Application of XGBoost and Stacking Models in Elastic Optical Network Optimization

Robert N. Kanimba
Department of Systems and Computer Networks
Wrocław University of Science and Technology
Wrocław, Poland
276884@student.pwr.edu.pl

Yusuf Yeşilyurt
Department of Systems and Computer Networks
Wrocław University of Science and Technology
Wrocław, Poland
277034@student.pwr.edu.pl

***Abstract*—The rapid expansion of internet usage, with over 70% of the global population connected in 2023, has escalated the demands on optical networks, particularly in the realm of Elastic Optical Networks (EONs). EONs, known for their dynamic resource allocation capabilities, face the critical challenge of efficiently managing Routing and Spectrum Allocation (RSA). This paper addresses this challenge by applying advanced machine learning techniques, specifically focusing on the stacking of XGBoost models, to optimize RSA in EONs. Our research presents a comprehensive analysis of these techniques, highlighting their effectiveness in predicting network performance metrics with higher accuracy. We compare our approach with existing methods, such as deep learning and deep reinforcement learning frameworks, underscoring the advantages of XGBoost stacking in handling complex data structures and enhancing prediction accuracy. The findings demonstrate the superiority of ensemble methods, especially stacking with linear and ridge regression as meta-learners, in terms of predictive capabilities. The trade-off between computational efficiency and model accuracy is also examined, emphasizing the strategic use of ensemble learning for network performance optimization in EONs. The paper concludes with recommendations for further research in machine learning applications to improve EON management, offering insights into the potential advancements in network design and operation efficiency. The paper concludes by advocating for ongoing advancements in machine-learning techniques to optimize Elastic Optical Networks.**

I. Introduction

Optical networking leverages light to transmit data across various global locations, finding applications in fields like fiber optic communications and cable television. The design and management of optical networks are crucial due to the surging traffic demand, evidenced by the 70% of the world population being internet users in 2023 [1]. In Elastic Optical Networks, spectrum slots accommodate fine granularity, significantly enhancing network efficiency compared to conventional Wavelength Division Multiplexing (WDM) systems. The flexibility in EONs, often termed as 'elastic,' manifests in the network's ability to adjust resources such as data rate, channel wavelength, and bandwidth dynamically while maintaining the Quality of Transmission (QoT) for both newly assigned and previously established connections [2]. Studies like those by Zhang et al. have delved into high-speed transmission technologies and spectrum management techniques, respectively, for EONs. A critical aspect of EON management and planning is the Routing and Spectrum Allocation (RSA) problem, crucial for accommodating a specified set of static demands within the network. RSA entails the delineation of link sequences between source and destination nodes for a given request and the allocation of optical spectrum resources for executing the request. The RSA problem in EONs is critical due to the need for efficient path computation and spectrum assignment that adapts to variable bandwidth demands, ensuring high data throughput and minimal spectral wastage [3].

This paper delves into the application of advanced machine learning techniques, with a specific focus on the stacking of XGBoost models, to tackle the challenges in RSA for EONs. Our work presents an in-depth analysis of these techniques, assessing their effectiveness in enhancing the prediction accuracy of network performance metrics. This approach represents a novel contribution to the field of EON management and optimization. The structure of the paper is as follows: Section II offers a review of the literature, placing our research within the broader scope of existing studies. Section III details our methodology, focusing on the implementation of the stacking of XGBoost models. In Section IV, we discuss the results and analyze the performance of our models. Section V provides a comprehensive analysis of these results in the context of optical networking. The paper concludes in Section VI, summarizing our key findings and outlining potential avenues for future research.

II. Related Works

Machine learning applications in routing and resource allocation within Elastic Optical Networks (EONs) have seen varied approaches. Yu et al. [5] proposed a deep learning-based method for Routing and Spectrum Assignment (RSA), optimizing RSA for specific network configurations and demonstrating effectiveness in reducing spectrum fragmentation and blockages. Salani et al. [6] introduced a machine learning approach for Quality of Transmission (QoT) estimations in RSA, combining ML-based QoT estimations with Integer Linear Programming to iteratively solve RSA problems. Chen et al. [7] developed DeepRMSA, a deep reinforcement learning framework for routing, modulation, and spectrum assignment in EONs, employing deep neural networks for online RMSA policies and demonstrating advancements in network operation optimization. Musumeci et al. [9] provided a comprehensive overview of machine learning applications in optical communications and networking. In their work, they classified and surveyed relevant literature, presenting an introductory tutorial on machine learning for both researchers and practitioners in the field.

Our research diverges from these methodologies by focusing on the stacking of XGBoost models. Unlike the aforementioned approaches, we leverage the stacking of XGBoost due to its efficacy in processing complex datasets and its robustness in handling diverse data structures. This approach allows for high-accuracy predictions and efficient data processing, offering a novel and effective solution for RSA optimization in EONs.

III. Problem Definition

This research addresses the optimization of Routing and Spectrum Allocation (RSA) in Elastic Optical Networks (EONs) through the application of machine learning. The primary goal is to create a predictive model capable of estimating essential network performance metrics, thus offering a viable alternative to conventional, computationally intensive RSA techniques. The model is designed to predict four key metrics: the highest occupied slot in the network, the average highest occupied slot on links, the sum of all occupied slots, and the number of transceivers in use. These metrics are critical for assessing the network's operational efficiency and performance.

The model utilizes input features such as source and destination nodes, Bit Rate, and, where relevant, the physical distance between nodes. The selection of these features is based on their relevance to the RSA process and the potential to enhance the model's accuracy in reflecting network conditions. Training on historical network data, the model aims to establish a relationship between these inputs and the network's performance metrics, thereby streamlining the RSA decision-making process.

IV. Proposed Algorithm

In this study, we adopt gradient boosting as a foundation for our ensemble learning strategy combined with stacking. Gradient boosting is a powerful technique that sequentially enhances weak learners to construct a robust predictive model while adeptly managing the bias-variance trade-off [10].

A. Extreme Gradient Boosting (XGBoost)

Extreme Gradient Boosting (XGBoost) is an instantiation of gradient boosting that excels in terms of computational speed, scalability, and model performance. It is noteworthy that the development of XGBoost as a scalable system for tree boosting has been acknowledged for its significant performance across various machine learning challenges, particularly for its ability to efficiently handle large-scale datasets in both distributed and constrained memory environments [11]. XGBoost constructs an ensemble of decision trees in sequence, with each tree addressing the residuals left by the previous ones. It utilizes gradient optimization to minimize a loss function, making it proficient for various machine-learning tasks.

XGBoost comprises the following elements:

- **Decision Trees**: XGBoost leverages decision trees as base learners due to their capability to model complex interactions among features. Each tree is constructed by partitioning the data to minimize the loss function.
- **Gradient-based Optimization**: XGBoost refines the model by calculating the gradient of the loss function with respect to predictions, guiding the model towards reduced error.
- **Regularization**: XGBoost incorporates regularization, defined as supplementary techniques for better model generalization, by penalizing unnecessary complexity in the loss function [8].
- **Shrinkage**: Also known as the learning rate, shrinkage is a parameter that controls the influence of each tree and helps prevent overfitting.
- **Feature Importance**: XGBoost provides a score for each feature's importance, assisting in understanding each variable's influence on the predictive model.

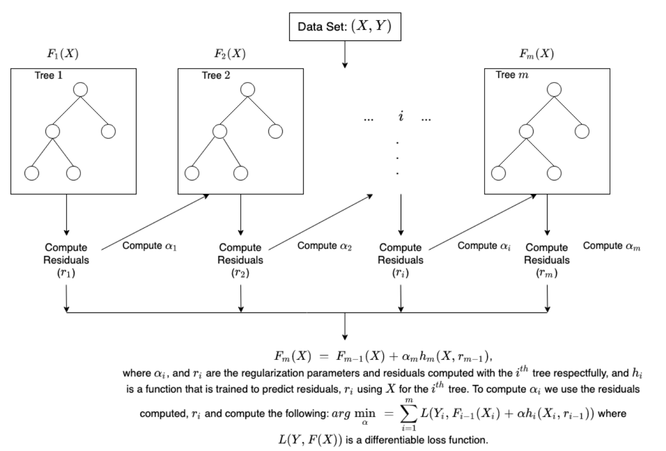

Fig. 1. Pseudocode for Gradient Boosting Trees, adapted from [15].

B. Stacking with XGBoost and Random Forest

To enhance the predictive power of our ensemble model, we introduce stacking, combining the strengths of XGBoost and Random Forest. Stacking is an ensemble learning technique that leverages multiple base models and a meta-learner to make predictions [12]. It is observed that the introduction of a stacking ensemble model, particularly one that utilizes the attention mechanism and combines models like RF, AdaBoost, and XGBoost, has been shown to notably enhance accuracy in runoff forecasting, underscoring the effectiveness of stacking approaches in improving prediction models [13]. In our case, the base models are XGBoost and Random Forest, and we employ meta-learners such as linear regression, ridge regression, and support vector machine (SVM) to optimize the final predictions.

C. Meta Learners

Our stacking approach involves three different base learners: linear regression, ridge regression, and support vector machine (SVM).

1) Linear Regression: Linear regression is a straightforward regression algorithm that seeks to establish a linear relationship between input features and the target variable. It minimizes the sum of squared errors to fit a linear equation to the data.

2) Ridge Regression: Ridge regression is a regularized linear regression technique that adds a penalty term to the linear regression objective function to prevent overfitting. It introduces L2 regularization, which helps control the model's complexity.

3) Support Vector Machine (SVM): Support Vector Machine (SVM) is a versatile machine learning algorithm that can be used for both classification and regression tasks. It aims to find the hyperplane that best separates data points while maximizing the margin [14].

By utilizing these three base learners with the power of XGBoost and Random Forest through stacking, we aim to create a robust ensemble models for our Routing and Spectrum Allocation (RSA) optimization problem.

V. Dataset Description

The dataset for this study comprises simulation outputs from two different Network topologies: 'euro28' and 'us26'. Each topology is represented by 100 simulations, reflecting a diverse set of network scenarios. For each simulation, the dataset contains information of the network request characterized by features such as source, destination, and bitrate. Corresponding to each set of requests, are four key performance metrics: the highest occupied slot in the network, the average highest occupied slot on links, the sum of all occupied slots, and the number of transceivers used which we aim to predict.

Preliminary exploratory data analysis was conducted to understand the underlying structure and relationships within the data. Figure 2 and Figure 3 present the correlation matrices for the euro28 and us26 topologies, respectively. These heatmaps provide insights into the linear relationships between features and targets, highlighting potential multicollinearity or independence among variables.

Moreover, the distributions of the features and targets are depicted in Figure 4 and Figure 5. These plots illustrate the spread and skewness of the data, which are critical factors to consider before applying any machine learning algorithms.

VI. Experimental Setup

Our experimental framework evaluates various machine learning models on their ability to predict key performance metrics in Elastic Optical Networks (EONs). We employed a set of individual models, such as XGBoost, RandomForest, and AdaBoost, and ensemble methods using stacking techniques with different meta-learners: linear regression, SVM, and ridge regression. The computational efficiency of each approach was recorded, providing insights into the trade-off between prediction accuracy and resource consumption.

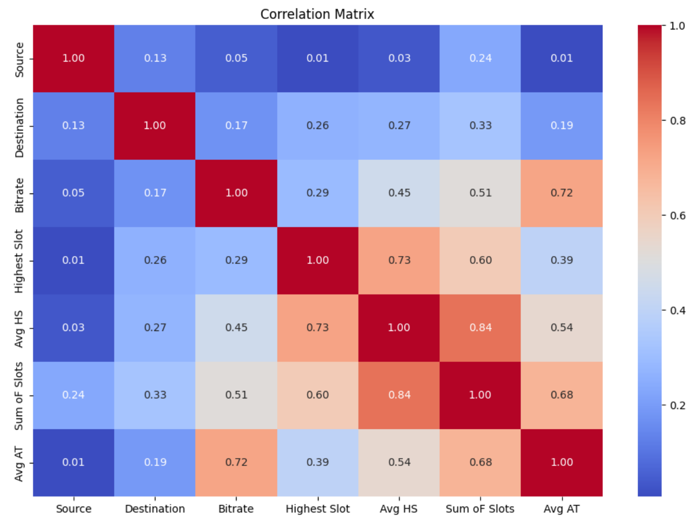

Fig. 2. Correlation matrix heatmap for the euro28 topology.

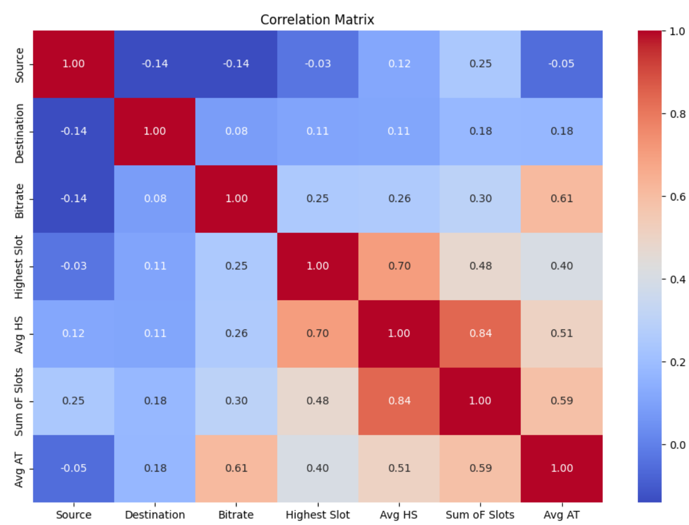

Fig. 3. Correlation matrix heatmap for the us26 topology.

VII. Results

The experimental results for the two network topologies, 'euro28' and 'us26', are summarized in the tables below. The models were evaluated based on their performance metrics such as Mean RMSE, Standard Deviation of RMSE, and Mean MAPE.

VIII. Analysis

The analysis of the experimental results from the 'euro28' and 'us26' network topologies offers significant insights into the performance of various machine learning models, particularly the stacking methods. The stacking approaches, combining XGBoost and Random Forest with meta-learners such

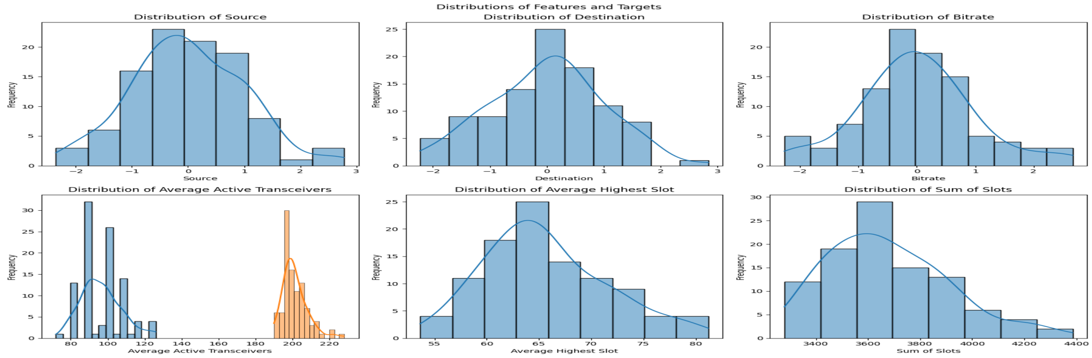

Fig. 4. Distributions of features and targets for the euro28 topology.

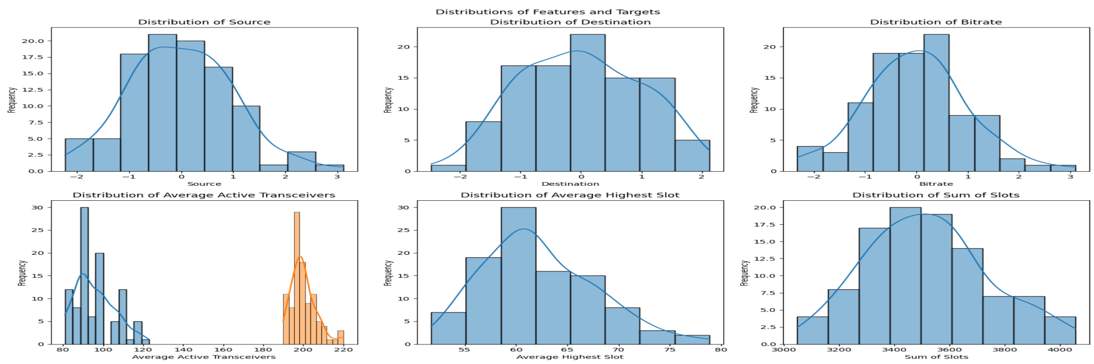

Fig. 5. Distributions of features and targets for the us26 topology.

TABLE I
PERFORMANCE METRICS FOR EURO28 TOPOLOGY

Model	Mean RMSE	Std RMSE	Mean MAPE
XGBoost	61.80	6.99	0.0627
RandomForest	56.23	6.44	0.0553
AdaBoost	58.01	7.41	0.0564
XGB & RF w/ LinearReg	55.89	7.10	0.0552
XGB & RF w/ SVM	64.45	8.92	0.0574
XGB & RF w/ Ridge	56.01	7.24	0.0552

TABLE II
PERFORMANCE METRICS FOR US26 TOPOLOGY

Model	Mean RMSE	Std RMSE	Mean MAPE
XGBoost	62.55	8.46	0.0595
RandomForest	55.73	6.78	0.0544
AdaBoost	59.52	6.65	0.0566
XGB & RF w/ LinearReg	56.10	7.47	0.0552
XGB & RF w/ SVM	60.04	8.49	0.0559
XGB & RF w/ Ridge	56.52	7.68	0.0553

as Linear Regression, SVM, and Ridge Regression, have demonstrated their effectiveness in this context.

In both ‘euro28‘ and ‘us26‘ topologies, the stacking methods with Linear Regression and Ridge Regression as meta-learners consistently outperformed the individual models (XGBoost, RandomForest, AdaBoost) in terms of Mean RMSE. This indicates a superior predictive accuracy, highlighting the strength of the stacking approach in integrating diverse learning algorithms.

However, the stacking model with SVM as a meta-learner tended towards higher RMSE values, suggesting that the choice of meta-learner is crucial in optimizing the stacking model's performance. This variance in performance underscores the need for careful selection of the meta-learner in stacking methods.

The Mean MAPE values across all models were comparable, indicating a consistent level of percentage error across different models and network topologies. This consistency is essential for reliable network performance predictions, especially in varied network scenarios.

Moreover, the analysis reveals a trade-off between computational efficiency and predictive accuracy. While the stacking methods, especially those involving complex meta-learners, demand more computational resources, their enhanced accuracy and robustness make them a valuable tool in scenarios where predictive precision is paramount.

Overall, the results show potential for the use of stacking methods in machine learning tasks within the domain of RSA optimization in EONs. Their ability to outperform single models and provide more accurate predictions makes them particularly suitable for complex network environments. Future research could focus on further optimizing these stacking methods, exploring different combinations of base learners and meta-learners to enhance both efficiency and accuracy.

Additionally, the analysis of computational efficiency reveals a noticeable variation among the models. While ensemble methods, especially those involving stacking techniques, generally required more computational resources, the increase in computation time was balanced by their improved accuracy and robustness. On the other hand, simpler models like AdaBoost, while being computationally more efficient, did not achieve the same level of accuracy as the more complex models.

IX. Conclusions

This study demonstrates the efficacy of various machine learning models in optimizing Routing and Spectrum Allocation (RSA) in Elastic Optical Networks (EONs). Ensemble methods, particularly stacking with linear and ridge regression as meta-learners, emerged as superior in their predictive capabilities. The RandomForest model, with its consistent performance and relatively efficient computation, proved to be a strong contender, particularly in terms of Mean Absolute Percentage Error (MAPE).

The findings also highlight the trade-off between computational efficiency and model accuracy. While the ensemble methods required more computational resources, their enhanced accuracy and robustness justified the additional computational expense. This observation is crucial for applications where the balance between computational efficiency and predictive accuracy is a key consideration.

Overall, the study advocates for the strategic use of ensemble learning techniques in predicting network performance metrics in EONs. Future research could explore further optimization of these models, potentially enhancing their efficiency and accuracy. The study's implications extend to the design and operation of EONs, where machine learning can significantly contribute to network performance optimization.

Acknowledgment

The authors would like to thank...

References

[1] Cisco, "Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2015–2020," Feb. 2016.

[2] D. Sharma et al., "An overview of elastic optical networks and its enabling technologies," International Journal of Engineering and Technology, vol. 9, no. 3, pp. 1643–1649, 2017.

[3] Mukherjee D et al., "Core Netw Part B," in Handbook of Optical Networks, Springer, p. 395.

[4] J. Mata et al., "Artificial intelligence (AI) methods in optical networks: A comprehensive survey," Optical Switching and Networking, vol. 28, pp. 43–57, 2018.

[5] J. Yu et al., "A deep learning based RSA strategy for elastic optical networks," in 2019 18th International Conference on Optical Communications and Networks (ICOCN), pp. 1–3, 2019.

[6] M. Salani et al., "Routing and spectrum assignment integrating machine-learning-based QoT estimation in elastic optical networks," in IEEE INFOCOM 2019 - IEEE Conference on Computer Communications, pp. 1738–1746, 2019.

[7] X. Chen et al., "DeepRMSA: A Deep Reinforcement Learning Framework for Routing, Modulation and Spectrum Assignment in Elastic Optical Networks," 2019.

[8] Kukačka, Jan, Vladimir Golkov, and Daniel Cremers. "Regularization for deep learning: A taxonomy." arXiv preprint arXiv:1710.10686 (2017).

[9] Musumeci, Francesco, et al. "An overview on application of machine learning techniques in optical networks." IEEE Communications Surveys & Tutorials 21.2 (2018): 1383-1408.

[10] Natekin, Alexey, and Alois Knoll. "Gradient boosting machines, a tutorial." Frontiers in neurorobotics 7 (2013): 21.

[11] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016, pp. 785–794.

[12] Pavlyshenko, Bohdan. "Using stacking approaches for machine learning models." 2018 IEEE Second International Conference on Data Stream Mining & Processing (DSMP). IEEE, 2018.

[13] Lu, Mingshen, et al. "A Stacking Ensemble Model of Various Machine Learning Models for Daily Runoff Forecasting." Water 15.7 (2023): 1265.

[14] Noble, William S. "What is a support vector machine?." Nature Biotechnology 24.12 (2006): 1565-1567.

[15] J. H. Friedman, "Stochastic gradient boosting," Computational Statistics & Data Analysis, vol. 38, no. 4, pp. 367–378, 2002.