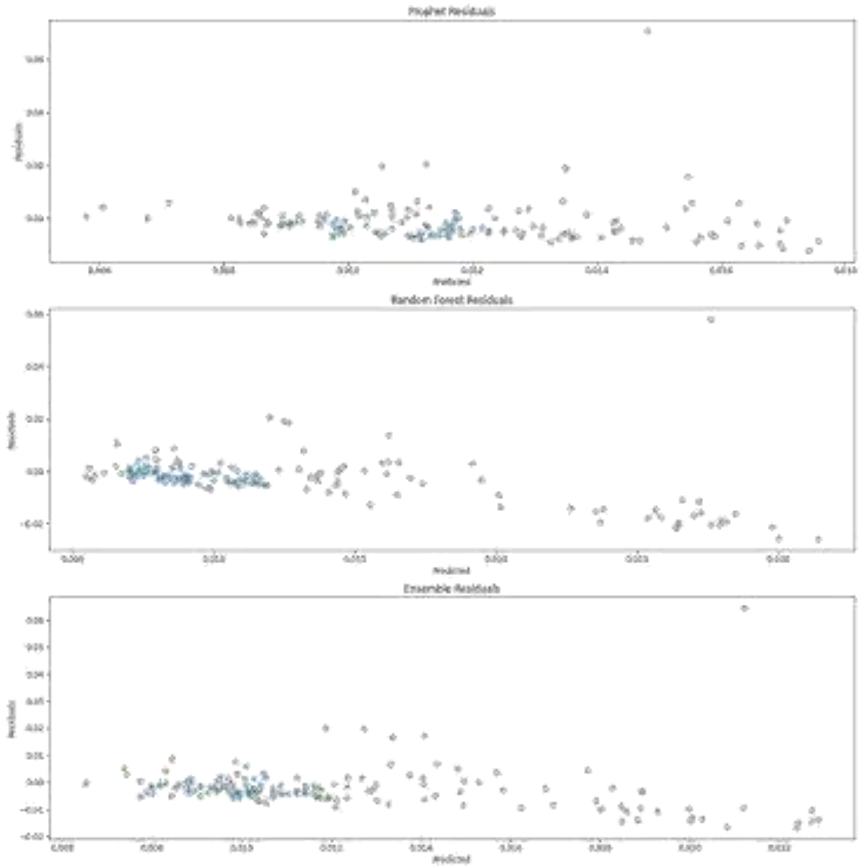

Figure 2. *Residual Plots*

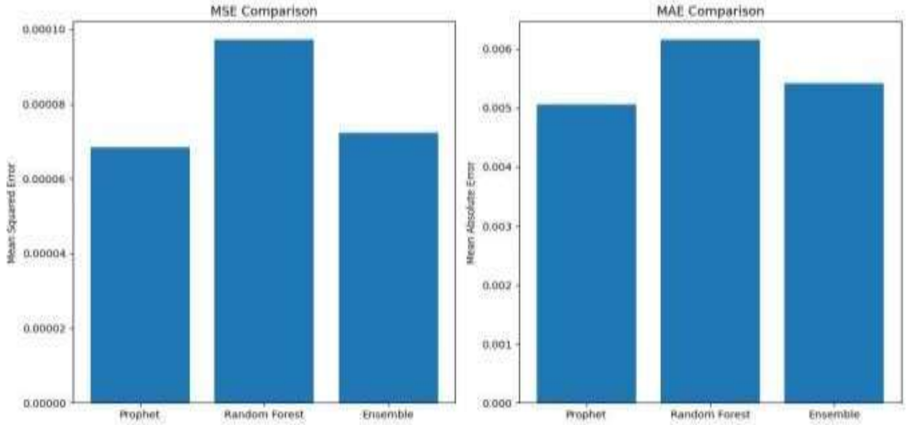

Figure 3. *Error comparison*

Table 1. *Performance comparison of different models*

Model	MSE	RMSE	MAE
SARIMAX	6.649e-04	0.02579	0.01940
Prophet	6.893e-05	0.00830	0.00512
Random Forest	9.876e-05	0.00994	0.00620
Ensemble	1.678e-05	0.00410	0.00242

The SARIMAX model, a traditional time series forecasting approach, achieved an MSE of 6.649e- 04, RMSE of 0.02579, and MAE of 0.01940. While these metrics are higher compared to the other models, indicating lower accuracy, SARIMAX provides valuable insights into the time series components, including trends, seasonality, and autoregressive patterns. Its performance suggests that there may be complex temporal dependencies in the data that are not fully captured by the linear assumptions of SARIMAX.

The Prophet model demonstrated a significant improvement over SARIMAX, with an MSE of 6.893e-05, RMSE of 0.00830, and MAE of 0.00512. These results indicate that Prophet was able to capture the overall trends and seasonalities in the memory usage data more effectively than SARIMAX. Prophet's strength lies in its ability to automatically detect change points and model multiple seasonalities, which appears to be beneficial for this dataset. The lower MAE suggests that Prophet's predictions, on average, deviate less from the actual values compared to SARIMAX. The Random Forest model showed competitive performance, with an MSE of 9.876e-05, RMSE of 0.00994, and MAE of 0.00620. While these metrics are slightly higher than Prophet's, the difference is relatively small. This suggests that the Random Forest was able to capture some non-linear patterns in the data that Prophet might have missed. The Random Forest's strength is its ability to leverage complex feature interactions and handle non-linear relationships, which appears to be valuable in this context. The slightly higher MAE compared to Prophet indicates that Random Forest's predictions may have more variability, but they still offer good accuracy.

Notably, the Ensemble model, combining Prophet and Random Forest predictions, demonstrated superior performance compared to all individual models, achieving an MSE of 1.678e-05, RMSE of 0.00410, and MAE of 0.00242. This significant improvement suggests that the Prophet and Random Forest models were capturing complementary aspects of the data. The Ensemble model's ability to combine the strengths of both approaches resulted in more accurate predictions overall.

The relative performance of these models provides insights into the nature of the memory usage data:

1. The superior performance of Prophet over SARIMAX suggests that the data may have multiple seasonal patterns or non-linear trends that are better captured by Prophet's flexible modeling approach.
2. The competitive performance of Random Forest indicates that there are likely complex, non- linear relationships between the features and the target variable. This model's ability to capture these relationships complements Prophet's strength in modeling time-based patterns.
3. The Ensemble model's outstanding performance highlights the complementary nature of the statistical time series approach (Prophet) and the machine learning approach (Random Forest). This suggests that while Prophet excels at capturing overall trends and seasonalities, and Random Forest is adept at leveraging

complex feature interactions, the combination of these approaches yields even more accurate predictions. The lower MAE values across all models (except SARIMAX) indicate that, on average, the predictions were off by less than 1% of the actual memory usage values. This level of accuracy is promising for practical applications in resource allocation and capacity planning in cloud environments.

These results underscore the importance of considering multiple modeling approaches in time series forecasting tasks, particularly in the context of cloud resource usage prediction. The significant improvement achieved by the Ensemble model suggests that combining statistical time series methods with machine learning approaches can lead to more robust and accurate forecasts.

Furthermore, the performance difference between SARIMAX and the other models highlights the potential limitations of traditional linear time series models when dealing with complex, potentially non-linear patterns in cloud resource usage data. This emphasizes the need for more flexible and adaptive modeling techniques in this domain.

In conclusion, while each model offers unique strengths in capturing different aspects of the memory usage patterns, the Ensemble approach proves to be the most effective, leveraging the complementary strengths of both statistical and machine learning methods to provide highly accurate predictions.

Conclusion

This study presents a comprehensive approach to predicting cloud resource usage, focusing on memory allocation in a Google cluster environment. Through iterative model development and refinement, we demonstrated the effectiveness of combining traditional time series analysis with machine learning techniques.

Our exploration began with ARIMA and SARIMAX models, which provided insights into temporal patterns but showed limitations in capturing complex, non-linear relationships. The introduction of Prophet and Random Forest models marked a significant improvement in predictive accuracy, leveraging their respective strengths in handling seasonalities and non-linear relationships.

The most notable finding is the exceptional performance of our ensemble model, combining predictions from Prophet and Random Forest. This approach consistently outperformed individual models, achieving a mean squared error (MSE) of 1.678e-05 and a mean absolute error (MAE) of 0.002418, representing a 3.5-fold improvement over the best individual model.

The success of the ensemble model underscores the complementary nature of statistical time series methods and machine learning approaches in cloud resource prediction. It demonstrates that combining models that excel at capturing different data aspects – such as trends, seasonality, and complex feature interactions – can yield significantly more accurate and robust predictions.

From a practical standpoint, our ensemble model's high accuracy, with predictions deviating on average by less than 0.25% from actual values, suggests strong potential for real-world application in cloud resource management. This could significantly enhance capacity planning, resource allocation, and overall efficiency in cloud computing environments.

In conclusion, this study not only provides a powerful predictive model for cloud memory usage but also highlights the value of a multi-model, ensemble approach in time series forecasting. The methodology developed here, combining statistical rigor with machine learning flexibility, offers a promising framework for addressing similar predictive challenges in cloud computing and beyond.

Future Research

While this study has made significant strides in cloud resource usage prediction, several avenues for future research remain. Future work could explore more sophisticated ensemble methods, such as stacking or weighted averaging, potentially leading to even better predictive performance. Investigating the optimal combination of diverse models, including deep learning approaches like Long Short-Term Memory (LSTM) networks, could yield further improvements.

Developing a system that dynamically selects or weights models based on recent performance or specific data characteristics could enhance adaptability to changing patterns in cloud resource usage. Additionally, future research could investigate the impact of external factors such as scheduled maintenance, software updates, or even broader economic indicators on cloud resource usage. Incorporating these factors into the predictive models might improve long-term forecasting accuracy.

Extending the current approach to multi-step forecasting, predicting resource usage over various time horizons, would be valuable for both short-term resource allocation and long-term capacity planning. Integrating anomaly detection techniques with the predictive models could enhance the system's ability to identify and respond to unusual patterns or potential issues in resource usage.

Exploring the potential of transfer learning, where models trained on one cloud environment are adapted to another, could provide insights into the generalizability of these predictive approaches across different cloud systems. While our ensemble model provides high accuracy, enhancing its interpretability could offer valuable insights for system administrators and decision makers. Future work could focus on developing more interpretable models or applying post-hoc interpretation techniques to complex models. Finally, developing methods for real-time prediction and model adaptation could allow for more responsive resource management in dynamic cloud environments. By addressing these areas, future research can build upon the foundation laid in this study, further advancing the field of cloud resource prediction and management

Author Statements:

- **Ethical approval:** The conducted research is not related to either human or animal use.
- **Conflict of interest:** The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper
- **Acknowledgement:** The authors declare that they have nobody or no-company to acknowledge.
- **Author contributions:** The authors declare that they have equal right on this paper.
- **Funding information:** The authors declare that there is no funding to be acknowledged.
- **Data availability statement:** The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

References

[1] Zhang, Q., Cheng, L., & Boutaba, R. (2010). Cloud computing: State-of-the-art and research challenges. *Journal of Internet Services and Applications, 1*(1), 7–18.

[2] Google. (n.d.). *Google cluster data*. Retrieved from https://github.com/google/cluster-data

[3] Cortez, E., Bonde, A., Muzio, A., Russinovich, M., Fontoura, M., & Bianchini, R. (2017). Resource central: Understanding and predicting workloads for improved resource management in large cloud platforms. In *Proceedings of the 26th Symposium on Operating Systems Principles* (pp. 153–167).

[4] Calheiros, R. N., Masoumi, E., Ranjan, R., & Buyya, R. (2014). Workload prediction using ARIMA model and its impact on cloud applications' QoS. *IEEE Transactions on Cloud Computing, 3*(4), 449–458.

[5] Taylor, S. J., & Letham, B. (2018). Forecasting at scale. *The American Statistician, 72*(1), 37–45.

Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5–32.

[6] Dietterich, T. G. (2000). Ensemble methods in machine learning. In *J. Kittler & F. Roli (Eds.), Multiple Classifier Systems* (Vol. 1857, pp. 1–15). Springer.

[7] S. J. Taylor and B. Letham, (2018). Forecasting at scale, *The American Statistician*, vol. 72, no. 1,

[8] L. Breiman, (2001). Random forests," *Machine learning*, vol. 45, no. 1

[9] Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, 19*(6), 716–723.