Journal of Computing & Biomedical Informatics
ISSN: 2710 - 1606
Volume 09 Issue 02
2025
Research Article
https://doi.org/10.56979/902/2025

Modeling BBR v3 Congestion Control Behavior Using Supervised ML Techniques

Muhammad Ahsan[1*], Ashfaq Ahmad[2], Raybal Akhtar[3], and Iqra Javed[4]

[1]Department of Software Engineering, University of Management and Technology, Lahore, Punjab, Pakistan.
[2]Department of Artificial Intelligence, University of Management and Technology, Lahore, Punjab, Pakistan.
[3]Department of Informatics and Systems, University of Management and Technology, Lahore, Punjab, Pakistan.
[4]Department of Computer Science, University of Management and Technology, Lahore, Punjab, Pakistan.
[*]Corresponding author: Muhammad Ahsan. Email: mahsan@alumni.usc.edu

Received: June 29, 2025 **Accepted:** August 23, 2025

Abstract: BBR v3 has recently emerged as the most sophisticated model-based congestion control. It measures different network parameters such as bottleneck bandwidth, round-trip time (RTT), packet loss rate, and explicit congestion notification (ECN) to get the true picture of the available network bandwidth and then sets its pacing rate. The goal is to operate near Kleinrock's optimal operating point to prevent excessive queue formation in case of large buffers and to prevent overreacting in case of shallow buffers. However, there are still limitations that exist in properly setting up the pacing rate that matches the delivery rate at the receiver's side. BBR v3 generally sets the pacing rate relatively high, as the pacing gain values are generally fixed in its probing for the bandwidth phase. In this paper, we have evaluated this issue using machine learning regression algorithms such as XGBoost, Random Forest, Neural Networks, Linear Regression, Support Vector Regression, Gradient Boosting, and Decision Tree by training these models and predicting its pacing rate. The machine learning (ML) regression models are then evaluated using various metrics such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R^2 score. Our results show that Linear Regression and XGBoost provide the best results in terms of lowest error and superior predictability of BBR v3's pacing rate.

Keywords: BBR; Congestion Control; XGBoost; Neural Network; MAE; RMSE; R^2

1. Introduction

The network's complexity is increasing with the passage of time, and more bandwidth is now available for it. This needs a rapid change in congestion control mechanisms, also, that can gauge this huge amount of bandwidth that is being made available by technologies like 5G, wired, and wireless broadband. The links are now more stable, especially in the case of wired scenarios, and the error rates have gone down. The traditional loss-based congestion control algorithms are not coping with this increased bandwidth, and they exhibit delays and bufferbloat issues. TCP BBR (Bottleneck Bandwidth and RTT), developed by Google [1], represents a significant shift in congestion control philosophy by estimating bottleneck bandwidth and round-trip time to regulate packet pacing. BBR v3 [2] is the latest version that has introduced refined mechanisms for probing, pacing, and fairness, making it a promising candidate to be used as the default congestion control for servers as well as end systems.

BBR which is model based, so it tries to build a model for the network pipe to gauge the available bandwidth in the shortest amount of time. It then controls its pacing rate of the packets being sent by the client side into the network pipe to ensure the data in-flight is optimum to keep the pipe just full but no fuller. It has always been difficult to model this pacing behavior analytically due to its dynamic response to varying network conditions. It gets more challenging in wireless environments, where variability in link quality and contention can significantly affect pacing decisions. To tackle this, data-driven approaches such as machine learning offer a compelling alternative for modeling and predicting pacing rate behavior.

In this study, we investigate the use of supervised machine learning techniques to predict the pacing rate of TCP BBR v3. The dataset was collected from a real-time physical testbed based on Wi-Fi 4 wireless technology, configured with TCP BBR v3 and four concurrent upload streams. To reduce noise and isolate stream-specific behavior, only metrics from a single stream were used for training and evaluation.

We apply and compare several regression models, including Linear Regression, Decision Tree, Random Forest [3], Gradient Boosting, Support Vector Regression, Neural Network [4], and XGBoost [5] to assess their effectiveness in predicting pacing rate. The models are evaluated using standard metrics such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Coefficient of Determination (R^2). Our results demonstrate that ensemble and neural models outperform traditional approaches, offering accurate and interpretable insights into BBR v3's pacing dynamics.

The contribution of this paper is that it predicts an important functionality of BBR v3, which is the pacing rate. The rate at which it sends packets into the network. This is dependent on pacing gain as well as the current bottleneck bandwidth. We have tried to predict this pacing rate using different machine learning regression techniques and shared which techniques provide the least error and superior predictability. This paper highlights how machine learning can be used with model-based congestion controls to predict their pacing rate.

2. Related Literature

BBR has been in continuous development since its release in 2016. The scientific community is working meticulously in analyzing its fairness with loss-based congestion controls such as TCP Reno and TCP CUBIC [6, 7]. BBR v3, the latest iteration, introduces refined pacing and probing mechanisms aimed at improving throughput and latency in diverse network environments. Several studies have explored the behavior of BBR in wired and wireless networks [8-11], but very little work has been done in exploring congestion controls with machine learning. Google's original BBR papers [12] laid the foundation for bandwidth and RTT-based control, while subsequent works have examined its fairness, responsiveness, and interactions with other flows [13]. More recent efforts have focused on BBR v2 and v3, analyzing their performance under varying queue disciplines, buffer sizes, and link conditions [9].

Machine learning [14] has increasingly been applied to networking problems, including traffic classification, anomaly detection, and congestion control modeling. Prior work has demonstrated the utility of supervised learning for predicting network metrics such as RTT, throughput, and queue occupancy. Ensemble models like Random Forest and XGBoost have shown promise due to their robustness and interpretability.

However, to date, there is no work done that has applied machine learning specifically to model pacing rate behavior in BBR v3. This study addresses that gap by leveraging stream-specific metrics from a real-time wireless testbed based on Wi-Fi 4 to generate the dataset on which different regression models are trained for their evaluation in predicting the pacing rate for BBR v3. The focus on stream isolation and pacing rate prediction offers a novel perspective on understanding and optimizing BBR's behavior in practical deployments.

3. Methodology

The dataset used in this study was collected from a real-time physical testbed based on Wi-Fi 4 wireless technology using Flent [15]. The testbed was configured to run TCP BBR v3 with four concurrent upload streams. To ensure precision and reduce inter-stream noise, only metrics from a single stream were selected for modeling. This allowed for a more focused analysis of pacing behavior in BBR v3.

From the collected dataset, features corresponding to stream 1 were extracted. These included pacing gain, bandwidth estimation, delivery rate, RTT, congestion window size, and other TCP metrics. Columns prefixed with TCP upload::1:: were used, excluding the target variable TCP upload::1::tcp_pacing_rate.

The dataset was cleaned by removing rows and columns with missing values. Features were standardized using z-score normalization to ensure uniform scaling across models, particularly for Support Vector Regression and Neural Network models. Seven regression models were trained and evaluated. Each model was trained using an 80/20 train-test split. Hyperparameters were kept at default settings for baseline comparison, with random seeds fixed for reproducibility. Model performance was assessed using metrics Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Coefficient of Determination (R^2), and Relative Absolute Error (RAE). Feature importance was extracted from the Linear Regression model to identify the most influential predictors of pacing rate. Metrics such as bandwidth estimation (tcp_bbr_bw), delivery rate, and RTT were found to be dominant.

4. Results and discussion

To evaluate the effectiveness of various regression models in predicting the pacing rate of BBR v3, we trained and tested seven widely used techniques using stream 1 features extracted from a multi-stream TCP upload dataset. The models include Linear Regression, Decision Tree, Random Forest, Gradient Boosting, Support Vector Regression (SVR), Neural Network (MLP), and XGBoost. Figs 1- 3 show the average magnitude of errors through MAE, RMSE chart that penalizes large errors more heavily, and the coefficient of determination (R^2) that indicates how well the model explains variance in pacing rate. Feature importance analysis from the linear regression model, Fig 4, revealed that metrics such as tcp_bbr_pacing_gain, tcp_bbr_mrtt, tcp_bbr_cwnd_gain, were among the most influential in predicting pacing rate. This highlights the relevance of bandwidth estimation and delay metrics in BBR v3's pacing behavior. The Linear regression model explains 98 % of the variance in the pacing rate data Fig 5, which is considered extremely accurate for regression tasks.

Finally, the pacing rate prediction for each regression model is shown in Figs. 6 - 12.

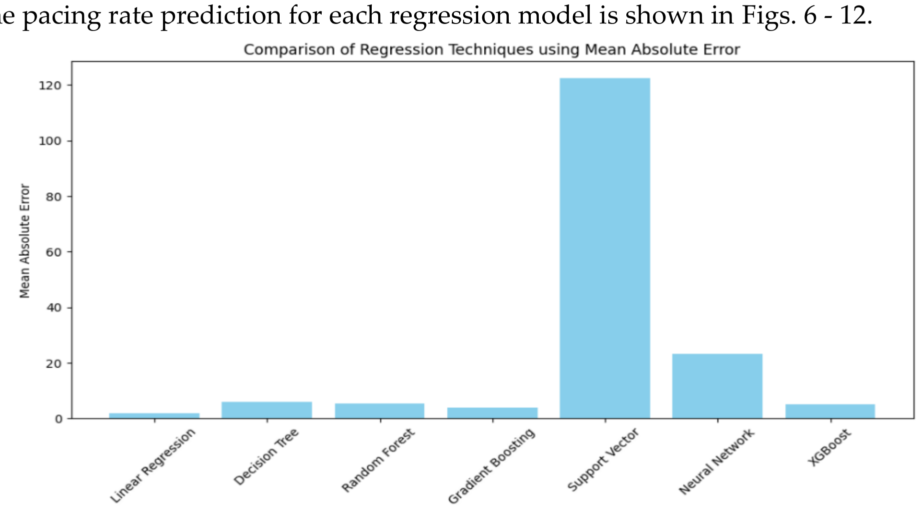

Figure 1. Mean Absolute Error

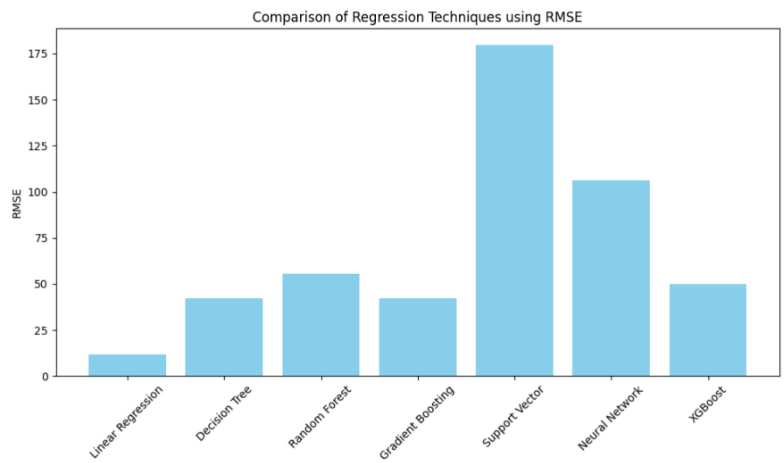

Figure 2. Root Mean Squared Error

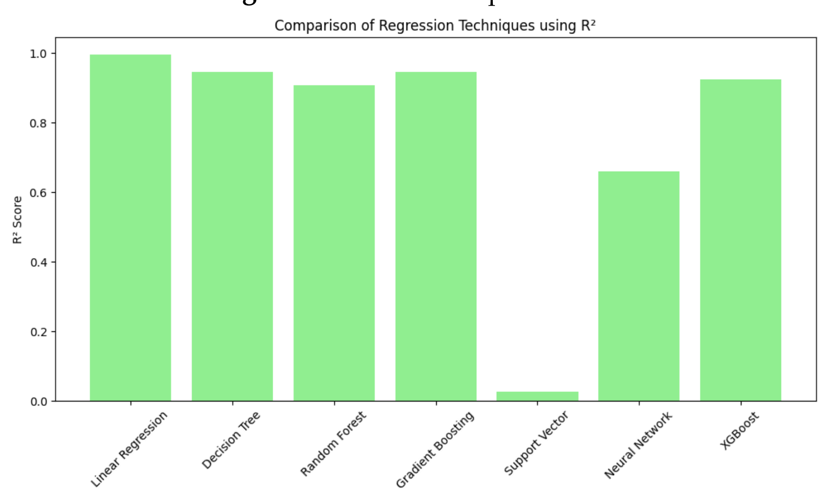

Figure 3. Coefficient of Determination

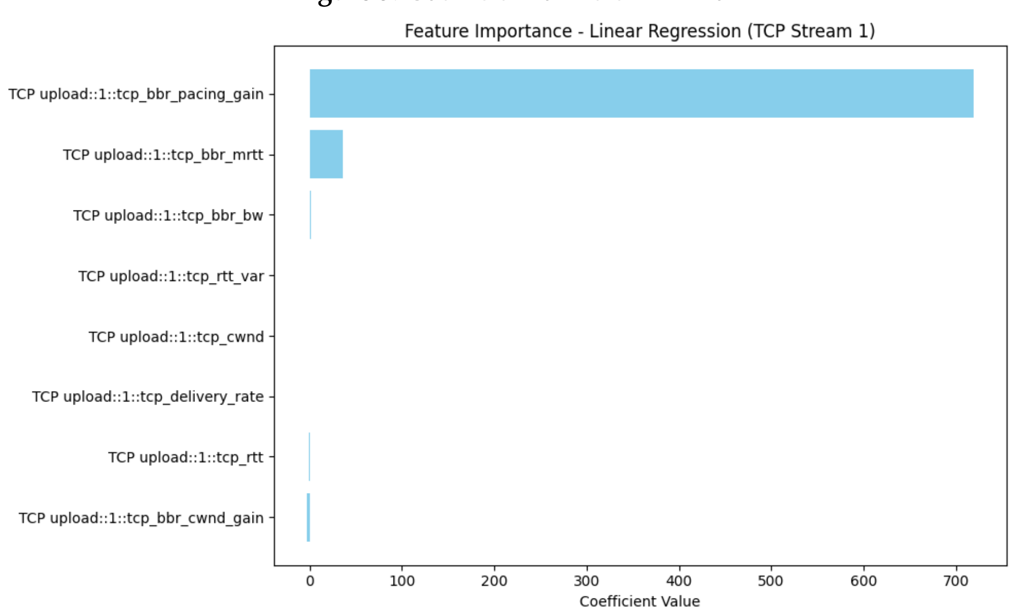

Figure 4. Linear Regression model trained to predict TCP pacing rate for stream 1

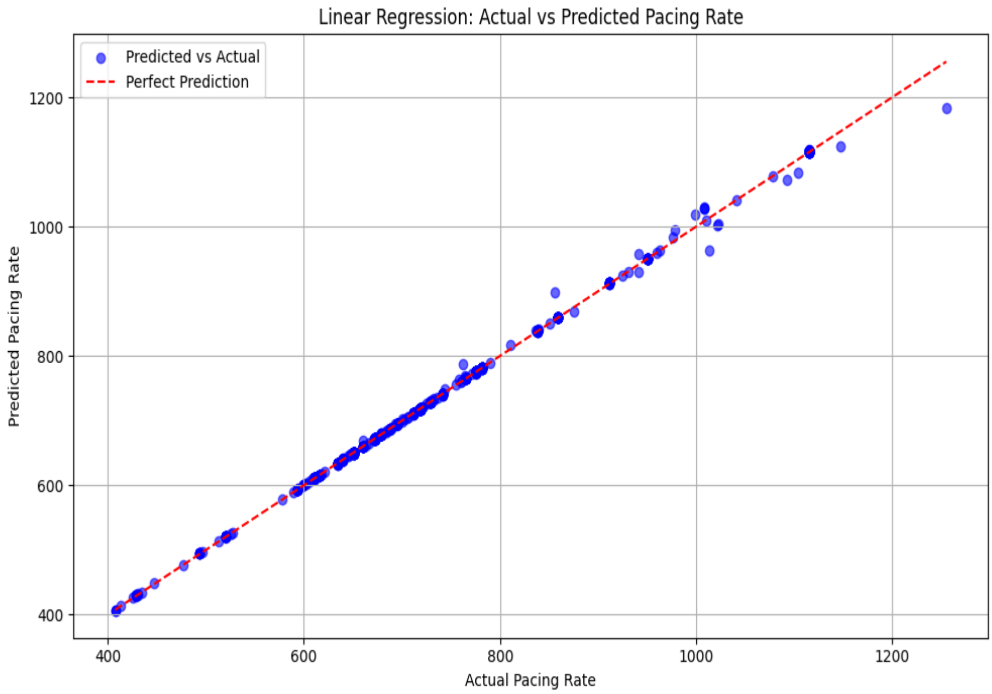

Figure 5. Linear Regression with 98 % accuracy

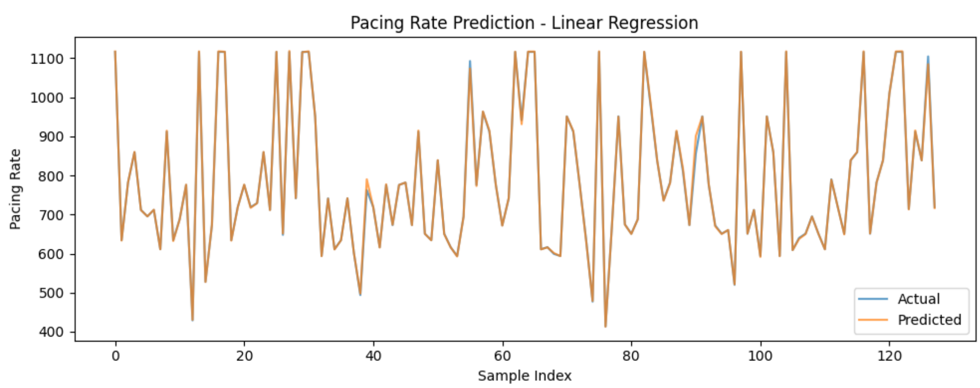

Figure 6. Linear Regression

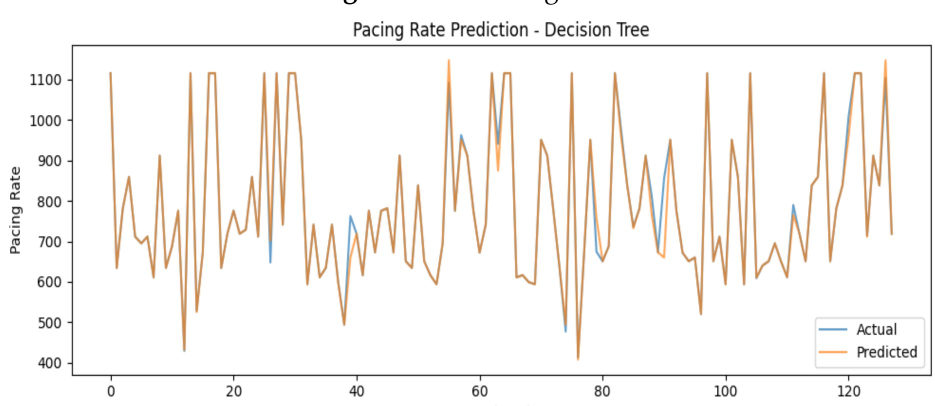

Figure 7. Decision Tree Prediction

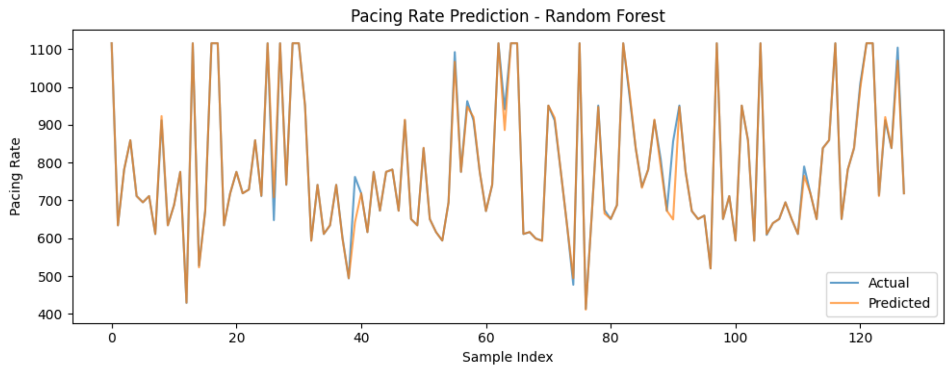

Figure 8. Random Forest Prediction

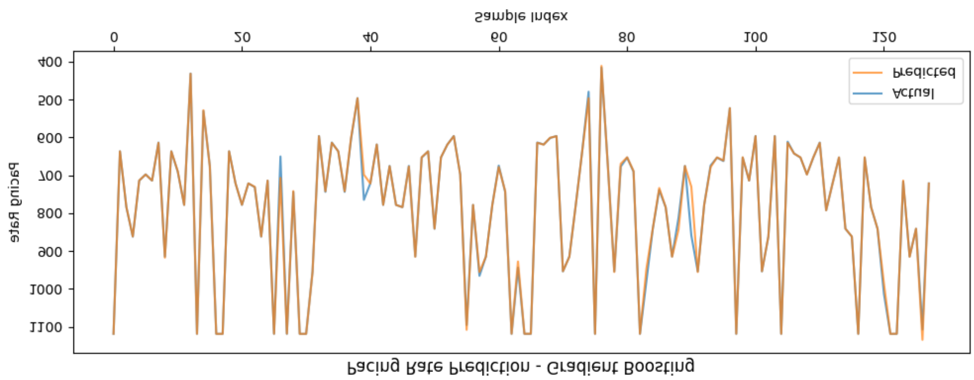

Figure 9. Gradient Boosting Prediction

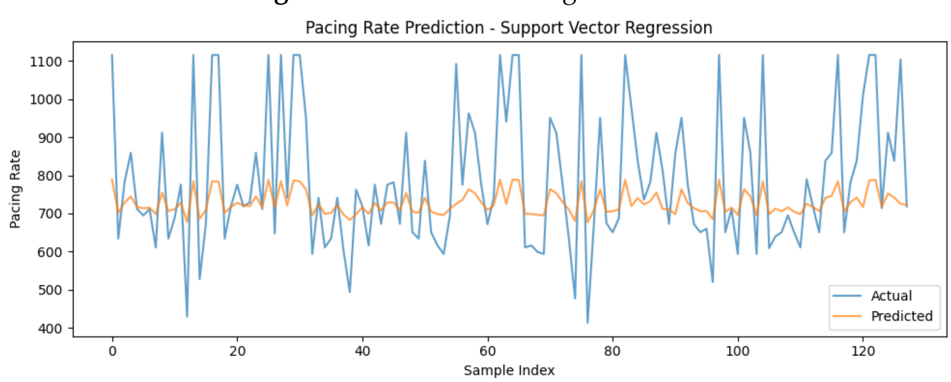

Figure 10. Support Vector Prediction

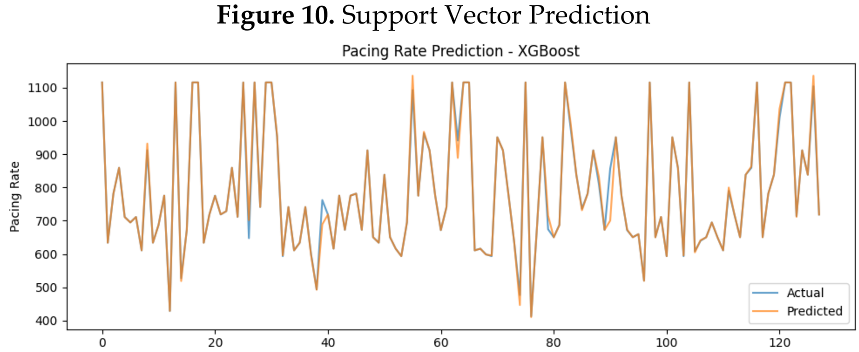

Figure 11. XGBoost Prediction

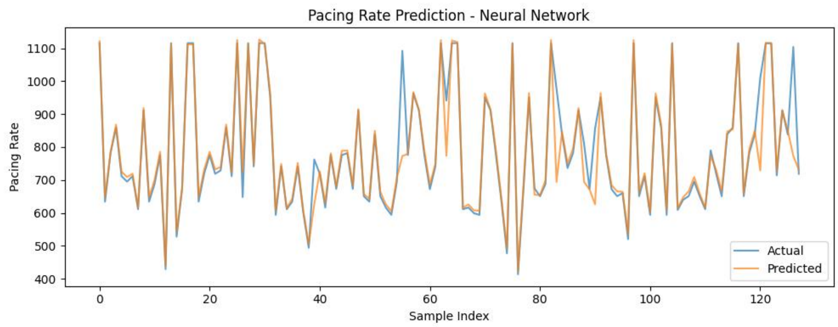

Figure 12. Neural Network Prediction

4.1. Results Summary

Table 1. Evaluation results of various models

Model	MAE	RMSE	R^2
Linear Regression	2	10	0.98
Decision Tree	10	40	0.92
Random Forest	10	55	0.93
Gradient Boosting	2	35	0.95
Support Vector	120	175	0.02
Neural Network (MLP)	30	90	0.65
XGBoost	2	30	0.92

Table 1 shows the results for MAE, RMSE, and R^2 analysis, with Linear Regression leading and XGBoost taking the second spot. From Figs 1-3, we see that linear regression achieved the lowest MAE and RMSE, indicating superior accuracy and robustness in predicting pacing rate. The Linear Regression also showed the highest R^2 score, suggesting strong generalization and the ability to capture non-linear relationships. Random Forest demonstrated consistently strong performance across all metrics, with a good balance between accuracy and interpretability. Its feature importance analysis revealed key pacing-related metrics such as bandwidth estimation (tcp_bbr_bw) and delivery rate as dominant predictors. Gradient Boosting closely followed Random Forest and XGBoost, offering competitive performance with slightly higher computational cost. Linear Regression, simple and interpretable, outperformed ensemble and non-linear models, indicating that pacing rate prediction benefits from simple rather than complex modeling. Decision Tree and Support Vector Regression offered moderate performance, with Decision Tree outperforming SVR and Neural Networks in terms of MAE and RMSE.

5. Conclusion

In this paper, we have presented a machine learning-based approach that can be used to predict pacing rate for BBR v3. The collected four stream datasets for Wi-Fi 4 based traffic in the uplink were trained using both traditional as well as modern machine learning regression models. In order to minimize noise due to redundancy, a single stream was extracted from the dataset. Multiple regression techniques were applied, including ensemble methods and neural networks. Among these, Linear Regression and XGBoost Regression demonstrated superior performance, achieving the lowest Mean Absolute Error (MAE) and highest R^2 scores. The results confirm that pacing rate in BBR v3 can be accurately predicted using stream-specific features, with simple as well as ensemble models offering both precision and interpretability.

This work highlights the potential of data-driven modeling in understanding and optimizing congestion control algorithms. Future research may explore cross-stream interactions, time-series modeling, and real-time adaptive pacing strategies using online learning techniques.

References

1. Cardwell N, Cheng Y, Gunn CS, Yeganeh SH, Jacobson V. Bbr: Congestion-based congestion control: Measuring bottleneck bandwidth and round-trip propagation time. Queue. 2016;14(5):20-53.
2. Cardwell N, Cheng Y, Yang K, Morley D, Yeganeh SH, Jha P, et al. BBRv3: Algorithm bug fixes and public internet deployment. Proceedings of the IETF-117 Internet Engineering Task Force, San Francisco, CA, USA. 2023;26.
3. Mubashar R, Awan MJ, Ahsan M, Yasin A, Singh VP. Efficient residential load forecasting using deep learning approach. International Journal of Computer Applications in Technology. 2022;68(3):205-14.
4. Fahs W, Chbib F, Rammal A, Khatoun R, El Attar A, Zaytoun I, et al. Traffic congestion prediction based on multivariate modelling and neural networks regressions. Procedia Computer Science. 2023;220:202-9.
5. Cherif IL, Kortebi A, editors. On using extreme gradient boosting (XGBoost) machine learning algorithm for home network traffic classification. 2019 Wireless Days (WD); 2019: IEEE.
6. Ahsan M, Awan MJ, Yasin A, Bahaj SA, Shehzad HMF. Performace evaluation of TCP cubic, compound TCP and NewReno under Windows 20H1, via 802.11 n Link to LTE Core Network. Annals of the Romanian Society for Cell Biology. 2021;25(6):5357-69.
7. Ha S, Rhee I, Xu L. CUBIC: a new TCP-friendly high-speed TCP variant. ACM SIGOPS operating systems review. 2008;42(5):64-74.
8. Ahsan M, Muhammad SS. TCP BBR-n: Increased throughput for wireless-AC networks. Plos one. 2023;18(12):e0295576.
9. Ahsan M, Muhammad SS. TCP BBR-n interplay with modern AQM in Wireless-N/AC networks: Quest for the golden pair. Plos one. 2024;19(9):e0304609.
10. Ahsan M, Nabeel M, Ahmad A, Javed I, Ashiq W. A Real-World Testbed Evaluation of BBR-n+ with Other Major TCP Congestion Control Algorithms. Journal of Computing & Biomedical Informatics. 2025.
11. Ahsan M, Nabeel M, Khan MMA, Hasan MZ, Ashiq W. Performance Evaluation of BBR-v3 with Cubic and Reno in a Ubiquitous Wired/Wi-Fi Channel. Journal of Computing & Biomedical Informatics. 2025.
12. Cardwell N, Cheng Y, Gunn CS, Yeganeh SH, Jacobson V. BBR: Congestion-based congestion control. Communications of the ACM. 2017;60(2):58-66.
13. Xu J, Pan W, Tan H, Cheng L, Li X, Li X. A Modified TCP BBR to Enable High Fairness in High-Speed Wireless Networks. Future Internet. 2024;16(11):392.
14. Jiang H, Li Q, Jiang Y, Shen G, Sinnott R, Tian C, et al. When machine learning meets congestion control: A survey and comparison. Computer Networks. 2021;192:108033.
15. Høiland-Jørgensen T, Grazia CA, Hurtig P, Brunstrom A, editors. Flent: The flexible network tester. Proceedings of the 11th EAI International Conference on Performance Evaluation Methodologies and Tools; 2017.