MULTI-STEP COMPUTATIONAL GRAPH PREDICTION FOR CLOUD WORKFLOWS: A COMPARATIVE STUDY OF COMMON MACHINE LEARNING AND DEEP LEARNING METHODS

Anonymous authors
Paper under double-blind review

ABSTRACT

This paper explores the application of traditional machine learning models and deep learning models in a cloud computing environment. For the first time, it advances cloud computing by performing joint multi-step prediction of tasks and their subsequent tasks within cloud workflows. We evaluated the performance of six benchmark models (LR, SVM, XGBoost, LightGBM, CNN, and GCN) in multi-step prediction tasks. The experimental results indicate that each model has its own strengths at different prediction lengths. LR and SVM models perform well across all prediction lengths, making them suitable for tasks requiring stability and consistent performance. XGBoost and LightGBM models excel in accuracy, making them ideal for tasks demanding high accuracy. Although CNN and GCN models exhibit significant fluctuations in performance across different prediction lengths, they have notable advantages in handling complex data structures and capturing the intricate relationships between tasks. In the future, we will explore more deep learning models suitable for cloud workflow prediction tasks and apply these models in fields such as finance, healthcare, and the Internet of Things to verify their effectiveness and feasibility in various application scenarios.

1 INTRODUCTION

The era of cloud computing has arrived, with growing focus on cloud technology Günther & Praeg (2023); Alzoubi et al. (2024); Gao et al. (2021b); Aslam (2023). As a key technology supporting big data processing and analysis, cloud computing allows users to deploy advanced technologies through providers like Amazon, Google, Microsoft, and Alibaba, enhancing business flexibility Yu et al. (2022a). These providers enable processing of large data workloads on cloud servers, benefiting from scalable and cost-effective computing resources over the internet Parappagoudar et al. (2023); Das & Dash (2021).

With the proliferation of cloud services, enterprises and research institutions increasingly rely on cloud computing for complex computational tasks and data analysis. This includes not only flexible hardware scheduling but also automated software deployment and management. Workflow Management Systems (WMS) in the cloud integrate various services and resources, managing tasks from data input to processing and output. This makes cloud workflows ideal for large-scale tasks requiring efficient data processing and real-time analysis in fields such as finance, healthcare, and research Belgacem & Beghdad-Bey (2022).

Cloud workflows involve a series of tasks executed on distributed computing resources, essential for applications like data processing, scientific simulations, and business processes. Efficient resource management is a fundamental challenge in these systems, particularly in dynamic environments like cloud data centers Maiyza et al. (2023). Tasks are often interdependent, with each task's output potentially serving as input for the next. Effective workflow management directly impacts performance, cost, and scalability of cloud services, including predicting task sequences and resource requirements (CPU, memory, storage) Gao et al. (2021a); Luo et al. (2018).

Flexible scheduling is required to adapt to varying workloads and ensure efficient resource utilization in cloud workflows. Multi-step prediction helps ensure resources are available when needed, reducing idle time and bottlenecks, crucial for maintaining Quality of Service (QoS). This optimizes cloud infrastructure utilization, minimizes energy consumption, and promotes data center sustainability. For end-users, predicting resource demands and potential network changes based on historical data allows for accurate task scheduling, cost estimation, and completion time forecasting, enhancing budgeting and planning for more predictable and user-friendly cloud services.

Currently, the main challenge in cloud computing-related predictions is that many cloud computing prediction tasks require real-time processing and forecasting. However, the primary research focus is on the performance prediction of workflows. There is a significant gap in predicting the relationships between workflow tasks based on time series, which has sparked our interest. In this context, computational graph prediction becomes particularly important. A computational graph is a structured representation used to describe and manage the computation processes in neural networks Ward et al. (2022); Khemani et al. (2024). It is widely used in deep learning for efficiently handling and optimizing complex computational tasks. We regard cloud workflows as computational graphs, which not only clearly show the data flow and control flow between tasks but also provide an intuitive perspective for analyzing the properties of workflows. Prediction based on computational graphs is a technique for processing and analyzing graph-structured data, making it especially suitable for multi-step predictions in cloud workflows. This involves predicting the outcomes of tasks over multiple time steps, which is crucial for ensuring the continuous accuracy and stability of the predictions. Given the structural information and complexity of cloud workflows, we employ classical deep learning methods to automatically extract features of different tasks in the workflow and capture the complex relationships and graph structure between different tasks, comparing them with classical machine learning methods. Our contributions can be summarized as follows:

- **Achieving the first joint multi-step prediction of task features and their subsequent tasks in cloud computing workflows.** This innovative approach captures the complex relationships between task features and subsequent tasks. By predicting the state of tasks over multiple time steps, it provides more comprehensive insights. This has important practical significance for resource scheduling, performance optimization, and fault prediction in cloud workflows, helping to achieve more efficient management and optimization in cloud computing environments.
- **In this paper, we introduce various existing advanced machine learning and deep learning models as baseline models to comprehensively evaluate their performance on the studied problem.** These baseline models include Logistic Regression (LR) Tang et al. (2022); Jia & Gong (2018), Support Vector Machine (SVM) Haghani & Keyvanpour (2019), Extreme Gradient Boosting (XGBoost) Li et al. (2020); Haghani & Keyvanpour (2019); Chen & Guestrin (2016); Chen et al. (2015), Light Gradient Boosting Machine (LightGBM) Ke et al. (2017); Al Daoud (2019), Convolutional Neural Network (CNN) Wang et al. (2023), and Graph Convolutional Network (GCN) Wang & Le (2020); Cai et al. (2019); Zhang et al. (2019); Coşkun & Koyutürk (2021).

The remainder of this paper is structured as follows. In the section 2, we review related work in this field. In the section 3, we introduce the methods used. In the section 4, we briefly describe the experimental setup. In the section 5, we discuss the experimental results. In the section 6, we conclude the paper and look forward to future work. Additionally, in the section A, we have included some supplementary experimental results.

2 Related Works

This section briefly reviews research work related to cloud workflow prediction, machine learning, and deep learning. By analyzing existing literature, we summarize the current progress, identify existing gaps.

Cloud workflow prediction is a significant research direction in the field of cloud computing, with many scholars conducting extensive studies in this area. For example, Yu et al.Yu et al. (2022b) used machine learning and deep learning methods to predict performance in cloud workflows. Gupta et al.Gupta et al. (2018) employed BiLSTM to predict CPU usage in cloud workflows. Zhong et al.Zhong et al. (2018) improved load prediction accuracy using a weighted wavelet support vector machine. However, these methods mainly focus on performance prediction of cloud workflows. There is a noticeable gap in research on link prediction within these workflows, and most studies concentrate on single-step prediction, failing to effectively capture the complex relationships between tasks.

This paper aims to capture the complex relationships between tasks in cloud workflows. By performing joint multi-step prediction of task characteristics and their subsequent tasks, we comprehensively evaluate the performance of various machine learning and deep learning models on this research problem.

3 METHODOLOGY

This section describes the specific methods we used in multi-step prediction tasks within a cloud workflow. We employed traditional machine learning models and deep learning models, conducting a comprehensive evaluation of their performance.

- **LR:** A classical linear model that estimates outputs by fitting a linear equation, suitable for regression and classification tasks. The LR model is simple and efficient, commonly used for basic predictions and feature analysis.
- **SVM:** A model based on the principle of maximum margin. SVM improves model performance by finding the optimal hyperplane or regression line to separate data points or make regression predictions.
- **XGBoost:** A boosting tree model known for its efficient training speed and performance. XGBoost reduces prediction errors by gradually building a collection of decision trees, excelling in handling large-scale data.
- **LightGBM:** A distributed gradient boosting framework based on decision tree algorithms. LightGBM enhances model efficiency and accuracy with faster training speed and lower memory consumption, suitable for processing large datasets.
- **CNN:** A model designed for processing data with a grid-like topology, excelling in feature extraction and pattern recognition. CNN extracts local features from data through convolutional and pooling layers.
- **GCN:** A deep learning model specifically for handling graph-structured data. GCN captures complex relationships between nodes and edges by performing convolution operations on graph structures, making it particularly suitable for processing graph data.

4 EXPERIMENTS

4.1 DATASET SELECTION AND PROCESSING

The Alibaba Cluster-trace-v2018 dataset Alibaba cluster trace (2020) captures the operational characteristics of Alibaba's production clusters. It not only includes detailed records of large-scale cluster computing tasks from the real world, reflecting potential real-world scenarios and issues more accurately, but also provides fine-grained resource usage records, covering various types of tasks and resource usage. Therefore, this paper selects Alibaba's Cluster-trace-v2018 as the benchmark dataset, as detailed in Figure 1.

This paper focuses on the information of DAGs, hence the emphasis is solely on the "Task" dimension. According to the research by Yu et al. (2022a), we selected workflows containing 7 tasks for our study due to their sufficient samples and diverse structures. We preprocessed the dataset by removing redundant information and determining the window size for processing the graph structure. The final dataset includes task information with window size dimensions: task start time (start_time), task end time (end_time), CPU usage (plan_cpu), memory usage (plan_mem), and whether there is a link between the current task and historical tasks, with a value of 1 if a link exists, otherwise 0.

	task_name	inst_name	task_type	job_name	status	start_time	end_time	plan_cpu	plan_mem
0	M1	1.0	j_1	1	Terminated	419912	419912	100.0	0.20
1	R2_1	1.0	j_2	1	Terminated	87076	87086	50.0	0.20
2	M1	1.0	j_2	1	Terminated	87076	87083	50.0	0.20
3	R6_3	371.0	j_3	1	Terminated	157297	157325	100.0	0.49
4	J4_2_3	1111.0	j_3	1	Terminated	157329	157376	100.0	0.59
5	R2_1	1.0	j_3	1	Terminated	157322	157328	100.0	0.39
6	J8_3_7	1111.0	j_3	1	Terminated	157331	157376	100.0	0.59
7	M3	12846.0	j_3	1	Terminated	157213	157295	100.0	0.30
8	R9_8	1.0	j_3	1	Terminated	157376	157381	100.0	0.39
9	R1_3	371.0	j_3	1	Terminated	157297	157322	100.0	0.49
10	R5_4	1.0	j_3	1	Terminated	157376	157381	100.0	0.39

Figure 1: Description of the Cluster-trace-v2018 Dataset. A "Job" consists of multiple "Tasks," with dependencies between "Tasks" represented by a Directed Acyclic Graph (DAG). Each "Task" comprises multiple "Instances," and a "Task" is only considered complete when all its "Instances" are finished.

4.2 Prediction Objectives

The aim of this paper is to predict tasks and their features and link relationships with all subsequent tasks over multiple time steps. Specifically, the prediction objectives include the following two aspects:

- **Prediction Objective 1:** Predict the feature values of tasks, including start_time, end_time, plan_cpu, and plan_mem.
- **Prediction Objective 2:** Predict the link relationships between tasks.

In terms of selecting the number of steps (or prediction length), we use a step-by-step prediction approach to ensure continuity and accuracy. The specific steps are as follows:

- **Initial Prediction:** First, predict the current task to obtain the next step's task features and link relationships.
- **Update Dataset:** Add the obtained prediction results to the dataset, updating it to reflect the new current task status.
- **Repeat Prediction:** Continue predicting the next step based on the updated dataset. This process is repeated until the prediction of the entire cloud workflow is complete.

We selected [2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20] as the prediction steps (or prediction lengths). This multi-step prediction method ensures that each prediction result is dynamically integrated into the dataset, progressively advancing the prediction of the entire task chain, and ultimately completing the comprehensive prediction of the cloud workflow.

4.3 Quantitative Evaluations of Joint Multi-Step Predictions

In our experiments, we used the following evaluation metrics: for predicting task features in cloud workflows, which is essentially a regression problem, we selected two of the most popular evaluation metrics: Mean Squared Error (MSE) and Coefficient of Determination (R^2) to measure the model's prediction performance. MSE quantifies the average squared difference between the predicted values and the actual values, while R^2 reflects the proportion of the variance in the dependent variable that is predictable from the independent variables Zhang et al. (2021); Nie et al. (2023). For predicting subsequent tasks in cloud workflows, which is essentially a binary classification problem, we chose Binary Cross-Entropy Loss (BCE) and Accuracy as evaluation metrics. BCE measures the difference between the predicted probability distribution and the actual distribution, while Accuracy indicates the proportion of correct predictions made by the model on the test data Yu et al. (2022a); Shakibian & Moghadam Charkari (2017); Zhu et al. (2023); Yuan et al. (2019). These evaluation metrics comprehensively reflect the model's performance in predicting task features and subsequent tasks in workflows, providing strong support and validation for our research.

4.4 ABLATION STUDY

This paper is the first to use traditional machine learning models and deep learning models for joint multi-step prediction of tasks and their subsequent tasks in cloud workflows, providing an innovative perspective to advance cloud computing. To comprehensively evaluate the performance of these models, we conducted ablation experiments with different data split ratios, such as [60:20:20], [70:15:15], and [80:10:10]. Ultimately, we selected the [60:20:20] data split ratio for all subsequent experimental result presentations.

4.4.1 COMPARISON OF R^2 METRIC ACROSS DIFFERENT MODELS, DATA SPLIT RATIOS, AND PREDICTION LENGTHS

As illustrated in Figure 2,3,4,5,6 and 7 in the section A, we can observe the following:

- In most models and prediction steps, the R^2 performance with the [60:20:20] split ratio is stable and consistent. Specifically, in the CNN model, compared to the [70:15:15] and [80:10:10] split ratios, although the R^2 with [60:20:20] shows some fluctuations across different prediction lengths, it generally demonstrates stability. In the GCN model, despite being slightly inferior to [70:15:15] and [80:10:10] in long-term predictions (such as 15 and 20), it performs well in short-term predictions (such as 2, 3, 4, and 5). In the XGBoost and LightGBM models, the R^2 with the [60:20:20] split ratio is close to the highest level across all prediction steps, showing very high stability. In the SVM model, the R^2 with the [60:20:20] data split ratio is slightly better, demonstrating its consistency and reliability.
- Although the R^2 of the LR model with the [60:20:20] split ratio is slightly inferior to other split ratios, [60:20:20] overall performs the best, consistently providing reliable prediction results across different models and tasks. This robustness suggests that [60:20:20] is a balanced and effective choice for diverse predictive scenarios.
- Unlike the [70:15:15] and [80:10:10] split ratios, which exhibit fluctuations in R^2 across different prediction steps, [60:20:20] demonstrates excellent adaptability when handling various models and tasks. It consistently maintains high performance across a wide range of scenarios, indicating that it can effectively manage the complexities inherent in different prediction tasks. This stability makes [60:20:20] a preferred choice for achieving reliable and accurate predictive outcomes in both short-term and long-term forecasting.

4.4.2 COMPARISON OF $Accuracy$ METRIC ACROSS DIFFERENT MODELS, DATA SPLIT RATIOS, AND PREDICTION LENGTHS

As illustrated in Figure 8,9,10,11,12 and 13 in the section A, we can observe the following:

- When using a [60:20:20] split ratio, the $Accuracy$ of various models (LR, SVM, XGBoost, LightGBM, CNN, and GCN) shows a high level of consistency across all prediction lengths (2 to 20). This consistency indicates that this split ratio can stably provide high-quality prediction results.
- Although the [70:15:15] and [80:10:10] split ratios also exhibit relatively high $Accuracy$, except for the GCN model where the $Accuracy$ is slightly higher with the [70:15:15] and [80:10:10] split ratios for prediction lengths [2, 3, 4, 5, 6, 7, 8, 9, 10, 20], the $Accuracy$ of other models across various prediction lengths is lower with these split ratios compared to [60:20:20]. At the same time, the [60:20:20] split ratio does not show significant performance degradation across all models and prediction lengths, indicating that this split ratio ensures the robustness of the models.
- Whether for short-term [2, 3, 4, 5, 6, 7, 8, 9] or long-term [10, 15, 20] predictions, the models maintain a high level of $Accuracy$ using the [60:20:20] split ratio. This suggests that this split ratio is suitable for multi-step prediction tasks, effectively capturing the complex relationships between task features and subsequent tasks.

In summary, the [60:20:20] data split ratio was selected for all subsequent experimental results, primarily due to its superior performance in terms of stability, consistency, overall performance, and adaptability. This split ratio has proven to be the optimal choice for various prediction tasks.

5 EVALUATING MODEL PREDICTION QUALITY ACROSS VARIOUS MODELS, DATA SPLIT RATIOS, AND PREDICTION LENGTHS

To evaluate the prediction performance of different models on the Cluster-trace-v2018 dataset, this paper compares six benchmark models. We experimentally analyze the performance of these models in multi-step prediction tasks and summarize their performance at different prediction lengths. The results are presented in Tables 1, 2 in the section 5, and in Tables 3, 4, 5, 6 in the section A, detailing each model's MSE, R^2, BCE, and $Accuracy$. These results validate the effectiveness and stability of various machine learning and deep learning methods in handling cloud workflow prediction tasks.

Table 1: LR Prediction Performance

PREDICTION LENGTHS	AVG_MSE	AVG_R^2	AVG_BCE	AVG_ACCURACY
2	0.0188	0.7577	0.7938	0.9770
3	0.0188	0.7577	0.7938	0.9770
4	0.0188	0.7577	0.7938	0.9770
5	0.0188	0.7577	0.7938	0.9770
6	0.0188	0.7577	0.7938	0.9770
7	0.0188	0.7577	0.7938	0.9770
8	0.0188	0.7577	0.7938	0.9770
9	0.0188	0.7577	0.7938	0.9770
10	0.0188	0.7577	0.7938	0.9770
15	0.0188	0.7577	0.7938	0.9770
20	0.0188	0.7577	0.7938	0.9770

Table 1 summarizes the performance of the LR model across different prediction lengths: AVG_MSE: Remains consistent across all prediction lengths at 0.0188, indicating the model's stability in controlling prediction error. AVG_R^2: Consistent at 0.7577 across all prediction lengths, showing high explanatory power, indicating that the model can explain the data's variance well. AVG_BCE: Fixed at 0.7938, indicating the model's stability in classification tasks. $AVG\_ACCURACY$: Accuracy is consistently 0.9770 across all prediction lengths, showing the model's stable performance across different prediction lengths.

Table 2: SVM Prediction Performance

PREDICTION LENGTHS	AVG_MSE	AVG_R^2	AVG_BCE	AVG_ACCURACY
2	0.0190	0.7550	0.9603	0.9722
3	0.0190	0.7551	0.9589	0.9722
4	0.0190	0.7551	0.9547	0.9724
5	0.0190	0.7551	0.9550	0.9724
6	0.0190	0.7551	0.9529	0.9724
7	0.0190	0.7551	0.9533	0.9724
8	0.0190	0.7551	0.9520	0.9724
9	0.0190	0.7551	0.9524	0.9724
10	0.0190	0.7551	0.9514	0.9725
15	0.0190	0.7551	0.9511	0.9725
20	0.0190	0.7551	0.9510	0.9725

Table 2 summarizes the performance of the SVM model across different prediction lengths: AVG_MSE: It remains at 0.0190 across all prediction lengths, slightly higher than the LR model. AVG_R^2: It fluctuates slightly across prediction lengths, around 0.7550 to 0.7551, showing similar explanatory power to the LR model. AVG_BCE: As the prediction length increases, the binary cross-entropy decreases slightly, indicating that the model's performance on classification tasks improves slightly with longer prediction lengths. $AVG\_ACCURACY$: The accuracy remains between 0.9722 and 0.9725 across all prediction lengths, slightly lower than the LR model but still performing well.

Table 3 summarizes the performance of the XGBoost model across different prediction lengths: AVG_MSE: Ranges from 0.0107 to 0.0108 across all prediction lengths, indicating lower prediction errors compared to LR and SVM models. AVG_R^2: Slightly decreases with increasing prediction length, from 0.7349 to 0.7002, but remains relatively high overall. AVG_BCE: Consistently around 0.2296 to 0.2297, indicating stable performance in classification tasks. $AVG\_ACCURACY$: Maintains a high accuracy of 0.9918 across all prediction lengths, demonstrating excellent classification performance.

Table 4 summarizes the performance of the LightGBM model across different prediction lengths: AVG_MSE: Consistently at 0.0126 across all prediction lengths, slightly higher than XGBoost but lower than LR and SVM models. AVG_R^2: Approximately 0.7757 to 0.7765 across all prediction lengths, indicating good explanatory power. AVG_BCE: Fixed at 0.0627, showing stable performance in classification tasks. $AVG\_ACCURACY$: Maintains an accuracy of 0.9726 across all prediction lengths, slightly lower than XGBoost but still performing well.

Table 5 summarizes the performance of the CNN model across different prediction lengths: AVG_MSE: Varies from 0.0155 to 0.0269, showing some fluctuation. AVG_R^2: Fluctuates significantly across prediction lengths, from 0.7181 to 0.8359, indicating sensitivity to data changes. AVG_BCE: Ranges from 0.5874 to 1.0013, indicating variability in classification performance across different prediction lengths. $AVG\_ACCURACY$: Maintains a high accuracy between 0.9932 to 0.9970 across all prediction lengths, showing excellent classification performance despite fluctuations.

Table 6 summarizes the performance of the GCN model across different prediction lengths: AVG_MSE: Ranges from 0.0158 to 0.0363, indicating some fluctuation. AVG_R^2: Fluctuates significantly across prediction lengths, from 0.6327 to 0.8329, indicating sensitivity to data changes. AVG_BCE: Ranges from 0.2146 to 2.3441, indicating significant variability in classification performance across different prediction lengths. $AVG\_ACCURACY$: Ranges from 0.9867 to 0.9988, showing high classification accuracy despite fluctuations.

Through the analysis of the six models above, we can draw the following comprehensive conclusions:

- LR and SVM Models: Both models exhibit stable performance across all prediction lengths, with high $Accuracy$. Particularly, the LR model maintains consistent performance across all metrics, indicating its reliability in prediction tasks. The high R^2 value and MSE of the LR model demonstrate its advantages in explaining data variations and controlling prediction errors.
- XGBoost and LightGBM Models: These two models show excellent performance in terms of $Accuracy$, especially XGBoost, which maintains an $Accuracy$ of 0.9918 across all prediction lengths. Additionally, the LightGBM model exhibits consistency across all metrics. The low MSE and high $Accuracy$ of the XGBoost model highlight its outstanding performance in prediction accuracy and classification tasks.
- CNN and GCN Models: Both models show significant fluctuations in performance across different prediction lengths, but their $Accuracy$ remains high. The CNN model performs exceptionally well at some prediction lengths, while the GCN model performs best at shorter prediction lengths. The high R^2 value and classification accuracy of the CNN model demonstrate its advantages in handling complex data structures, while the good performance of the GCN model at short prediction lengths indicates its capability in capturing data relationships.

In conclusion, each model has its own strengths across different prediction lengths. LR and SVM models perform well, making them suitable for tasks requiring stability and consistent performance. XGBoost and LightGBM models excel in accuracy, making them ideal for tasks demanding high accuracy. Although CNN and GCN models exhibit significant fluctuations, their performance at specific prediction lengths is valuable, especially for tasks requiring the handling of complex data structures.

6 Conclusion and Future Directions

This paper explores the application of traditional machine learning models and deep learning models in a cloud computing environment through joint multi-step prediction of tasks and their subsequent tasks in cloud workflows. Six benchmark models (LR, SVM, XGBoost, LightGBM, CNN, and GCN) were evaluated for their performance in multi-step prediction tasks through experiments. The results demonstrate that each model has its own advantages at different prediction lengths.

LR and SVM Models: These traditional machine learning models perform excellently across different prediction lengths, exhibiting high stability and consistency. The LR model, in particular, maintains stable performance across all metrics, indicating its high reliability in prediction tasks. Similarly, the SVM model also shows stable performance, slightly inferior to the LR model but still demonstrating reliable predictive capabilities in most cases. Therefore, these models are particularly suitable for tasks requiring stable and consistent performance.

XGBoost and LightGBM Models: These tree-based ensemble learning models excel in accuracy. The XGBoost model demonstrates outstanding performance in prediction accuracy and classification tasks. The LightGBM model exhibits high consistency across all metrics, further proving its advantages in handling large-scale datasets and high-complexity tasks. For tasks demanding high accuracy and strong predictive capabilities, XGBoost and LightGBM are ideal choices.

CNN and GCN Models: Despite fluctuations in performance across different prediction lengths, these deep learning models maintain high prediction accuracy. The CNN model performs particularly well at specific prediction lengths, showcasing its advantage in capturing local features and patterns in the data. The GCN model performs best at shorter prediction lengths, indicating its powerful ability to handle complex relationships and graph-structured data. These models demonstrate their strengths in applications requiring the handling of complex data structures and capturing intricate task relationships.

Although this paper has achieved some significant results in multi-step prediction tasks for cloud workflows, there are still many directions for further exploration. In the future, we will explore more deep learning models suitable for cloud workflow prediction tasks to better address various challenges in cloud computing environments. This will lay the foundation for more efficient and intelligent cloud computing management and optimization. Secondly, applying these models to other fields, such as finance, healthcare, and the Internet of Things, will be explored to verify their effectiveness and feasibility in different application scenarios.

Acknowledgments

The research was carried out within the financial support for the autonomous non-profit organization “Analytical Center for the Government of the Russian Federation” (Agreement No. 70-2023-001321 dated December 27, 2023, id: 000000D730324P540002) and the China Scholarship Council (award to Ruimin Ma study abroad at the Saint Petersburg University).

References

Essam Al Daoud. Comparison between xgboost, lightgbm and catboost using a home credit dataset. *International Journal of Computer and Information Engineering*, 13(1):6–10, 2019.

Alibaba cluster trace. cluster-trace-v2018. `https://github.com/alibaba/clusterdata/blob/master/cluster-trace-v2018`, 2020.

Yehia Ibrahim Alzoubi, Alok Mishra, and Ahmet Ercan Topcu. Research trends in deep learning and machine learning for cloud computing security. *Artificial Intelligence Review*, 57(5):132, 2024.

F. Aslam. Role of cloud computing for big data. *J. Res. Com. Sci.*, 8:1440–1436, 2023. doi: 10.5281/zenodo.8311108.

Ali Belgacem and Kadda Beghdad-Bey. Multi-objective workflow scheduling in cloud computing: trade-off between makespan and cost. *Cluster Computing*, 25(1):579–595, 2022.

Ling Cai, Bo Yan, Gengchen Mai, Krzysztof Janowicz, and Rui Zhu. Transgcn: Coupling transformation assumptions with graph convolutional networks for link prediction. In *Proceedings of the 10th international conference on knowledge capture*, pp. 131–138, 2019.

Tianqi Chen and Carlos Guestrin. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pp. 785–794, 2016.

Tianqi Chen, Tong He, Michael Benesty, Vadim Khotilovich, Yuan Tang, Hyunsu Cho, Kailong Chen, Rory Mitchell, Ignacio Cano, Tianyi Zhou, et al. Xgboost: extreme gradient boosting. *R package version 0.4-2*, 1(4):1–4, 2015.

Mustafa Coşkun and Mehmet Koyutürk. Node similarity-based graph convolution for link prediction in biological networks. *Bioinformatics*, 37(23):4501–4508, 2021.

Madhusmita Das and Rasmita Dash. Role of cloud computing for big data: A review. *Intelligent and Cloud Computing: Proceedings of ICICC 2019, Volume 2*, pp. 171–179, 2021.

Ming Gao, Yuchan Li, and Jixiang Yu. Workload prediction of cloud workflow based on graph neural network. In Chunxiao Xing, Xiaoming Fu, Yong Zhang, Guigang Zhang, and Chaolemen Borjigin (eds.), *Web Information Systems and Applications*, pp. 169–189, Cham, 2021a. Springer International Publishing.

Ming Gao, Yuchan Li, and Jixiang Yu. Workload prediction of cloud workflow based on graph neural network. In *Web Information Systems and Applications: 18th International Conference, WISA 2021, Kaifeng, China, September 24–26, 2021, Proceedings 18*, pp. 169–189. Springer, 2021b.

Jochen Günther and Claus-Peter Praeg. Bedeutung und management von cloud computing, multi-cloud und cloud brokerage in unternehmen. *HMD Praxis der Wirtschaftsinformatik*, 60(5):959–974, 2023.

Shaifu Gupta, Aroor Dinesh Dileep, and Timothy A Gonsalves. A joint feature selection framework for multivariate resource usage prediction in cloud servers using stability and prediction performance. *The Journal of Supercomputing*, 74:6033–6068, 2018.

Sogol Haghani and Mohammad Reza Keyvanpour. A systemic analysis of link prediction in social network. *Artificial Intelligence Review*, 52:1961–1995, 2019.

Jinyuan Jia and Neil Zhenqiang Gong. {AttriGuard}: A practical defense against attribute inference attacks via adversarial machine learning. In *27th USENIX Security Symposium (USENIX Security 18)*, pp. 513–529, 2018.

Guolin Ke, Qi Meng, Thomas Finley, Taifeng Wang, Wei Chen, Weidong Ma, Qiwei Ye, and Tie-Yan Liu. Lightgbm: A highly efficient gradient boosting decision tree. *Advances in neural information processing systems*, 30, 2017.

Bharti Khemani, Shruti Patil, Ketan Kotecha, and Sudeep Tanwar. A review of graph neural networks: concepts, architectures, techniques, challenges, datasets, applications, and future directions. *Journal of Big Data*, 11(1):18, 2024.

Kuanyang Li, Lilan Tu, and Lang Chai. Ensemble-model-based link prediction of complex networks. *Computer Networks*, 166:106978, 2020.

Haoyu Luo, Jin Liu, Xiao Liu, and Yun Yang. Predicting temporal violations for parallel business cloud workflows. *Software: Practice and Experience*, 48(4):775–795, 2018.

Aya I Maiyza, Noha O Korany, Karim Banawan, Hanan A Hassan, and Walaa M Sheta. Vtgan: hybrid generative adversarial networks for cloud workload prediction. *Journal of Cloud Computing*, 12(1):97, 2023.

Renjuan Nie, Guoyin Wang, Qun Liu, and Chengxin Peng. Link prediction for attribute and structure learning based on attention mechanism. In *Rough Sets*, pp. 580–595, Cham, 2023. Springer Nature Switzerland.

Dr. Sachin K. Parappagoudar et al. Basic understanding cloud computing. *J. Res. Com. Sci.*, 8: 220–226, 2023. doi: 10.5281/zenodo.7840938.

Hadi Shakibian and Nasrollah Moghadam Charkari. Mutual information model for link prediction in heterogeneous complex networks. *Scientific reports*, 7(1):44981, 2017.

Rui Tang, Xingshu Chen, Chuancheng Wei, Qindong Li, Wenxian Wang, Haizhou Wang, and Wei Wang. Interlayer link prediction based on multiple network structural attributes. *Computer Networks*, 203:108651, 2022.

Chenxu Wang, Xin Wang, Zhao Li, Zirui Chen, and Jianxin Li. Hyconve: A novel embedding model for knowledge hypergraph link prediction with convolutional neural networks. In *Proceedings of the ACM Web Conference 2023*, pp. 188–198, 2023.

Hui Wang and Zichun Le. Seven-layer model in complex networks link prediction: A survey. *Sensors*, 20(22):6560, 2020.

Isaac Ronald Ward, Jack Joyner, Casey Lickfold, Yulan Guo, and Mohammed Bennamoun. A practical tutorial on graph neural networks. *ACM Computing Surveys (CSUR)*, 54(10s):1–35, 2022.

Jixiang Yu, Ming Gao, Yuchan Li, Zehui Zhang, Wai Hung Ip, and Kai Leung Yung. Workflow performance prediction based on graph structure aware deep attention neural network. *Journal of Industrial Information Integration*, 27:100337, 2022a.

Jixiang Yu, Ming Gao, Yuchan Li, Zehui Zhang, Wai Hung Ip, and Kai Leung Yung. Workflow performance prediction based on graph structure aware deep attention neural network. *Journal of Industrial Information Integration*, 27:100337, 2022b.

Weiwei Yuan, Kangya He, Donghai Guan, Li Zhou, and Chenliang Li. Graph kernel based link prediction for signed social networks. *Information Fusion*, 46:1–10, 2019.

Si Zhang, Hanghang Tong, Jiejun Xu, and Ross Maciejewski. Graph convolutional networks: a comprehensive review. *Computational Social Networks*, 6(1):1–23, 2019.

Yuyi Zhang, Ruimin Ma, Jing Liu, Xiuxiu Liu, Ovanes Petrosian, and Kirill Krinkin. Comparison and explanation of forecasting algorithms for energy time series. *Mathematics*, 9(21):2794, 2021.

Wei Zhong, Yi Zhuang, Jian Sun, and Jingjing Gu. A load prediction model for cloud computing using pso-based weighted wavelet support vector machine. *Applied Intelligence*, 48:4072–4083, 2018.

Junxi Zhu, Fang Dai, Fengqun Zhao, and Wenyan Guo. Integrating node importance and network topological properties for link prediction in complex network. *Symmetry*, 15(8):1492, 2023.

A APPENDIX

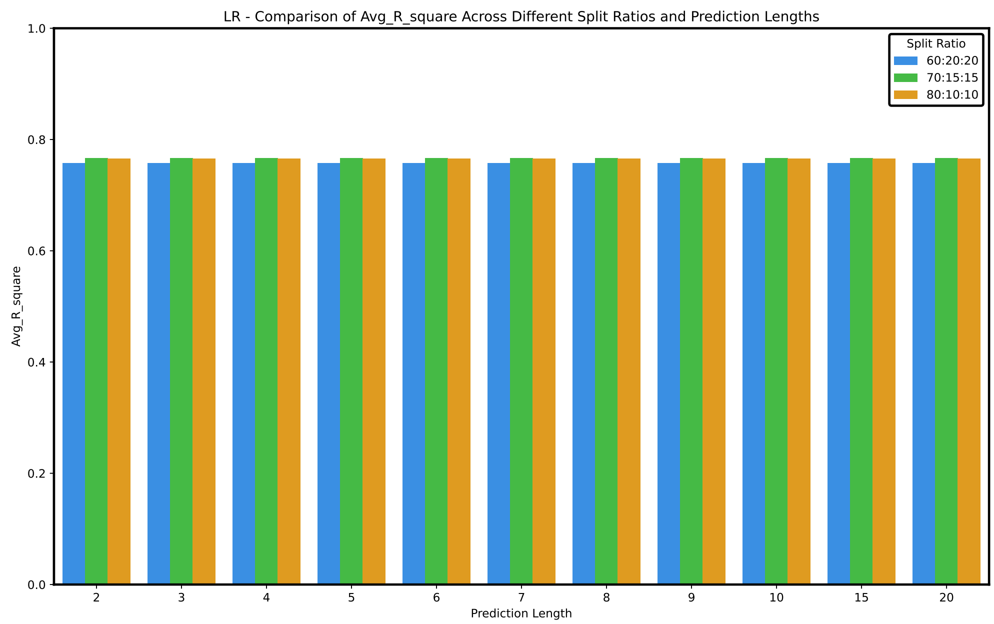

Figure 2: Comparison of Avg_R^2 Across Different Split Ratios and Prediction Lengths Using LR Model

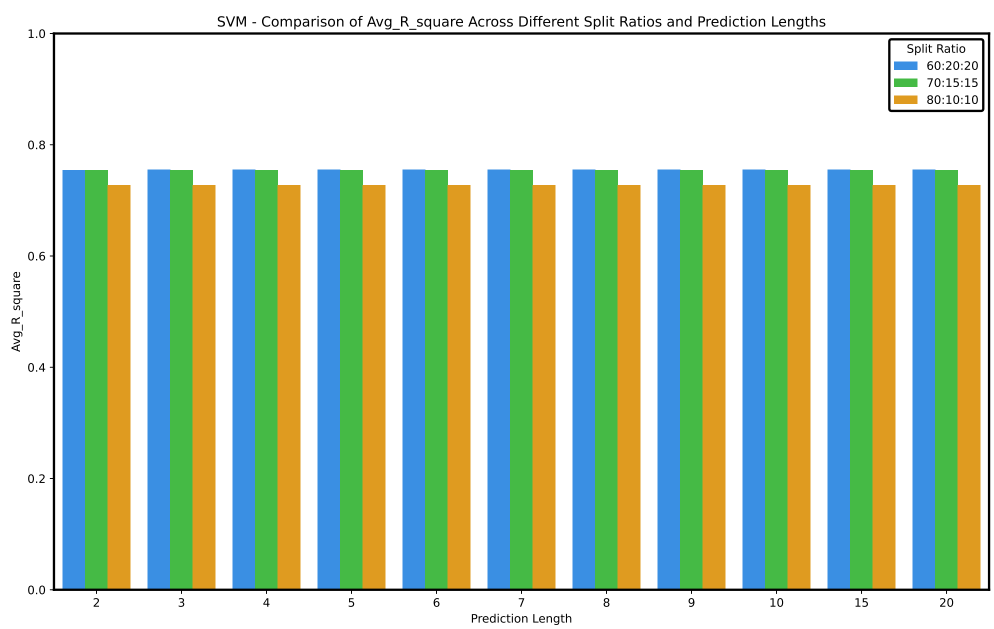

Figure 3: Comparison of Avg_R^2 Across Different Split Ratios and Prediction Lengths Using SVM Model

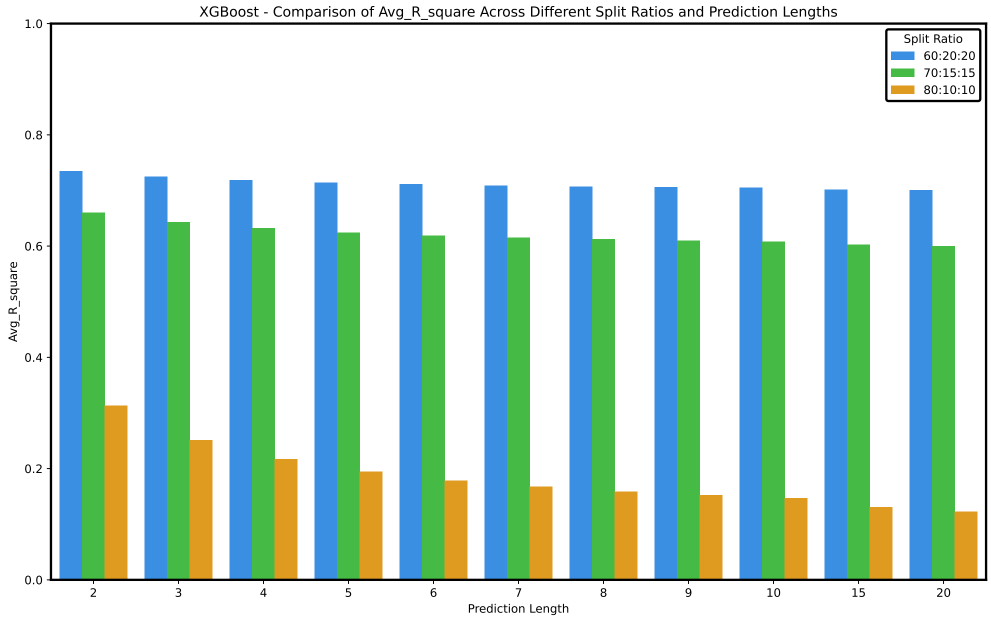

Figure 4: Comparison of Avg_R^2 Across Different Split Ratios and Prediction Lengths Using XGBoost Model

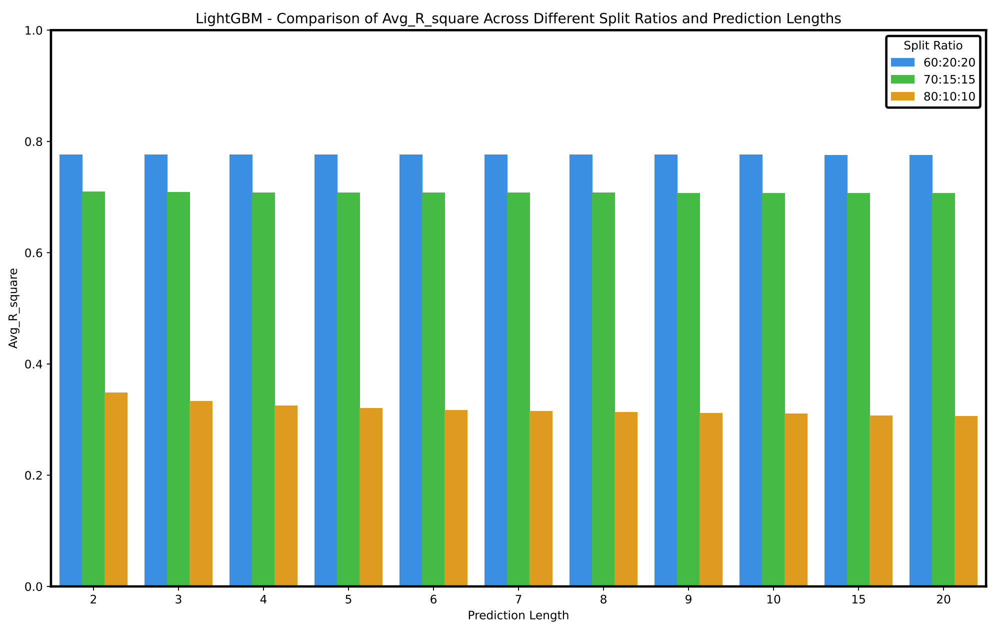

Figure 5: Comparison of Avg_R^2 Across Different Split Ratios and Prediction Lengths Using LightGBM Model

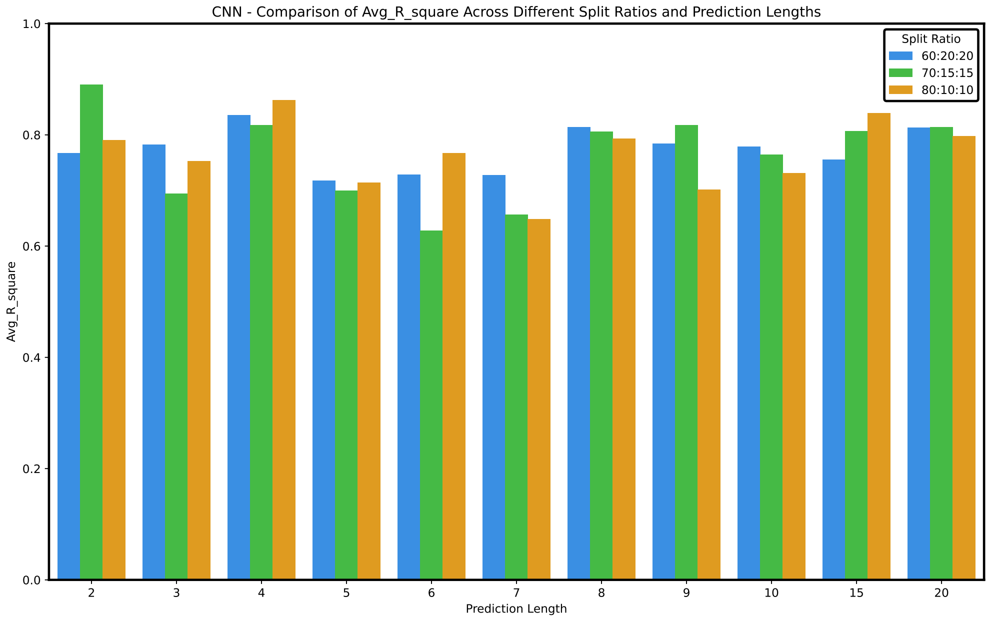

Figure 6: Comparison of Avg_R^2 Across Different Split Ratios and Prediction Lengths Using CNN Model

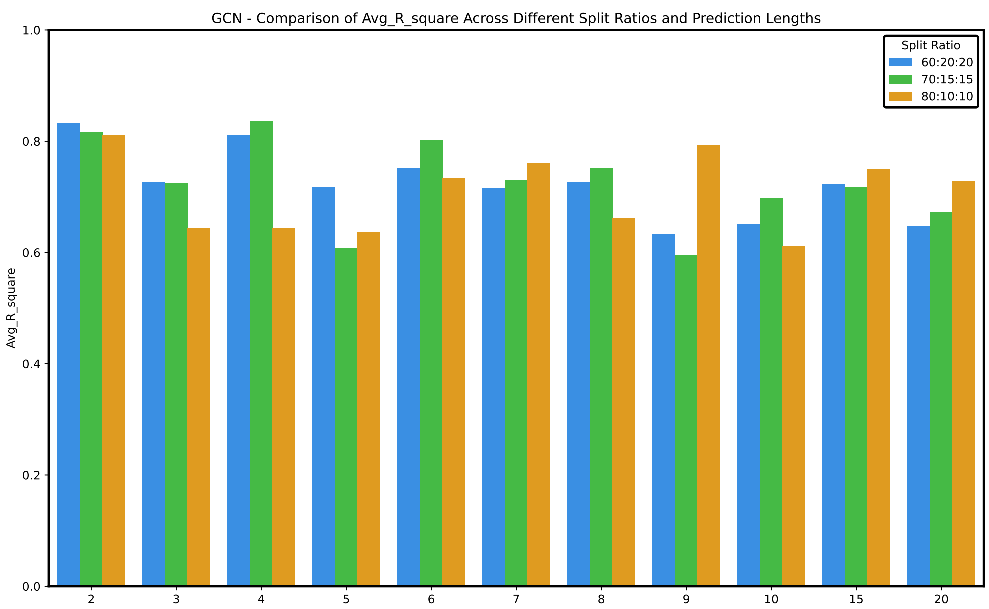

Figure 7: Comparison of Avg_R^2 Across Different Split Ratios and Prediction Lengths Using GCN Model

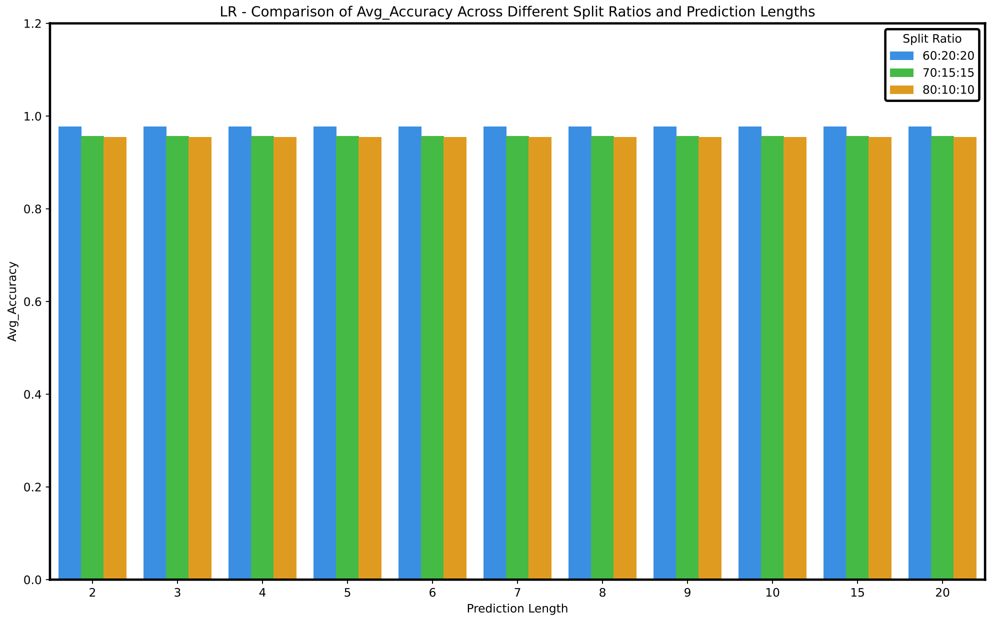

Figure 8: Comparison of $Avg\_Accuracy$ Across Different Split Ratios and Prediction Lengths Using LR Model

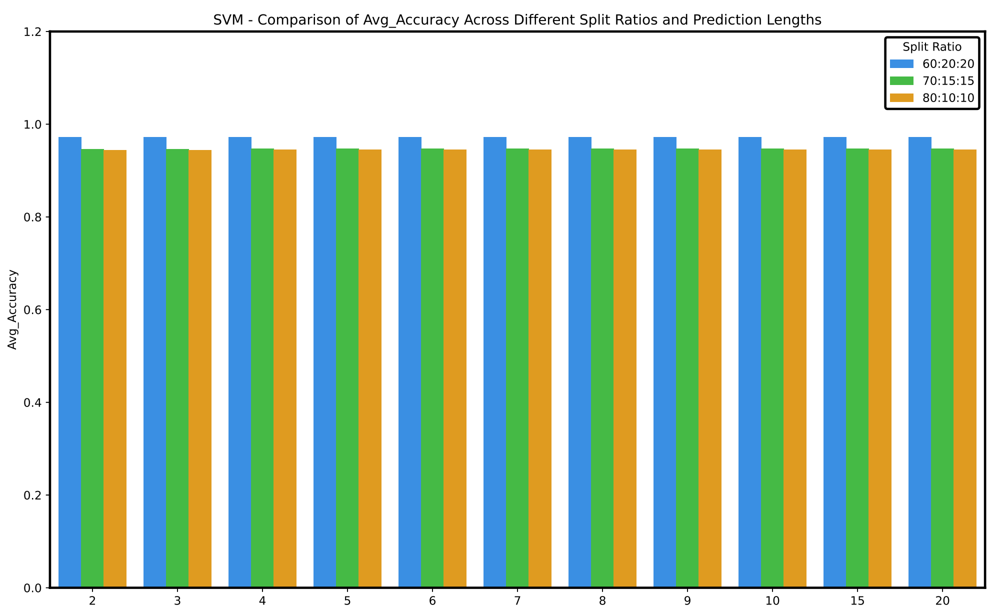

Figure 9: Comparison of $Avg\_Accuracy$ Across Different Split Ratios and Prediction Lengths Using SVM Model

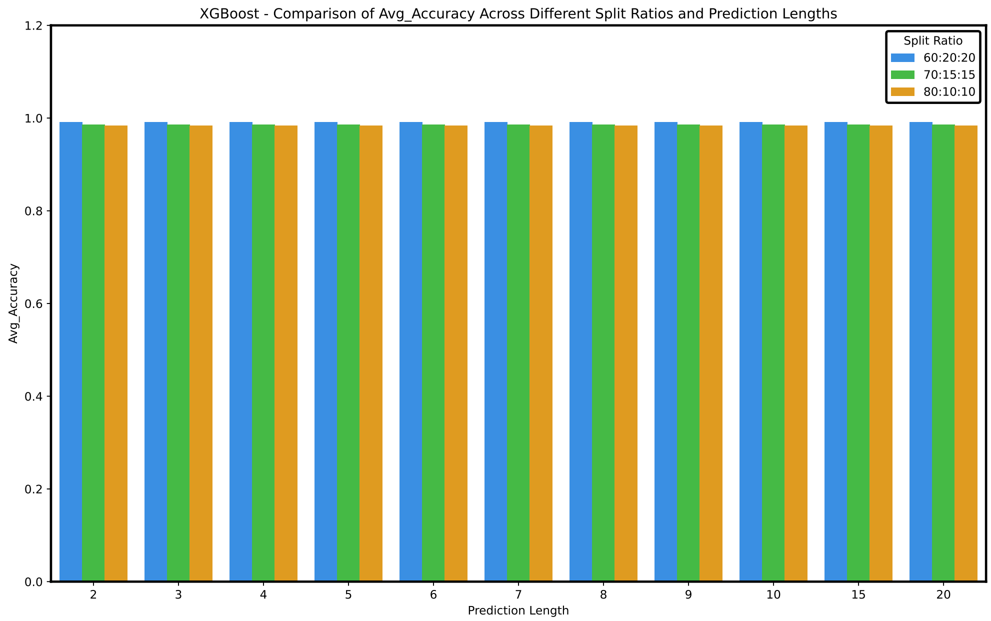

Figure 10: Comparison of $Avg\_Accuracy$ Across Different Split Ratios and Prediction Lengths Using XGBoost Model

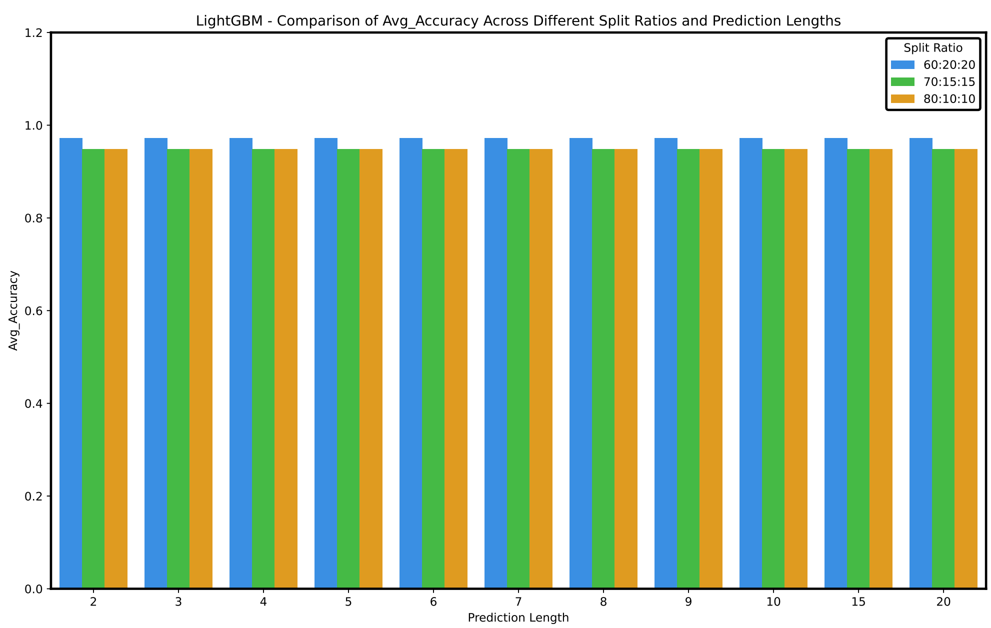

Figure 11: Comparison of $Avg\_Accuracy$ Across Different Split Ratios and Prediction Lengths Using LightGBM Model

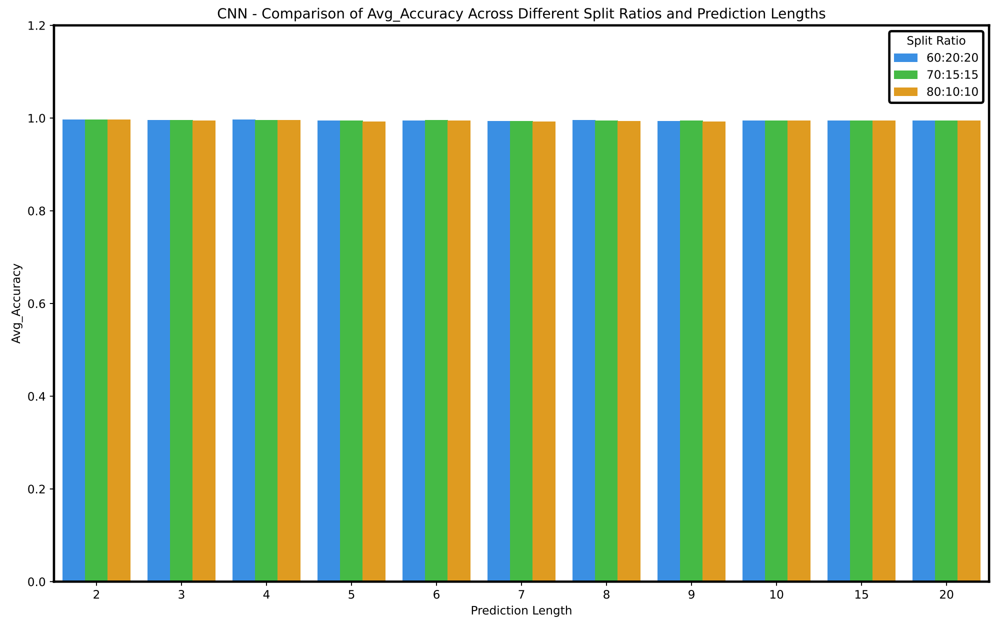

Figure 12: Comparison of $Avg\_Accuracy$ Across Different Split Ratios and Prediction Lengths Using CNN Model

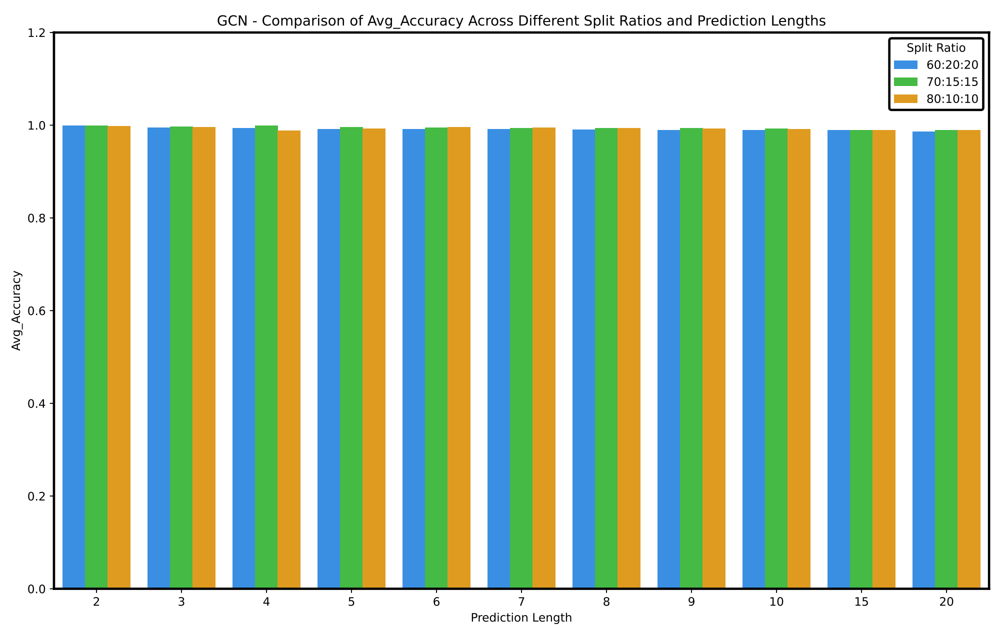

Figure 13: Comparison of $Avg\_Accuracy$ Across Different Split Ratios and Prediction Lengths Using GCN Model

Table 3: XGBoost Prediction Performance

PREDICTION LENGTHS	AVG_MSE	AVG_R^2	AVG_BCE	AVG_ACCURACY
2	0.0107	0.7349	0.2296	0.9918
3	0.0107	0.7251	0.2296	0.9918
4	0.0107	0.7188	0.2296	0.9918
5	0.0108	0.7143	0.2296	0.9918
6	0.0108	0.7113	0.2296	0.9918
7	0.0108	0.7090	0.2296	0.9918
8	0.0108	0.7073	0.2296	0.9918
9	0.0108	0.7060	0.2296	0.9918
10	0.0108	0.7050	0.2296	0.9918
15	0.0108	0.7017	0.2297	0.9918
20	0.0108	0.7002	0.2297	0.9918

Table 4: LightGBM Prediction Performance

PREDICTION LENGTHS	AVG_MSE	AVG_R^2	AVG_BCE	AVG_ACCURACY
2	0.0126	0.7765	0.0627	0.9726
3	0.0126	0.7762	0.0627	0.9726
4	0.0126	0.7760	0.0627	0.9726
5	0.0126	0.7759	0.0627	0.9726
6	0.0126	0.7758	0.0627	0.9726
7	0.0126	0.7758	0.0627	0.9726
8	0.0126	0.7758	0.0627	0.9726
9	0.0126	0.7758	0.0627	0.9726
10	0.0126	0.7758	0.0627	0.9726
15	0.0126	0.7757	0.0627	0.9726
20	0.0126	0.7757	0.0627	0.9726

Table 5: CNN Prediction Performance

PREDICTION LENGTHS	AVG_MSE	AVG_R^2	AVG_BCE	AVG_ACCURACY
2	0.0218	0.7672	0.5874	0.9970
3	0.0206	0.7820	0.6847	0.9959
4	0.0155	0.8359	0.6803	0.9963
5	0.0269	0.7181	0.8214	0.9948
6	0.0262	0.7284	0.8463	0.9947
7	0.0266	0.7276	1.0013	0.9932
8	0.0182	0.8136	0.7925	0.9955
9	0.0214	0.7838	0.9427	0.9938
10	0.0213	0.7788	0.8574	0.9947
15	0.0234	0.7555	0.8298	0.9951
20	0.0188	0.8130	0.9214	0.9945

Table 6: GCN Prediction Performance

PREDICTION LENGTHS	**AVG_MSE**	**AVG_R^2**	**AVG_BCE**	**AVG_ACCURACY**
2	0.0158	0.8329	0.2146	0.9988
3	0.0263	0.7265	0.9022	0.9948
4	0.0177	0.8110	0.9110	0.9940
5	0.0265	0.7176	1.3577	0.9916
6	0.0238	0.7521	1.3626	0.9920
7	0.0280	0.7159	1.2843	0.9917
8	0.0268	0.7267	1.4857	0.9912
9	0.0363	0.6327	1.7288	0.9894
10	0.0337	0.6502	1.6726	0.9895
15	0.0267	0.7220	1.6420	0.9901
20	0.0355	0.6469	2.3441	0.9867