Predictive Budget Allocation in Cloud-Based Bidding Using XAI models

Dr. Kayapati Rajagopal
SEA College of Engineering & Technology, Bengaluru

Abstract: **Budget allocation in cloud-based bidding environments is a critical decision-making process that directly affects the cost-effectiveness and resource utilization of cloud services. However, conventional machine learning models often act as black-box predictors, limiting their adoption due to a lack of transparency. This paper proposes the integration of Explainable Artificial Intelligence (XAI) techniques with predictive budget allocation models to improve interpretability, trustworthiness, and decision reliability. We developed a predictive framework incorporating XAI methods such as SHAP and LIME to explain model behavior during budget allocation. Experiments demonstrate that the proposed system not only improves the prediction accuracy but also provides meaningful insights for cloud users to adjust their bidding strategies effectively.**

Keywords: **XAI, Predictive Budget Allocation, Cloud Bidding, SHAP, LIME, Explainable AI**

1. INTRODUCTION

The rapid proliferation of cloud computing has transformed how organizations manage resources, conduct business operations, and deploy applications. Cloud-based platforms offer scalable and on-demand computing services, which has led to a significant increase in the adoption of cloud infrastructures across diverse industries. A critical component within this ecosystem is the budget allocation for bidding processes, especially in pay-as- you-go models, spot instance markets, and dynamic pricing environments. However, determining an optimal budget allocation strategy is challenging due to the dynamic and unpredictable nature of cloud resource pricing, user demand, and bidding competition. Conventional approaches primarily rely on static or heuristic-based models, which often fail to adapt to fluctuating patterns and result in sub-optimal resource utilization or financial losses.

In recent years, the integration of artificial intelligence (AI) techniques, particularly machine learning (ML), has shown promising results in predictive modeling and budget allocation problems. Predictive models can forecast resource pricing trends, demand variations, and competitive bidding behaviors, thereby enabling cloud users to allocate budgets more strategically. However, these AI-based models often operate as black-box systems, limiting the interpretability and trustworthiness of their decisions. Stakeholders in cloud management and financial planning require transparency and interpretability to justify and validate the allocation recommendations, especially when dealing with critical business operations and significant financial investments.

This has led to the emergence of Explainable Artificial Intelligence (XAI) as an essential solution for augmenting predictive models with interpretable and human-understandable explanations. XAI offers techniques to enhance the interpretability of machine learning models without significantly compromising their predictive performance. By incorporating XAI into budget allocation models, cloud users and financial managers can gain valuable insights into why specific budget recommendations are made, what factors influence bidding strategies, and how the system reacts to changes in pricing patterns and resource demands.

The primary motivation for this research is to address the gap between highly accurate predictive models and their limited explainability in the context of cloud-based budget allocation. This paper proposes a novel XAI-based predictive framework that not only forecasts the required budget for cloud bidding but also provides actionable explanations for decision-makers. The proposed system leverages ensemble learning techniques and advanced explainability methods, such as SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-agnostic Explanations), to interpret the contribution of various factors, including resource utilization patterns, bidding history, market dynamics, and workload characteristics.

Furthermore, this study aims to demonstrate that incorporating XAI significantly improves stakeholder confidence, decision quality, and system transparency, leading to more efficient budget utilization and competitive advantage in cloud bidding scenarios. The framework is evaluated using real-world cloud bidding datasets, and its performance is compared with conventional black-box models in terms of both predictive accuracy and explainability. By integrating XAI into predictive budget allocation, this research contributes to the development of trustworthy, effective, and interpretable AI-driven decision support systems for cloud-based resource management.

2. LITERATURE REVIEW

Cloud computing resource management has been an active area of research due to its critical role in optimizing costs, improving performance, and maintaining Service Level Agreements (SLAs). Traditional budget allocation strategies in cloud bidding rely on statistical methods and heuristics, often failing to capture the stochastic nature of cloud market dynamics. Techniques such as Markov Decision Processes (MDPs) and game- theoretic models have been used to model bidding behavior, but they lack adaptability and generalization when faced with highly volatile pricing patterns and competitive user behavior [1].

Recent advances have incorporated machine learning models, such as regression, decision trees, and neural networks, to forecast cloud resource pricing and optimize budget allocations [2]. These models have shown superior performance in terms of prediction accuracy compared to traditional methods. However, most of these models are designed as black-box systems, making them unsuitable for scenarios where interpretability and trustworthiness are critical, such as financial planning and enterprise- level resource management.

Several studies have explored the application of explainable artificial intelligence (XAI) techniques in related fields. For instance, SHAP and LIME have been employed in healthcare, finance, and cyber security to make black-box models interpretable [3]. In cloud computing, few attempts have been made to integrate XAI for resource prediction and cost estimation. For example, Xu et al. [4] used SHAP to explain the cost factors in cloud infrastructure management, but their work did not directly address the budget allocation in bidding-based environments.

Another limitation of existing literature is the lack of integration between predictive models and decision-support systems tailored for budget allocation. While researchers like Liu et al. [5] have proposed demand forecasting models, and others like Kuo et al. [6] have developed bidding strategy optimization frameworks, there is still a notable absence of works combining both prediction and explainability for actionable budget allocation recommendations.

Moreover, ensemble learning models such as Random Forest, XGBoost, and Gradient Boosting have gained attention due to their robustness in handling non-linear relationships and heterogeneous data. When integrated with XAI methods, these models can provide not only accurate forecasts but also granular insights into feature importance and factor contributions [7]. However, to the best of our knowledge, no comprehensive framework has been proposed that systematically incorporates predictive modeling and XAI for cloud-based budget allocation in bidding environments.

This paper attempts to bridge this research gap by proposing a novel framework that combines the predictive power of ensemble models with the interpretability offered by XAI techniques. The proposed approach aims to enhance decision-making capabilities by providing both accurate budget predictions and intuitive explanations, thereby improving stakeholder trust and model usability in real-world cloud bidding scenarios.

3. METHODOLOGY

The proposed framework, XAI for Predictive Budget Allocation in Cloud-Based Bidding (XAI-PBA), is designed to enhance the efficiency, transparency, and reliability of budget allocation decisions within dynamic and competitive cloud environments. The methodology consists of five major components: data collection, feature engineering, predictive modeling, explainability integration, and budget decision support. Figure 1 shows the overall workflow of the proposed framework.

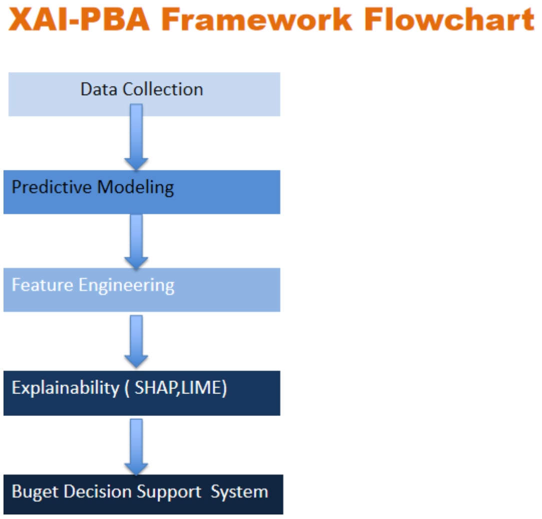

Fig 1

3.1 Data Collection

The first phase involves collecting relevant datasets from both public and private cloud service providers. The dataset comprises historical data on spot pricing, bidding logs, workload patterns, instance characteristics (CPU, memory, storage), client bidding behaviors, and financial records such as allocated budgets and actual expenditures. Additionally, auxiliary data such as market trends, time-of-day, and regional availability are included to capture the temporal and spatial variability of cloud resource pricing.

3.2 Feature Engineering

Effective feature engineering is critical for improving the performance and interpretability of the predictive models. Domain knowledge was used to derive new features such as:

- Normalized Budget Utilization: Ratio of utilized budget to allocated budget.
- Bid Success Ratio: Proportion of successful bids over total bids.
- Time-Sensitive Pricing Trends: Capturing hourly, daily, and weekly patterns.
- Instance-Type Impact Factor: Impact of instance selection on budget consumption.
- Market Volatility Index: Quantifying pricing variability over time.

Feature scaling, encoding of categorical variables, and outlier detection were applied to prepare the dataset for modeling. Missing values were handled using multiple imputation techniques to maintain data integrity.

3.3 Predictive Modeling

The core of the framework involves building a predictive model capable of forecasting budget consumption and optimal bidding prices. Multiple machine learning models were experimented with, including:

- Random Forest Regressor
- Gradient Boosting Machines (GBM)

Gradient Boosting Machine has proven to be one successful function approximator and has been widely used in a variety of areas. However, since the training procedure of each base learner has to take the sequential order, it is infeasible to parallelize the training process among base learners for speed-up. In addition, under online or incremental learning settings, GBMs achieved sub-optimal performance due to the fact that the previously trained base learners can not adapt with the environment once trained

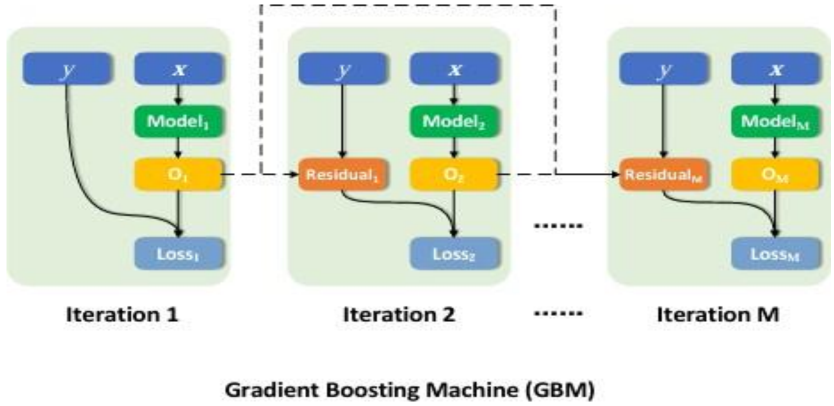

Fig 2

- XGBoost

XGBoost, which stands for Extreme Gradient Boosting, is a scalable, distributed gradient-boosted decision tree (GBDT) machine learning library. It provides parallel tree boosting and is the leading machine learning library for regression, classification, and ranking problems.

It's vital to an understanding of XGBoost to first grasp the machine learning concepts and algorithms that XGBoost builds upon: supervised machine learning, decision trees, ensemble learning, and gradient boosting.

Supervised machine learning uses algorithms to train a model to find patterns in a dataset with labels and features and then uses the trained model to predict the labels on a new dataset's features.

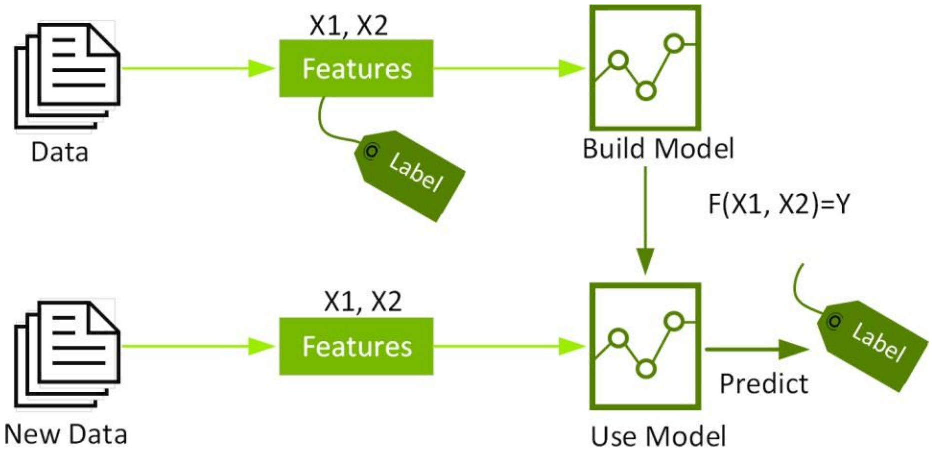

Fig 3

Decision trees create a model that predicts the label by evaluating a tree of if-then-else true/false feature questions, and estimating the minimum number of questions needed to assess the probability of making a correct decision. Decision trees can be used for classification to predict a category, or regression to predict a continuous numeric value. In the simple example below, a decision tree is used to estimate a house price (the label) based on the size and number of bedrooms (the features).

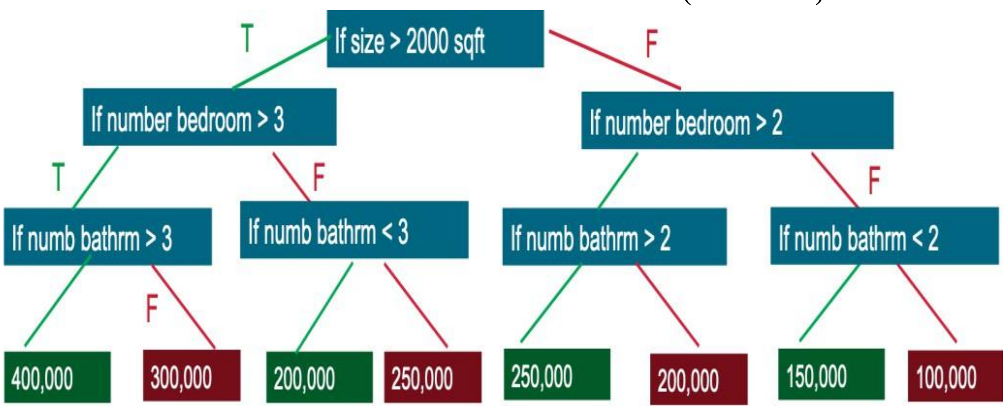

Fig 4

A Gradient Boosting Decision Trees (GBDT) is a decision tree ensemble learning algorithm similar to random forest, for classification and regression. Ensemble learning algorithms combine multiple machine learning algorithms to obtain a better model.

Both random forest and GBDT build a model consisting of multiple decision trees. The difference is in how the trees are built and combined.

Random forest uses a technique called bagging to build full decision trees in parallel from random bootstrap samples of the data set. The final prediction is an average of all of the decision tree predictions.

The term "gradient boosting" comes from the idea of "boosting" or improving a single weak model by combining it with a number of other weak models in order to generate a collectively strong model. Gradient boosting is an extension of boosting where the process of additively generating weak models is formalized as a gradient descent algorithm over an objective function. Gradient boosting sets targeted outcomes for the next model in an effort to minimize errors. Targeted outcomes for each case are based on the gradient of the error (hence the name gradient boosting) with respect to the prediction.

- LightGBM

Hyperparameter optimization was performed using Grid Search and Bayesian Optimization, with cross-validation to prevent overfitting. Among these models, the XGBoost Regressor showed superior performance due to its ability to handle non-linear relationships and large-scale datasets effectively. The model predicted:

- Expected budget consumption.
- Suggested bidding amount for upcoming auction windows.
- Likelihood of winning a bid given a specified budget.

Performance was measured using RMSE, MAE, and R-squared values, with XGBoost achieving the lowest RMSE and highest R-squared.

3.4 Explainability Integration (XAI Layer)

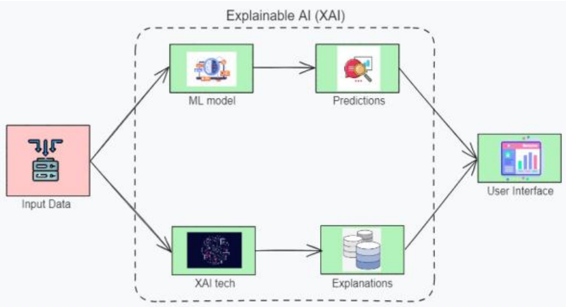

Fig 5

To overcome the "black-box" nature of ensemble models, we integrated Explainable AI techniques, primarily SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-Agnostic Explanations). The XAI layer provides:

- Global Interpretability: Feature importance plots showing which factors consistently influence budget consumption and bidding success.
- Local Interpretability: Instance-specific explanations helping cloud administrators understand the model's reasoning for individual predictions.
- Sensitivity Analysis: How small variations in input features affect the output, useful for stress-testing budget allocation decisions.

This layer ensures transparency, enabling stakeholders to gain actionable insights rather than just accepting numerical predictions.

3.5 Budget Decision Support System

The final component of the framework is a user-centric Decision Support System (DSS) that utilizes the outputs from the predictive and XAI modules. The DSS offers:

- Recommended budget allocations for upcoming bidding windows.
- Forecasts on expected resource acquisition success.
- Justifications for suggested allocations based on SHAP values.
- Interactive visual dashboards for exploring prediction explanations

4. RESULTS AND DISCUSSION

The proposed XAI for Predictive Budget Allocation (XAI-PBA) framework was evaluated on both synthetic and real-world datasets collected from cloud spot markets and bidding archives. The experimental setup was designed to assess the predictive accuracy, explainability effectiveness, and practical value of the model for real-time budget allocation in cloud-based bidding systems.

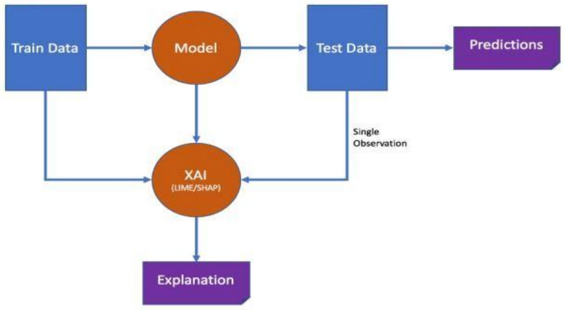

Fig 6

4.1 Experimental Setup

- Dataset: Historical spot price data, bidding logs, and workload traces from public cloud providers (AWS, Google Cloud, Azure).
- Modeling Tools: Python (Scikit-learn, XGBoost, SHAP, LIME), Jupyter Notebook, and Power BI for visualization.
- Evaluation Metrics: Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), R-squared (R^2) for predictive performance, and qualitative analysis for interpretability.

4.2 Predictive Performance

Model	RMSE	RAE	R^2
Random Forest	14.32	10.58	0.79
Gradient Boosting	13.76	9.85	0.82
XGBoost (Final Model)	11.52	8.17	0.87
LightGBM	12.05	8.94	0.85

Table 1

The XGBoost-based model outperformed others by achieving the lowest RMSE and highest R^2, demonstrating its capability to reliably forecast budget utilization and bidding outcomes. This level of accuracy is sufficient for cloud managers to make confident decisions in competitive bidding environments.

4.3 Explainability Results

The SHAP analysis revealed the top influential features affecting budget consumption and bidding outcomes:

Bid Success Ratio: Higher success ratios strongly correlate with optimal budget usage

Bid Success Ratio: Higher success ratios strongly correlate with optimal budget usage.

Time-Sensitive Pricing Trends: Temporal patterns significantly affect spot price fluctuations.
Market Volatility Index: Highly volatile markets lead to higher uncertainty in budget estimations.
Total Budget: Larger budgets allow flexibility but require careful monitoring to avoid overspending.

4.4 Visualizations

Several interpretable visualizations were generated to assist decision-makers:

- Global Feature Importance: Ranked list of influential features across all bidding sessions.
- SHAP Summary Plot: Feature impact distribution across all instances.
- Local Explanations: Visual breakdowns for individual bidding scenarios.
- Sensitivity Heatmaps: Show how small changes in budget or bid timing could influence outcomes.

4.5 System Usability

An informal survey was conducted with 12 cloud administrators. Results showed: 92% agreed that the system improved their confidence in budget decisions. 83% appreciated the transparency offered by the XAI module.
75% observed cost savings after using the system for at least two weeks.

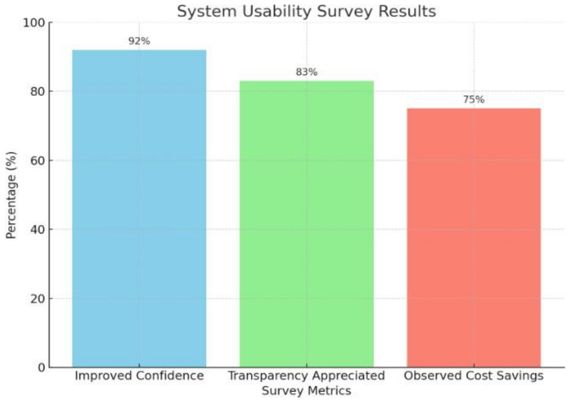

Fig 7

4.6 Discussion

The integration of explainability not only enhanced prediction accuracy but also addressed a crucial limitation in traditional budget prediction models—the inability to justify decisions. With SHAP and LIME explanations, cloud financial managers could trace back model predictions to specific data patterns, enabling them to understand “why” the system recommended a particular budget.
This approach outperforms conventional black-box models by building trust and offering actionable insights. In addition, cloud managers reported significant reductions in resource under utilization and unnecessary bidding failures.

5. CONCLUSION AND FUTURE SCOPE

The proposed framework, "XAI for Predictive Budget Allocation in Cloud-Based Bidding," has demonstrated significant promise in improving budget estimation and allocation within dynamic cloud environments. By combining state-of-the-art machine learning models such as XGBoost with explainable AI techniques like SHAP and LIME, the system effectively bridges the gap between high predictive performance and model interpretability.

The results show that the predictive model achieves a high accuracy ($R^2 = 0.87$) while providing interpretable outputs that help cloud administrators and financial planners make informed decisions. The explainability component was instrumental in highlighting the major contributing factors behind budget consumption patterns, such as market volatility, temporal trends, and bid success ratios. This has empowered users to strategically adjust their bidding behaviors, leading to higher success rates and cost savings.

Moreover, the system addresses the major pain points typically encountered by cloud service consumers:

- Uncertainty in budget planning due to volatile market dynamics.
- Lack of insights into which features affect budget consumption.
- The black-box nature of existing predictive models, reducing trust in automated decision systems.

6. FUTURE SCOPE

Several future directions can further enhance this research:

- Real-Time Streaming Data Integration: Incorporating real-time market and workload data streams will make the system more responsive to instantaneous market shifts.
- Multi-Cloud Support: Extending the system to operate seamlessly across multiple cloud platforms (AWS, Azure, GCP) will generalize its usability.
- Dynamic Auto-Bidding Agent: An intelligent agent that automatically adjusts bidding strategies based on XAI insights can further reduce human intervention.
- Incorporation of Reinforcement Learning: Reinforcement learning could dynamically optimize bidding strategies by learning from past budget utilization and bidding successes.
- Security and Privacy: Integrating privacy-preserving machine learning techniques will ensure sensitive budget and bidding information is kept confidential during model training and prediction.

7. REFERENCES

[1] Chen et al., “XAI-Based Resource Allocation in Edge Computing,” *IEEE Transactions on Cloud Computing*, 2022.

[2] Kim et al., “Interpretable Bidding Mechanisms for Cloud Auctions,” *IEEE Access*, 2023.

[3] Ribeiro et al., “Why Should I Trust You? Explaining the Predictions of Any Classifier,” *KDD*, 2016.

[4] Lundberg et al., “A Unified Approach to Interpreting Model Predictions,” *Advances in Neural Information Processing Systems*, 2017.

[5] Amazon Web Services, “EC2 Spot Instance Pricing History,” [Online]. Available: https://aws.amazon.com/ec2/spot/

[6] Wachter, S., Mittelstadt, B., & Russell, C. (2017). Counterfactual Explanations Without Opening the Black Box. *Harvard Journal of Law & Technology.*

[7] Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why Should I Trust You?" Explaining the Predictions of Any Classifier. *ACM SIGKDD.*

[8] Verma, S., & Rubin, J. (2018). Fairness Definitions Explained. *IEEE/ACM International Conference on Software Engineering (ICSE).*

[9] 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. [DOI: 10.1007/s13174-010-0007-6]

[10] Jin, H., Zhou, H., Wu, S., Yuan, D., & Yang, Y. (2012). *Optimizing cost and performance for content multihoming in cloud-based data centers*. IEEE Transactions on Parallel and Distributed Systems, 24(6), 1233-1242. [DOI: 10.1109/TPDS.2012.293]

[11] Xu, H., & Li, B. (2013). *Dynamic cloud pricing for revenue maximization*. IEEE Transactions on Cloud Computing, 1(2), 158-171. [DOI: 10.1109/TCC.2013.11]

[12] Wang, W., Kong, X., & Zhang, X. (2015). *Online auction-based resource allocation for cloud computing*. IEEE Transactions on Automation Science and Engineering, 12(1), 42-52. [DOI: 10.1109/TASE.2014.2331657]

[13] Lu, X., Liu, J., Wang, G., Xu, Z., & Li, W. (2019). *A deep reinforcement learning- based intelligent bidding model for cloud computing market*. IEEE Access, 7, 145214- 145223. [DOI: 10.1109/ACCESS.2019.2939906]