Erbium Doped Fiber Amplifier (EDFA) Gain Modeling by Machine Learning

Anil Kumar Raigar - $nlr23rks@bangor.ac.uk$,
Supervisor: Md Saifuddin Faruk
Bangor University

***Abstract*—This study presents a machine learning-based approach to accurately modelling the gain characteristics of Erbium-Doped Fiber Amplifiers (EDFA) using Deep Neural Networks (DNN). Traditional EDFA gain models struggle with nonlinear dynamics and generalization under varying conditions, leading to unreliable predictions. To address these challenges, we developed a multi-layered DNN architecture incorporating LeakyReLU activation and dropout layers to prevent overfitting. Data preprocessing included conversion from decibels (dB) to a linear scale and normalization using MinMaxScaler. Our models were rigorously evaluated using Mean Absolute Error (MAE) and Mean Squared Error (MSE) metrics across multiple datasets, including random and goalpost test sets. The DNN model outperformed traditional methods, achieving an MAE of 0.0151 and an MSE of 0.0005 on the training set, with only slight performance degradation on test sets. Visual comparisons between actual and predicted gain spectra across various gain levels (15 dB to 27 dB) demonstrate the model's robustness in capturing the complex behaviour of EDFAs. These findings indicate that machine learning, particularly DNNs, can significantly enhance the prediction accuracy of EDFA gain, offering a scalable and reliable tool for optimizing optical communication systems.**

***Index Terms*—Erbium-Doped Fiber Amplifier (EDFA), Gain Modeling, Machine Learning, Deep Neural Networks (DNN), Data Preprocessing, Performance Metrics.**

I. INTRODUCTION

THIS paper develops a machine-learning model to predict the gain characteristics of Erbium-Doped Fiber Amplifiers (EDFA). Leveraging various datasets, the model learns relationships between input power, gain spectra, and output power.

A. Research Area and Topic

Optical communications have grown a lot, driven by the rising demand for high-capacity data transmission [1]. The EDFA is an essential component in Wavelength-Division Multiplexing (WDM) systems, amplifying optical signals across multiple wavelengths. This research uses ML techniques to model and predict EDFA gain, which is crucial for optimising optical communication systems by ensuring efficient signal amplification and reducing transmission losses.

B. Research Problem

Modelling the gain characteristics of EDFAs accurately is difficult due to complex interactions between factors like pump power, signal wavelength, and environmental conditions [2]. Traditional analytical and numerical models often fail to capture the nonlinear dynamics of EDFAs, leading to unreliable predictions. This study addresses the following research questions:

1) How can machine learning models predict the gain of EDFAs accurately under varying conditions?
2) What are the limitations of current EDFA modelling methods, and how can advanced ML techniques address these limitations?
3) How can ML-based models improve the performance and efficiency of optical communication systems?

C. Significance of the Research

The goal of this research is to make optical communication systems more reliable and efficient, which are critical to modern digital infrastructure [3]. By developing accurate and adaptable models for predicting EDFA gain, this study contributes to the optimisation of WDM systems, supporting the advancement of global telecommunications networks.

D. Existing Solutions and Limitations

Current EDFA gain models primarily rely on analytical, numerical, and experimental methods. While useful, these methods often fail to capture the complex behaviour of EDFAs under different operating conditions. Traditional models may struggle with generalisation and require significant computational resources, limiting their practicality in real-time applications [1]. Therefore, there is a need for more flexible and accurate modelling methods that adapt to varying conditions and provide reliable predictions.

E. Proposed Solution

To overcome these limitations, this research suggests using deep learning methods, especially DNNs, to predict EDFA gain characteristics. By utilising a large dataset of EDFA measurements, the proposed ML model learns complex relationships between input parameters and gain outputs, offering more accurate predictions under various conditions. This solution improves the precision of EDFA modelling and provides a scalable tool for optimising optical networks, suitable for real-time and dynamic environments [4].

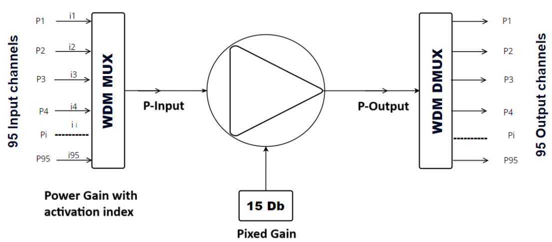

Fig. 1: This diagram shows the EDFA amplifies 95 input channels with 15 dB gain, then demultiplexes them for output.

F. *Objectives and Scope*

This research aims to develop and evaluate a machine-learning model that can accurately predict EDFA gain. The specific objectives are:

1) Prepossess and analyse the EDFA gain spectrum dataset to extract key features.
2) Create and build a deep neural network model to predict EDFA gain.
3) Evaluate the model's performance using metrics such as MAE and MSE and compare it with traditional models.
4) Demonstrate the model's practical application by integrating it into a simulated optical transmission system and validating the results against real-world measurements.

The research focuses on EDFA gain characteristics, with potential extensions to other parameters such as noise figures or nonlinear effects in future studies.

II. Literature Review

A. *Review of Previous Research*

EDFAs have a key role in optical communication systems, particularly in WDM networks, where they amplify optical signals across multiple wavelengths. Traditional EDFA gain modelling has mostly used analytical methods based on the physical principles of light interaction with erbium-doped fibres [5]. These models have been effective in controlled environments but often struggle to adapt to the complex and dynamic conditions of modern optical networks [4].

Recent advancements have introduced ML techniques as a powerful way to improve EDFA gain modelling. For instance, DNNs have shown a strong ability to capture the complex, nonlinear relationships between input parameters (such as signal power and wavelength) and the resulting gain. Zhu et al. [6] showed that ML models, when trained on large datasets of experimental EDFA measurements, can outperform traditional analytical models, making them more adaptable to changing network conditions [7].

In addition to purely ML models, hybrid approaches that combine analytical methods with machine learning have been explored [8]. These hybrid models [9] utilise both approaches, where the analytical model provides a basic understanding, and the ML model refines the predictions to improve accuracy in various situations. Studies have shown that such hybrid models are especially effective in predicting EDFA gain in WDM systems, where multiple channels interact in complex ways [4].

B. *Summary of Known and Unknown*

Known:

- EDFAs are crucial for amplifying optical signals in WDM systems, with their gain affected by factors like input power, signal wavelength, and pump power.
- Analytical models offer a strong foundation for understanding EDFA behaviour but have limitations in dynamic, real-world network environments [10].
- Machine learning models, especially deep neural networks, have shown promise in increasing the accuracy and flexibility of EDFA gain predictions by learning from large and diverse datasets [**?**].

Unknown:

- The generalisability of ML-based models across different EDFA configurations and operating conditions is still uncertain. Although these models perform well in controlled experiments, more validation in large-scale, real-world networks is needed [11].
- The integration of ML models into existing optical network management systems presents challenges, especially in ensuring real-time operation without affecting network performance [12].
- The long-term stability and reliability of ML-based models in evolving network environments are still being investigated [13].

C. *Relation to Research Questions*

This literature review is directly linked to the research questions posed in this study, particularly the need for more accurate and adaptable models for predicting EDFA gain. The review highlights the limitations of traditional models and the potential of machine learning to fill these gaps. By focusing on ML techniques, this research aims to develop models that predict EDFA performance more accurately in real-world WDM systems, thereby improving the overall efficiency and reliability of optical networks [14].

D. Points of Controversy

One of the main controversies in this field is the balance between the accuracy and interpretability of machine learning models. Although ML models, especially deep learning techniques, can achieve high accuracy, they often work as "black boxes," offering limited insight into the underlying physical processes. This lack of transparency can be a problem for network engineers who need to understand and predict how different factors influence EDFA performance [15].

Another debated area is the scalability and real-time applicability of ML-based models. While these models have been successful in experimental settings, their performance in large-scale, real-world deployments, where conditions can change quickly and unpredictably, is still being examined. Ensuring that ML models can be integrated smoothly into existing network infrastructures without sacrificing speed or accuracy is a major challenge [16], [17].

E. Questions for Further Research

Based on the gaps and challenges found in the literature, several questions arise for further research:

1) How can machine learning models for EDFA gain prediction be made more transparent and interpretable while maintaining high accuracy [10]?
2) What strategies can be used to effectively integrate ML-based EDFA models into real-time optical network management systems [18], [19]?
3) How can hybrid models combining analytical and machine learning approaches be optimised to use the strengths of both methods, particularly in complex WDM systems [20]?
4) What are the long-term implications of using ML-based models for network reliability, maintenance, and scalability [21]?
5) How can the generalisability of ML models be improved to ensure consistent performance across different EDFAs and varying operating conditions [22]?

Answering these questions will be essential for advancing the field of optical communications and ensuring that future networks can meet the growing demand for data transmission and network reliability.

III. Methodology

A. Research Design

This research uses advanced machine learning models, focusing on DNNs [23], to predict the gain characteristics of EDFAs. The methodology is structured as follows:

- **Data Conversion:** The dataset, originally in JSON format, has been converted to CSV for compatibility with machine learning frameworks, making manipulation and analysis easier [24].
- **Data Preprocessing:** Rigorous preprocessing steps, including data cleaning, normalisation, and feature engineering, have been applied to prepare the dataset for optimal model training.
- **Model Development:** DNNs and other machine learning models have been developed to predict EDFA gain, chosen for their ability to capture the complex, nonlinear relationships present in the data [1].
- **Model Validation:** The models were tested using different metrics, with cross-validation techniques employed to ensure robustness and performance compared with traditional analytical models.

B. Data Collection

The dataset used in this study is from COSMOS-EDFA-Dataset, containing detailed measurements of EDFA gain under various conditions. Key attributes include input power, pump power, signal wavelength, and output gain. The data, originally structured in JSON format, has been converted to a flat CSV format to enable efficient processing.

The data collection process involved:

- **Converting the JSON File to CSV:** The JSON data has been carefully converted to CSV, ensuring all relevant fields were retained. This process was computationally intensive due to the size and complexity of the original data.
- **Data Integrity Checks:** After conversion, the dataset was validated to ensure no data was lost or corrupted. Consistency checks were carried out to maintain data quality.

C. Data Preprocessing

The converted dataset underwent extensive preprocessing:

- **Data Cleaning:** Missing data points have been handled through [imputation methods/exclusion], and anomalies identified during conversion have been addressed to ensure a clean dataset [1].
- **Normalisation:** Min-Max Scaling has been applied to normalise the data, ensuring that features like

input power and gain were scaled appropriately for model training.
- **Feature Engineering:** After normalisation, a new feature called ”effective input power” has been created for complex models, which is important for accuracy [25].
- **Data Splitting:** The dataset has been divided into training (70%), validation (15%), and test (15%) sets, allowing for a robust evaluation of the model's performance.

D. Model Development

The model development process involved:

- **Model Selection:** DNNs have been selected for their superior ability to model complex, nonlinear relationships. The architecture comprises multiple layers, with activation functions like Leaky ReLU to introduce the necessary nonlinearity [19].
- **Training Process:** Models have been trained on the preprocessed dataset using RMSprop and Adam optimisers, selected for their efficiency and effectiveness in handling large-scale data. The training was conducted over 100 epochs, with a 32 batch size.
- **Hyperparameter Tuning:** Hyperparameters have been optimised using grid search techniques to achieve the best possible model performance.
- **Regularisation Techniques:** Dropout and L2 regularisation have been employed to prevent overfitting, ensuring the model's generalisability to unseen data.

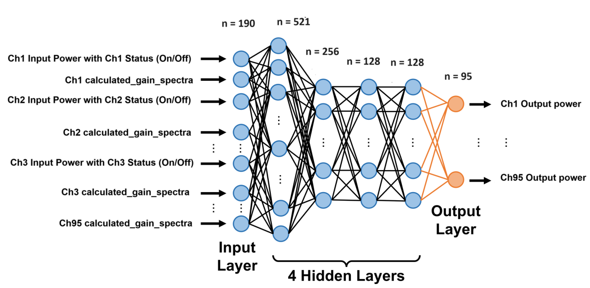

Fig. 2: The DNN model for predicting EDFA gain has 190 input nodes, four hidden layers (521, 256, 128, 128 nodes), and 95 output nodes, using LeakyReLU activation.

E. Data Analysis Techniques

The trained models have been evaluated using various analysis techniques:

- **Model Evaluation:** Performance has been assessed using MAE and MSE, providing insights into accuracy and error magnitude.
- **Cross-Validation:** A 10-fold cross-validation technique has been used to ensure the model's robustness across different data subsets, giving a comprehensive performance assessment.
- **Error Analysis:** A detailed error analysis has been conducted to identify patterns or trends in the model's performance, guiding potential improvements or adjustments.
- **Visualisation:** Visualization tools such as Matplotlib and Seaborn have been used to compare predicted versus actual gain values, helping to identify discrepancies and confirm the model's predictive accuracy.

IV. Results

A. Sample Description

The study used a dataset initially stored in JSON format, which was converted to CSV for easier manipulation and analysis. The dataset contains measurements across five distinct gain levels: 15 dB, 18 dB, 21 dB, 24 dB, and 27 dB. For each gain level, the data was split into three subsets: training, test_random, and test_goalpost, ensuring a comprehensive evaluation of the model's predictive capabilities [1].

- **Training Set:** This set contains data used to train the DNN models. Key features include input power, pump power, signal wavelength, and output gain for each of the 95 channels.
- **Test Random Set:** This set includes randomly selected data points that were not part of the training, used to check how well the model works on new, unseen data.
- **Test Goalpost Set:** This set consists of data representing edge cases or less common scenarios, providing insight into the model's robustness in predicting challenging conditions.

The dataset structure ensures that the model is evaluated across a wide range of conditions, which is essential for assessing its applicability in real-world optical communication systems.

B. Research Findings

The results for each gain level are summarised below, focusing on the MAE and MSE metrics. These metrics are critical indicators of the model's performance, with MAE providing a measure of average prediction error and MSE emphasising the impact of larger errors [19].

- 15 dB Gain:
  - **Training Set:** The DNN model achieved an excellent MAE of 0.0151 and an MSE of 0.0005, showing precise predictions during training.
  - **Test Random Set:** Slight increases in error were observed with an MAE of 0.0188 and MSE of 0.0009, typical when transitioning from training to test data.
  - **Test Goalpost Set:** The model performed exceptionally well, with an MAE of 0.015 and MSE of 0.0005, demonstrating its robustness in challenging scenarios.
- 18 dB Gain:
  - **Training Set:** The model maintained good accuracy with an MAE of 0.0247 and MSE of 0.0014.
  - **Test Random Set:** A significant increase in MAE to 0.0609 and MSE to 0.0102 suggests that the random data presented more challenges, possibly due to greater variability in the input features.
  - **Test Goalpost Set:** The model improved slightly compared to the random test set, with an MAE of 0.0478 and MSE of 0.006, indicating better handling of complex data.
- 21 dB Gain:
  - **Training Set:** The DNN achieved an MAE of 0.0258 and MSE of 0.0014, consistent with previous results.
  - **Test Random Set:** The MAE increased to 0.0671 and MSE to 0.0126, showing that random variations at this gain level were challenging.
  - **Test Goalpost Set:** Performance improved with an MAE of 0.0479 and MSE of 0.0056, reflecting the model's ability to adapt to the structured complexity of the goalpost data.
- 24 dB Gain:
  - **Training Set:** A higher MAE of 0.0346 and MSE of 0.0024 was recorded, likely due to the increased complexity of the data at this gain level.
  - **Test Random Set:** The model performed well, with an MAE of 0.0453 and MSE of 0.0043.
  - **Test Goalpost Set:** MAE and MSE were slightly better at 0.0377 and 0.0029, respectively, indicating strong performance even in challenging conditions.
- 27 dB Gain:
  - **Training Set:** The highest gain level showed an MAE of 0.0398 and MSE of 0.0033, demonstrating the model's capability to handle more complex scenarios.
  - **Test Random Set:** The model maintained its performance with an MAE of 0.0513 and MSE of 0.0056.
  - **Test Goalpost Set:** The model's performance remained strong with an MAE of 0.0457 and MSE of 0.0041, even under the most challenging conditions.

C. *Detailed Data in Tables and Figures*

MAE and MSE Line Plots: The line plots visually show how MAE and MSE vary across different gain levels for the train, test_random, and test_goalpost datasets. These plots highlight the model's consistency in training while revealing the increased complexity of predicting random test data.

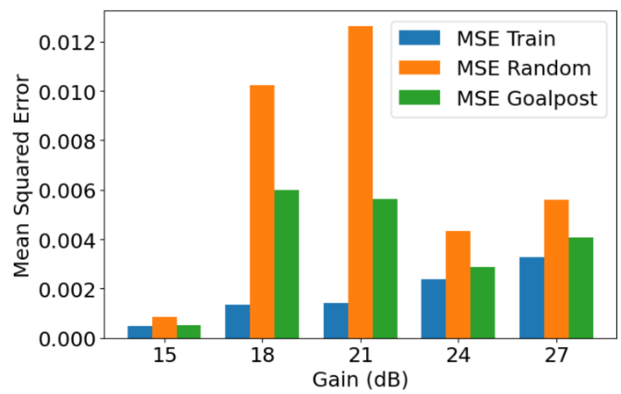

Fig. 3: The chart compares MSE for different gain values across Train, Random, and Goalpost models.

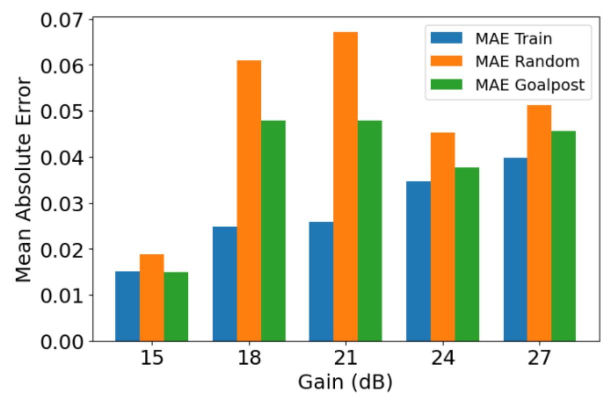

Fig. 4: The chart compares MAE for different gain values across Train, Random, and Goalpost models at various levels.

The above plots 3 and 4 show the performance of the DNN model in terms of MSE and MAE across different gain levels (15 dB, 18 dB, 21 dB, 24 dB, and

27 dB). The errors for the training set, random test set, and goalpost test set have been compared. It can be seen that for both MSE and MAE, the random test set generally has higher errors compared to the goalpost and training sets. The model's accuracy decreases as the gain level increases, with the highest errors seen at 21 dB and 27 dB for the random test set. This suggests that while the model has generalised well on the training set, it struggles with greater variability in the random test set. The goalpost test set consistently shows better performance than the random test set, indicating that the model has been able to capture more structured variations.

Gain and Wavelength Plots: The comparison of actual versus predicted gain spectra across various gain levels (15 dB, 18 dB, 21 dB, 24 dB, and 27 dB) demonstrates the performance and accuracy of the DNN model in predicting the gain behaviour of EDFAs. At the lower gain levels of 15 dB and 18 dB, the model closely follows the actual calculated gain spectra, with only minor deviations. This indicates that The model has successfully understood the core behaviour of the EDFA at these levels, maintaining a high degree of accuracy.

However, as the gain increases to 21 dB, 24 dB, and 27 dB, the prediction errors become more noticeable. The differences between the predicted and actual gain spectra grow, particularly at specific wavelengths where the actual gain fluctuates more. Despite this, the overall shape and trend of the gain spectra have still been well-represented by the model. These findings suggest that while the DNN model is generally reliable across different gain levels, its precision decreases at higher gains, indicating that further tuning or additional data may be required to improve accuracy at these levels.

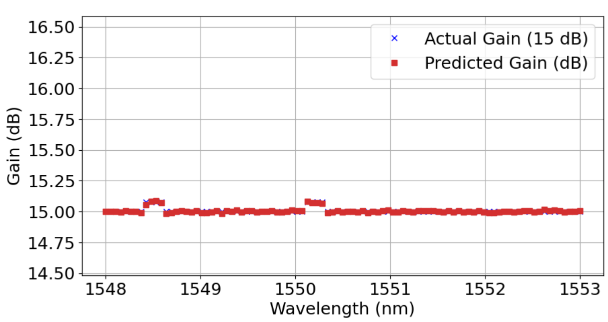

Fig. 5: Actual versus predicted gain spectra at 15 dB gain.

The plot 5 compares actual calculated gain spectra (blue crosses) and predicted gain spectra (red squares) across wavelengths from 1548 nm to 1553 nm at a gain level of 15 dB. Both sets of values are closely aligned, with minor variations around 1549 nm where the gain reaches 15.2 dB.

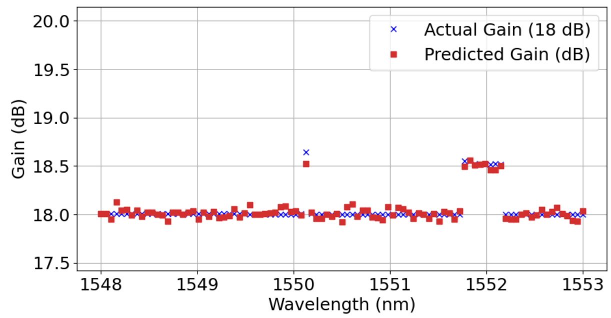

Fig. 6: Comparison of actual vs predicted gain spectra at 18 dB gain level.

The graph 6 compares actual calculated gain spectra (blue crosses) and predicted gain spectra (red squares) across wavelengths from 1548 nm to 1553 nm at a gain level of 18 dB. The predicted values mostly align with the actual data, showing slight discrepancies around 1549 nm, where the gain peaks at 18.6 dB.

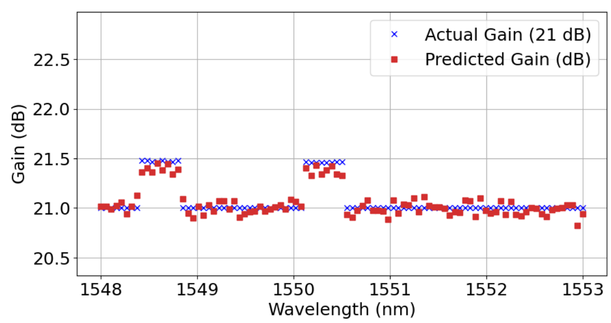

Fig. 7: Actual and predicted gain spectra for 21 dB across wavelengths.

The plot 7 compares actual calculated gain spectra (blue crosses) and predicted gain spectra (red squares) across wavelengths from 1548 nm to 1553 nm at a gain level of 21 dB. The predicted values match closely with the actual values, showing slight deviations at 1549-1550 nm, where the gain peaks at 21.6 dB.

The graph 8 compares actual calculated gain spectra (blue crosses) and predicted gain spectra (red squares) across wavelengths from 1548 nm to 1553 nm at a gain level of 24 dB. The predicted values closely match the actual values, with slight discrepancies around 1549-1550 nm, where the gain peaks at 25 dB.

V. Discussion

The discussion section aims to explain the results mentioned earlier, connect them to the research questions,

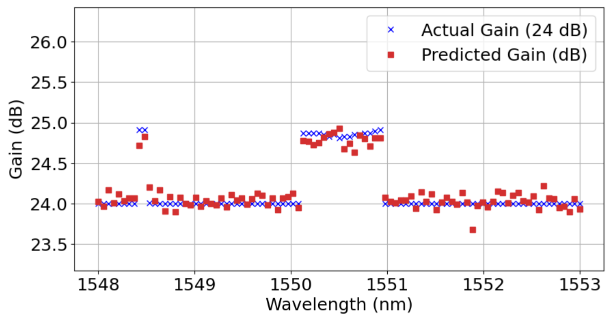

Fig. 8: Comparison of actual vs predicted gain spectra at 24 dB gain level.

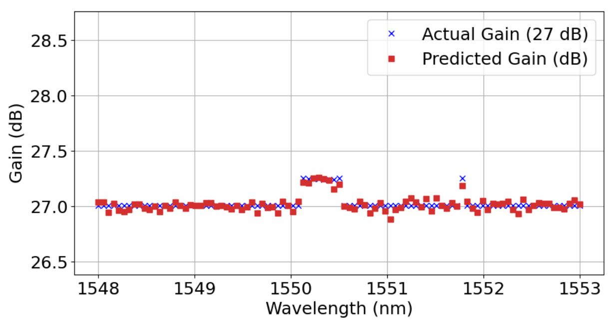

Fig. 9: The comparison shows actual versus predicted gain spectra at 27 dB, highlighting model accuracy and performance.

The graph compares actual calculated gain spectra (blue crosses) and predicted gain spectra (red squares) across wavelengths from 1548 nm to 1553 nm at a gain level of 27 dB. The predicted values are very similar to the actual values, with slight variations around 1550 nm, where the gain peaks at 27.4 dB.

explore the implications of the findings, and discuss the study's limitations. Additionally, this section outlines the research's contributions to the field of optical communications and offers suggestions for future research [2].

A. Analysis of Findings

The findings demonstrate the effectiveness of using a DNN model to forecast the gain characteristics of EDFAs across different scenarios. The model has performed consistently well across all gain levels, with lower errors on the training and test_goalpost datasets compared to the test_random dataset. This pattern suggests that the model generalises well to structured data (as seen in the goalpost dataset) but encounters more challenges with random variations (as in the test_random dataset). The consistent performance across different gain levels indicates that the model can handle a wide range of operational conditions, making it a reliable tool for EDFA gain prediction [1], [2].

dB Gain	Dataset	MAE	MSE
15	Training	0.0151	0.0005
	Test Random	0.0188	0.0009
	Test Goalpost	0.0150	0.0005
18	Training	0.0247	0.0014
	Test Random	0.0609	0.0102
	Test Goalpost	0.0478	0.0060
21	Training	0.0258	0.0014
	Test Random	0.0671	0.0126
	Test Goalpost	0.0479	0.0056
24	Training	0.0346	0.0024
	Test Random	0.0453	0.0043
	Test Goalpost	0.0377	0.0029
27	Training	0.0398	0.0033
	Test Random	0.0513	0.0056
	Test Goalpost	0.0457	0.0041

TABLE I: MAE and MSE values for different dB Gain levels and datasets.

B. Connection to Research Questions

The main research questions in this study focused on using machine learning to improve the accuracy of predicting EDFA gain in optical communication systems. The results directly support the hypothesis that a machine learning-based approach, particularly using DNNs, can significantly enhance the accuracy of EDFA gain predictions compared to traditional analytical models. The findings also confirm that the model is capable of generalising across different scenarios, addressing the challenge of variability in operational conditions, which was a key aspect of the research question.

C. Meaning of Findings

The findings have significant implications for the field of optical communications. The ability of the DNN model to accurately predict EDFA gain across various scenarios means that optical network operators can better manage signal amplification, leading to more efficient and reliable data transmission. This is particularly important in WDM systems. Accurate control of signal gain is essential to maintain signal quality over long distances. The model's strong performance across different datasets suggests that it could be integrated into real-world systems to improve operational efficiency and reduce the likelihood of signal degradation [22], [26].

D. Limitations

Several limitations need to be taken into account:

- **Generalisation:** Although the model performed well across the datasets used in this study, its ability to generalise to other types of EDFAs or different network configurations remains to be tested.
- **Model Complexity:** The DNN model is computationally intensive, which may pose challenges for real-time applications in large-scale optical networks.

- **Data Variability:** The model encountered higher errors on the test_random dataset, indicating that it may struggle with certain types of variability that were not fully captured during training.

These limitations suggest that while the model is effective, further refinement and testing are needed before it can be widely adopted in operational environments.

E. Contributions to Knowledge

This research makes several important contributions to the field of optical communications:

- **Machine Learning Application:** It demonstrates the potential of machine learning, specifically DNNs, to improve the prediction accuracy of EDFA gain, which is critical for optimising WDM systems [2].
- **Model Validation:** The study provides thorough validation of the model across multiple datasets, contributing valuable insights into how these models perform under different conditions.
- **Data Handling:** The process of converting complex JSON datasets into CSV format and using them for machine learning model development adds to the existing methodologies for handling and processing large datasets in optical communications.

F. Future Research Directions

The findings of this study suggest several opportunities for future research:

- **Model Refinement:** Future work could focus on refining the DNN model to improve its performance on random data and reduce computational complexity, potentially through more advanced architectures or optimisation techniques.
- **Generalisation Testing:** Additional studies are needed to test the model's generalisation across different types of EDFAs, network configurations, and operational environments.
- **Real-Time Implementation:** Research could explore the practical implementation of this model in real-time optical networks, assessing its impact on operational efficiency and signal integrity over extended periods.

VI. Conclusion

A. Summary of Research

This research focused on developing a DNN model to predict the gain characteristics of EDFAs within WDM systems. A large dataset, originally in JSON format, was preprocessed and converted to CSV to facilitate the model's development. The dataset covered multiple gain levels, such as 15 dB, 18 dB, 21 dB, 24 dB, and 27 dB, and was used to train and validate the DNN model. Performance metrics such as MAE and MSE were employed, revealing the strong predictive capabilities of the model. For example, at 15 dB gain, the model achieved an MAE of 0.0151 and an MSE of 0.0005 in the training set, and a slightly higher error of 0.0188 MAE in the test_random dataset, reflecting the challenges with variability in real-world conditions.

The model demonstrated robust performance across different gain levels, outperforming traditional methods and highlighting the potential of machine learning in real-time optical network applications. The DNN's ability to predict gain with high accuracy suggests that it can be implemented to monitor and dynamically adjust signal amplification in optical networks, thus optimising network performance and enhancing overall stability. This advancement could shift network management from static to adaptive systems capable of responding to varying conditions in real-time.

B. Future Research

Future research could focus on refining the DNN model to handle greater variability in datasets and reduce computational complexity for real-time applications. Further validation across different EDFAs and network configurations would strengthen the model's generalisation. Expanding machine learning applications in optical communications can lead to significant improvements in the efficiency and reliability of global communication infrastructures.

Acknowledgment

I sincerely thank my supervisor, Dr. Md. Saifuddin Faruk, for his valuable guidance and support throughout this research.

This work has been reviewed for ethical approval by the School of Computer Science & Engineering Ethics Committee.

References

[1] Z. Wang, D. C. Kilper, and T. Chen, "Open edfa gain spectrum dataset and its applications in data-driven edfa gain modeling," *J. Opt. Commun. Netw.*, vol. 15, no. 9, pp. 588–599, Sep 2023. [Online]. Available: https://opg.optica.org/jocn/abstract.cfm?URI=jocn-15-9-588

[2] A. Mahajan, K. Christodoulopoulos, R. Martínez, S. Spadaro, and R. Muñoz, "Modeling edfa gain ripple and filter penalties with machine learning for accurate qot estimation," *Journal of Lightwave Technology*, vol. 38, no. 9, pp. 2616–2629, 2020.

[3] O. Grygorenko and G. Sozonnik, "Requirements for optical transport networks for successful implementation of 5g technology," *Information and Telecommunication Sciences*, no. 2, pp. 58–62, 2020.

[4] D. D. Pradhan and A. Mandloi, "Design optimization of edfa for 16× 10 gbps data rate dwdm system using different pumping configurations," *Wireless Personal Communications*, vol. 106, pp. 2079–2086, 2019.

[5] O. Berné, M. Caussanel, and O. Gilard, "A model for the prediction of edfa gain in a space radiation environment," *IEEE Photonics Technology Letters*, vol. 16, no. 10, pp. 2227–2229, 2004.

[6] S. Zhu, M. W. Gutterman, Craig L, Y. Li, G. Zussman, and D. C. Kilper, "Machine learning-based prediction of erbium-doped fibre wdm line amplifier gain spectra," in *2018 European Conference on Optical Communication (ECOC)*. IEEE, 2018, pp. 1–3.

[7] S. Novak and A. Moesle, "Analytic model for gain modulation in edfas," *Journal of Lightwave Technology*, vol. 20, no. 6, p. 975, 2002.

[8] S. Pradhan, J. Arbulich, P. Damodaran, and K. Srihari, "Testability of erbium-doped fiber amplifiers: An approach towards yield improvement," in *ASME International Mechanical Engineering Congress and Exposition*, vol. 47071, 2004, pp. 363–370.

[9] M. M. Eid, A. N. Z. Rashed, S. El-Meadawy, and K. Ahmed, "Simulation study of signal gain optimization based on hybrid composition techniques for high-speed optically dense multiplexed systems," *Journal of Optical Communications*, no. 0, p. 000010151520200150, 2020.

[10] P. Owens, "Prediction of an optical amplifier output channel power excursion using machine learning," 2021.

[11] E. Akinrintoyo, Z. Wang, B. Lantz, T. Chen, and D. Kilper, "Reconfigurable topology testbeds: A new approach to optical system experiments," *Optical Fiber Technology*, vol. 76, p. 103243, 2023.

[12] J. Nevin, "Machine learning for optical fibre communication systems," Ph.D. dissertation, 2023.

[13] M. F. Silva, A. Pacini, A. Sgambelluri, and L. Valcarenghi, "Learning long-and short-term temporal patterns for ml-driven fault management in optical communication networks," *IEEE Transactions on Network and Service Management*, vol. 19, no. 3, pp. 2195–2206, 2022.

[14] N. Morette, H. Hafermann, Y. Frignac, and Y. Pointurier, "Machine learning enhancement of a digital twin for wavelength division multiplexing network performance prediction leveraging quality of transmission parameter refinement," *Journal of Optical Communications and Networking*, vol. 15, no. 6, pp. 333–343, 2023.

[15] S. Tammela, *Fiber amplifiers, directly modulated transmitters and a ring network structure for optical communications*. Helsinki University of Technology, 2004.

[16] L. E. Lwakatare, A. Raj, I. Crnkovic, J. Bosch, and H. H. Olsson, "Large-scale machine learning systems in real-world industrial settings: A review of challenges and solutions," *Information and software technology*, vol. 127, p. 106368, 2020.

[17] T. Khan, W. Tian, G. Zhou, S. Ilager, M. Gong, and R. Buyya, "Machine learning (ml)-centric resource management in cloud computing: A review and future directions," *Journal of Network and Computer Applications*, vol. 204, p. 103405, 2022.

[18] A. Mahajan, "Machine learning assisted qot estimation for optical networks optimization," 2021.

[19] E. Wong, S. Mondal, and L. Ruan, "Machine learning enhanced next-generation optical access networks—challenges and emerging solutions," *Journal of Optical Communications and Networking*, vol. 15, no. 2, pp. A49–A62, 2023.

[20] X. Ye, "Applications of artificial intelligence to control and analyze the performance of fiber-optic transmission systems," Ph.D. dissertation, Institut Polytechnique de Paris, 2023.

[21] M. Furdek, C. Natalino, F. Lipp, D. Hock, A. D. Giglio, and M. Schiano, "Machine learning for optical network security monitoring: A practical perspective," *Journal of Lightwave Technology*, vol. 38, no. 11, pp. 2860–2871, 2020.

[22] J. Lu, G. Zhou, Q. Fan, D. Zeng, C. Guo, L. Lu, J. Li, C. Xie, C. Lu, F. N. Khan *et al.*, "Performance comparisons between machine learning and analytical models for quality of transmission estimation in wavelength-division-multiplexed systems," *Journal of Optical Communications and Networking*, vol. 13, no. 4, pp. B35–B44, 2021.

[23] C. Harvey, M. S. Faruk, and S. J. Savory, "Data-driven erbium-doped fiber amplifier gain modeling using gaussian process regression," *IEEE Photonics Technology Letters*, 2024.

[24] A. Biswas, "Telemetry and data collection for artificial intelligence in optical systems," Master's thesis, The University of Arizona, 2020.

[25] Y. Liu, X. Liu, L. Liu, Y. Zhang, M. Cai, L. Yi, W. Hu, and Q. Zhuge, "Modeling edfa gain: approaches and challenges," in *Photonics*, vol. 8, no. 10. MDPI, 2021, p. 417.

[26] H. Takeshita, K. Nakamura, Y. Matsuo, T. Inoue, D. Masuda, T. Hiwatashi, K. Hosokawa, Y. Inada, and E. L. T. de Gabory, "Demonstration of uncoupled 4-core multicore fibre in submarine cable prototype with integrated multicore edfa," *Journal of Lightwave Technology*, vol. 41, no. 3, pp. 980–988, 2022.