On the Application of Explainable Artificial Intelligence to Lightpath QoT Estimation

Omran Ayoub[1], Andrea Bianco[2], Davide Andreoletti[1], Sebastian Troia[3], Silvia Giordano[1], Cristina Rottondi[2]

[1] *University of Applied Sciences of Southern Switzerland, Switzerland*
[2] *Politecnico di Torino, Italy* [3] *Politecnico di Milano, Italy*

{*omran.ayoub,davide.andreoletti,silvia.giordano*}*@supsi.ch,*{*cristina.rottondi,andrea.bianco*}*@polito.it, sebastian.troia@polimi.it*

Abstract: We demonstrate the potentialities of explainable AI when applied to distill knowledge from a trained supervised machine learning model for lightpath quality of transmission estimation in optical networks, with synthetic datasets. © 2021 The Author(s)

1. Introduction

In the last few years, the application of Machine Learning (ML) technologies has been widely investigated in several fields of optical communications and networking, including e.g. lightpath Quality of Transmission (QoT) estimation, automated fault management, routing and wavelength/spectrum assignment [1]. However, the vast majority of ML approaches proposed in the scientific literature are used as "black boxes", i.e., they do not expose their internal mechanics nor the decisional processes adopted to associate the produced outputs with the set of feature values provided as input. This hinders the interpretability of the models and prevents the extraction of useful insights that could be leveraged to better understand the nature of the problem at hand.

To overcome these shortcomings, the research field of eXplainable Artificial Intelligence (XAI) focuses on the development of frameworks aimed at making decisions taken by a ML algorithm understandable by humans [2], with the final goal of improving transparency and increasing trust in ML-based systems. XAI-oriented models have already been applied in several domains such as transportation, healthcare and finance.

In this paper, we make a first attempt to demonstrate the benefits of applying XAI in optical networking, as it represents a promising tool to make network operators capable of distilling precious knowledge from the analysis of their network monitoring data, e.g. the identification of the main driving factors that lead to ML decisions and of the causes of wrong predictions. In particular, we focus on lightpath QoT estimation and consider the SHapley Additive exPlanations (SHAP) XAI framework [3]. Due to the difficulty of collecting data from real deployments, in this study we consider two synthetic datasets recently made available in [4]. Though the specific conclusions drawn from the application of XAI techniques strictly depend on the intrinsic characteristics of those datasets, the nature of the obtainable outcomes can be generalized to any optical network setting. We therefore expect that this preliminary showcase will pave the way for further investigation on the application of XAI to QoT estimation and to other optical networking problems.

2. Explainable AI for QoT Estimation

Problem Statement: We model the lightpath QoT estimation task as a classification problem, which consists in predicting if the value of the Bit Error Rate (BER) associated with the transmission along a perspective lightpath will be below or above a reference acceptability threshold. The input to the classification algorithm is a set of features that characterize the ligthpath itself (e.g. length, spectral width, number of spans and traversed nodes, amount of carried traffic, modulation format) and its spectral proximity (e.g., the overall spectral occupation of the traversed links and the features of the spectrally adjacent lightpaths).

We exploit three supervised ML models: Random Forests (RF), Artificial Neural Networks (ANN) and Extreme Gradient Boosting (XGB). By applying XAI, we aim at addressing the following research questions: *Q1)* Which features are the most relevant for lightpath QoT estimation? Are relevant features common across different datasets and across different classification models? *Q2)* How features contribute to estimating lightpath QoT? Can we extract problem insights by examining features' contributions?

Adopted XAI Framework: To explain pre-trained "black-box" models with the objective of understanding their reasoning, we use an XAI framework in a *post-hoc* manner, i.e., after the model has taken its decision. XAI frameworks can be either model-specific, i.e., their application is restricted to specific ML models, or model-agnostic, i.e., they can be applied to any ML model. In our model, we employ a model-agnostic framework,

Table 1: Classification results of RF, ANN and XGB for the two datasets. Numbers in brackets are obtained after selecting as feature set the eight most important features for each case.

	Dataset 1				Dataset 2			
Model	Accuracy	Precision	Recall	F1-Score	Accuracy	Precision	Recall	F1-Score
RF	0.997 (0.987)	0.997 (0.966)	0.998 (0.989)	0.998 (0.985)	0.996 (0.987)	0.994 (0.981)	0.994 (0.988)	0.994 (0.984)
ANN	0.994 (0.988)	0.992 (0.985)	0.992 (0.986)	0.992 (0.985)	0.988 (0.988)	0.98 (0.98)	0.986 (0.985)	0.984 (0.983)
XGB	0.997 (0.994)	0.997 (0.994)	0.997 (0.993)	0.997 (0.994)	0.996 (0.993)	0.994 (0.991)	0.995 (0.991)	0.995 (0.991)

Table 2: The eight most important features according to SHAP for each model and for each dataset considered.

Dataset	Model	Feature Rank 1	2	3	4	5	6	7	8
1	XGB	**Num Spans**	**Mod Order**	**Link Occ**	**Path Len**	Freq	Avg Link Occ	Max BER Right	**Num Links**
	RF	**Mod Order**	**Num Spans**	**Path Len**	**Num Links**	LP Linerate	Conn Linerate	**Link Occ**	Min Link Len
	ANN	**Mod Order**	**Num Spans**	**Path Len**	**Num Links**	**Link Occ**	Avg Link Occ	Max Link Len	Std Link Occ
2	XGB	**Mod Order**	**Num Spans**	**Path Len**	Avg Link Occ	Freq	Max BER Right	Min Mod Order Left	Min Mod Order Right
	RF	**Mod Order**	**Num Spans**	**Path Len**	LP Linerate	Conn Linerate	Num Links	Link Occ	Max Link Len
	ANN	**Mod Order**	**Num Spans**	**Path Len**	Num Links	Link Occ	Avg Link Occ	Freq	Avg Link Len

namely, SHAP [3]. SHAP is based on *Shapley Value*, a method from coalitional game theory, which calculates and assigns fair payouts, i.e., Shapley values, to players based on their contribution to the total payout. In the context of XAI, the method is used to calculate the marginal contribution of each input feature of an observation (i.e., the average influence each feature has) on the prediction. Following this method, SHAP explains decisions taken by a ML model by computing the contribution that each feature (and its value in an observation) has on the model's decision. Note that explanations can be global, i.e., they explain the behavior of the whole model, or local, i.e., they provide explanations for a specific observation. In our analysis, we exploit SHAP to build a global description of model's behavior by aggregating Shapley values of features of observations. The goal is to understand which features are most influential and which values of features drive the predictions to the target class and which not.

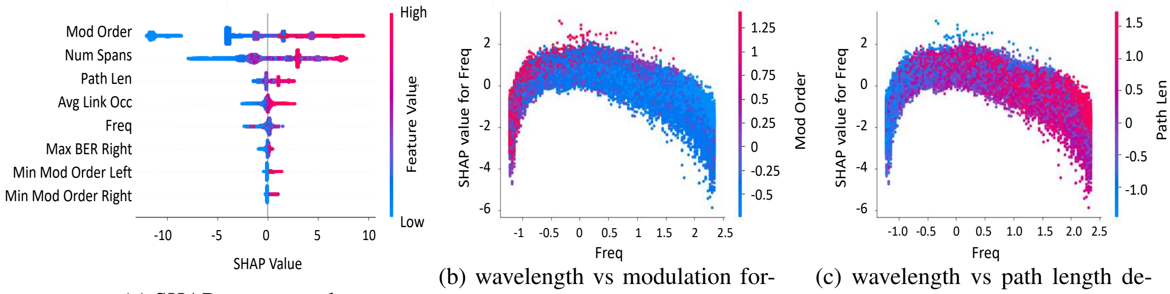
3. Numerical Results and Discussion

Datasets: Results have been obtained using two of the four lightpath datasets available in [4]. Datasets contain about three million samples, each one consisting in a vector of 32 features (see [4] for details) associated with a binary label, depending on whether the lightpath belongs to the positive class (class P, i.e., below BER threshold) or negative class (class N, i.e., above BER threshold).

Supervised ML models: The combinations of hyperparameters of RF, XGB and ANN have been evaluated as follows. For RF, we consider different amounts of trees, ranging between 10 and 300 with step of 10, and splitting criteria, chosen between *Gini Index* and *Cross-Entropy*. For XGB, we vary both *eta parameter* and *subsample* between 0.1 and 1 with step 0.1, and *max depth* between 1 and 10 with step 1. As for ANN, we consider multi-layer ANNs with number of hidden layers ranging between 1 and 10, number of neurons per layers in the set {10, 20, 50, 100}, and activation function among {*Identity*, *Sigmoidal*, *Tanh*, *Relu*}. The combination of hyperparameters were selected based on the best performance obtained on dataset 1, and are as follows. For RF, 290 trees and *Cross-Entropy* splitting criterion. For XGB, values of *eta*, *subsample* and *max depth* are 0.5, 0.8 and 9, respectively. For ANN, 4 hidden layers with 10 neurons each and *tanh* activation function were considered. After selecting the hyperparameters, the obtained models are trained on 90% of the dataset and the performance of each model was evaluated on the remaining 10% of the dataset.

Performance evaluation: Numerical results are shown in Tab.1. We evaluated the algorithms over the two datasets in two cases: *a*) when samples include all 32 features and *b*) when samples include the 8 most important features, as identified by SHAP for each case (see next paragraph). For both datasets, in case *a*) RF and XGB show an almost-identical performance (classification metrics ranging between 99.4% and 99.7%) with a slight advantage over ANN (metrics ranging between 98% and 99.4%). In case *b*) (results in brackets in Tab.1), performance degradation is limited (up to 1% accuracy reduction for RF and ANN and up to 0.3% for XGB), which validates the effectiveness of the XAI-based feature selection procedure.

Feature Selection: Tab.2 shows the 8 most important features of each model and for each dataset, selected by SHAP. In all cases, *Mod Order*, *Num Spans* and *Path Len* are the three most important features, except for *XGB* in dataset 1, which has *Link Occ* as third while *Path Len* is fourth. This shows the relevant impact that modulation order, number of spans and path length have on estimating the lightpath's QoT in both datasets. We also see that the models share 5 features among the 8 most important ones for dataset 1, while only 3 out of 8 for dataset 2 (common features among the three models for each of the datasets are highlighted in bold in Tab.2). More precisely, for dataset 1, all models significantly rely on the number of links traversed by the lightpath (*Num Links*) and on the total link occupation (*Link Occ*), while in dataset 2, for *XGB*, the two features do no appear among the 8 most important ones. Instead, *XGB* relies on *Min Mod Order Left* and *Min Mod Order Right*. Interestingly, *XGB*



(a) SHAP summary plot (b) wavelength vs modulation format dependency plot (c) wavelength vs path length dependency plot

Fig. 1: SHAP summary plot and examples of dependency plots for class N obtained with *XGB* using dataset 2.

is the best performing model, which shows that leveraging features that characterize spectrally adjacent channels benefits the classification accuracy.

Features' contribution: To discuss the relation between feature values and model's decisions, we focus on the *XGB* model for dataset 2. Fig.1(a) shows the *SHAP summary plot* for the 8 most important features and their contribution towards the decision for class N. A summary plot combines feature importance and feature effects to explain the model's behavior. The y-axis lists features according to their importance, and each point on the plot represents a given feature and a given data point, positioned on the basis on its Shapley value. Each point has also a color, which qualitatively represents the value of the feature in a low-to-high scale. Finally, each feature can either contribute towards (positive Shapley value) or against (negative Shapley value) the class in consideration.

We note that, for *Mod Order* and *Num Spans*, high/medium-to-high values (red and purple points) contribute significantly towards class N (high positive SHAP values), with some exceptions for a relatively low number of samples with medium values of both features which weakly contribute against class N. In contrast, low values of each of the features show a direct relationship with acceptable QoT estimation, i.e., they show a negative contribution towards class N. We now focus on the *Min Mod Order Left* and *Min Mod Order Right* features, i.e., the minimum modulation format order of the left (respectively right) neighboring lightpaths. We first notice that the relationship between these two features and the model's decision is asymmetrical: high values show a low positive SHAP value towards class N, whereas low values give no contribution (zero SHAP value). This suggests that the model operates as follows: if the minimum modulation order of the left/right neighboring lightpaths have high values (i.e., *all* the ligthpaths spectrally adjacent to the one under consideration use high-order modulation formats), the features drive the decision towards classifying the lightpath as belonging to class N. Instead, if either feature has a low value, the information extracted does not impact the decision in any way. A network operator could therefore leverage such insights and opt for a low-order modulation format when deploying a new lightpath, if both *Min Mod Order Left* and *Min Mod Order Right* exhibit high values.

Finally, Figs.1(b)-(c) report two examples of *SHAP dependency plots* for an individual feature (in this case *Freq*, i.e., the ligthpath's wavelength in the WDM grid). Each point on the plot represents a given sample, the x-axis shows its normalized *Freq* value while the y-axis shows its Shapley value towards class N. Each point is also colored depending on the value of a second feature (respectively the modulation format order in Fig.1(b) and the lightpath length in Fig.1(c)), using a low-to-high scale. Thanks to this plot, it is possible to visualize the combined impact of two features on the model's decisions. From the two figures it is possible to gain insigths on the impact of the spectrum allocation policy (the dataset was generated using a First Fit (FF) strategy and contains instances for 4 different traffic load conditions). We observe that at the two extremes of the spectrum band, samples always exhibit negative Shapley values (i.e. against class N), due to the fact that in those spectrum regions lightpaths suffer less interference from adjacent channels. Conversely, in the center-left frequency range, most points have positive Shapley values (i.e., in favor of class N). Indeed, the FF policy starts scanning the spectrum from left to right, so the left side of the spectrum has on average higher occupation than the right side. Moreover, it also emerges that, due to fragmentation effects, short lightpaths with high-order modulation formats are typically allocated on the left spectrum side, whereas long lightpaths with low-order modulation formats are allocated on the right side. This way, operators may visually inspect how the combined effect of path length and choice of wavelength/modulation format influence the likelihood that the perspective lightpath will exhibit unacceptable QoT.

References

[1] F. Musumeci et al. "An overview on application of machine learning techniques in optical networks." IEEE Comm. Surveys and Tutorials 21.2 (2018): 1383-1408.

[2] A. Adadiand et al. "Peeking inside the black-box: a survey on explainable artificial intelligence." IEEE access 6 (2018): 52138-52160.

[3] Lundberg, Scott M., and Su-In Lee. "A unified approach to interpreting model predictions." Proc. of the 31st int. conf. on neural information processing systems, 2017.

[4] P. Safari et al. "Deep convolutional neural network for network-wide QoT estimation," OFC, 2021.