

Avaliação de Algoritmos de Aprendizado de Máquina para Predição de QoE em Redes 6G

Felipe S. Dantas Silva^{1,2}, Mathews Lima^{1,2}, Charles H. F. Santos^{1,2}, Augusto Neto²

¹LaTARC Research Lab
Instituto Federal de Educação, Ciência e Tecnologia do Rio Grande do Norte (IFRN)
Natal-RN, Brasil

²Universidade Federal do Rio Grande do Norte (UFRN)
Natal-RN, Brasil

Abstract. *This paper evaluates the performance of Machine Learning (ML) algorithms to provide Quality of Experience (QoE) user decisions in multimedia services. A specialized dataset was built for mapping network Quality of Service (QoS) Key Performance Indicators (KPIs) with video quality metrics to perform QoE predictions. An evaluation is then carried out considering the main regression models to support the development of QoE-aware systems and meet the critical requirements of applications characteristic of 6G scenarios.*

1. Introdução

As prospecções sobre o desenvolvimento e operação do ecossistema da sexta geração (6G, do inglês *Sixth Generation*) de redes móveis indicam um significativo progresso sobre os padrões de cobertura e desempenho da rede, bem como na performance das aplicações subjacentes [Wang et al. 2023]. Para prover suporte a tais perspectivas de evolução, modelos de aprendizado de máquina (ML, do inglês *Machine Learning*) necessitarão ser amplamente explorados para prover uma alta qualidade de experiência ao usuário. Ao permitir melhores capacidades de predição com ML ubíqua é possível, por exemplo, adaptar melhor os recursos da rede as condições de densidade de tráfego e de usuários no contexto da crescente demanda de serviços especializados, à exemplo daqueles baseados no fatiamento de recursos de rede (*network slicing*) [Wu et al. 2022].

Um aspecto importante em direção as redes inteligentes orientadas à qualidade se refere à predição de métricas de Qualidade de Experiência (QoE, do inglês *Quality of Experience*) dos usuários finais. Em aplicações multimídia, por exemplo, é possível determinar a performance do desempenho de uma transmissão por meio das métricas objetivas de qualidade de vídeo (VQM, do inglês *Video Quality Metrics*), tal como o método de similaridade estrutural (SSIM, do inglês *Structural Similarity*). Entretanto, a extração de métricas em tempo real não é viável devido à necessidade da existência de vídeos de referência, além do alto custo computacional do processo de extração. Para tratar tal desafio, diversos estudos na literatura buscaram estabelecer o relacionamento do QoE com indicadores chave de desempenho (KPI, do inglês *Key Performance Indicator*) baseados nas estatísticas de Qualidade de Serviço (QoS, do inglês *Quality of Service*).

Embora estes estudos sejam importantes para estabelecer um mapeamento entre aspectos de QoS/QoE, eles se encontram datados devido ao avanço da qualidade das mídias, materializado pelo aumento de resolução e taxa de quadros por segundo. Além

disso, no melhor do nosso conhecimento, e confirmado em levantamento recente da literatura, há uma carência de estudos sobre a aplicação de modelos de ML para predição dos níveis de qualidade de vídeos quando submetidos a diferentes condições de rede. Somente através de um estudo desta natureza será possível estabelecer novos critérios para o desenvolvimento de sistemas inteligentes capazes de suportar as especificidades de casos de uso de missão crítica demandados em redes 5G e além do 5G (B5G, do inglês *Beyond 5G*) [Silva et al. 2022].

Neste trabalho oferecemos uma avaliação de desempenho de algoritmos de ML para suportar a aplicação de QoE fim-a-fim em cenários 6G. A avaliação leva em consideração as principais métricas de avaliação de performance empregadas em problemas de regressão. Adicionalmente, são apresentados os desempenhos individuais de cada algoritmo em termos dos tempos de treinamento e de predição. As contribuições fornecidas no presente estudo pavimentam o desenvolvimento de novos mecanismos para o gerenciamento de redes 6G ao promover, de forma inteligente, decisões de controle orientadas à satisfação do usuário das aplicações envolvidas. Como resultado, os blocos estruturais 6G poderão se beneficiar de novas habilidades para tomada de decisões que vão além das capacidades estáticas atualmente fornecidas pelo estado da arte.

1.1. Organização do artigo

O restante deste artigo está organizado da seguinte maneira: a Seção 2 descreve abordagens de trabalhos relacionadas recentes e faz uma comparação com a proposta apresentada pelo presente estudo. A Seção 3 apresenta a metodologia empregada na avaliação. A Seção 4 apresenta os resultados em termos de performance dos modelos de ML. Finalmente, a Seção 5 apresenta as conclusões do artigo e faz apontamentos para trabalhos futuros.

2. Trabalhos Relacionados

Levantamento recente da literatura constatou a existência de diversos trabalhos anteriores comprometidos em prover comparativos entre algoritmos de ML em vários domínios de problema em redes 5G e B5G. Os trabalhos existentes empregam distintas métricas de avaliação de performance com o objetivo de expor as capacidades de cada uma das técnicas de aprendizado sob variadas situações demandadas por cada domínio de problema sob investigação. A Tabela 1 sumariza alguns dos principais estudos em termos do (i) domínio do problema sob investigação, (ii) algoritmos de ML empregados nas avaliações, (iii) métricas de desempenho adotadas, e (iv) perspectivas das capacidades em termos do tempo de treinamento e do tempo de predição.

Conforme apresentado na Tabela 1, nenhum dos estudos anteriores oferece uma avaliação de desempenho dos algoritmos de ML com perspectivas à predição de QoE através de métricas objetivas. Desta maneira, este trabalho promove avanço no estado da arte da predição de QoE em sistemas B5G ao oferecer uma avaliação de performance de algoritmos de ML compatível com os cenários típicos de sistemas 6G.

3. Metodologia

Esta seção descreve a metodologia empregada para condução dos experimentos, incluindo a geração da base de dados (*dataset*) com KPIs de QoS e os indicadores de QoE.

Tabela 1. Comparativo entre trabalhos existentes e que realizam avaliação de algoritmos de ML através de métricas de desempenho.

Proposta	Domínio do problema	Algoritmos	Métricas de desempenho	Tempo de treinamento/predição
[Gebrie et al. 2019]	Predição de mobilidade baseada em economia de energia	XGBoost, SVM, DNN e Semi-Markov	Precisão	✓
[Jardim et al. 2021]	Encadeamento de funções de rede em aplicações e-health	MLP, LSTM e CNN	MSE, MAE, RMSE, MedAE	✓
[Sharma et al. 2021]	Predição de vazão	Random Forest e SVM	RMSE, precision, recall, F1, R ²	✗
[Khan et al. 2021]	Localização de UEs	KNN, Logistic Regression, Random Forest e Decision Tree	Matriz de confusão	✗
[Liu and Huang 2023]	Predição de sinal em ondas milimétricas	GRU e LSTM	MAE, RMSE e MAPE	✗
Este trabalho	Predição objetiva de QoE em transmissões multimídia	Gradient Boosting, XGBoost, Linear Regression, SVM, Random Forest, Kernel Ridge, Bayesian Ridge, Stochastic Gradient, LightGBM e ANN	MAE, MSE, RMSE, R ² Score	✓

3.1. Criação do *dataset*

O objetivo do processo de extração de dados é obter o mapeamento entre os valores médios de QoS e a qualidade do vídeo resultante para um determinado intervalo de tempo. O *dataset* foi gerado em uma infraestrutura de rede emulada sob a plataforma Mininet-WiFi¹. O ambiente emulado foi configurado em um desktop Dell Optiplex 7080 com sistema operacional Linux Ubuntu 20.04, Intel Core i7-10700T CPU e 32GB de memória RAM.

A Tabela 2 detalha as características do *dataset*, considerando quatro execuções (i.e. medições de SSIM) por instância. Os parâmetros de QoS considerados foram: (i) atraso (*delay*), (ii) variação do atraso (*jitter*) e (iii) perda de pacotes (*packet loss*). A variação dos parâmetros de QoS na rede foi realizada através da manipulação de configurações de controle de tráfego com o utilitário tc^2 , com o objetivo de submeter as transmissões à diferentes condições de tráfego.

Tabela 2. Descrição do *dataset* utilizado para o mapeamento entre os parâmetros de QoS fim-a-fim e seu respectivo valor de QoE resultante.

Parâmetros de QoS			Parâmetros de vídeo			Classe do <i>dataset</i>
Delay (ms)	Jitter (ms)	Loss (%)	Taxa de bits (Mbps)	Taxa de quadros (fps)	Resolução (px)	SSIM resultante
[0 - 30]	[0 - 6]	[0 - 3]	[4 - 8]	[24 - 50]	[1280x720 - 1920x1080]	[0 - 1]

Os seguintes intervalos foram aplicados para cada uma das KPIs de QoS: 0ms e 30ms para *delay*, com variação de 1ms em cada instância; 0ms e 6ms para *jitter*, com variação de 0.5ms, e 0% e 3% de *packet loss*, com variação em 0.2%. Cinco sequências de vídeos com resolução Full HD (1920x1080 pixels) e HD (1280x720 pixels) foram utilizadas para se obter o SSIM, variando a taxa de bits entre 4Mbps e 8Mbps, e com taxa de quadros entre 24fps e 50fps. Para cada instância, o valor de SSIM foi calculado utilizando a média interquartil entre os valores resultantes dos quatro experimentos.

3.2. Execução dos experimentos

Uma vez constituído, o *dataset* (originalmente composto por 12.815 instâncias) foi pré-processado com a finalidade de identificar relações incoerentes entre os atributos que

¹<https://mininet-wifi.github.io/>

²<https://man7.org/linux/man-pages/man8/tc.8.html>

pu dessem afetar o desempenho dos preditores. Para tanto, considerou-se um limiar baseado no erro absoluto médio (MAE, do inglês *Mean Absolute Error*) global igual ou maior a 0.02 (ou seja, valores de MAE a partir de 0.02 foram excluídos da amostragem). Após a etapa de pre-processamento o *dataset* resultante consistiu em 60.5% (i.e., 7.753 instâncias) do conjunto de dados original. A significativa redução na quantidade dos dados foi necessária para garantir maior consistência nas predições.

Os dados pré-processados foram treinados utilizando os principais algoritmos de aprendizado supervisionado por regressão. Para isso, um conjunto de 10 algoritmos de ML foram implementados em linguagem Python3 com o uso das bibliotecas scikit-learn³, nomeadamente (i) Gradient Boosting, (ii) XGBoost, (iii) Linear Regression, (iv) Support Vector Machine (SVM), (v) Random Forest, (vi) Kernel Ridge, (vii) Bayesian Ridge, (viii) Stochastic Gradient Descent, (ix) LightGBM e (x) Artificial Neural Networks (ANN). Tais preditores foram avaliados de acordo com as métricas comuns a problemas de regressão, como o MAE, erro quadrático médio (MSE, do inglês *Mean Squared Error*), raiz quadrada do erro quadrático médio (RMSE, do inglês *Root Mean Squared Error*) e coeficiente de determinação (R^2 score) e em relação aos tempos de treinamento e predição.

Os experimentos foram conduzidos levando em consideração um total de 70% do *dataset* para treinamento e 30% para testes. Cada algoritmo de ML foi executado 20 vezes e, ao final das execuções foi calculada a média interquartil entre os valores resultantes. A próxima seção apresenta detalhes dos resultados obtidos nas experimentações.

4. Resultados

Nesta seção são apresentados os resultados das avaliações sob a perspectiva das diferentes técnicas de desempenho adotadas. A Figuras 1 e 2 apresentam os resultados das avaliações dos métodos em termos das métricas MAE, MSE, RMSE e R^2 score, respectivamente.

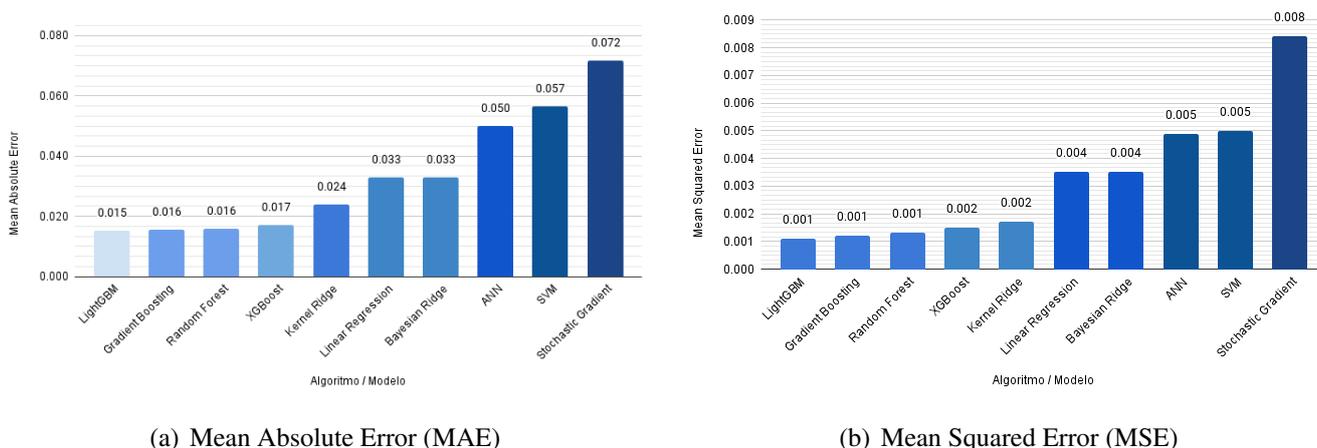
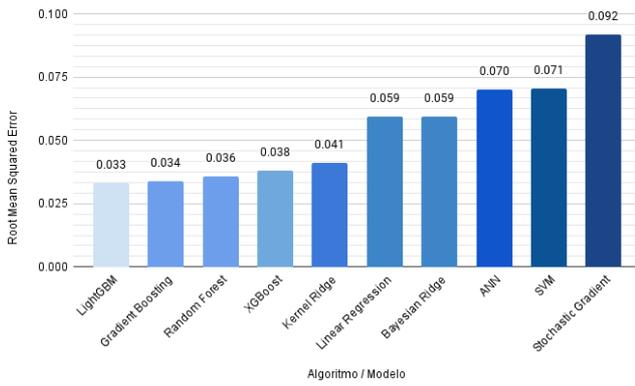


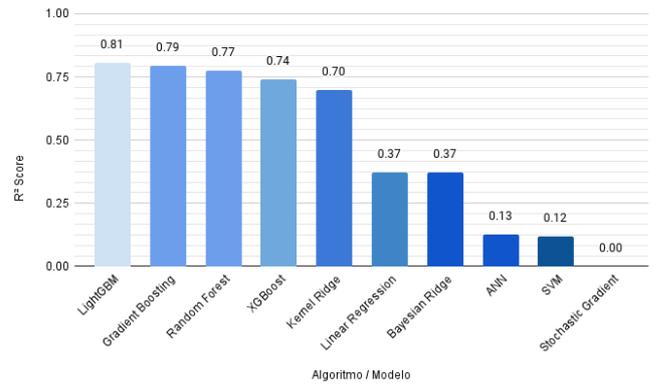
Figura 1. Resultados das avaliações: MAE e MSE.

Em termos gerais, de acordo com as Figuras 1 e 2, o algoritmo LightGBM apresentou melhores habilidades de predição, seguido pelo Gradient Boosting, Random Forest

³<https://scikit-learn.org/>

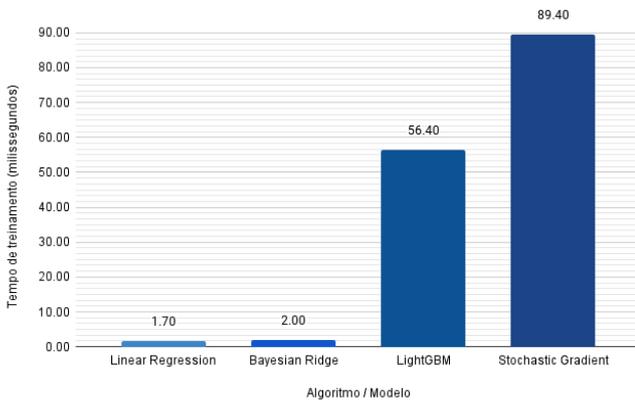


(a) Root Mean Squared Error (RMSE)

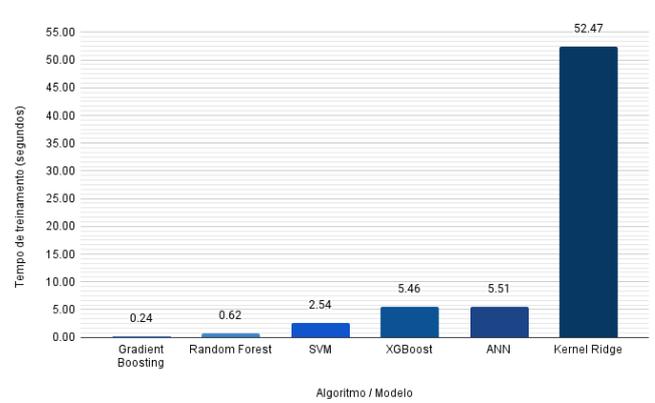


(b) R² Score

Figura 2. Resultados das avaliações: RMSE e R² Score.

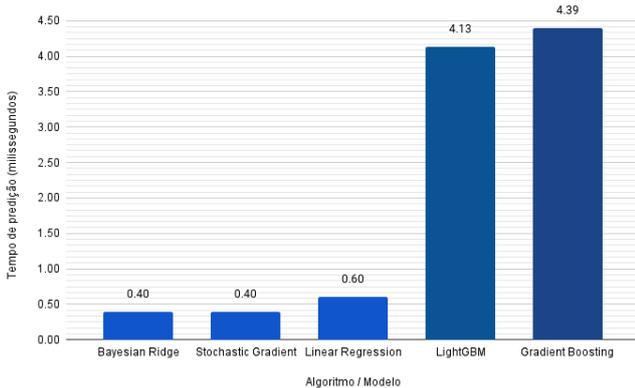


(a) Tempo de treinamento dos algoritmos até 90ms

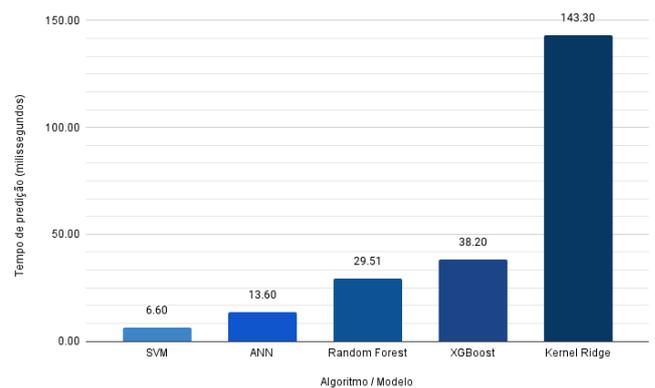


(b) Tempo de treinamento dos algoritmos a partir de 0.24s

Figura 3. Tempos de treinamento.



(a) Tempo de predição dos algoritmos até 5ms



(b) Tempo de predição dos algoritmos a partir de 6ms

Figura 4. Tempos de predição.

e XGBoost, respectivamente. De acordo com os resultados, estes quatro algoritmos demonstraram perspectivas semelhantes de predição em termos das métricas de erro (i.e., MAE, MSE e RMSE). Já sob a perspectiva do coeficiente de determinação (R² score) a diferença entre o LightGBM e o XGBoost chegou a 10%. As Figuras 3 e 4 detalham

o desempenho individual de cada um dos algoritmos em relação aos tempos de treinamento e predição, respectivamente. Para fins de melhor visualização, os resultados foram organizados em escalas de milissegundos e segundos.

Em termos gerais, as avaliações demonstraram uma não linearidade nos aspectos reportados nos resultados das métricas de erro, coeficiente de determinação e tempos de treinamento e predição. No entanto, uma análise sensitiva dos resultados é determinante para identificar que alguns algoritmos, à exemplo do LightGBM e Gradient Boosting, apresentaram regularidade de desempenho nas métricas avaliadas.

5. Conclusões e Trabalhos Futuros

Este estudo avaliou o desempenho dos principais métodos de regressão para a predição de QoE a partir de KPIs de QoS da rede. Os algoritmos avaliados tiveram suas performances discutivas em termos de métricas de avaliação de desempenho, tempo de treinamento e tempo de predição.

Em trabalhos futuros serão realizados novos conjuntos de experimentações para avaliar o impacto das predições de QoE no problema de decisão de *handover* em cenários com alta demanda de mobilidade e habilitados por fatiamento de recursos. Nestes cenários, as capacidades de predição serão avaliadas sob a perspectiva de entrega fim-a-fim em serviços de rede e nuvem.

Referências

- Gebrie, H., Farooq, H., and Imran, A. (2019). What machine learning predictor performs best for mobility prediction in cellular networks? In *2019 IEEE International Conference on Communications Workshops (ICC Workshops)*, pages 1–6.
- Jardim, S., Dantas Silva, F. S., Neto, A., Bustos, H., Immich, R., and Fontes, R. (2021). Analysis of ml algorithms to support elastic service chaining in ehealth vertical applications. In *2021 International Wireless Communications and Mobile Computing (IWCMC)*, pages 336–341.
- Khan, W., Raza, A., Kuusniemi, H., and Elmusrati, M. (2021). Comparison of machine learning algorithms for venue presence with inclusion of neighbours. In *2021 29th Telecommunications Forum (TELFOR)*, pages 1–4.
- Liu, J. and Huang, C.-T. (2023). Evaluation of different time series forecasting models for 5g v2v networks. In *2023 International Conference on Computing, Networking and Communications (ICNC)*, pages 278–282.
- Sharma, A., Pandit, S., and Talluri, S. R. (2021). A comparative study to classify and predict the throughput of fifth generation wireless technology using supervised machine learning algorithms. In *2021 Sixth International Conference on Image Information Processing (ICIIP)*, volume 6, pages 288–292.
- Silva, F. S. D., Schneider, L. M., Rosário, D., and Neto, A. V. (2022). Network slicing mobility aware control to assist handover decisions on e-health 5g use cases. In *2022 International Wireless Communications and Mobile Computing (IWCMC)*, pages 1034–1039.
- Wang, C.-X., You, X., Gao, X., Zhu, X., Li, Z., Zhang, C., Wang, H., Huang, Y., Chen, Y., Haas, H., Thompson, J. S., Larsson, E. G., Renzo, M. D., Tong, W., Zhu, P., Shen, X., Poor, H. V., and Hanzo, L. (2023). On the road to 6g: Visions, requirements, key technologies and testbeds. *IEEE Communications Surveys & Tutorials*, pages 1–1.
- Wu, W., Zhou, C., Li, M., Wu, H., Zhou, H., Zhang, N., Shen, X. S., and Zhuang, W. (2022). Ai-native network slicing for 6g networks. *IEEE Wireless Communications*, 29(1):96–103.