Explainable Artificial Intelligence-Guided Optimization of ML-Based Traffic Prediction

Aleksandra Knapińska[1], Omran Ayoub[2], Cristina Rottondi[3], Piotr Lechowicz[1,4], Krzysztof Walkowiak[1]

[1]*Wrocław University of Science and Technology, Poland* [2]*University of Applied Sciences of Southern Switzerland, Switzerland*
[3]*Politecnico di Torino, Italy* [4]*Chalmers University of Technology, Sweden*

Abstract*—Traffic prediction is an evergreen research topic in networking, with modern allocation algorithms often utilizing forecasts for optimized decisions. However, the employed machine learning (ML) models are usually operated as black boxes – without any insight into their internal operations. Such an approach creates a risk of using excessive input features, unnecessarily expanding the model complexity. In this work, we extract insights into the operation of traffic prediction models using explainable artificial intelligence (XAI) tools. We explore the impact of literature-proposed features on various traffic types, sampling rates, and ML algorithms. We identify the common trends and dependencies regarding the most relevant features depending on traffic fluctuation levels and aggregation type. We discover how only a subset of inputs contributes meaningfully to the final model decision, as opposed to the conventional approach of only analyzing the resulting prediction quality after adding new features. We demonstrate how training and inference times can be significantly reduced by exploiting the obtained knowledge without degrading prediction quality and bandwidth blocking.

***Index Terms*—Traffic Prediction; Machine Learning; Explainable Artificial Intelligence; Feature Selection;**

I. Introduction

Network traffic forecasting has proven to be essential for improving network operation in terms of, e.g., operational costs, energy efficiency and stability [1], [2]. Operators can exploit the knowledge obtained from traffic prediction to make informed decisions for various networking problems such as dynamic routing and spectrum allocation (RSA) [1], [2].

Among the various network segments, traffic forecasting within the backbone segment, which serves as a large-scale aggregation of a multitude of individual connections, receives particular emphasis, as it empowers the operators to preemptively identify potential congestion points at a large scale.

The traffic within backbone networks exhibits strong seasonality. Consequently, clear patterns and trends can be extracted in the daily utilization of specific network-based services and applications [3] or the overall traffic crossing internet exchange points [4]. Nevertheless, network traffic forecasting is a challenging task and sophisticated methods are adopted to obtain accurate predictions [5], [6], including advanced statistical methods and machine learning (ML)-based approaches [5], [7]. While it is possible to achieve satisfactory performance with the former, previous studies (e.g., [5], [6]) show that methods based on ML yield better quality forecasts than those based solely on time-series analysis.

Exploiting the predictive power of ML, however, comes at the expense of two aspects. First, and in contrast to the use of statistical methods which use plain traffic measurements, ML-based approaches require the processing of raw traffic traces to engineer input features, usually not given up-front, for training the learning algorithms. A typical way to proceed is to characterize traffic time series via a number of temporal or statistical indicators that are expected to be meaningful for the prediction task. Feature engineering is usually performed utilizing the observed trends, seasonality, and expert knowledge [8]–[10]. In more detail, the prediction model features can be composed of significant past observations (e.g., traffic measurements a day/week before), growth rate, or other statistics [10]. No thought is, however, given to the actual importance and contribution of the utilized inputs. Such a workflow generates the risk of unnecessarily enlarging the models and increasing their complexity by including features of no significant impact on the model's predictions.

The second aspect is the fact that ML models operate as black boxes, i.e., they provide their traffic forecasts without offering any indications (or reasons) behind their decisions. Without clear explanations of the model's predictions, network operators are left in the dark regarding the factors influencing the forecasts. This lack of interpretability hampers the decision-making process, making it challenging to validate the reliability of predictions, understand the model's limitations, and incorporate human expertise in refining forecasts. Another critical issue stems from the inability to extract meaningful insights or reasons behind the model's predictions, which may help engineer features for traffic prediction and better understand the problem at hand.

In this work, we aim to address these two aspects by exploiting explainable artificial intelligence (XAI) techniques for generating explanations that allow to quantify features' contribution to model's decisions, with the goals of enhancing the transparency of the employed ML model and refining the model development process with insights for feature engineering and selection. Specifically, we explore how ML-based traffic forecasting methods operate and which features meaningfully contribute to their predictions. To reach this aim, we employ a XAI framework, namely, Shapley Additive Explanations (SHAP) [11] to gain insights into the mechanics

This work was supported by National Science Center, Poland under Grant 2019/35/B/ST7/04272, the NAWA STER Programme Internationalisation of Wroclaw University of Science and Technology Doctoral School and by the European Union under the Italian National Recovery and Resilience Plan (NRRP) of NextGenerationEU, partnership on "Telecommunications of the Future" (PE00000001 - program "RESTART").

of the forecasters. We examine literature-based representatives of various types of ML algorithms (a neural network, an ensemble method, and a simple single predictor) on multiple real and semi-synthetic traffic datasets with different sampling rates to identify the most common trends and dependencies. We show how various models rely almost solely on *temporal* features when making their predictions. Then, we use this knowledge to reduce the size of the employed feature sets and demonstrate that such reduction has almost no impact on the prediction quality while significantly decreasing the model training and inference times. We finally show the practical application of our study in multilayer network operation.

The remainder of this paper is organized as follows. In Sec. II we discuss the related work. In Sec. III we describe the problem, datasets used and our approach. In Sec. IV we discuss findings from applying SHAP to developed ML models. In Sec. V we leverage insights from the previous section to train ML models using a limited number of features and then examine the practical application of our approach. Sec. VI concludes this work.

II. Related Work

Numerous works have investigated network traffic forecasting: advanced statistical and ML-based methods were proposed to accurately predict the upcoming traffic evolution [5], [7]. However, for the successful deployment of network traffic forecasters, a particular focus needs to be put on practical aspects, including data preparation and feature engineering. In particular, features are necessary for the models to learn the relationship between their inputs and the target and thus make accurate forecasts. Usually, the raw traffic measurements are the only data available; various time series characteristics can then be crafted as features, including the highly correlated past measurements (e.g., traffic recorded the day before), information about the growth rate, statistics regarding the date and time of the measurements, or traffic evolution within a period [8]–[10]. If additional information, such as road conditions or newly generated flows, is also available – they can be turned into helpful features as well [12], [13].

Apart from the feature engineering process, the development and choice of the ML model to be adopted for traffic forecasting takes solely into consideration the model's predictive power, overlooking other aspects such as the actual contribution of each feature. In turn, the black-box models can grow unnecessarily large and complex because no analysis is performed to verify which of their inputs are the main decision factors and if there are any unnecessary ones. In our work, we aim to investigate this issue and, for the first time, explore the contribution of multiple previously proposed features for the network traffic prediction task to gain insights into their importance and impact through XAI.

XAI is one of the research fields that enables examining ML models to improve their transparency and trust in their outputs [14]. In the networking community, XAI is recently gaining attention to better understand and improve ML models solving various tasks. The prime example is the quality of transmission

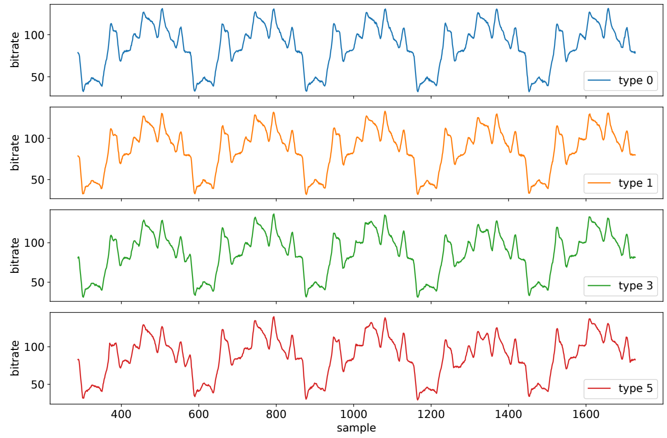

Fig. 1: Examples of traffic types *0, 1, 3, 5* (top to bottom).

estimation, where XAI has been employed for quantifying the contribution of features, simplifying the models, or validating their uncertain decisions [15]. Other XAI use cases in optical networks include failure management, specifically its detection, localization, and cause identification [16], [17]. In the context of network traffic prediction, the application of XAI is a rather unexplored field. In particular, explainability was employed to gain insights into video quality classification [18]. Still, the specific task of traffic forecasting in optical networks remains open. To the best of our knowledge, this is the first study to explore feature engineering for traffic prediction using XAI tools. We believe that the gained knowledge will provide viable insights on how different ML algorithms predict various traffic types and thus encourage the operators to employ AI-based traffic prediction tools.

III. Problem Description and Approach

A. Problem Description

The underlying objective of this work is short-term network traffic prediction. Given a series of historical traffic measurements, our goal is to predict the bitrate for the following sample. Additionally, as black-box ML models are employed for traffic forecasting using various input features, our primary goal is to gain insights into models' behavior by computing the importance of these features and identify the most contributing ones. We aim to determine common trends and dependencies for different traffic datasets and ML algorithms. We aim to examine how the behavior of distinct models differ when trained on the same dataset, and how the behavior of the models vary across different datasets or when raw data is processed differently. Finally, we wish to exploit this knowledge to identify a minimal yet informative set of features that can be used for model training. We examine the impact of this in a practical implementation of ML-based network traffic forecasting application by analyzing, in addition to the quality of predictions, the training and inference times and the blocking probability experienced when adopting ML models with a reduced set of features identified using SHAP.

B. Traffic Data Generation

We use four semi-synthetic datasets, in which real data patterns are used as a base for a generator adding desired

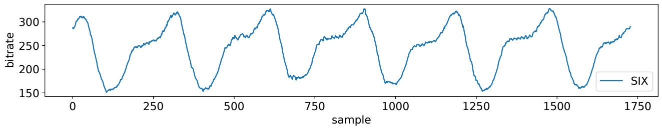

Fig. 2: Example of real traffic from the SIX [4].

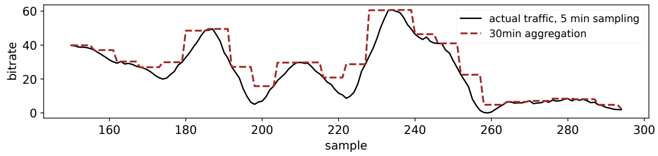

Fig. 3: Illustration of the 30-min aggregation.

characteristics (e.g., noise), and a real dataset from Seattle Internet Exchange Point (SIX) [4] to cross-verify our findings. The semi-synthetic traffic datasets contain patterns of various applications available in [3], which is a good representation of diverse traffic in application-aware networks. As in our previous work [2], we use our custom approximation and noising algorithm (available as the Traffic Weaver package [19]) to transform the provided bar plots with average traffic in each hour into a continuous signal capturing the characteristics of each traffic type. We build four datasets with a 5-minute sampling for 14 days. The datasets, referred to as *Type 0*, *Type 1*, *Type 3*, and *Type 5*, differ in the level of the added fluctuation. Specifically, in traffic *Type 0*, the pattern repeats daily with added Gaussian noise. The subsequent types (*Type 1*, *Type 3*, *Type 5*) are characterized by additional 1%, 3%, and 5% of fluctuation in terms of day-to-day bitrate and shape changes within days. Each dataset contains a collection of 50 distinct traces following a pattern of one network-based service or application from [19]. Fig. 1 shows an example plot of each traffic type present in the data. As for the real dataset [4], it includes 5-minute sampled measurements collected over an approximately 2-month-long period. It is a widely used data source in networking research, including previous works [10], [20]. Refer to a representative fragment in Fig. 2.

In addition to the 5-minute sampling, we resample all datasets used in our study following a 30-minute maximum aggregation (see Fig. 3 for an example). Our rationale is that, although network traffic changes dynamically, allocation algorithms often perform resource assignment for more extended periods for the sake of stability, even if it implies overprovisioning (e.g., [20]). Therefore, a longer period might be desirable to ensure versatility of the identified trends.

C. Data Processing and Feature Extraction

The highly seasonal nature of backbone network traffic is one of the foundations for feature engineering for the considered problem [8]–[10]. In particular, clear daily and weekly patterns emerge when analyzing it over an extended period (e.g., see [4] and the example plotted in Fig. 2). In turn, features can be crafted using significant, highly correlated past samples or statistics describing their neighborhood and evolution. Therefore, in previous research [10] we proposed several ways of creating input features from raw data. We showed how information extracted from traffic traces can be turned into three groups: *statistical* (sin and cos components of date and time, hour window percentiles, etc.), *growth rate* (growth trend in significant past moments), and *temporal* (highly correlated past samples). The experimental evaluation on various ML algorithms revealed that the addition of each group of features increases the prediction quality for different traffic types. The final recommended model, which we consider as the base of this study, contains seventeen (17) features from the three aforementioned groups. The feature names are given on the plots in Sec. IV, and for more details regarding their creation, we refer to [10].

D. SHAP for Feature Contribution

To conduct our experiments, we first need to employ ML models for the task at hand. To this end, we rely on ML models that have proven their efficacy in this task in our previous work [10]. Specifically, we consider three diverse ML algorithms, namely, Multilayer Perceptron (MLP) with one hidden layer of twenty-five neurons, with the *ReLU* activation function and *adam* optimizer, Random Forest (RF) with 75 trees and a Linear Regression (LR). We follow the 5x2 cross-validation (for details on experimental protocol refer to [10]).

To quantify the contribution of the features (i.e., how features impact the model's outcome), we employ SHAP [11], a model-agnostic XAI framework that explains the output of ML models by estimating each feature's contribution to the model's outcome in a post-hoc manner (i.e., after models are trained). SHAP adopts a game theoretic approach exploiting Shapley values to quantify each feature's contribution [11]. For regression problems, the SHAP value associated to a feature indicates its numerical contribution to the model's prediction. A positive (resp. negative) SHAP value indicates that the feature has a positive (resp. negative) outcome on the model's prediction, i.e., it increases (resp. reduces) the prediction. To compute the SHAP values for a particular model, SHAP takes as input the trained ML approach and the training dataset.

IV. Analyzing Feature Contribution of ML Models for Traffic Prediction

A. Performance of ML Models for Traffic Prediction

We measure the performance of the ML models across the five datasets and considering the 5-min sampling and 30-min aggregation using the mean absolute percentage error (MAPE)[1]. We report the MAPE achieved by the models in Tab. I (please see the cases corresponding to *all* features). The prediction accuracy of all models, and in all cases, is generally satisfactory, with MAPE never exceeding 5%. As expected, for both the 5-min sampling and 30-min aggregation, the models show best performance (lowest MAPE) with the least fluctuating traffic *Type 0*. MAPE then shows a slight upward trend as the datasets become more intricate with increased noise levels. Comparing the performance of the different models, we notice that they exhibit a comparable performance, and that there is no discernible trend indicating one model consistently outperforming the others.

[1]As a percentage measure, the MAPE allows for a direct comparison of the prediction quality between datasets with different traffic volumes.

TABLE I: Comparison of average MAPE and the training and inference times for the considered datasets when using all input features or only the most contributing, temporal ones.

Model	Features	5-min Sampling					30-min Aggregation					Training Time (s)	Inference Time (s)
		Type 0	Type 1	Type 3	Type 5	Real	Type 0	Type 1	Type 3	Type 5	Real		
LR	all	0.0054	0.0144	0.0184	0.0183	0.0042	0.0060	0.0116	0.0266	0.0402	0.0125	0.00133	0.00033
	temporal only	0.0047	0.0118	0.0216	0.0241	0.0044	0.0060	0.0114	0.0260	0.0397	0.0151	0.00083	0.00027
RF	all	0.0049	0.0109	0.0165	0.0214	0.0050	0.0054	0.0083	0.0171	0.0261	0.0140	0.41515	0.01223
	temporal only	0.0051	0.0112	0.0189	0.0242	0.0052	0.0055	0.0085	0.0183	0.0290	0.0170	0.15200	0.01216
MLP	all	0.0069	0.0125	0.0236	0.0301	0.0055	0.0072	0.0126	0.0273	0.0408	0.0155	0.55563	0.00065
	temporal only	0.0061	0.0130	0.0231	0.0272	0.0044	0.0072	0.0125	0.0279	0.0412	0.0156	0.19137	0.00051

B. *Feature Contribution Analysis*

We now examine the models' behavior in terms of features' contributions (their SHAP values). Our analysis is concentrated on specific facets of the problem, and consequently, on a set of selected cases. More specifically, we aim to examine: i) the impact of data fluctuations on the models' behavior, ii) the impact of considering a relatively large aggregation period of data traffic, and iii) if, and in case, to which extent, does the models' behavior change with real data in respect to the case with semi-synthetic data.

a) Impact of Data Fluctuations on Feature Contribution: Fig. 4 shows SHAP summary plots for RF and MLP for the least fluctuating traffic type in our semi-synthetic data traffic, i.e., *Type 0*. The summary plot can be read as follows. The primary y-axis reports the top-10 features ranked from most (top) to least (bottom) contributing to the model's decision. The x-axis illustrates each feature's impact on the model's output (i.e., how much the value of a given feature pushes the model's decision in the positive or negative direction). The secondary y-axis reports a color scale for feature values. Finally, for a given feature, each point represents the SHAP value assigned to the feature for a particular prediction query. The plot shows, for both models, that only the *temporal* features (namely, "day_ago_value" and "week_ago_value" for RF, and additionally "previous_value" for MLP) have a significant impact on the models' predictions. In particular, high values of such features (purple-red points) increase the value of the prediction (positive SHAP value), whereas low feature values (blue points) tend to decrease models' outcomes (negative SHAP value). We further note that feature contribution plots of other ML models considered in our study for traffic *Type 0* and *Type 1* (not reported for the sake of conciseness) show similar patterns. *These findings indicate that, regardless of the ML algorithm, the models create an internal prediction function that is highly (and only) correlated to past measurements for traffic characterized by light fluctuations*. This observation also suggests that features pertaining to both *statistical* and *growth rate* do not contribute with any additional knowledge to the model beyond what is already provided by *temporal* features.

We now analyze feature impact on the models' decisions considering the ML models trained for forecasting more fluctuating traffic types. Fig. 5 shows SHAP summary plot for the RF and MLP for traffic *Type 5*. The plots show that, in both cases, the "previous_value" feature dominates the models' predictions and exhibits significant feature importance. This is also the case for traffic *Type 3* (not shown in figure due to space limitations); however, the "day_go_value" and

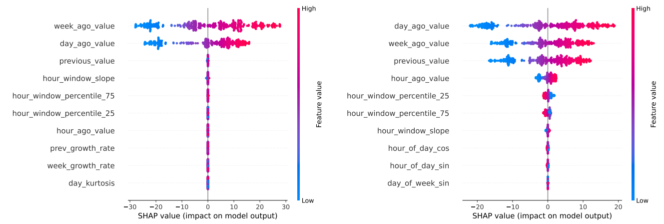

Fig. 4: SHAP summary plots for traffic *Type 0*, examples of the RF (left) and MLP (right) regressors.

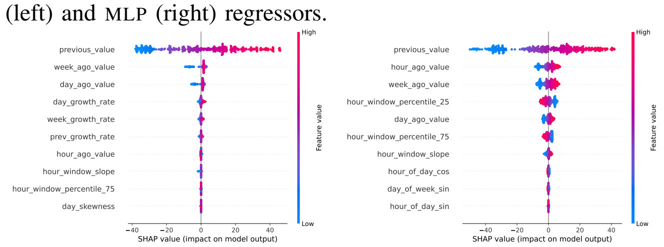

Fig. 5: SHAP summary plots for traffic *Type 5*, examples of the RF (left) and MLP (right) regressors.

"week_ago_value" features still play a slightly more significant role than in traffic *Type 5*. Such a trend is expected as we increase the noise levels in the subsequent datasets. With the relatively high sampling rate of 5 minutes, the models learn to rely on neighboring samples while only scarcely using direct seasonality information. For the rest of the features, similar to the case of traffic *Type 0*, the plots show they do not make any significant contributions to the models' decision-making process, irrespectively of the ML algorithm employed. *We can conclude that, regardless of the ML algorithm employed, in case of fluctuating traffic types, the models construct their forecasts mostly based on the directly preceding samples, without relying on features relative to other past observations.* Note that literature surrounding ML-based traffic prediction has always assumed measurements (and hence, features) pertaining for that extend beyond the immediate preceding samples. Therefore, our findings introduce a new direction for investigation for feature engineering for ML-based traffic prediction.

b) Impact of the 30-minute aggregation of traffic: We now focus on the impact of using a 30-minute aggregation of traffic data on features' contribution to the models' decisions. In this case, the subsequent samples are much less correlated as they describe broader periods (see illustrative example in Fig. 3), which might present a challenge for the prediction models. Consequently, the feature "previous_value" represents the maximum traffic from the past 30 minutes, as opposed to the one taken 5 minutes before in the previous case. Fig. 6 shows two examples of SHAP summary plots for the resampled dataset of traffic *Type 5*. Similar trends

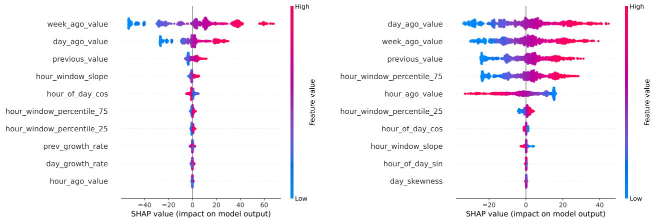

Fig. 6: SHAP summary plots for traffic *Type 5* after the 30-min aggregation, examples of the RF (left) and MLP (right) regressors.

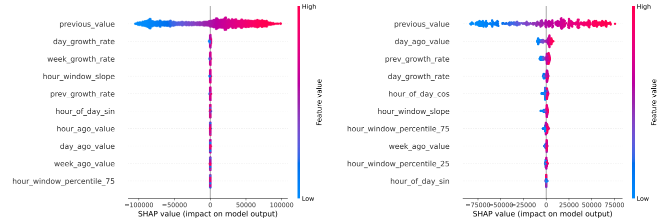

Fig. 7: SHAP summary plots of RF regressor obtained with real traffic with 5-min sampling (left) and with 30-min aggregation (right).

are visible for all the remaining ML models and datasets (omitted here). We notice that the models once again rely almost solely on the *temporal* features. Interestingly, the feature ranking does not change significantly among datasets, and the "day_go_value" and "week_ago_value" are the most important ones despite the higher fluctuation levels. The only discrepancy we notice is that the MLP relies significantly on "hour_window_percentile_75". This example illustrates how distinct types of learning algorithms have the capability to extract knowledge and subsequently establish correlations in unique ways (in fact, the MLP outperforms, albeit slightly, RF in terms of MAPE (see Tab. I)). Despite this difference, the top-3 and the 5$^{\text{th}}$ most important features pertain to the *temporal* set of features. *This once again shows the highly informative value of the temporal features, which are the basis for the predictions of various* ML *models regardless of the operated traffic sampling rate.*

The knowledge extracted from analyzing the models' behavior should be exploited during the feature engineering process for traffic forecasting and, together with the quantitative evaluation, should be examined when deciding on suitable ML model to employ in a given scenario.

c) Models' Behavior on Real Data: We now examine features' contribution when models are trained for the traffic prediction task using real data. The objective of this analysis is to investigate whether the found trends and dependencies would hold in real-world settings to confirm the versatility of our study. Fig. 7 shows SHAP summary plots for the 5-minute sampled and 30-minute aggregated data obtained using the RF regressor[2]. The plots show that "previous_value" dominates the model's predictions in both cases, with no contribution by other features in the case of 5-minute sampling and very minimal contributions by other features in the case of 30-minute aggregation of data traffic. These findings prove that the model can only rely on the last seen observation ("previous_value") to make its predictions. This behavior is consistent with that of the semi-synthetic datasets with traffic *Type 5*, which indicates the highly variable nature of real traffic. However, similar to our previous analysis, the most impacting contribution to the model output is generated by features from the *temporal* group. This confirms the versatility of our study – the identified feature contribution trends hold for different fluctuation levels and sampling/aggregation rates for semi-synthetic and real data.

[2] SHAP summary plots obtained from the remaining regressors (omitted here due to space limitations) show similar trends.

V. Predictive Performance and Practical Implementation

In Sec. IV, we showed how only a subset of features (mainly from the *temporal* group) contribute to the models' predictions. Among those, the exact feature choice depends on the fluctuation levels of the traffic to be forecasted. We now leverage this knowledge to train models using those inputs, namely, "day_ago_value," "week_ago_value" and "previous_value", and examine the implications of employing an ML model trained using this set of features in a practical implementation.

A. MAPE, Inference and Training Times

Table I reports the average MAPE computed across all datasets, the inference and training times of all three models considering 5-minute samples and 30-minute aggregation when training using all 17 features and when only using the *temporal* features selected based on the findings of the SHAP analysis. In more detail, based on our investigation, for the semi-synthetic data, the most-contributing features are the "day_ago_value," "week_ago_value," and "previous_value," so those three were used. However, for the real data, the first two had a less meaningful impact, so we used two features describing samples closer in time: "previous_value" and "hour_ago_value." For the 30-minute aggregation, we used all four *temporal* features for all the traffic types.

The results show that, regardless of the ML model and traffic type, there is no impact on prediction quality despite using 3 or 4 features instead of 17. In many cases, the performance in terms of MAPE improves when using only the selected feature set. In particular, considering the original 5-minute sampling, for the most repetitive traffic *Type 0*, the LR and MLP obtained better prediction quality when using fewer features. RF yielded slightly worse but very comparable results. With the more fluctuating traffic types, there are still instances with lower MAPE for models when relying only on the *temporal* features chosen with SHAP. In other cases, the differences are marginal. Furthermore, the above-described trends also hold for the real data and after the 30-minute aggregation, confirming the versatility of the found trends and dependencies. We can conclude that using only the few selected features does not have a notable impact on the prediction quality and, in many cases, can even slightly improve it.

Another advantage of reducing the number of features and thus using more lightweight models is their improved training and inference speed, quantified in Tab. I. For conciseness, we report the averaged results for models with the 5-minute

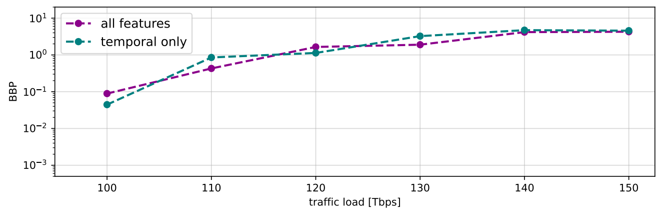

Fig. 8: Bandwidth blocking probability (BBP), averaged across ten simulation on US26 topology.

sampling for traffic *Type 0*, but the trends still hold for other traffic types and after aggregation. The training and inference times are clearly affected by the number of features in all the ML algorithms, showing an evident advantage of using fewer of them. On average, the training time is reduced by 56%, while the inference time is 13% shortened. Although one can argue that the training time does not have a viable real-world impact, the inference time is an important issue when using traffic forecasts in practice. Obtaining the predictions faster allows more time for their processing and utilization in allocation algorithms.

B. Bandwidth Blocking Probability

We now examine whether the slightly higher MAPE experienced when using only the temporal features impacts network operation in terms of *bandwidth blocking probability* (BBP). To this end, we utilize our recently proposed RSA algorithm for multilayer networks with time-varying traffic, operating with traffic predictions [2]. The connection requests take the form of *intents* – continuous signals of various network-based applications and services with bitrate varying throughout the day (see Fig. 1), as in [21]. The algorithm dynamically adapts to changing conditions and uses traffic grooming for better utilization of spare bandwidth. The knowledge coming from traffic prediction enables more informed decisions in the long term, i.e., for each intent with an increasing bitrate trend, the algorithm makes all the allocation decisions using its maximum predicted bitrate from the upcoming period, preventing grooming tightly fitting requests to avoid their frequent reallocations and increase network stability. For more details about the network model and algorithm, we refer to [2].

Fig. 8 shows the BBP with respect to traffic load. We repeated the simulations using the bitrate predictions for all intents coming from models using all features and then models using the *temporal* features only. Results reveal that there is no noticeable difference between the two cases. Any blocking appears at the same traffic load in both cases, and the difference in the amount of blocked bandwidth is marginal. In other words, there is no notable BBP change when using smaller prediction models, utilizing only the *temporal* features. Simultaneously, simplifying the models decreases the inference time and thus allows quicker forecasting, enabling the algorithm to use them and react to bitrate changes proactively in practice. Additionally, training and storing smaller models aligns with the green networking paradigm.

VI. Conclusions

In this work, we employed XAI to examine the behavior of traffic prediction models for backbone optical networks and extract insights into their operation that can serve their design process. We examined the contribution of input features from the literature on the outcome of different ML models designed for traffic forecasting. Through a broad analysis with various real and semi-synthetic datasets with different sampling rates, we discovered that the predictors base their decisions almost solely on the *temporal* features describing the traffic in highly correlated past samples. We then verified this conclusion by retraining all models with only those features and confirmed it by showing marginal changes in the prediction quality. At the same time, we demonstrated multiple benefits coming from the newly-shrunken feature sets, including significant problem simplification and shorter model training and inference times. Based on our study, we thus recommend training traffic prediction models for backbone optical networks using *temporal* input features as they enable fast and accurate forecasting. Finally, we demonstrated the practical application of the conducted analysis on network operation. We showed how using smaller prediction models has almost no impact on the BBP for the dynamic RSA problem. In the future, we plan to further investigate XAI for the optimization of various ML-assisted networking tasks.

References

[1] A. Valkanis *et al.*, “A traffic prediction assisted routing algorithm for elastic optical networks,” CCCI 2021.

[2] A. Knapińska *et al.*, “On advantages of traffic prediction and grooming for provisioning of time-varying traffic in multilayer networks,” ONDM 2023.

[3] Sandvine, “The mobile internet phenomena report,” May 2021.

[4] Seattle Internet Exchange, https://www.seattleix.net/statistics/.

[5] I. Lohrasbinasab *et al.*, “From statistical- to machine learning-based network traffic prediction,” *Trans. on Emerg. Telecommun. Technol.*, 2022.

[6] N. Krishnaswamy *et al.*, “Data-driven learning to predict WAN network traffic,” SNTA 2020.

[7] G. O. Ferreira *et al.*, “Forecasting network traffic: A survey and tutorial with open-source comparative evaluation,” *IEEE Access*, 2023.

[8] S. Narejo *et al.*, “An application of internet traffic prediction with deep neural network,” *Multidiscipl. Approaches to Neural Computing*, 2018.

[9] Q. Zhuo *et al.*, “Long short-term memory neural network for network traffic prediction,” ISKE 2017.

[10] A. Knapińska *et al.*, “On feature selection in short-term prediction of backbone optical network traffic,” ONDM 2022.

[11] S. M. Lundberg *et al.*, “A unified approach to interpreting model predictions,” NeurIPS 2017.

[12] S. Li *et al.*, “Network traffic prediction based on the feature of newly-generated network flows,” IFIP Networking 2022.

[13] S. S. Sepasgozar *et al.*, “Network traffic prediction model considering road traffic parameters using artificial intelligence methods in VANET,” *IEEE Access*, 2022.

[14] A. Adadi *et al.*, “Peeking inside the black-box: a survey on explainable artificial intelligence (XAI),” *IEEE Access*, 2018.

[15] O. Ayoub *et al.*, “Towards explainable artificial intelligence in optical networks: the use case of lightpath qot estimation,” *JOCN*, 2022.

[16] O. Karandin *et al.*, “If not here, there. Explaining machine learning models for fault localization in optical networks,” ONDM 2022.

[17] O. Ayoub *et al.*, “On using explainable artificial intelligence for failure identification in microwave networks,” ICIN 2022.

[18] A. Morichetta *et al.*, “Explain-it: Towards explainable AI for unsupervised network traffic analysis,” Big-DAMA 2019.

[19] P. Lechowicz *et al.*, “Traffic Weaver: semi-synthetic time-varying traffic generator based on averaged time series,” 2024.

[20] A. Włodarczyk *et al.*, “An algorithm for provisioning of time-varying traffic in translucent SDM elastic optical networks,” ICTON 2020.

[21] L. Velasco *et al.*, “Intent-based networking and its application to optical networks,” *JOCN*, 2022.