POLITECNICO DI TORINO
Repository ISTITUZIONALE

Using SHAP Values to Validate Model’s Uncertain Decision for ML-based Lightpath Quality-of-Transmission Estimation

Original
Using SHAP Values to Validate Model’s Uncertain Decision for ML-based Lightpath Quality-of-Transmission Estimation / Houssiany, Hadi; Ayoub, Omran; Rottondi, Cristina; Bianco, Andrea. - ELETTRONICO. - (2023), pp. 1-5. (Intervento presentato al convegno 23rd International Conference on Transparent Optical Networks (ICTON) tenutosi a Bucharest, Romania nel 2-6 July 2023) [10.1109/ICTON59386.2023.10207178].

Availability:
This version is available at: 11583/2981205 since: 2023-08-23T13:09:08Z

Publisher:
IEEE

Published
DOI:10.1109/ICTON59386.2023.10207178

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright
IEEE postprint/Author's Accepted Manuscript

©2023 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collecting works, for resale or lists, or reuse of any copyrighted component of this work in other works.

(Article begins on next page)

08 December 2024

Using SHAP Values to Validate Model's Uncertain Decision for ML-based Lightpath Quality-of-Transmission Estimation

Hadi Houssiany[1], Omran Ayoub[2], Cristina Rottondi[3], Andrea Bianco[3]

[1]*Lebanese University {hadi.houssainy@st.ul.edu.lb},*

[2]*University of Applied Sciences and Arts of Southern Switzerland {omran.ayoub@supsi.ch},*

[3]*Politecnico di Torino {cristina.rottondi@polito.it, andrea.bianco@polito.it}*

ABSTRACT We apply Quantile Regression (QR) for lightpath quality-of-transmission (QoT) estimation with the aim of identifying uncertain decisions and then exploit Shapley Additive Explanations (SHAP) to quantify lightpath features' importance by means of SHAP values and validate the model's decisions in a post-processing phase. Numerical results show that our approach can eliminate more than 98% of false predictions and that SHAP values can validate up to 90% of the model's uncertain decisions.

Keywords: Lightpath QoT Estimation; Uncertainty Quantification; Explainable Artificial Intelligence.

1. Introduction

Quality-of-Transmission (QoT) estimation of unestablished lightpaths is a fundamental problem in optical networks planning [1]. The QoT estimation task is often formalized as a regression problem, which consists of predicting the value of the Bit Error Rate (BER) or Signal-to-Noise Ratio (SNR) associated with each lightpath. Furthermore, the predicted value is associated with a binary class label that indicates whether the BER of the lightpath is above or below the system's acceptability threshold [2].

Machine Learning (ML)-based regression is commonly used to build a QoT estimation model. To this aim, the regression minimizes a least squares loss function, which approximates the relationship between the input variables (i.e., lightpath's features) and the output variables (i.e., QoT values). The output of the QoT model is an estimate, i.e., an approximation; therefore, it contains some uncertainty, resulting from inaccuracies in the model and noise in the input data and reflected in the variability around the mean response value. To achieve confident decisions, it is necessary to go beyond a point estimate, as it may lead to the establishment of lightpaths with unacceptable BER/SNR values, and construct a Prediction Interval (PI) that measures the level of certainty of an estimate. A PI is a range of values in which an observation falls with a certain probability.

Quantifying the uncertainty associated with predictions has recently attracted attention within the research community of optical networking, with several works focusing, in particular, on lightpath QoT estimation [3-6]. In Refs. [3, 4] authors investigate the application of Deep Quantile Regression (DQR), showing that DQR accurately infers QoT bounds with reduced margins, which allows for more confident decision-making in the planning of optical networks. Ref. [5] compares two uncertainty representation frameworks for deep QoT estimation models, based on DQR and on the Monte Carlo dropout technique, respectively. Results show that both frameworks improve classification accuracy of unestablished lightpaths. Ref. [6] investigates the estimation of uncertainty in deep learning models for time series forecasting, such as Long Short-Term Memory and MultiLayer Perceptron (MLP), and compares the proposed models to a Naive model. Authors implement Monte Carlo dropout and Quantile Regression (QR) on the MLP model to measure uncertainty and confidence levels and perform uncertainty calibration to establish reliable uncertainty estimation in time series predictions. Similarly, we focus on the problem of ML-assisted lightpath QoT estimation for achieving confident decision making by quantifying the model's uncertainty, and, in a post-processing step, we validate the uncertain decisions of the model using feature importance extracted with eXplainable Artificial Intelligence (XAI) techniques. More specifically, we propose a QR model encapsulating three Histogram-based Gradient Boosting Regression Trees (HGBRT), one using a squared error loss function and the other two using a quantile loss function, to obtain a 95% PI. Then, we propose exploiting Shapley Additive Explanations (SHAP) [7], an XAI framework which permits extracting features' contribution towards model's decisions, to validate uncertain predictions. We train a ML model that makes use of the SHAP values as features to re-evaluate, in a post-hoc phase, the uncertain decisions. The rationale behind this approach is that XAI frameworks can be used to identify the most influential lightpath features for QoT estimation [8, 9] and may therefore be exploited to guide a post-hoc validation process.

2. Problem Statement and Research Questions

We model the problem of ML-assisted lightpath QoT estimation with uncertainty quantification as a QR problem. The problem consists in predicting i) if the value of the transmission BER along a prospective lightpath will be below (class 0) or above (class 1) a reference acceptability threshold T, and ii) the upper and lower band values of the PI. The input to the regressor is a set of lightpath features that characterize the lightpath itself (e.g. length, modulation format, spectral width, number of spans) and its spectral proximity (e.g., the overall spectral occupation of the traversed links and the lightpath features of the spectrally-adjacent lightpaths). We consider model's decisions whose PI spans across the BER acceptability threshold T to be uncertain, and we aim to reduce the number of false predictions of the models by rejecting uncertain decisions. In addition to this objective, we aim at addressing the following research questions: Q1) Are the explanations (i.e., the combinations of lightpath feature importance values) based on SHAP correlated with the model's outcome (class 0 or class 1)? Q2) Can we use SHAP values to validate, in a post-processing phase, the model's decisions? In other words, can we use SHAP values to identify whether a decision is true or false?

3. Methodology

Fig. 1 shows the overall scheme of our proposed approach, consisting of two models, namely, Model 1 (M1), which performs QoT estimation with QR, and Model 2 (M2), which uses SHAP values to validate uncertain decisions[1]. The dataset used is divided into train and test sets following an 80-20% split. The train set is further divided into two halves, the first is used for training for M1 while the second is fed as a test set to M1 and the outcome of M1 in this case is fed as a train set for M2. The test set is used for testing the whole pipeline. Sec. 3.1 and 3.2 describe M1 and M2, respectively.

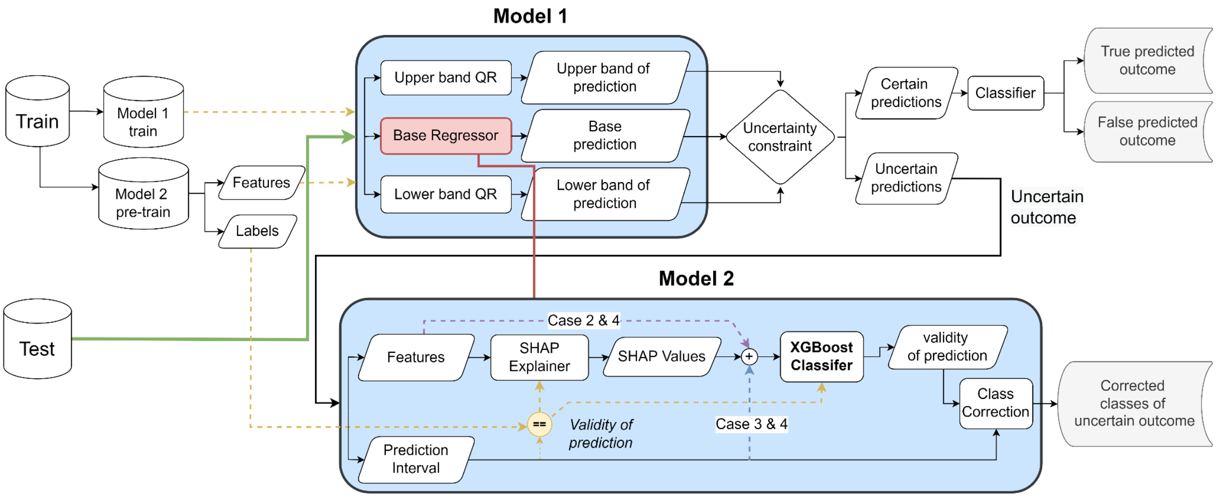


Figure 1. Schematic representation of the overall framework.

3.1 Model 1: Lightpath QoT Estimation with QR

M1 takes as input the set of features characterizing the lightpath and produces as output a PI, consisting of a base value (mean), upper and lower band value. We train a HGBRT to predict i) the BER, ii) the lower bound and iii) the upper bound of the BER PI. The PI is obtained using a quantile loss function with quantiles 0.975 for the upper band and 0.025 for the lower band, to obtain a 95% PI. The output of the model, i.e., the PI, is then evaluated based on an uncertainty policy. We consider the model's outcome to be uncertain if the PI spans across the acceptability threshold T, independently of its width. In other words, even if the PI is relatively short (model is likely certain) but crosses T, the prediction is considered uncertain. On the contrary, even if the PI is large but does not cross T (i.e., its range of values is entirely above or entirely below T), the decision is considered certain. Following this approach, only certain decisions are labeled/classified as either lightpath with acceptable (class 0) or unacceptable (class 1) QoT, while uncertain (but labeled) predictions are post-processed in a later stage (discussed in detail in M2).

3.2 Model 2: Validating Model's Uncertain Decisions Using SHAP Values

M2 works only on the uncertain predictions obtained from M1. It takes as input a set of features of the uncertain decision, and produces as output a label that corresponds to whether the decision of M1 is true or false. We model this task as a supervised ML problem where each sample is given a label of either true or false decision by M1 and then train and fine-tune an eXtreme Gradient Boosting (XGB) classifier algorithm.

We use SHAP, a model-agnostic XAI framework that interprets predictions based on Shapley values used in cooperative game theory, to explain the model's decisions. SHAP computes an explanation by calculating the importance value, also referred to as SHAP value, for each lightpath feature in a given prediction. SHAP is fed with two main inputs, a trained model and a dataset composed of N samples, each represented by K features. The trained model in our case is M1, which is used for obtaining the PIs for a given data point. The output of SHAP is a matrix with N rows (i.e., one for each data sample) and K columns (i.e., one for each lightpath feature), and the value of a lightpath feature for a particular data sample represents the SHAP value of the lightpath feature for that sample. The SHAP value, which can be either positive or negative, indicates how much a lightpath feature

[1] Source code is publicly available on https://github.com/0had0/QoT-Estimation.

contributed to the model's prediction, by either increasing or decreasing the predicted BER value. As set of input features to M2, we consider: 1) only SHAP values (i.e., one SHAP value for each lightpath feature), 2) SHAP values and the original set of lightpath features, 3) SHAP values and the BER PI and 4) SHAP values, the PI and the original set of lightpath features.

4. Results

We consider a synthetic dataset described in detail in [9,10] (we use dataset 01 of the Continental Core Optical Network of the United States network topology). We divide the dataset into train and test sets as explained in Sec. 3. We first compare the performance of the HGBRT regressors with QR to that without QR. The aim of this comparison is to quantify the capability of the proposed model in identifying possible misclassifications by rejecting samples whose PI crosses the BER threshold T. Tab. 1 reports the classification performance of the two approaches, namely, the base regressor and our proposed approach (M1). Both approaches achieve a near-optimal performance (False Positives (FP) + False Negatives (FN) are relatively low with respect to True Positives (TP) + True Negatives (TN)) with M1 showing a slightly better performance. Results show that M1 rejects 11,200 samples, reducing FP from 760 to 17 and FN from 757 to 14, in addition to rejecting TN samples. Overall, M1 rejects 98.6% of the false predictions and 3.5% of the true predictions of the base regressor, showing the effectiveness of the approach in eliminating false predictions.

Table 1. Classification performance of our proposed approach (M1) and baseline approach (base regressor)

	TP	TN	FP	FN	Uncertain Samples
Base Regressor	175872	73594	760	757	-
Model 1	171198	68554	17	14	11200 (98.6% of False and 3.5% of True predictions)

We now use SHAP to explain the model's decisions for the three sets of outcomes, namely, the true predictions, the false predictions and the uncertain predictions. Fig.2 shows SHAP summary plots for true predictions (Fig. 2(a)), false predictions (Fig. 2(b)) and uncertain predictions (Fig. 2(c)). A summary plot combines lightpath feature importance and effects to explain the model’s behavior. The y-axis lists lightpath features according to their importance, and each point on the plot represents a given lightpath feature and a given data point, positioned on the basis of its SHAP value, and the color of a point reflects lightpath feature values in a low-to-high scale. A lightpath feature can either contribute towards increasing the model's estimation (positive Shapley value) or decreasing the model's estimation (negative Shapley value). We show the plots considering the 8 most important lightpath features. The plots show that i) the most influential lightpath features are the same among the different outcomes., and ii) that the same lightpath features show a similar contribution to the model's outcome in terms of SHAP values regardless of the decision value. For instance, high values (red points) of *Mod Order*, *Num Spans* and *Sum Link Occ* show positive SHAP values, i.e., they increase BER predictions, whereas low values (blue points) decrease BER predictions. The plots also show that predictions are mainly driven by *Mod Order* and *Num Spans*, which exhibit a significantly larger range of SHAP values than all the other lightpath features (see Ref. [8] for a detailed description of features’ impact).

Addressing Q1: We now investigate whether the combinations of SHAP values of the various lightpath features are correlated with the model's outcome. We first reduce the dimensionality of the SHAP values extracted from the model's explanations to a 2-dimensional feature space using Uniform Manifold Approximation and Projection (UMAP) and then cluster them considering their label (class 0 for samples with sufficient QoT and class 1 for samples with insufficient QoT). We perform clustering separately for each set of decisions using Density-Based Spatial Clustering of Applications with Noise. Fig. 3(a) shows the clustering of the explanations, (i.e., the set of SHAP values) of the true predictions. The clustering shows that, except for some outliers, two clusters can be identified, which divide the explanations based on their label. Fig. 3(b) shows the clustering of the explanations of the false predictions. Similarly to the case of true predictions, the clustering shows that two clusters can be identified. Unlike for the other cases, in the case of uncertain predictions (Fig. 3(c)), the clustering does not show a clear distinction between the model's decisions and the 2-d representation of the explanations. Yet, the resulting clusters show some separation between the 2-d representation of the explanations, demonstrating a non-marginal correlation between SHAP values and predicted QoT class.

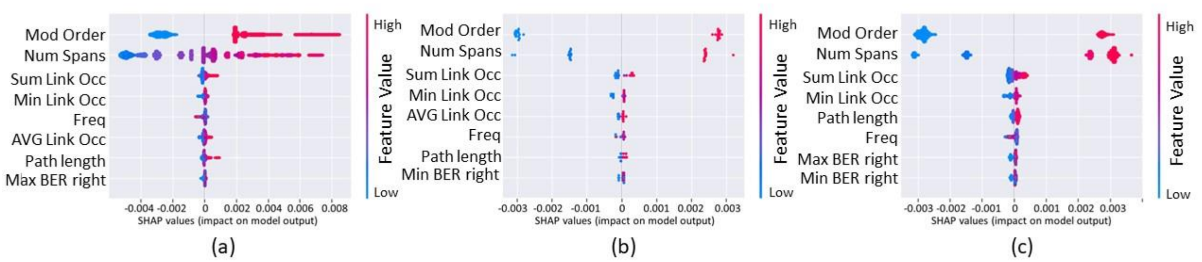
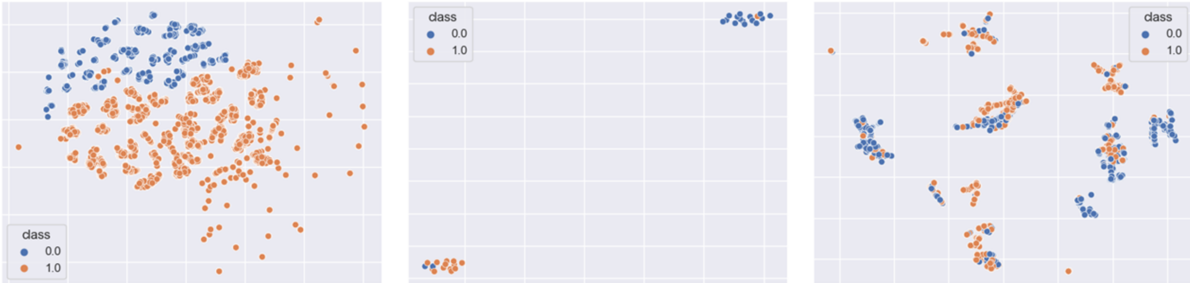


Figure 2. SHAP summary plots for (a) true predictions, (b) false predictions and c) uncertain predictions



(a) (b) (c)

Figure 3. Clustering outcome of the 2-dimensional representation of model's explanations for a) true predictions, b) false predictions and c) uncertain predictions

Addressing Q2: Based on the findings of Q1, we now investigate whether the SHAP values can be used to validate decisions, i.e., to identify if an uncertain decision is true or false. To this end, we compare the performance of the XGB classifier trained considering the four cases discussed in Sec.3, which differ in the set of features used in input. Tab. 2 reports the Recall, F1 score and Accuracy of the model for the various cases. Results show a high accuracy, ranging between 0.867 (using only SHAP values) and 0.893 (using SHAP values, lightpath features and PI). These results reveal that a high correlation exists between SHAP values and the model's outcome (i.e., whether the outcome is true or false). Moreover, when features and model's PI are considered in addition to SHAP values, an accuracy of around 0.9 and a recall and F1 score of 0.89 can be achieved, revealing that the SHAP values can be used to validate decisions in a post-processing phase.

Table 2. Classification performance of the XGB model trained to validate uncertain decisions.

	SHAP values		SHAP values + features		SHAP values + PI		SHAP values + features + PI	
	class 0	class 1	class 0	class 1	class 0	class 1	class 0	class 1
Recall	0.871	0.862	0.890	0.884	0.871	0.862	0.893	0.892
F1	0.871	0.862	0.891	0.884	0.871	0.862	0.896	0.890
Acc.	0.867		0.888		0.867		0.893	

5. Conclusion

We utilize Quantile Regression (QR) to estimate the quality-of-transmission (QoT) of prospective lightpaths and employ Shapley Additive Explanations (SHAP) to extract SHAP values, which quantify the importance of input features and validate model decisions during post-processing. Our approach successfully eliminated over 98% of false predictions and was able to validate up to 90% of uncertain model decisions using SHAP values.

ACKNOWLEDGEMENTS

This work was partially supported by the European Union under the Italian National Recovery and Resilience Plan (NRRP) of NextGenerationEU, partnership on "Telecommunications of the Future" (PE00000001 - program "RESTART").

REFERENCES

[1] Pointurier, Yvan. "Machine learning techniques for quality of transmission estimation in optical networks." Journal of Optical Communications and Networking 13.4 (2021): B60-B71.
[2] Rottondi, Cristina, et al. "Machine-learning method for quality of transmission prediction of unestablished lightpaths." Journal of Optical Communications and Networking 10.2 (2018): A286-A297.
[3] Panayiotou, Tania, et al., "Deep quantile regression for QoT inference and confident decision making." 2021 IEEE Symposium on Computers and Communications (ISCC). IEEE, 2021.
[4] Maryam, Hafsa et al., "Learning quantile QoT models to address uncertainty over unseen lightpaths." Computer Networks 212 (2022).
[5] H. Maryam et al., "Representing Uncertainty in Deep QoT Models," 2022 20th Mediterranean Communication and Computer Networking Conference (MedComNet), Pafos, Cyprus, 2022, pp. 113-121, doi: 10.1109/MedComNet55087.2022.9810426
[6] Yousefi, Somaiieh, et al., "Forecasting lightpath quality of transmission and implementing uncertainty in the forecast models." Journal of Lightwave Technology (2023).
[7] Lundberg, Scott M., and Su-In Lee. "A unified approach to interpreting model predictions." Advances in neural information processing systems 30 (2017).
[8] Ayoub, Omran, et al. "Quantifying Features' Contribution for ML-based Quality-of-Transmission Estimation using Explainable AI." 2022 European Conference on Optical Communication (ECOC). IEEE, 2022.
[9] Ayoub, Omran, et al. "Towards explainable artificial intelligence in optical networks: the use case of lightpath QoT estimation." Journal of Optical Communications and Networking 15.1 (2023): A26-A38.
[10] Bergk, Geronimo, et al. "ML-assisted QoT estimation: a dataset collection and data visualization for dataset quality evaluation." Journal of Optical Communications and Networking 14.3 (2022): 43-55.