# **ONT | Task-7**

## **Explainable AI**

## **Introduction**

This report provides an in-depth analysis of Explainable AI (XAI) techniques applied to optical communication networks and compares the performance of regression models on the EU dataset. We focused on the Global Signal-to-Noise Ratio (GSNR) in optical networks and assessed how various factors impact GSNR using XAI methods. Additionally, we evaluated regression models trained on the EU dataset to understand their performance and interpretability.

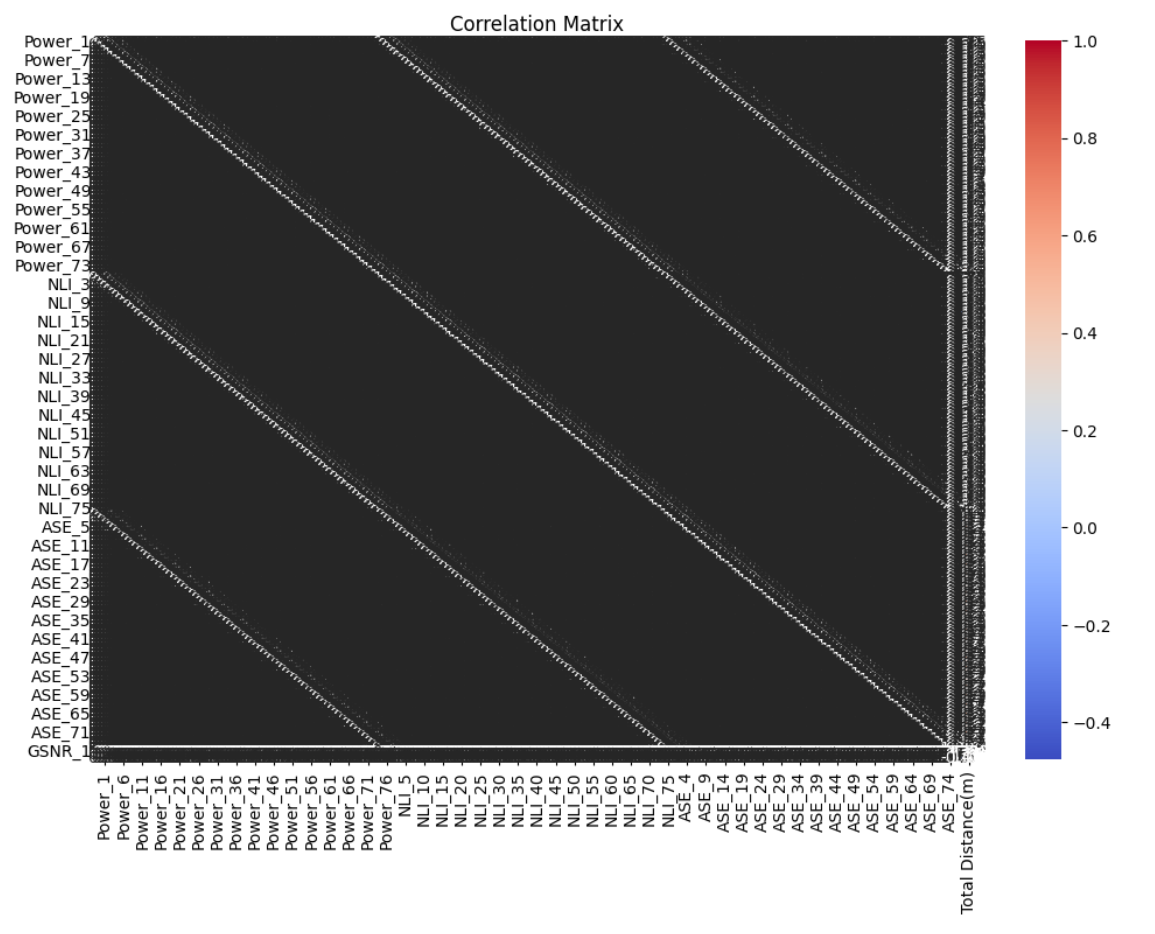
## **EU Dataset Analysis**

**Data Preparation**

The EU dataset consisted of numeric data tailored for regression tasks, featuring 230 features and 15,000 data points for training, with an additional 3,000 points reserved for testing. The data was preprocessed to handle missing values and normalized using the MinMaxScaler technique. This normalization ensured that all features were scaled between 0 and 1, improving model convergence during training.

The correlation matrix below provides an overview of the relationships between different features in the dataset. The diagonal elements represent perfect correlation (1.0), while the off-diagonal elements offer insight into the pairwise correlations between features.

**Correlation Matrix:**



## **Model Training & Evaluation**

Several regression models were trained on the EU dataset to compare performance. The models used include:

* Linear Regression
* Random Forest Regressor
* Gradient Boosting Regressor
* Decision Tree Regressor

Each model was trained on the MinMax normalized data, and the performance was evaluated using the Mean Squared Error (MSE) metric. Below is the summary of the model training results:

## **Model Performance on EU Dataset:**

* **Linear Regression:** MSE = 0.046565
* **Random Forest Regressor:** MSE = 0.000035
* **Gradient Boosting Regressor:** MSE = 0.000028
* **Decision Tree Regressor:** MSE = 0.000071

As seen in the results, the Gradient Boosting Regressor and Random Forest Regressor performed exceptionally well, with minimal errors. However, understanding the decision-making processes behind these models is crucial for ensuring transparency and trustworthiness, which is where Explainable AI (XAI) techniques come into play.

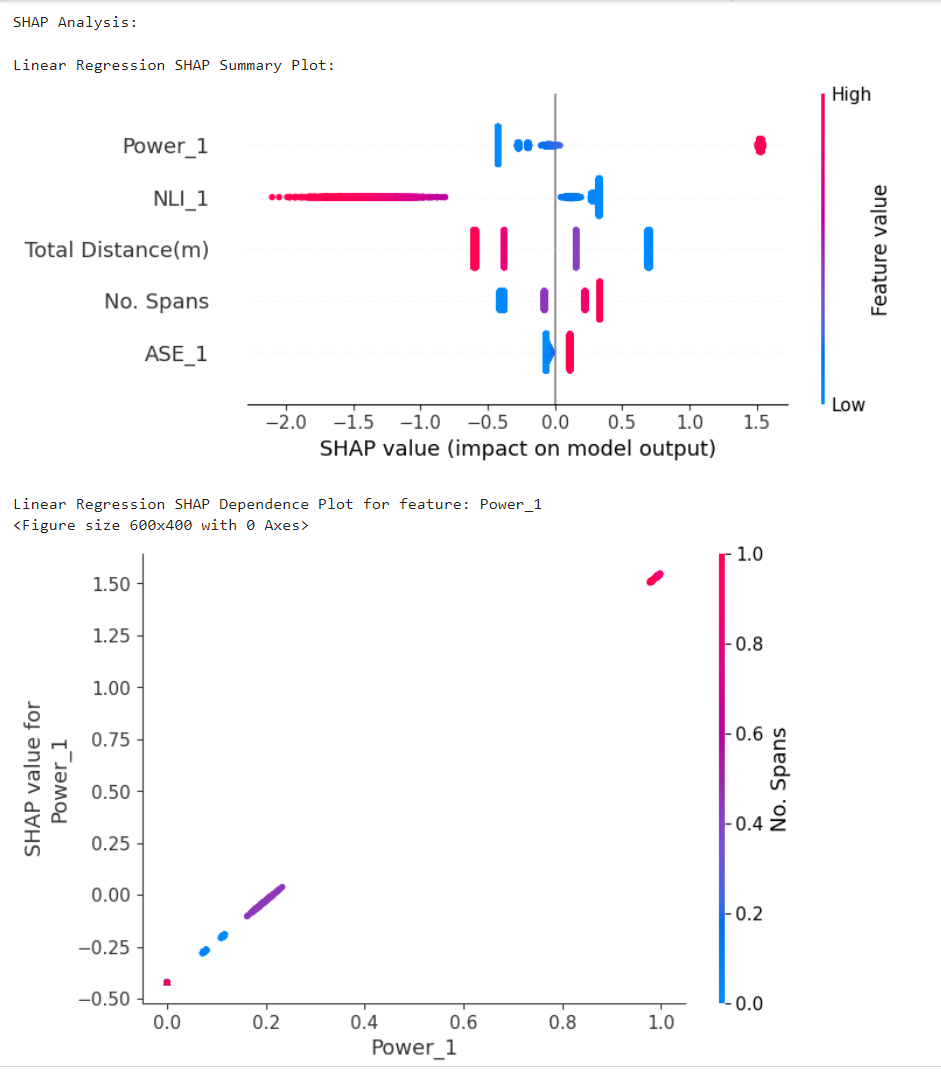
## **Explainable AI (XAI) Implementation**

To interpret the models trained on the EU dataset, various XAI techniques were applied. These methods provided insights into how features influence model predictions and allowed for a comparison of these insights with the original regression model predictions.

The following XAI methods were utilized:

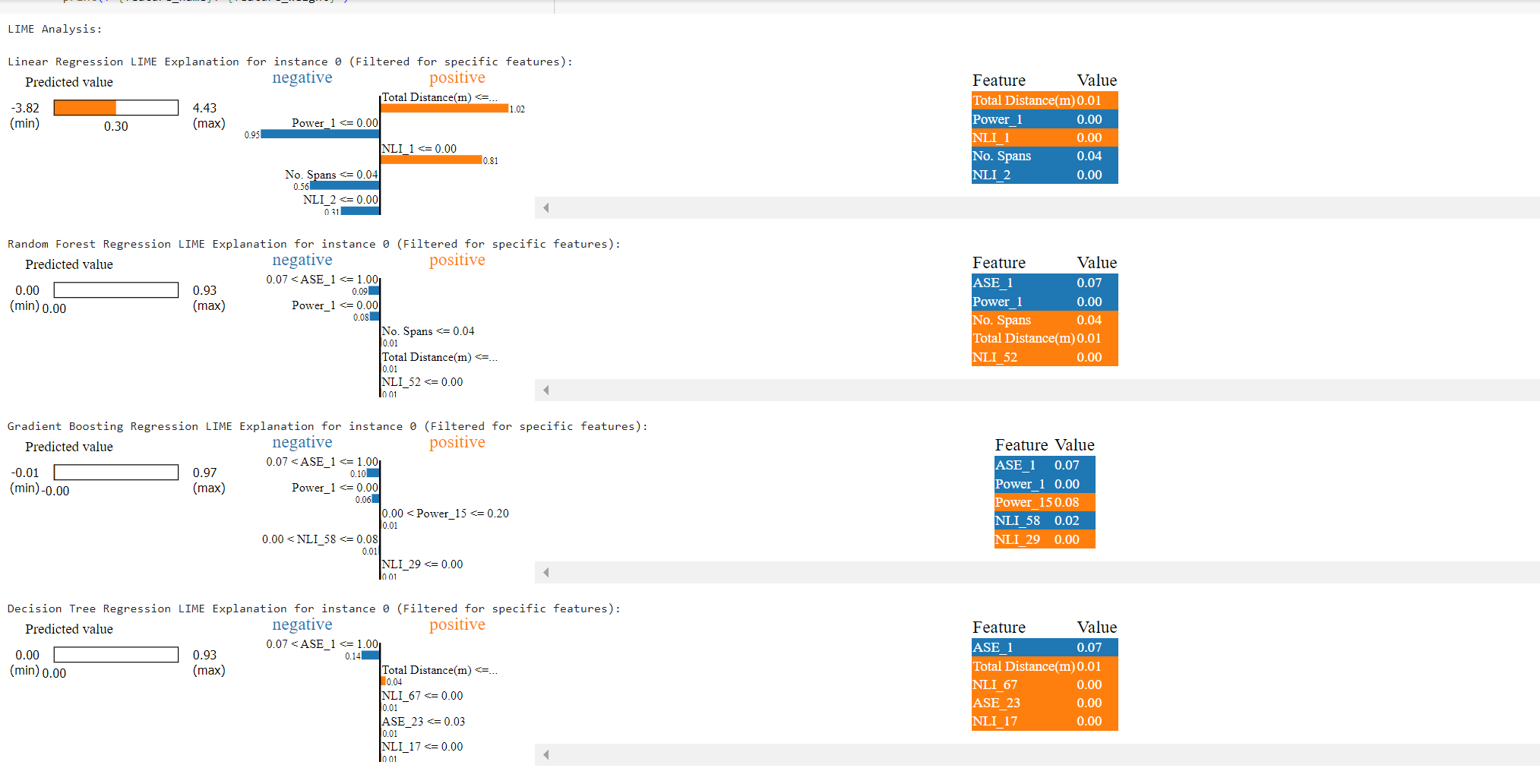
* **SHAP (SHapley Additive exPlanations):** This technique explains individual predictions by assigning each feature an importance value.
* **LIME (Local Interpretable Model-agnostic Explanations):** LIME builds local surrogate models to approximate and explain complex models.
* **Partial Dependence Plots (PDPs):** PDPs show the marginal effect of a feature on the predicted outcome.
* **Individual Conditional Expectation (ICE) Plots:** ICE plots display the relationship between a feature and the predicted outcome for individual instances.

**SHAP Values:**



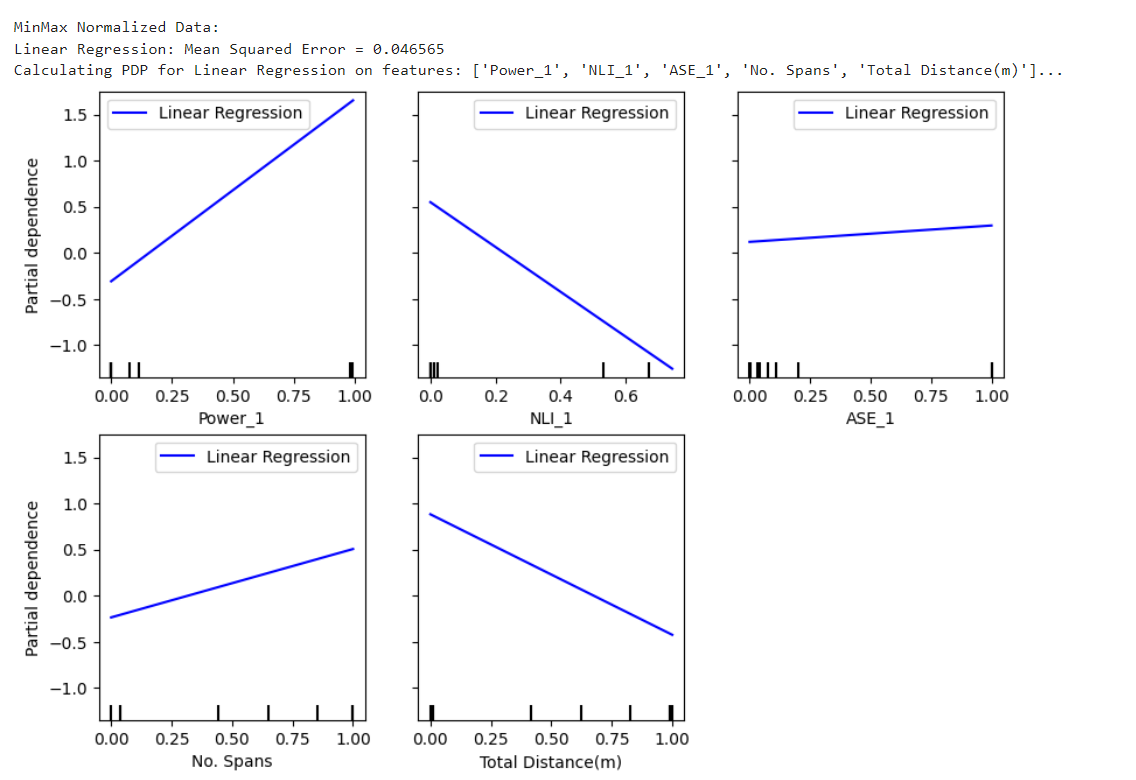
**Analysis:** The SHAP values highlighted that certain features had a significant impact on the model’s predictions. These insights align with the domain knowledge of the EU network.

**LIME Explanations:**



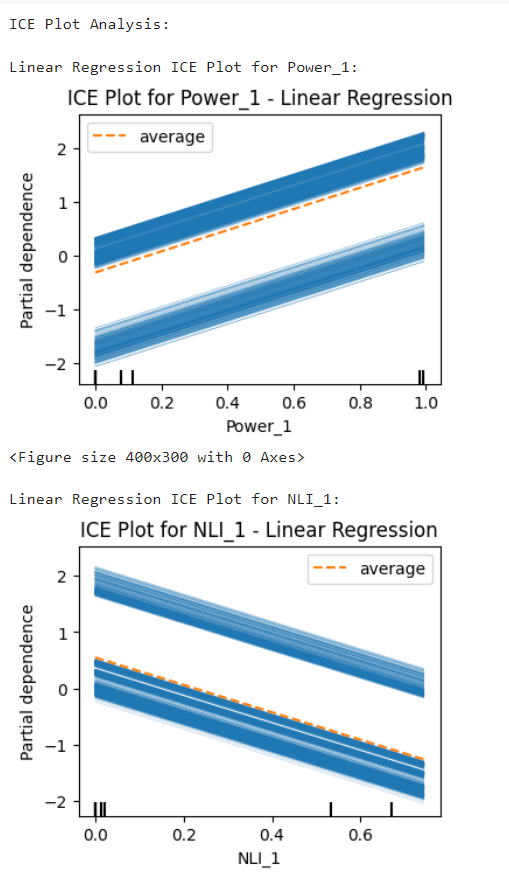
**Analysis:** LIME explanations reinforced the importance of the top features identified by SHAP. The local surrogates provided additional clarity on model behavior for specific instances.

**Partial Dependence Plots (PDPs):**



**Analysis:** The PDPs illustrated that the model’s predictions were sensitive to changes. This understanding can be vital for feature engineering and further model refinement.

**Individual Conditional Expectation (ICE) Plots:**



**Analysis:** ICE plots revealed the variability in individual predictions, offering insights into how different data points respond to changes in a particular feature.

## **Impact of Key Factors on GSNR in Optical Communication Networks**

This analysis delved into the impact of key factors such as Power, Amplified Spontaneous Emission (ASE), Nonlinear Interference (NLI), Number of Hops, and Total Distance on the Global Signal-to-Noise Ratio (GSNR) in optical communication networks. By leveraging various interpretability techniques, including SHAP, Partial Dependence Plots (PDPs), LIME, and Individual Conditional Expectation (ICE) plots, we gained insights into how these factors influence GSNR across different models like Linear Regression, Random Forest, Gradient Boosting, and Decision Trees.

**Key Takeaways:**

* **Power:** An increase in power generally led to an increase in GSNR, though the effect was sometimes nonlinear, with optimal power levels being more beneficial before nonlinear effects like NLI start to degrade the signal.
* **Amplified Spontaneous Emission (ASE):** ASE consistently showed a significant impact across models. Higher ASE levels tended to degrade GSNR, particularly in Random Forest and Gradient Boosting models.
* **Nonlinear Interference (NLI):** NLI was critical, especially in scenarios with higher power or longer distances. Higher NLI levels generally reduced GSNR, aligning with theoretical expectations.
* **Number of Hops:** More hops generally reduced GSNR due to cumulative noise and nonlinear effects. However, mid-range values for hops showed higher GSNR in some models, suggesting an optimal balance.
* **Total Distance:** Longer distances typically reduced GSNR due to attenuation, ASE, and NLI. Some models showed that mid-range distances sometimes offered better GSNR than very short or very long paths.

**Insights from Model Interpretations:**

* **SHAP Values:** Highlighted the importance and direction of each feature's impact on GSNR. Power, ASE, and NLI were the most significant factors, with ASE often showing the strongest negative impact.
* **PDP and ICE Plots:** Confirmed the non-linear relationships and the importance of finding optimal settings for Power and NLI.
* **LIME Explanations:** Showed ASE as a critical factor across different models, with varying impacts depending on the specific model.

## **Results & Insights**

The XAI techniques provided a deeper understanding of the models' behavior on the EU dataset and GSNR in optical networks. SHAP and LIME offered granular explanations, while PDPs and ICE plots highlighted feature interactions and their effects on predictions. These findings emphasize the need for transparency in model predictions and optimization strategies, ensuring that models can be trusted and effectively deployed in real-world scenarios.

### **Note: A similar analysis was conducted on the USA dataset, which also involved evaluating the impact of key factors on the Global Signal-to-Noise Ratio (GSNR) in optical communication networks. The USA dataset, with its distinct values and distribution, revealed comparable insights into how factors like Power, Amplified Spontaneous Emission (ASE), Nonlinear Interference (NLI), Number of Hops, and Total Distance affect GSNR. For instance, the relationship between Power and GSNR, while consistent in its general trend, exhibited differences in optimal levels and nonlinear effects due to the unique characteristics of the USA dataset. Similarly, ASE and NLI demonstrated significant impacts, with variations in their influence reflecting regional differences in data. Overall, while the fundamental patterns were aligned with those observed in the EU dataset, the specific values and interactions provided nuanced insights relevant to the USA dataset's context.**