FEDERAL UNIVERSITY OF PARÁ
INSTITUTE OF TECHNOLOGY
GRADUATE PROGRAM IN ELECTRICAL ENGINEERING

UNSUPERVISED LEARNING ALGORITHMS FOR DATA-DRIVEN FAULT MANAGEMENT IN OPTICAL NETWORKS

ANDREI NOGUEIRA RIBEIRO

DM 30/2024

UFPA / ITEC / PPGEE
Guamá University Campus
Belém-Pará-Brazil
2024

FEDERAL UNIVERSITY OF PARÁ
INSTITUTE OF TECHNOLOGY
GRADUATE PROGRAM IN ELECTRICAL ENGINEERING

ANDREI NOGUEIRA RIBEIRO

UNSUPERVISED LEARNING ALGORITHMS FOR DATA-DRIVEN FAULT MANAGEMENT IN OPTICAL NETWORKS

Master dissertation submitted to the Examining Board of the Graduate Program in Electrical Engineering from the Federal University of Pará to obtain the Master's Degree in Electrical Engineering, Area of Concentration in Applied Computing.

DM 30/2024

UFPA / ITEC / PPGEE
Guamá University Campus
Belém-Pará-Brazil
2024

UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

"ALGORITMOS DE APRENDIZADO NÃO-SUPERVISIONADO PARA O GERENCIAMENTO DE FALHAS ORIENTADO POR DADOS EM REDES ÓPTICAS"

AUTOR: **ANDREI NOGUEIRA RIBEIRO**

DISSERTAÇÃO DE MESTRADO SUBMETIDA À BANCA EXAMINADORA APROVADA PELO COLEGIADO DO PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA, SENDO JULGADA ADEQUADA PARA A OBTENÇÃO DO GRAU DE MESTRE EM ENGENHARIA ELÉTRICA NA ÁREA DE COMPUTAÇÃO APLICADA.

APROVADA EM: 09/12/2024

BANCA EXAMINADORA:

Prof. Dr. João Crisóstomo Weyl Albuquerque Costa
(Orientador - PPGEE / ITEC / UFPA)

Prof. Dr. Fabrício Rossy de Lima Lobato
(Coorientador - PPGEE / ITEC / UFPA)

Dr. Moisés Felipe Mello da Silva
(Avaliador Interno - PPGEE / ITEC / UFPA - LANL)

Dr. Carlos Natalino da Silva
(Avaliador Externo - CHALMERS UNIVERSITY OF TECHNOLOGY

VISTO:

Prof. Dr. Diego Lisboa Cardoso
(Coodernador do PPGEE / ITEC / UFPA)

Dados Internacionais de Catalogação na Publicação (CIP) de acordo com ISBD
Sistema de Bibliotecas da Universidade Federal do Pará
Gerada automaticamente pelo módulo Ficat, mediante os dados fornecidos pelo(a) autor(a)

R484a Ribeiro, Andrei Nogueira.
Algoritmos de aprendizado não-supervisionado para o gerenciamento de falhas em redes ópticas orientados por dados / Andrei Nogueira Ribeiro. — 2024.
53 f. : il. color.

Orientador(a): Prof. Dr. João Crisóstomo Weyl Albuquerque Costa
Coorientador(a): Prof. Dr. Fabrício Rossy de Lima Lobato
Dissertação (Mestrado) - Universidade Federal do Pará, Instituto de Tecnologia, Programa de Pós-Graduação em Engenharia Elétrica, Belém, 2024.

1. Redes ópticas. 2. Aprendizado de máquina. 3. Tolerância à falhas. 4. Aprendizado não-supervisionado. I. Título.

CDD 621.3

In honor of my family, especially my mother Carolina Nogueira.
Your efforts to ensure the best for your children are a true example of bravery.
Thank you for all the support you've given me throughout my life.

Acknowledgements

I want to dedicate this dissertation to my mother, Carolina Nogueira, who supports me unconditionally under all circumstances. Her efforts made all this work possible.

To my grandmothers, Edviges and Graça, for raising me and teaching me valuable principles about simplicity. Thank you for your efforts to provide me with good food throughout this dissertation.

To my grandfathers, Anselmo and Milton, who have always been willing to help me financially or with favors. Thank you for all the lessons learned. I inherited my perseverance from you.

To my uncle, Carlos Nogueira, who has been and continues to be my inspiration in life. Thank you for all your support.

To my sister Cecília Nogueira, who brings me joy every day. Thank you for being my emotional outlet when I'm tired and unhappy.

To my father, Milton Ribeiro, for all the lessons learned and the support given throughout my life.

To Priscila Kubota, whose sweetness and love make my life more beautiful and joyful. Thank you for all the moments when you gave me the courage to never give up.

To my advisors, João Costa, Fabrício Lobato, and Moisés Silva, for all the lessons learned, for the patience, and for allowing me to find my professional vocation. Also, I'd like to thank Carlos Natalino for the collaboration and suggestions that certainly enriched this work.

Last but not least, I thank my friends Bernardo Duarte, Ronald Oliveira, Victor Cardoso, and Cristiano Campos for all the happy and sad moments that helped my personal development. Love y'all.

"Knowing yourself is the beginning of all wisdom."
Aristotle

Abstract

Over the past years, the emergence of more complex and bandwidth-hungry applications has charged efforts to ensure the reliability of optical networks. The occurrence of faults, for instance, can directly affect the quality of transmission of these optical systems, leading to several implications, including packet losses and service disruption. Hence, it is vital to mitigate faults in optical networks to guarantee the availability of the system and meet the service level agreement requirements. Moreover, as the complexity of optical networks evolves constantly, machine learning-based approaches have been proposed to deal with the system dynamics while providing automated fault management. In that regard, most proposed approaches are based on supervised learning (SL) models, which require large amounts of fault data to be properly trained. However, data from fault conditions are typically scarce in practical scenarios, which poses limitations for deploying SL-based models. Therefore, this work explores several unsupervised learning algorithms to perform fault management in optical networks. As fault data are absent in several real-world scenarios, unsupervised strategies trained with only data from normal operating conditions can be an effective alternative. These strategies disregard the need for data from abnormal network conditions and thus require much less data for model training. In this work, the fault detection and localization performances of cluster-based algorithms (K-means, Fuzzy C-means, Mahalanobis Squared-Distance-based model, and Gaussian Mixture Model) and dimensionality reduction-based approaches (Principal Component Analysis and Autoencoder) are compared leveraging a dataset derived from an optical testbed. The techniques are evaluated in terms of Type I (false-positive) and Type II (false-negative) error trade-offs. Ultimately, all techniques demonstrated satisfactory fault detection results when trained with only data from normal conditions, achieving an average accuracy of more than 90%. Such results suggest their applicability to real-world optical network fault management scenarios.

Key-words: optical networks, fault management, machine learning, unsupervised learning

Resumo

Ao longo dos últimos anos, o surgimento de aplicações mais complexas e exigentes em termos de largura de banda tem exigido esforços para garantir a confiabilidade das redes ópticas. A ocorrência de falhas, por exemplo, pode afetar diretamente a qualidade de transmissão destes sistemas ópticos, causando várias implicações, incluindo perdas de pacotes e interrupção do serviço. Assim, é vital mitigar falhas em redes ópticas para garantir a disponibilidade do sistema e cumprir os requisitos dos acordos de nível de serviço. Além disso, uma vez que a complexidade das redes ópticas evolui, abordagens baseadas em aprendizado de máquina têm sido propostas para lidar com a dinamicidade dos sistemas, proporcionando simultaneamente uma gestão automatizada das falhas. Grande parte destas abordagens baseia-se em algoritmos de aprendizagem supervisionada, os quais exigem grandes quantidades de dados de falhas para serem corretamente treinados. No entanto, dados de condições falhosas são tipicamente escassos em cenários práticos, impondo limitações para a aplicação de modelos supervisionados. Portanto, neste trabalho, são explorados vários algoritmos de aprendizagem não-supervisionada para efetuar a gestão de falhas em redes ópticas. Uma vez que os dados de falhas são ausentes em vários cenários do mundo real, estratégias não-supervisionadas utilizando apenas dados de condições normais podem ser uma alternativa eficaz. Tais estratégias não necessitam de dados de falha, por conseguinte, requerem muito menos dados para o treinamento dos modelos. Os desempenhos da detecção e localização de falhas de algoritmos baseados em agrupamentos (*K-means*, *Fuzzy C-means*, modelo basedo em Distância de Mahalanobis Quadrática, e modelo de Mistura Gaussiana) e de algoritmos baseados em reducção de dimensionalidade (Análise de Componentes Principais e *Autoencoder*) são comparados neste trabalho baseados em um conjunto de dados derivado de um *testbed* óptico. As técnicas são avaliadas em termos de erros do Tipo I (falsos positivos) e do Tipo II (falsos negativos). Em última análise, todas as técnicas demonstraram resultados satisfatórios na detecção de falhas, mesmo que treinadas apenas com dados de condições normais, atingindo um acurácia média acima de 90%. Tais resultados sugerem sua aplicabilidade em cenários reais de gestão de falhas em redes ópticas.

Palavras-chave: redes ópticas, gestão de falhas, aprendizado de máquina, aprendizado não-supervisionado

List of figures

List of tables

List of abbreviations and acronyms

QoT	Quality of Transmission
SLA	Service-level Agreements
TRX	Optical transponders
WSS	Wavelength selective switches
OA	Optical amplifiers
DWDM	Dense Wavelength Division Multiplexing
BER	Bit error rate
FEC	Forward error correcting
EONs	Elastic optical networks
SDN	Software-defined network
OSA	Optical spectrum analyzers
OSNR	Optical signal to noise ratio
BL	Bit-rate distance product
LED	Light-emitting diodes
SMF	Singe-mode fiber
WDM	Wavelength-division multiplexing
FDM	Frequency-division multiplexing
ROADMs	Reconfigurable add-drop multiplexers
EDFA	Erbium-doped fiber amplifier
SDM	Space-division multiplexing
MIMO	Multiple inputs, multiple outputs
FMF	Few-mode fiber
DSP	Digital signal processing

API	Application programming interface
ASE	Amplifier spontaneous emission
EMP	Electromagnetic pulse
DDoS	Distributed denial of service
SL	Supervised learning
UL	Unsupervised learning
ML	Machine learning
AI	Artificial intelligence
NN	Neural network
ANN	Artificial neural network
SVM	Support vector machine
RF	Random forest
GAN	Generative adversarial networks
VAE	Variational autoencoder
AE	Autoencoder
LSTnet	Long- and short-term time-series network
PCA	Principal component analysis
FCM	Fuzzy C-means
MSD	Mahalanobis squared-distance
GMM	Gaussian mixture model
DBSCAN	Density-based spatial clustering for applications with noise
AIC	Akaike information criteria
PC	Principal component
FI	Fault indicator
TP	True positive
FP	False positive

TN	True negative
FN	False negative
OPEX	Operational expenses
LLM	Large language model

Contents

1 Introduction

1.1 Context

Over the last few years, driven by the rapid development of communications technologies (i.e., 5G, Internet of Things, and cloud computing) and emerging network services (i.e., VR/AR, autonomous vehicles, and 4K video), the demand for high data transmission in optical communication networks has experienced exponential growth (GU; YANG; JI, 2020). Improving intelligence and versatility in optical networks is imperative to conquer this data rate as their complexity grows and new data-hungry applications emerge. Moreover, the interconnected nature of our world means that any faults or disruption to optical links affect the quality of transmission (QoT), resulting in data loss and leading to service-level agreements (SLA) no longer being met (KRUSE et al., 2024). Hence, promoting accurate and real-time fault management is paramount to enhancing the assurance of optical network transmissions. In the same sense, the fulfillment of SLAs can be efficiently provided by ensuring effective monitoring of the physical layer regarding the occurrence of faults or degradation in the network (RAK et al., 2021)

Most often, faults in optical networks can be classified into soft- and hard-failures. They can occur due to several factors, including component malfunctioning and aging (CHEN et al., 2018), attacks to the physical layer (MAS; TOMKOS; TONGUZ, 2005), and lack of maintenance (REJEB; LEESON; GREEN, 2006). Faults that cause the immediate connections to break, e.g., caused by fiber cuts, directly impact the network transmissions and are categorized as hard failures. On the other hand, soft failures are typically defined as anomalies that cause significant variations in network transmissions but not severe enough to trigger obvious alarms of service disruption (RAK et al., 2021). Soft failures can occur in several optical devices, such as Optical Transponders (TRX), Wavelength Selective Switches (WSS), and Optical Amplifiers (OA) (MAYER et al., 2021). Those failures may evolve into hard failures and compromise several optical connections supporting many network services. Therefore, detecting and localizing the fault is essential to facilitate network maintenance and reduce operational costs (BARZEGAR et al., 2020; VELA et al., 2018).

Regarding fault management methods, the conventional ones rely on simplified thresholds and probability models (WANG et al., 2022; MUSUMECI et al., 2018; MUSUMECI, 2021). However, these traditional methods are ineffective, error-prone, and labor-intensive for complex and dynamic cases, which need to be conducted by experienced technicians. Thus, effective threshold-based methods become difficult as current networks become highly dynamic with extensive optical devices (ZHANG et al., 2014; POINTURIER, 2017). In this scenario, providing large and fully automated networks capable of performing policy-driven self-diagnostics following the specific set of applications running over the network is the ultimate focus (CARROZZO

et al., 2020).

Following this goal, machine learning (ML) and artificial intelligence (AI) techniques have been increasingly investigated to address a variety of tasks related to automated optical network management and control (MUSUMECI et al., 2018). As simplified threshold-based methods can lead to low fault detection rates or trigger many false alarms (MUSUMECI, 2021), ML-based approaches have become a promising alternative to handle these limitations, as well as the need for automation. In this scenario, the system must be data-driven once ML performance often depends on the quality and amount of data for model training (CARROZZO et al., 2020; XIE et al., 2020). Also, as the network dynamics may vary, data-driven strategies can guarantee an agile adaptation to the new conditions of the system. With the proper incorporation of these solutions, network operators become able to realize knowledge-based autonomous service provisioning (CHEN et al., 2018), dealing with the dynamicity of the systems (e.g., resource allocation optimization, end-to-end QoT, traffic profile) and guaranteeing effective management of faults (PROIETTI et al., 2019; ODA et al., 2017; CHEN et al., 2017; LI et al., 2018; GUO; ZHU, 2018).

Regarding ML-based solutions for fault management in optical networks, most studies have focused on supervised learning (SL) algorithms in the last decade. Several techniques, such as Artificial Neural Networks (ANNs) (VELA et al., 2017; RAFIQUE et al., 2018), tree-based algorithms, and Support Vector Machines (SVMs) (SHAHKARAMI et al., 2018; MAYER et al., 2021) can work following an SL manner. In these cases, the models need prior knowledge of data labels, containing information about what the models should provide as the correct outputs based on those labels. However, in optical networks, the datasets collected from the practical operating system always present an imbalanced nature, where the volume of data under normal conditions, i.e., without faults, is larger than that of data under fault conditions. This imbalance leads to a slow development process, possible increase in budgetary demands, and misclassifications for models not trained with a considerable number of fault samples. Although ML model based on *prior-statistics* can be an option to create artificial failure data, their implementation can be critical due to the expensive and intensive efforts demanded for complex networks.

In this scenario, parallel to the fact that abnormal behaviors differ from normal network behaviors (VELA et al., 2017; SHIMIZU et al., 2020), Unsupervised Learning (UL) techniques have become a promising alternative to fault management in optical networks. In handling the problem of imbalanced data in SL models, only data from normal network conditions are required for model training, disregarding the acquisition of data from equipment undergoing failures or simulating failure conditions from statistical models. More specifically, UL models can learn hidden patterns from normal data by recognizing underlying information, which means that when anomalies (failures) are fed into the model, they can be identified on the basis of an outlier detection framework. Unlike the conventional supervised techniques, UL-based approaches simplify the data collection for training purposes (as data from regular network conditions are

more accessible to acquire) and accelerate model deployment.

Therefore, this dissertation compares the fault management (detection and localization) performance of several UL algorithms based on clustering and dimensionality reduction. These techniques are evaluated regarding false-positive (Type I errors), false-negative (Type II errors), and accuracy. Comparing the performance of these approaches can provide several insights into how they could be used in that application and how they differ from each other, highlighting their advantages and disadvantages.

1.2 Related work

In recent years, the scientific community has made several efforts to develop ML algorithms for managing optical network failures, including failure detection, identification, and localization. Hence, this section will discuss several works regarding ML-based approaches for failure management.

1.2.1 Supervised learning approaches for fault management

Vela et al. (VELA et al., 2017) proposed two different finite state machine algorithms to detect and identify the cause of several failures causing significant bit error rate (BER) degradations in optical connections. The first algorithm analyzes pre-FEC BER data to detect gradual changes with time that might derive into BER degradation and intolerable and anomalous BER values. The second is a probabilistic algorithm that returns the most likely failure among a set of failure classes, identifying the cause of the failure. Considering similar failure scenarios, the same authors also investigated ML-aided algorithms for soft failure localization in Elastic Optical Networks (EONs) (VELA et al., 2018). The authors find relevant points in the signal spectrum and analyze them to localize the failure and estimate its magnitude.

Similarly, Shahkarami et al. (SHAHKARAMI et al., 2018) compared the performance of various machine learning algorithms to detect and identify equipment failures using monitored BER data. Binary Support Vector Machine (SVM), Random Forest (RF), Multiclass SVM, and neural network (NN) with a single hidden layer are compared for failure detection in terms of complexity and accuracy. On the other hand, for failure identification, an NN with two hidden layers was exploited. Although good results were presented, data from failure and no-failure conditions were required for model training, as they follow a supervised learning approach.

All the previous works use BER characteristics to perform fault management in optical networks. However, some approaches use different optical parameters for the same task. For instance, Rafique et al. (RAFIQUE et al., 2018) use a neural network-based method using experimental data from optical power measurements under diverse failure modes. The core of this approach is to first identify potential failures by applying block-based deviation and categorization to be further evaluated by the neural network, which predicts true abnormal

behavior (if any). Furthermore, Mayer et al. (MAYER et al., 2020) proposed a supervised NN-based approach to localize soft-failures. The model evaluates the telemetry data collected from software-defined network (SDN) streams. However, similar to other supervised approaches, this work requires a numerical model of the physical network to generate simulated data from failure scenarios for training.

Using the optical spectrum captured by an optical spectrum analyzer (OSA), Velasco et al. (VELASCO et al., 2018) proposed a soft-failure identification and localization approach. First, a multi-class classifier in the form of a decision tree is used to detect filtering-related failures. Then, an SVM binary classifier predicts whether the failure is due to Filter Shift or Tight Filtering. In the end, three approaches based on feature extraction were investigated to cope with the filter cascading problem and improve the classifier's failure detection performances. In the same sense, Shimizu et al. (SHIMIZU et al., 2020) introduce a neural model trained to detect the presence (or not) of a failure in an optical link based on the optical signal to noise ratio (OSNR) values from multiple lightpaths, where the model output provides a true/false response. Although efficient, this technique is limited to only indicating the presence of a failure in the system, not concerning data imbalance and failure localization.

1.2.2 Unsupervised learning approaches for fault management

All the works mentioned above rely on supervised learning approaches. These algorithms must be trained with data from both failure and normal conditions. However, in real-world optical networks, obtaining data from failure conditions may be unfeasible and expensive as the design of optical links tends to be conservative and over-engineered. In this scenario, data under failure conditions become often rare in real systems, while data under normal conditions is abundant. Therefore, recent studies have focused on unsupervised learning (UL) algorithms, which only require data under normal conditions to be trained and thereby perform well with imbalanced data. For instance, Lun et al. (LUN et al., 2021) proposed an unsupervised approach based on generative adversarial networks (GANs) trained with electrical spectrum data to perform soft-failure detection and identification. Although the GAN model has been trained with only normal samples to perform failure detection, the proposed approach used supervised algorithms trained with failure samples to perform failure identification.

Similarly, authors in (CHEN et al., 2022) proposed a hybrid unsupervised/supervised fault detection framework that combines density-based clustering techniques and deep neural networks. First, the approach uses the clustering technique to learn hidden patterns from unlabeled data and classify them into clusters corresponding to fault and normal conditions. Finally, the deep neural network is fed with labeled data from the clustering technique to perform anomaly classification. However, the proposed approach needs failure samples for the model training. It can also not perform failure localization, which is fundamental in rapidly repairing eventual failures in optical networks.

Combining a variational autoencoder (VAE) and a GAN model, Kruse et al. (KRUSE et al., 2024) enable fault management that relies only on Euclidean distance to distinguish between normal and failure data. It utilizes a variational autoencoder-based generative adversarial network (VAE-GAN) running on optical spectral data obtained by optical spectrum analyzers. The framework can reliably run on a fraction of available training data and identify unknown failure types. However, this approach cannot perform fault localization as it is only trained for failure detection/identification, disregarding in-operation network reconfiguration enabled by fault localization. Similarly, Liu et al. (LIU et al., 2021) proposed an autoencoder-based approach to perform only failure detection. Although the approach uses only normal data for the training phase, a small percentage (less than 3%) of failure data was needed to validate the model.

Silva et al. (SILVA et al., 2022) used a deep learning-based approach enabling multivariate time-series modeling to perform complete fault management, including fault detection, localization, and forecasting. The proposed approach uses a large and complex architecture based on the LSTNet model composed of convolutional and attention layers to guarantee scalable forecasting performance as input time series and time steps increase. Although encompassing fault localization and forecasting, the proposed LSTNet-based approach relies on a computationally expensive model. In the same sense, concerning network scalability, Ribeiro et al. (RIBEIRO et al., 2024) detect faults while handling the scalability issue by exploiting a cluster-based algorithm assisted by a principal component analysis (PCA) technique. Despite enabling scalable failure detection, this approach does not provide fault localization.

Regarding PCA-based approaches, several works leverage the rotation invariant property of PCA to handle data security concerns. For instance, Silva et al. (SILVA et al., 2023) exploit PCA for privacy-preserving soft-failure detection by leveraging its ability to learn linear underlying information of the telemetry data, being able to detect abnormal behaviors (if any). For the same task, Sales et al. (SALES et al., 2024) proposed a disaggregated PCA approach that ensures data confidentiality by distributing the learning process within several local models. This approach prevents unauthorized access to confidential data and provides failure detection. However, although these PCA-based approaches concern data security, they disregard fault localization concerns.

1.3 Justification

Although several current studies for fault management in optical communication networks are based on unsupervised learning methods, most of the works in the literature still employ supervised learning-based approaches. Working in a supervised manner brings several limitations regarding the proper training of the model, as it needs a large amount of data under fault conditions. Unfortunately, fault data availability in operational networks often constitutes a practical limitation to deploying ML-based solutions due to the scarce availability of labeled

data that comprehensively models all possible fault types. One could purposely inject failures to collect training data, but this is time-consuming and undesirable for operators. In that sense, UL-based approaches present themselves as a great alternative to cope with the imbalanced nature of the data.

1.4 Motivation

Conventional methods, typically based on predefined fixed thresholds and/or decision rules, provide a limited yet effective capability in detecting faults in optical networks. However, as these methods still depend on network operator intervention, they can not cope with the increasing complexity of modern optical networks with numerous parameters. Hence, machine learning techniques emerge as a great alternative due to their inherent characteristics for scenarios where complexity and scalability issues are imperative. Generally, ML algorithms are divided into two big classes: Supervised learning and Unsupervised learning algorithms. Supervised models work by learning patterns of the data and consulting the label of the respective data, thereby correctly classifying or predicting new data when inserted into the model. These algorithms typically present an accurate classification performance for fault detection/localization in optical networks. However, they need a large amount of data corresponding to all the classes, including samples from the failure condition, which is scarce and unfeasible to collect in real-world optical networks. Thus, unsupervised algorithms emerge as techniques to handle this data scarcity. They disregard the need for labeled data and can successfully perform fault management by training with only data from regular and normal network operations. Also, UL-based approaches enable the detection of unknown failures, as they are not trained based on specific labeled failure samples. Taking advantage of that, some UL-based works have been proposed in the literature to manage faults in optical networks. However, no works directly compare the performance of such techniques. Therefore, the present work is motivated by the need to compare UL-based approaches that can detect and localize faults without the explicit acquisition of fault samples to be trained. Moreover, comparing these techniques can provide useful information on how they can be implemented in practical optical networks and the advantages and disadvantages of those methods for fault management.

1.5 Objectives

The main objectives of this work are to explore, apply, and compare several machine learning algorithms for fault management in optical communication networks. Moreover, the core of this work is concentrated on implementing algorithms that can learn the distributions of extracted failure-sensitive features of the telemetry data without the need for data from failure conditions, coping with the typical imbalanced nature of the data from practical optical networks. To achieve these goals, the particular objectives are listed below:

1. Overcome the limitations imposed by traditional methods based on simplified thresholds and supervised learning algorithms by applying and developing unsupervised learning approaches that can perform the same tasks as the traditional methods but address the imbalanced issue of the datasets.

2. Enrich the literature in the context of unsupervised learning algorithms used for fault management in optical networks by comparing these proposed methods and finding out how they differ from the traditional ones.

3. Compare different classes of unsupervised algorithms, including the cluster-based algorithms, namely K-means, Fuzzy C-means (FCM), Mahalanobis squared-distance model (MSD), Gaussian Mixture Model (GMM), and Density-based Spatial Clustering of Applications with Noise (DBSCAN); and the dimensionality reduction methods, namely PCA and Autoencoder.

4. Apply the proposed unsupervised methods in real-world telemetry data derived from an optical testbed susceptible to variations and noise in the data behavior, typically found in practical optical network operation.

1.6 Original contributions

Although most works in the context of fault management in optical networks still rely on supervised methods, several works have been proposed using unsupervised methods. However, there is still a gap in the literature regarding the direct comparison of the fault management performance of these methods submitted to equal scenarios (datasets). In that sense, the main original contributions of this work are the following:

1. Provide a robust comparison between several unsupervised learning-based approaches, submitting them to the same dataset, thereby clearly observing the differences inherent to each approach and how they affect the fault management performance. There is no direct comparison between different classes of unsupervised techniques in the literature.

2. Four cluster-based algorithms exploited in this work can be faced as the first application of K-means, FCM, MSD, and GMM for fault management in optical networks, intended to analyze the suitability and accuracy of simple techniques with low computational cost.

3. Compare the proposed approaches on the basis of Type I and Type II errors (false-positive and false-negative indications of failure, respectively) trade-off in a dataset collected from an experimental optical testbed.

The results indicate that the proposed unsupervised approaches achieved similar fault detection/localization performance regarding the evaluation metrics defined. Even though cluster-

based approaches have, on average, lower computational costs than the dimensionality reduction-based ones, the former ones can learn the normal variations in the data, distinguishing them from failures. Moreover, cluster-based techniques demonstrated to be able to give a clear physical visualization of the data distributions, providing model interpretability, which may be useful for the network operators. In fact, the dimensionality reduction-based approaches presented better results as they could learn the underlying and non-linear effects inherent in the dataset, typically presented in practical network conditions.

1.7 Organization of dissertation

The rest of this dissertation is structured as follows. Chapter 2 introduces the fundamentals regarding optical networks, specific concepts and parameters from the physical layer and the implications derived from the occurrence of failures in the network. Chapter 3 is devoted to summarizing the fundamentals of the proposed unsupervised learning techniques. Chapter 4 presents their particular application for fault management in optical networks. Chapter 4 details the experimental optical setup regarding network parameters and dataset, the fault detection/localization results from the compared algorithms, and discussions regarding those results. Finally, Chapter 5 summarizes the main conclusion and contributions of this dissertation. Also, directions for future research and published works are introduced.

2 Optical Communication Networks

This chapter discusses the fundamentals of optical communication networks. Learning the basic concepts about these systems is necessary for contextualization, providing a solid background about their components and functioning, and learning how network failures can negatively affect their well-functioning.

2.1 Brief historical perspective

Throughout the last decades, the emergence of optical communication networks has brought new perspectives to the ever-increasing data capacity demanded by various services and applications. Due to their particular capabilities, including high data transmission capacity over long distances, many consider optical networks the only future-proof technology capable of handling various data-hungry applications. Optical communication networks refer to fiber-optic-based systems that employ optical fibers for information transmission. Such systems transmit data through light pulses that travel along the fiber.

Optical networks emerged in 1966 with the paper entitled “Dielectric-fibre surface waveguides for optical frequencies” made by C. K. Kao as a solution to deal with the need for a higher bit rate-distance product, commonly denoted as BL, where B is the bit rate and L is the repeater spacing (KAO; HOCKHAM, 1966). The paper already contained the main elements of today’s optical fibers, including cylindrical geometry, with a higher refractive index core surrounded by a lower refractive index material. At this time, fiber optics suffered from the glass attenuation coefficient, around 200 dB/km (MELLO; BARBOSA, 2021). Only years later, researchers from NTT reached the levels of attenuation we know today, 0.2 dB/km (MIYA et al., 1979).

Besides providing large transmission capacity in the network with more BL, optical networks provide a common infrastructure for various bandwidth-hungry services. Unlike copper cable systems, these networks are also increasingly capable of delivering bandwidth flexibly where and when needed and are less susceptible to electromagnetic interference and other undesirable effects (RAMASWAMI; SIVARAJAN; SASAKI, 2009).

In that sense, these systems are the favorite for data transmission at a few tens of megabits per second over any distance of more than a kilometer. The same occurs for short-distance (a few meters to hundreds of meters) and high-speed (gigabits per second and above) interconnections inside large systems (RAMASWAMI; SIVARAJAN; SASAKI, 2009).

Optical networks have evolved throughout the past decades, being classified into different generations. The first generation used AlGaAs light-emitting diodes (LEDs) and lasers at 0.8 μm

operating in graded-index multi-mode fibers. However, the bit rate of these systems was limited to below 100 Mb/s because of modal dispersion in multi-mode fibers (LI, 1978; MAIONE; SELL; WOLAVER, 1978). The end of the first generation came with the development of single-mode fibers (SMFs) that could solve the problem of modal dispersion.

The second generation came with optical networks transmitting at 2 Gb/s over 44 km using a single-mode fiber (YAMADA; MACHIDA; KIMURA, 1981). By 1987, these networks were commercially available, operating at bit rates of up to 1.7 Gb/s with a repeater spacing of about 50 km. Correspondingly, the development of new InGaAsP sources moved the operating regime toward the so-called long wavelengths in the window near 1.3 μm, in which the fiber attenuation is lower (LI, 1983; IWAHASHI, 1981). However, the fiber losses (typically 0.5 dB/km) limited the repeater spacing of optical networks operating at this window. By 1979, a 0.2 dB/km loss was realized in the 1.55 μm spectral region. A drawback of third-generation optical networks is that the signal is regenerated periodically using electronic repeaters spaced apart, typically by 60-70 km (AGRAWAL, 2012). In fact, the challenges at this time moved from optical fiber to electronics (HENRY, 1985).

The fourth generation emerged with wavelength-division multiplexing (WDM) systems, in which different signals from different spectral segments share the same fiber (MELLO; BARBOSA, 2021). The core of WDM systems is to increase their capacity by transmitting more data using different frequencies/wavelengths. It can be faced as the same frequency-division multiplexing (FDM) technique widely used in radio systems. WDM-based systems make a single fiber look like multiple fibers. In this manner, the outputs of several transceivers operating in different wavelengths are optically multiplexed into a single fiber using a multiplexer. At reception, the wavelength-multiplexed signals are separated using a demultiplexer (MELLO; BARBOSA, 2021). In this scenario, an important component is noteworthy to mention: the Reconfigurable add-drop multiplexers (ROADMs). ROADMs are responsible for forwarding the light path through an optical network topology (AGRAWAL, 2012). The main building block in ROADMs is wavelength selective switches (WSSs). Working as a multiplexer, the WSS can select any set of wavelengths from its input ports and steer it to the output port (AGRAWAL, 2012). Working as a demultiplexer, the WSS can select any set of wavelengths from its input port and steer it to any of its output ports.

Although the WDM systems increase the data transmission rate using only a single fiber, it is still not scalable in terms of cost, as due to fiber attenuation, the signal must be regenerated periodically in the electronic domain to recover the transmitted bits (MELLO; BARBOSA, 2021). In that regard, efforts were made to overcome the need for regenerators for systems with long distances, where tens of regenerators will be required.

Erbium-doped fiber amplifiers (EDFAs) emerged as optical devices without opto-electro-optical conversion to deal with the need for regenerators, reducing the cost of long-haul optical networks and reducing even more when we deal with ultra-long-haul distances. EDFAs revolu-

tionized the optical communication landscape, and even today, these devices are widely used. For example, a WDM link of 20 wavelengths and 2000 km required $9 \times 40 = 360$ opto-electro-optical regenerators. Conversely, these 360 regenerators could be replaced by only 18 EDFAs, with each EDFA having a lower cost than a single regenerator.

The fifth generation (mid-2000s) emerged with a focus on expanding the wavelength range of WDM systems and increasing their spectral efficiency, opening the doors for L and S band transmission. The point is to employ advanced modulation formats in which information is encoded using the optical carrier's amplitude and phase (WINZER; ESSIAMBRE, 2006). Also, the fifth generation inaugurates the digital coherent optical systems. These are mainly deployed in high data rate core optical networks intended to provide internet traffic between geographically distant locations (MELLO; BARBOSA, 2021).

Although digital coherent systems appear in the fifth generation of optical systems, they date back to the 1980s and were first developed to increase spectral efficiency and sensitivity. However, with the astounding multiplication of capacity promoted by wavelength division multiplexing (WDM) in the 1990s, it stayed in the background for a long time until the mid-2000s (MELLO; BARBOSA, 2021). Only in 2004, now within the aforementioned fifth generation, digital coherent systems return. Firstly demonstrated by Taylor (TAYLOR, 2004), the return of these systems brings a front-end that converts the received optical signal into electric signals, followed by a high-speed oscilloscope capable of storing data windows for offline processing. In that sense, software-defined networking (SDN) emerges as systems that reduce network stiffness and ease work management. In the SDN paradigm, all network elements have standardized interfaces controlled by a single centralized control plane. Data and optical equipment from different vendors become "white boxes" whose main functionalities are managed by the control plane (MELLO; BARBOSA, 2021).

As of today, the state-of-the-art concepts and technologies regarding optical networks rely on space-division multiplexing (SDM) systems, which include fibers with multiple cores that can multiplicate the capacity of the traditional single-core fiber systems and the transmission of coupled signals but later separated by digital signal processing (DSP) with multiple outputs and multiple inputs (MIMO) (WINZER; FOSCHINI, 2011; HO; KAHN, 2011; MAROM et al., 2017). Few-mode fibers (FMF) that can transmit data through several, but few, modes alleviating the inter-modal interference, and multi-band systems, intended to transmit data through different bands, also appeared as current optical technologies that can likely serve as the base for the sixth generation of optical communication systems.

2.2 Software-defined optical networks

With the advent of the Internet, the amount of data transmitted worldwide has reached zettabytes and is still increasing (HUI, 2019). In that sense, it is imperative to improve the

automation and intelligence of optical communication networks in response to the rapid increase in network traffic and growing network operation complexity (GU; YANG; JI, 2020). Furthermore, as many users intend to connect simultaneously in various places, using several services and applications, controlling current optical networks becomes difficult. These network control problems become more dramatic when dealing with networks that handle massive amounts of data, posing challenges for network control and management (THYAGATURU et al., 2016). Also, the increasing Internet traffic and the several services that came with it require optical networks' capacity to expand accordingly.

In that scenario, to simplify and facilitate network management, the emergence of software-defined networking brings optical networks that can separate the data plane from the control plane and centralize network control in an SDN controller (THYAGATURU et al., 2016). Such capabilities make it possible to ensure and integrate intelligence into an optical network's tracking, administration, and management. Hence, the ultimate goal of SDN-based optical networks is to promote self-awareness and fully automated networks with minimum human interference capable of handling the demands of future optical networks.

SDN can bring several functionalities to optical networks. Nsafoa-Yeboah lists the following ones (NSAFOA-YEBOAH et al., 2022):

(i) Allocation of resources in primary and sublinks and cloud computers;

(ii) Flexible and resilient network and diverse network adaptability;

(iii) Service availability secures network stability and resilience;

(iv) Cross-layer management issues, such as debugging and checking.

By incorporating these SDN functions, optical communication networks become able to work with equipment derived from different vendors. This flexibility is particularly desirable as these networks constantly grow along with the number and variety of equipment. Moreover, leveraging an SDN protocol (e.g., OpenFlow protocol), it became possible to manage optical transceivers and utilize DSP techniques for recent generations of optical transceivers. This provides software monitoring of specific transceiver parameters, facilitating network management. Optical circuit switching, optical packet monitoring, and topology exploration are also useful features of incorporating an SDN controller. Essentially, an SDN architecture consists of three layers, namely the Infrastructure, Control, and Application layers, as shown in Figure 1.

The data is collected from the Infrastructure layer. The Control layer is regarded as the brain of the network, where the controller resides and governs the entire network. The controller conducts the overall operation of the network, such as routing packets, policy management, modification of flow rules, topology, and system configuration (RANGISETTI; TAMMA, 2017; ZAHMATKESH; KUNZ, 2017). The controller assigns flow rules for the new incoming packet

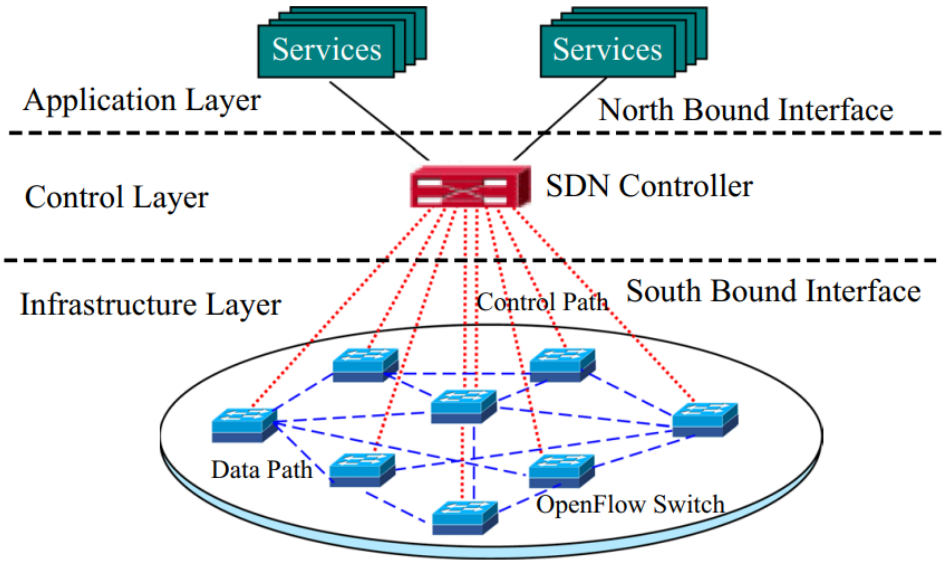


Figure 1 – Conventional SDN architecture comprising the Infrastructure, Control, and Application layers (SRIVASTAVA; PANDEY, 2020).

using the OpenFlow Protocol to the switch (FOERSTER; SCHMID; VISSICCHIO, 2018). The switch communicates or transfers data along the data path between themselves. The Application layer creates new rules using APIs for certain types of incoming packets passed to the controller when needed.

Furthermore, as depicted in Figure 1, one can note that an SDN-based architecture relies on a single SDN controller. For instance, when there are too many flow requests from switches, the single controller will not handle them at once, leading to a bottleneck situation that impacts the scalability performance of the single controller. In that case, there should be more than one single controller to avoid such issues and improve the network's robustness.

Optical network performance monitoring is also a desirable feature in the context of SDN-based networks. As optical transmission impairments, commonly divided into non-catastrophic and catastrophic impairments, pose concerns about the QoT, monitoring optical elements can extract useful features to maintain a well-functioning network, minimizing these impairments (THYAGATURU et al., 2016). Non-catastrophic impairments include photonic impairments that degrade the OSNR (e.g., forms of dispersion, cross-talk, and non-linear propagation effects) without disrupting the communication. Conversely, fiber cuts or malfunctioning switches can be faced as catastrophic impairments that have the potential to disrupt communication. An SDN-based optical network can properly avoid such impairments and guarantee service availability by monitoring information at the SDN controller and executing control algorithms.

Recently, from the state-of-art perspective, it can be inferred that the spotlight has been focused on algorithms and implementations that can provide self-awareness and complete control of software-defined optical networks. ML techniques emerge as one possible solution to the challenges posed by the current state of these networks in terms of complexity and operation.

As the demand for large amounts of heterogeneous data elevates the requirements for effective network management, such techniques become helpful due to their particular capabilities. From these, ML algorithms can solve complex problems and optimize the efficiency of optical systems through better performance in spectrum allocation tasks, traffic prediction, traffic classification, and QoT, among other parameters (VILLA et al., 2023).

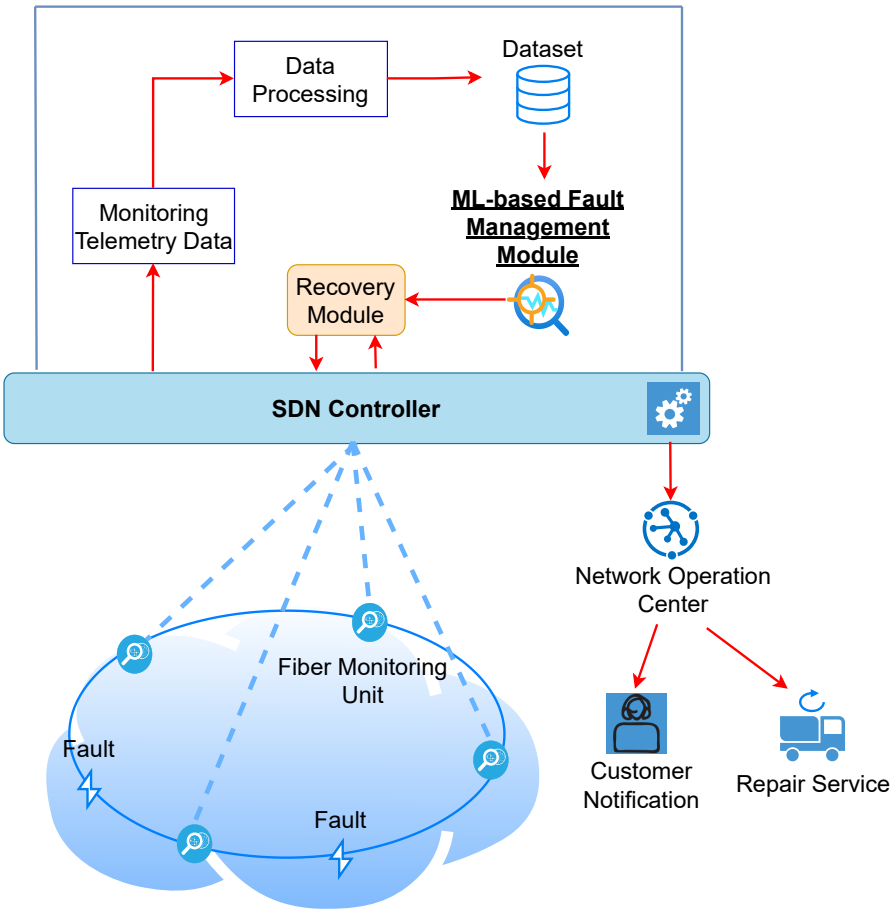


Figure 2 – Overview of an SDN-based optical network integrated with an ML-based fault management module.

For the particular optical network fault management case, ML-based approaches have been integrated into the SDN controller as an additional module responsible for analyzing the nature of the telemetry data. Figure 2 shows how an ML-based fault management module can be integrated into an SDN controller. The fundamental aspects of several machine learning algorithms will be discussed in the next chapter, as well as how they can be used for fault detection and localization in optical networks.

2.3 The occurrence of faults: Causes and implications

Generally, the ideal scenario when developing any system is to make these perform perfectly without any failure. However, in reality, it seems more like a utopian idea, as every system is susceptible to the occurrence of failures or faults. It is not different when dealing with optical communication networks. Such systems typically cover wide areas of thousands of kilometers and comprise a large number of components. Hence, the occurrence of faults may cause several prejudicial consequences that compromise the network's well-functioning.

Conventionally, faults in optical networks can be divided into two classes: soft failures and hard failures. Soft failures can be faced as gradual signal degradation that does not cause alarms or disrupt the signal but still compromises the connection. Device aging, equipment malfunctioning, misconfigurations, and physical-layer attacks can be categorized as soft failures (MAS; TOMKOS; TONGUZ, 2005; SKORIN-KAPOV et al., 2016). Conversely, hard failures refer to events that totally disrupt the link, putting the connection down immediately. Such failures include fiber cuts and breakdown of components (MAS; TOMKOS; TONGUZ, 2005).

Different components may be susceptible to the occurrence of faults. Figure 3 summarizes several optical components and the typical faults from different categories that can occur.

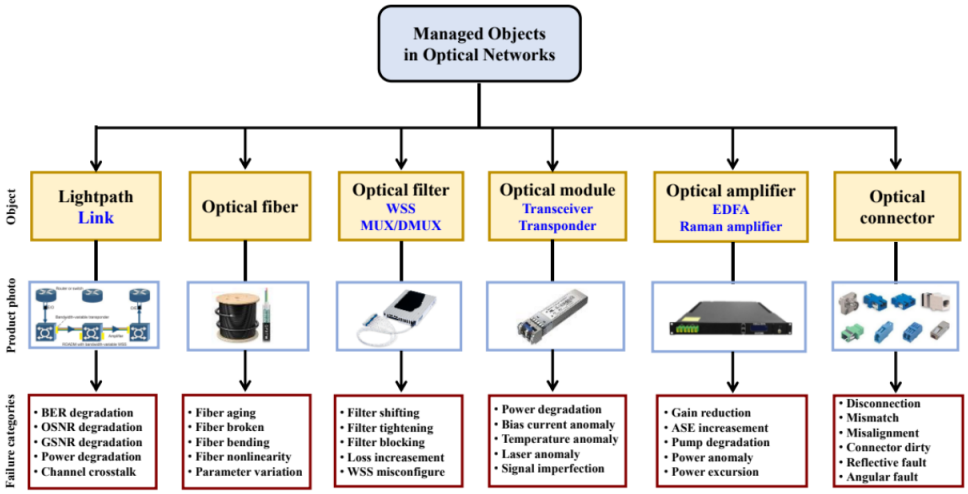


Figure 3 – Optical components with their respective product photos and the possible faults that may affect those products (WANG et al., 2022).

Optical lightpaths, optical filters, optical modules, and optical connectors are mainly affected by soft failures, such as BER, OSNR, and launch power degradations, misalignment, or disconnections. Optical fiber fault is the most typical fault source as various external factors, such as wildlife, natural disasters, or construction work, commonly damage the fiber (WANG et al., 2022). It makes fibers susceptible to several consequences like aging, bending, and breaking (MAS; TOMKOS; TONGUZ, 2005). Regarding optical amplifiers (commonly EDFAs), the typical failure types are gain reduction, output power anomaly, amplifier spontaneous emission

(ASE) noise increasing, aging of pump lasers, and incremental power excursion (WANG et al., 2022).

Malicious agents can also be the source of faults in optical networks. In these cases, the network is the primary target of an attacker aiming to cause the most severe damage possible. Electromagnetic pulse (EMP) attacks (FOSTER et al., 2004), distributed denial of service (DDoS) attacks (FURDEK et al., 2016), or attacks with massive weapons that indirectly affect the network are examples of malicious human-derived attacks.

All the faults mentioned before can lead to severe implications that may affect optical networks' reliability and make them not meet the customer's SLA requirements. Failures can lead to massive packet losses, directly compromising connections' integrity on a large scale. For instance, in 2008, an incident involving fiber cuts in The Mediterranean Sea caused the loss of 70% of Egypt's connection to the outside world and more than 50% of India's connectivity on the westbound route (BORLAND, 2008). Furthermore, the 7.1-magnitude earthquake in December 2006 in Taiwan disrupted seven submarine optical links and significantly degraded Internet connectivity between Asia and North America for weeks (MUKHERJEE; HABIB; DIKBIYIK, 2014; RAK et al., 2016). All users, including the government, private enterprises, financing institutions, the transportation industry, the manufacturing industry, and individuals, suffered heavy economic losses (WANG et al., 2022).

3 Unsupervised Learning

Machine learning algorithms have particular capabilities in solving complex and large-scale problems, learning patterns, and making predictions. These techniques are generally divided into two major categories that accomplish such capabilities: Supervised and Unsupervised learning algorithms.

Supervised learning, as the name says, refers to algorithms that need a "supervisor" to teach them how to learn patterns or characteristics related to the data based on labels. Each dataset sample must have a label that reveals its label, allowing the model to output the class of an unknown sample. Conversely, unsupervised learning algorithms learn without the need for labels simply by finding inherent characteristics from the data (LECUN; BENGIO; HINTON, 2015).

More specifically, unsupervised algorithms analyze several examples of a dataset x and attempt to implicitly or explicitly learn the probability distribution $p(x)$ or some properties of that distribution. At the same time, supervised algorithms analyze several examples of a dataset x and an associated value or vector y and learn to predict y from x, usually by estimating $p(y|x)$ (GOODFELLOW, 2016).

In this work, two classes of unsupervised learning algorithms are compared in the context of anomaly (faults) detection: cluster-based and dimensionality reduction-based algorithms. Unlike supervised algorithms, which need anomaly data to be appropriately trained, unsupervised algorithms can be trained with non-anomaly datasets. Hence, as mentioned in Chapter 1, in the case of optical fault detection, these techniques perform well trained with only data from normal conditions (non-failure), disregarding any faulty data, which are scarce and unfeasible to collect in practical scenarios. The theoretical background of the proposed algorithms of this work is presented in this chapter. Moreover, in the end, the particular manner for fault detection and localization using such algorithms is highlighted.

3.1 Cluster-based algorithms

Clustering refers to grouping data with similar characteristics. (RUSSELL; NORVIG, 2016). Clustering algorithms exploit the underlying structure of the data distribution and define rules for grouping the data with similar characteristics (JAIN; MURTY; FLYNN, 1999). This process results in partitioning a given dataset according to the clustering criteria without prior knowledge about the dataset. Ideally, each cluster consists of similar data instances entirely dissimilar from the instances in other clusters (AHMED; SERAJ; ISLAM, 2020). Generally, these algorithms are divided into centroid-based, distribution-based, and density-based. There

are other types, but only algorithms of these categories will be exploited for this work.

3.1.1 Centroid-based algorithms

3.1.1.1 K-means

K-means refers to a widely-used algorithm commonly applied to cluster large datasets. This algorithm is the simplest clustering technique but poses the potential to solve many problems. A plethora of applications using K-means can be found in the literature, ranging from face detection to text processing (SU; CHOU, 2001; ALHAWARAT; HEGAZI, 2018). It works by assigning each data sample to one of the K clusters generated by the method. Also, it is considered a hard-clustering algorithm, meaning each sample belongs to only one cluster with the smallest Euclidean distance from the respective sample, as shown in Figure 4.

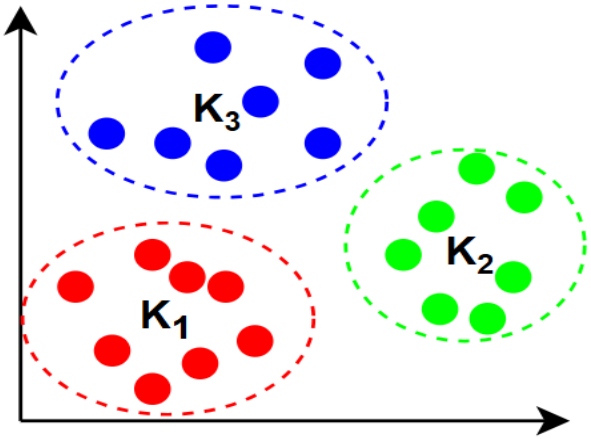


Figure 4 – K-means clustering forming three clusters and their respective centroids K.

The main goal of K-means is to update the centroids by calculating the mean of the samples belonging to the cluster and repeat the relocating and updating process until convergence criteria are satisfied. The user defines the number of clusters, and the initial positions of the clusters are generally random (MACQUEEN, 1967). The procedures are shown in the following pseudo-code:

Algorithm 1 K-means clustering algorithm

1: Initialize the cluster centers
2: Define the number of iterations L
3: **for** each iteration l **do**
4: Compute r_{nk}:
5: **for** each data point x_n **do**
6: Assign each data point to a cluster:
7: **for** each cluster k **do**
8: **if** $k == argmin\|x_n - \mu_k^{l-1}\|$ **then**
9: $r_{nk} = 1$
10: **else**
11: $r_{nk} = 0$
12: **end if**
13: **end for**
14: **end for**
15: **for** each cluster k **do**
16: Update cluster centers as the mean of each cluster:
17: $u_k^l = \frac{\sum r_{nk} x_n}{\sum r_{nk}}$
18: **end for**
19: **end for**

K-means aims to minimize the equation at line 7 in the above Algorithm 1, where x_n is a dataset sample, and μ_k is the centroid of the cluster k. Some iterations (L) are necessary until the algorithm reaches a local minimum. r_{nk} serves as a parameter that indicates when a sample belongs to a cluster.

The number of clusters defined by the user is a key factor that directly affects the performance of the K-means. Hence, many studies have been proposed approaching different methods to find the optimal number. The most conventional manner is to ask the user to input the number of clusters in advance, which needs expert domain knowledge of the underlying datasets. In contrast, several statistical criteria (e.g., Akaike information criteria (AIC)) have been investigated, providing an automatic selection of an appropriate number of clusters (KODINARIYA; MAKWANA et al., 2013).

3.1.1.2 Fuzzy C-means

Fuzzy C-means (FCM) is a centroid-based clustering algorithm similar to K-means (BEZDEK, 2013). Conversely, FCM is addressed in the soft-clustering (or fuzzy clustering) category. Unlike hard-clustering, which presents limitations when a single data point has patterns belonging to dissimilar clusters, soft-clustering algorithms reduce these limitations by providing more information about the memberships of the data points (NAYAK; NAIK; BEHERA, 2015). In the FCM method, the sample may belong to all clusters with a certain degree of fuzzy membership (FERREIRA; CARVALHO, 2014), as shown in Figure 5.

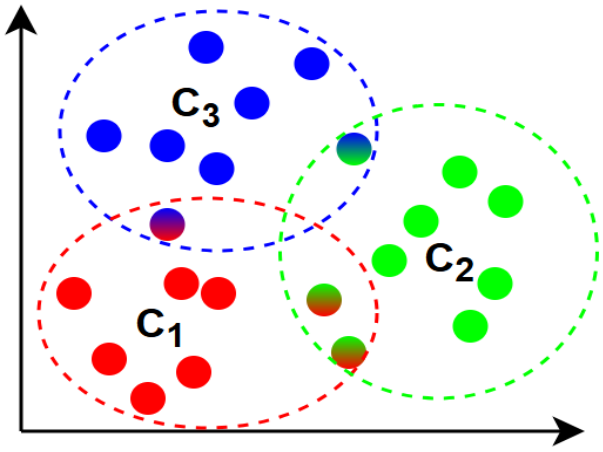


Figure 5 – FCM clustering forming three clusters and their respective centroids C. Note that some data points belong to two clusters.

The advantage of FCM is the generation of new clusters from the data points with close membership values to existing classes (ASYALI et al., 2006). In that regard, the equation for the centroids of each cluster refers to a weighted average, given by:

$$c_k = \frac{\sum_{i=1}^{N} w_{ik}^{p} x_i}{\sum_{i=1}^{N} w_{ik}^{p}}, \tag{3.1}$$

where x_i is the i-th samples of a dataset with N samples, w_{ki} is the degree of membership of the i-th sample to the K-th cluster, and p is the parameter that defines the fuzziness of the resulting clusters. Membership values are initially set randomly within the range the algorithm allows (0 to 1) and updated after calculating the cluster center. The updating of the membership values matrix is given by Equation 3.2 as follows:

$$w_{ik} = \frac{1}{\sum_{j=1}^{N} \left(\frac{\|x_i - c_k\|}{\|x_i - c_j\|} \right)^{\frac{2}{p-1}}}. \tag{3.2}$$

The FCM algorithm updates the membership values of C clusters, minimizing J in Equation 3.3 until it reaches a local minimum or a pre-defined number of iterations. Note that K-means aims to minimize the same equation, just restricting the membership values.

$$J = \sum_{i=1}^{N} \sum_{K=1}^{C} w_{ik}^{p} \|x_i - \mu_k\|^2. \tag{3.3}$$

The steps of the FCM algorithm can be summarized as follows:

Algorithm 2 FCM clustering algorithm

```
Initialize the centers of the C clusters.
Define the distance metric and the fuzziness parameter p.
for each iteration l do
    for each data point x_i do
        Assign each data point to a cluster:
        for each cluster k do
            Calculate the membership matrix W^l using Equation 3.2.
        end for
    end for
    for each cluster k do
        Update cluster centers using Equation 3.1
    end for
end for
```

3.1.2 Distribution-based algorithms

3.1.2.1 Mahalanobis squared-distance

The Mahalanobis squared-distance (MSD) is a well-known correlation-based statistical distance metric that measures the dissimilarity between two multivariate data sets (SARMADI et al., 2021; WORDEN; MANSON; ALLMAN, 2003). Unlike the Euclidean distance, it considers the correlation between the parameters, not depending on the scale of the features. However, such model assumes that data follows an unique multivariate Gaussian distribution, thereby the data can be modeled by a single Gaussian cluster. The Equation 3.4 shows the MSD calculation given a data point x_i from a dataset X with a mean μ and covariance matrix Σ^{-1}.

$$MSD(x_i) = (x_i - \mu)\Sigma^{-1}(x_i - \mu)^T \tag{3.4}$$

3.1.2.2 Gaussian mixture model

The Gaussian mixture model (GMM) is a model-based data clustering algorithm that overcomes the limitations of MSD. GMM allows the form of multiple ellipsoidal-shaped clusters or components based on probability density estimations using the Expectation-Maximization (EM) (PATEL; KUSHWAHA, 2020; DEMPSTER; LAIRD; RUBIN, 1977). Each cluster is modeled as a Gaussian distribution. Unlike K-means, which is based only on the mean, GMM works with the mean and the covariance, providing a better quantitative measure of fitness per number of clusters. Multivariate finite mixture models capture the main clusters. Hence, the GMM can learn non-linear relationships, assuming that the data can be modeled by a finite multivariate Gaussian distribution, as shown in Figure 6

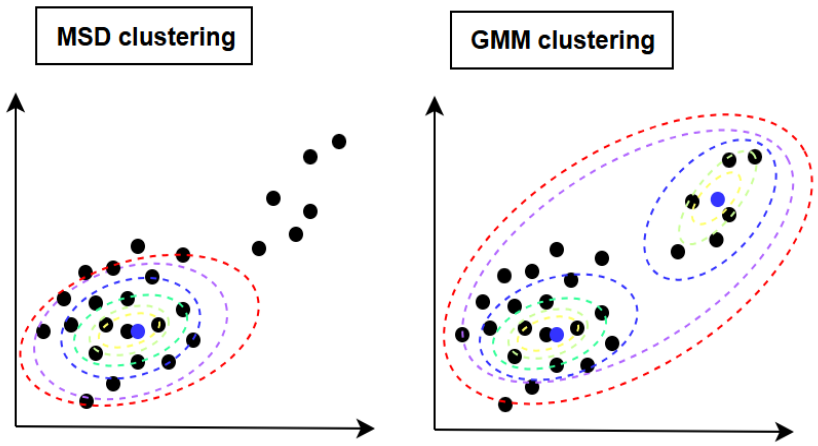


Figure 6 – Comparison between MSD and GMM clustering.

For a GMM, each component $g(x|\theta_k)$ is represented as a Gaussian distribution,

$$g(x|\theta_k) = \frac{exp\{-\frac{1}{2}(x-\mu_k)^T\Sigma_k^{-1}(x-\mu_k)\}}{(2\pi)^{m/2}\sqrt{det(\Sigma_k)}}, \tag{3.5}$$

being each component denoted by the parameters, $\theta_k = \{\mu_k, \Sigma_k\}$, composed by the mean vector, μ_k and the covariance matrix, Σ_k. Thus, a GMM is completely specified by a set of parameters $\Theta = \{\alpha_1, \alpha_2, ..., \alpha_K, \theta_1, \theta_2, ..., \theta_K\}$.

Hence, a finite mixture model, $g(x|\Theta)$, is the weighted sum of $K > 1$ components $g(x|\theta_k)$ in $\mathbb{R}^m$,

$$g(x|\Theta) = \sum_{k=1}^{K} \alpha_k g(x|\theta_k), \tag{3.6}$$

where α_k corresponds to the weight of each component. These weights are positive $\alpha_k > 0$ with $\sum_{k=1}^{K} \alpha_k = 1$.

The expectation–maximization (EM) local search method is one of the most used methods for estimating the GMM parameters (DEMPSTER; LAIRD; RUBIN, 1977; MCLACHLAN, 2000). This method consists of two steps: i) expectation and ii) maximization. For the log-likelihood (LogL), $log(g(X|\Theta)) = log(\prod_{i=1}^{n} g(x_i|\Theta))$, the two steps are applied alternately converge to a local optimum (DEMPSTER; LAIRD; RUBIN, 1977). The choice of initial parameters Θ directly impacts the performance of the EM algorithm, as a wrong choice of the initial parameter can result in many replications of this method during execution.

3.1.3 Density-based algorithms

3.1.3.1 Density-based spatial clustering of applications with noise

Unlike the previous algorithms, the Density-based spatial clustering of applications with noise (DBSCAN) forms clusters based on the density of the data (DENG, 2020). Datapoints with similar densities will belong to the same cluster. DBSCAN was first proposed by Ester et al.

in 1996 to cluster data of arbitrary shapes with noise in spatial and non-spatial high-dimensional databases (ESTER et al., 1996).

One of the major benefits of DBSCAN is that this algorithm does not require presenting the number of clusters, as the model finds this parameter by itself. Instead of centroids, DBSCAN forms clusters by finding the core points. Two user-chosen parameters ϵ and MinPts, which mean the radius of the neighborhood and minimum number of points in the ϵ-neighborhood of a core point, respectively, are the most relevant parameters for DBSCAN. The minimum number of neighbors within a radius ϵ (with a user-chosen distance measure) is used to estimate the minimum density level. Samples that satisfy the minimum density conditions are considered a core point. All neighbors within the ϵ radius are considered part of the same cluster as the central point. Points that are not center points and outside the radius of a center point are considered noise and do not belong to any cluster.

According to the distribution information of the data, this algorithm identifies regions with different densities and generates the cluster results (DENG, 2020). The DBSCAN algorithm steps are summarized as follows:

Algorithm 3 DBSCAN clustering algorithm

Initialize ϵ and $MinPts$.
Given a dataset $X_{i=1}^{N}$ with N data points.
for each x_i in X **do**
    if x_i do not belong to any cluster **then**
        Calculate the distances d between x_i and every other data points.
        if $d <= \epsilon$ **then**
            The data points are stored in a list L.
        else
            Pass to the next data point.
        end if
        if The number of elements in $L >= MinPts$ **then**
            Elements are marked as a cluster.
            x_i is defined as a core point.
        end if
    else if x_i is not a core point $\ \& \ \notin$ to any cluster **then**
        x_i is classified as a noise.
    end if
end for

3.2 Dimensionality reduction-based methods

In several scientific fields, such as image processing, time series analysis, recommendation systems, and text generation, databases have more variables or features than observations (SORZANO; VARGAS; MONTANO, 2014). Correspondingly, these high-dimensional data pose limitations in the performance of various ML algorithms, as these are sensitive to data with

a large number of dimensions. Sensitivity to high-dimensional data is characterized by cases where increasing the number of features degrades the model's accuracy (REDDY et al., 2020). In that sense, dimensionality reduction techniques emerge with the ability to analyze datasets with many dimensions and extract the smallest set that contains the most relevant information about the dataset (HINTON; SALAKHUTDINOV, 2006). PCA and autoencoder are examples of ML models that may be used to reduce the dimensionality of data. These techniques will be discussed in this section.

3.2.1 Principal component analysis

Principal component analysis (PCA) is a widely used technique in ML and statistics for dimensionality reduction and data visualization (KRAMER, 1991). PCA can reduce the data dimensionality by transforming it into a new coordinate system, where the variables are uncorrelated, orthogonal, and ordered by the amount of variance they capture (KURITA, 2019; ABDI; WILLIAMS, 2010). This transformation is achieved by finding the eigenvectors and eigenvalues of the covariance matrix.

In general, given a dataset X with N observations and m variables, the first step in PCA is to center the data. This is done by subtracting the mean of each variable from the respective variable values. This process ensures that the new coordinate system is aligned with the directions of maximum variance in the data rather than being centered on the means of the variables. Consequently, the arrangement of data points in this new coordinate system more accurately represents the underlying variance structure of the data. The centered data matrix is denoted as X'. Next, PCA computes the covariance matrix of the centered data X'. Each element of this matrix is calculated as the covariance between pairs of variables, and the matrix is then rearranged into a square symmetric matrix. Besides, the eigenvector decomposition is computed. The eigenvectors $\mathbf{v}$ of a covariance matrix $\mathbf{A}$ in PCA are found by solving the characteristic equation:

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v}, \tag{3.7}$$

where involves the following variables: $\mathbf{A}$, which represents the covariance matrix; $\mathbf{v}$, the eigenvector of the matrix $\mathbf{A}$; and λ, the eigenvalue corresponding to $\mathbf{v}$. To find the eigenvectors, one must compute the eigenvalues λ by solving the determinant equation:

$$\det(\mathbf{A} - \lambda\mathbf{I}) = 0, \tag{3.8}$$

where $\mathbf{I}$ is the identity matrix. After finding the eigenvalues, the corresponding eigenvectors are estimated by substitution into the characteristic equation. PCA then calculates eigenvectors $\mathbf{v_1}$, $\mathbf{v_2}$, ... $\mathbf{v_m}$ and eigenvalues $\lambda_1, \lambda_2, ..., \lambda_m$ of $\mathbf{A}$. These eigenvectors form the new basis vectors of the transformed coordinate system.

Finally, PCA selects a subset of the eigenvectors, called principal components (PC), based on the amount of variance they capture. The d PCs corresponding to the d largest eigenvalues represent the most important directions of variation in the data, as shown in Figure 7.

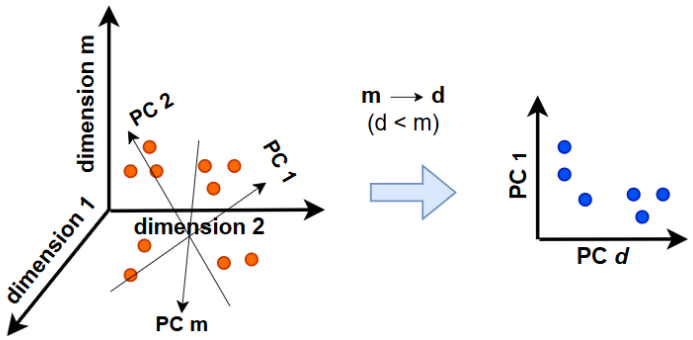


Figure 7 – Illustration of the PCA dimensionality reduction.

Specifically, the covariance matrix represents the relationships between the variables in the dataset. The eigenvectors and eigenvalues provide important information about the directions of maximum variance and the amount of variance captured by each direction, respectively. Leveraging such ability, PCA compresses the dimensions of the data into a lower one with the minimum losses possible, thereby being useful for scenarios where the number of features tends to be large.

3.2.2 Autoencoder

Autoencoders (AEs) refer to neural network-based models trained to copy their input to their output (GOODFELLOW, 2016). These models can reduce the dimensionality of the data by encoding the input features into a compressed representation comprising the most significant relationships within the data distribution. The same model can reverse the compressed data to the original space without significant information loss. Unlike PCA, AEs use non-linear mapping functions that better learn implicit information from the data in a latent dimensional space (GOODFELLOW, 2016).

In the most general form, AEs comprise an auto-associative neural network with three hidden layers, where the input features are the output target, and the middle layer (also referred to as bottleneck) is designed for learning the compressed representations of the input variables (BANK; KOENIGSTEIN; GIRYES, 2023). Although one may consider the ideal AEs to succeed in perfectly copying its inputs, it is not useful. In this case, the ultimate goal is to copy only the most valuable aspects of the input variables, constraining the number of hidden units to be less than the number of features, which allows the learning of valuable properties of the data.

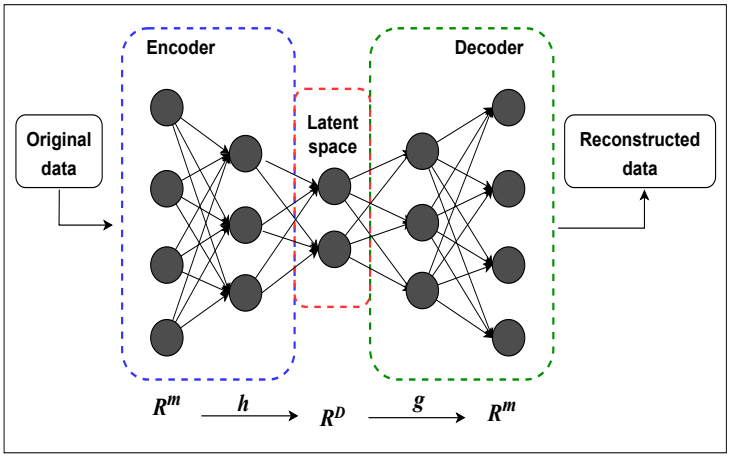


Figure 8 – Autoencoder architecture comprising the three main components (encoder, decoder, and latent space) and the input data reconstruction using the mapping functions.

Diving into the encoder-decoder process of the model, as shown in Fig. 8, the hidden layer describes a code used to represent the input by performing a mapping $\boldsymbol{h} : \Re^m \rightarrow \Re^d$ of the input $\boldsymbol{x}$. After, the data are reversed to the original space by a demapping operation $\boldsymbol{g} : \Re^d \rightarrow \Re^m$. The learning process aims to find the set of parameters $\Theta = \{\theta, \theta'\}$ that minimizes the loss function

$$L(\Theta) = \frac{1}{N} \sum_{\forall \boldsymbol{x} \in \boldsymbol{X}} \left\| \boldsymbol{x} - g_{\theta'}(h_\theta(\boldsymbol{x})) \right\|^2, \tag{3.9}$$

where $L(\cdot)$ is a loss function penalizing $g_{\theta'}(h_\theta(\boldsymbol{x}))$ for being dissimilar from $\boldsymbol{x}$, i.e, a mean square reconstruction error.

The most common approach for the encoder and decoder is through affine mappings and nonlinear functions:

$$\begin{aligned} h_\theta(\boldsymbol{x}) &= s_h(\theta \boldsymbol{x} + \boldsymbol{b}), \\ g_{\theta'}(\boldsymbol{x}) &= s_g(\theta^T \boldsymbol{x} + \boldsymbol{a}). \end{aligned} \tag{3.10}$$

Thus, the set of parameters turns out into $\Theta = \{\theta, \boldsymbol{b}, \boldsymbol{a}\}$, where $\boldsymbol{b}$ and $\boldsymbol{a}$ are the bias and θ is the weight matrix. This scheme of shared weights is called symmetric architecture (KRAMER, 1991).

One can note that by eliminating the mapping and demapping layers, the AE would have a combined network with only one hidden layer, i.e., the bottleneck layer between inputs and outputs. If the nodes of the bottleneck layer were linear, this would correspond similar to (linear) PCA. Moreover, the performance of an AE network with only one bottleneck layer of sigmoidal nodes is often no better than linear PCA, as by using a sigmoidal function, only linear combinations of the inputs compressed by the sigmoid into the range (- 1, 1) could be represented (BANK; KOENIGSTEIN; GIRYES, 2023).

Traditionally, AEs are used for dimensionality reduction or feature learning. However, learning a representation via different types of autoencoders, which can be modified or combined to create new models, can be used for classification, clustering, anomaly detection, or even generative modeling. In this case, an autoencoder learns to span a similar (not the same) subspace as to the principal components space (KRAMER, 1991).

Parallel to the advances in deep learning models, additional autoencoder-based approaches have been made by combining simple autoencoders to form deeper models. In this case, multiple autoencoders can be stacked on top of each other to compose a deep model formed by simple modules of classifiers (BANK; KOENIGSTEIN; GIRYES, 2023). These simple models can be independently trained to perform specific tasks instead of training an entire network to learn a global task.

4 Proposed approach

4.1 UL-based approach for fault detection and localization in optical networks

The steps for applying the proposed UL-based techniques for fault detection and localization are highlighted in this section. Figure 9 summarized the main steps. As all the techniques work in an unsupervised manner, only data from normal conditions is used for training. The telemetry is collected and pre-processed to be further fed into each model for training. Once the model is properly trained, Fault Indicators (FIs) are calculated for each sample in the training data. FIs are a key parameter for fault detection and localization, indicating whether a sample is derived from a fault or normal condition. Given a sample with m dimensions corresponding to the telemetry parameters, the respective FI for this sample is a vector with the m dimensions with values indicating whether a telemetry parameter has a faulty behavior. Leveraging this, fault localization can be performed as an eventual fault can be detected individually for each telemetry parameter.

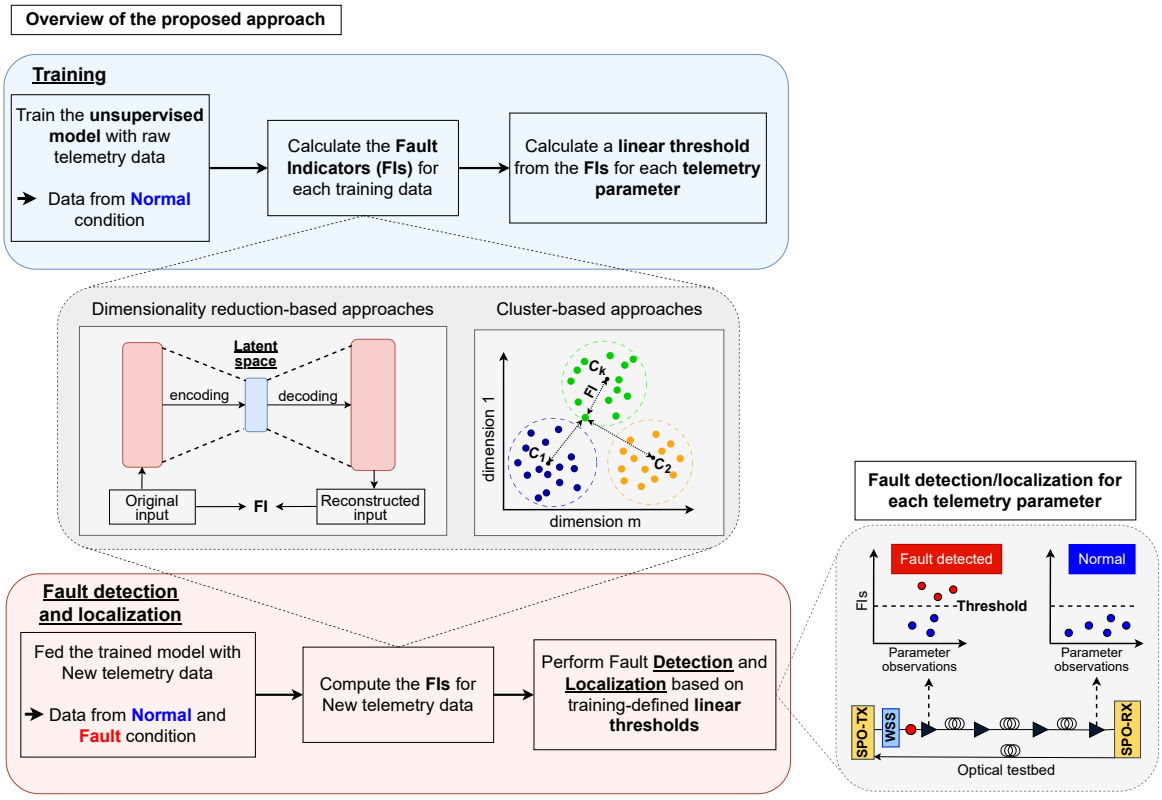


Figure 9 – Overview of the proposed UL-based approach for fault detection and localization.

As shown in Figure 9, the FIs calculation is different for each type of UL model. For the

cluster-based models, except for the distribution-based models, the FIs are calculated through the generated centroid values. For PCA and AE, the FIs are the reconstruction error (RE) generated from the input reconstruction at the output. The following topics highlight the procedures for FI calculation for each proposed method:

- ***FI for K-means, FCM, and DBSCAN.*** Considering a dataset $X_{n\times m}$, with n samples of m dimensions, the K-means FIs are given using the following Equation:

$$FI(x_i, \mu_j) = \sqrt{\sum_{p=1}^{m} (x_{ip} - \mu_{jp})^2}, \tag{4.1}$$

  where x_i is a sample from the dataset, and μ_j is the closest centroid to the sample.

  The FIs for the FCM are the same as those for K-means since both algorithms are based on the same clustering approach. For the FCM, FIs are given using the Equation 4.1. Similarly, the FIs calculation for the DBSCAN leverages the same equation with only one difference. The only difference is that this one uses the final core points instead of centroids. DBSCAN calculates the FIs with the following equation:

$$FI(x_i, x_j) = \sqrt{\sum_{p=1}^{m} (x_{ip} - c_{jp})^2}, \tag{4.2}$$

  where c_j refers to the closest core point to the dataset sample x_i. As the models are trained using only data from normal conditions, FIs calculated from the training set present small values, as the centroids or core points were generated based on such data. Conversely, fault samples will generate large FIs as these are expected to present larger distances from the centroids or core points.

- ***FI for MSD and GMM.*** The FIs for MSD and GMM differ from those of other algorithms presented. In this case, the Mahalanobis squared distance calculates the FI. Unlike the Euclidean distance, it takes into account the mean μ_k and covariance matrix Σ_k^{-1} related to the distribution of the data, for which the FIs are given using the following equation:

$$FI(x_i|\theta_k) = (x_i - \mu_k)\Sigma_k^{-1}(x_i - \mu_k)^T, \tag{4.3}$$

  where x_i is a sample from the dataset and θ_k is the closest component to the sample.

- ***FI for PCA and AE.*** Either PCA or AE can be used for anomaly detection tasks. In that case, the model leverages the reconstruction error (RE), which can be calculated using the Mean Squared Error (MSE) equation. Once the model is trained with only data from normal conditions, REs for such conditions will present small values compared to the ones computed from the data from fault conditions. Hence, REs act like FIs for the fault detection task. In practice, REs can be extracted from the loss function described in Eq.

3.9. Given an input datapoint $\boldsymbol{x}$ and a reconstructed input datapoint $\hat{\boldsymbol{x}}$ predicted at the output, the FI can be computed following the formula:

$$FI = \sum_{\forall \boldsymbol{x} \in \boldsymbol{X}} \|\boldsymbol{x} - \hat{\boldsymbol{x}}\|^2 . \tag{4.4}$$

Following the next procedures in Figure 9, once all the training-derived FIs are computed, linear thresholds are calculated for each telemetry parameter that composes the dataset. These thresholds are calculated based on a pre-defined percentile value from the FIs. Data from normal and fault conditions are used for the test phase. The proposed UL models aim to distinguish normal and fault samples accurately. Hence, the trained model computes new FIs for every sample in the testing dataset.

In the next step, the FIs derived from the testing dataset are evaluated under the training-defined thresholds. Note that each FI is a vector composed of variables corresponding to each telemetry parameter, as each FI is computed based on telemetry data collected from the optical network testbed. Thus, each feature value is analyzed for its respective threshold defined in the training phase. Hence, as shown in Figure 10, if a sample of a specific telemetry parameter eventually presents an FI value higher than the threshold, the sample is classified as a fault condition, indicating that the fault is located at the respective telemetry parameter; otherwise, it is normal.

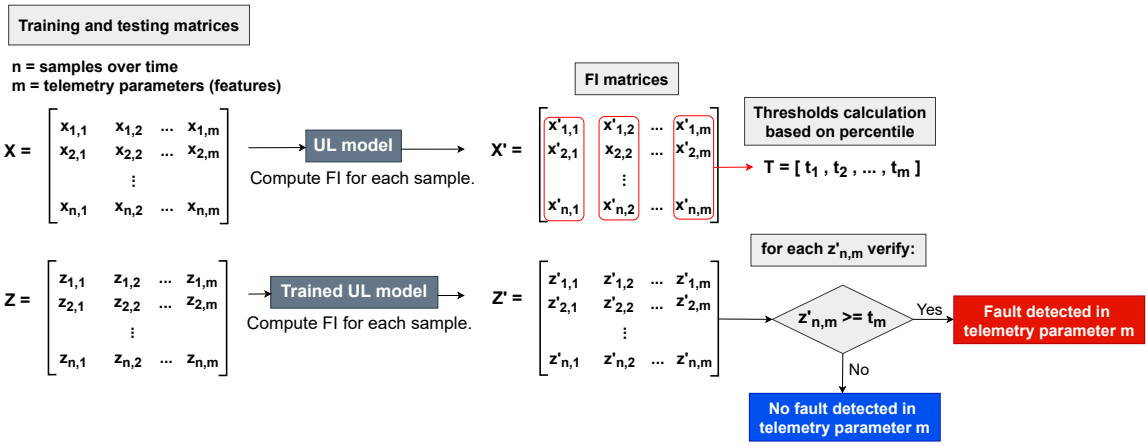


Figure 10 – Fault localization scheme of the proposed approach.

4.2 Model fine-tuning

Correctly mapping the normal condition of the network is important to ensure adequate fault management. For this reason, it is crucial to pass all the models through a fine-tuning process. In that case, this process aims to find the optimal hyperparameter values required to provide the best results in terms of accuracy and model computational complexity.

Different UL-based models use different methods to find their optimal hyperparameter values. For cluster-based models, an appropriate number of clusters is needed so that the models can better capture the normal behavior of the network. For MSD, no optimization is required as only one cluster or component can be generated in that approach. For DBSCAN, several training trials are conducted to evaluate different values of ϵ and $MinPts$. The same is made for PCA, varying the number of principal components. The best values are the ones that achieved the best fault detection performance during the training phase (only data from normal conditions are used in this phase).

For the rest of the models, a typical fine-tuning method can be used: the Akaike information criteria (AIC) (KRAMER, 1991; KODINARIYA; MAKWANA et al., 2013). AIC refers to a function expressing the trade-off between fitting accuracy and the number of adjustable parameters. For K-means, FCM, and GMM, the AIC uses the negative log-likelihood and adds a penalizing term associated with the number of parameters, penalizing more complex models. The AIC is given by:

$$AIC = -2L(A) + 2v(A), \tag{4.5}$$

where $L(A)$ is the log likelihood function and $v(A)$ is the number of free parameters in the clustering model A. Several trials with different numbers of clusters are conducted. The location of the minima in AIC defines the best model. The AIC equation is similar to the AE model, with a few changes. The AIC, in this case, is given by:

$$AIC = \ln(e) + \frac{2N_w}{X}, \tag{4.6}$$

where $N_w = (s + f + 1)(M_1 + M_2) + s + f$ is the number of weights, where s is the number of nodes (neurons) in the input and output layer, f is the number of nodes in the bottleneck layer (latent space), M_1 and M_2 are the number of nodes in mapping (encoding) and demapping (decoding) layers, respectively. $X = Nxm$ is the number of entries in the data matrix, and $e = E/(2X)$ is an average sum of squares error. Minimization of these functions identifies AE models that are neither over- nor under-parameterized. Several trials with different AE architectures (i.e., varying the number of nodes in the encoder, latent space, and decoder layers) are evaluated to find the best AE node architecture.

5 Results and Discussion

This chapter presents fault detection and localization results for each proposed approach. Information about the optical testbed used to generate the dataset evaluated in this work is also detailed. Moreover, the procedures for data pre-processing are highlighted, including selecting the telemetry parameters and splitting the data into training and testing sets.

Additionally, the optimization of the hyperparameters of each technique is carried out. The performance of the cluster-based approaches and the dimensionality reduction techniques are first discussed separately. Finally, for the overall analysis, all methods are compared on their performance in the evaluated scenario.

5.1 Experimental setup and data acquisition

In this work, the proposed approaches consider an experimental testbed to collect the required data. This testbed is part of the ARNO testbed at the Tecip Institute of Scuola Superiore Sat'Anna in Pisa, Italy. The considered testbed is shown in Fig. 11 and presents optical network devices dedicated to research activities for studying novel control, management, and monitoring schemes (INRETE, 2021).

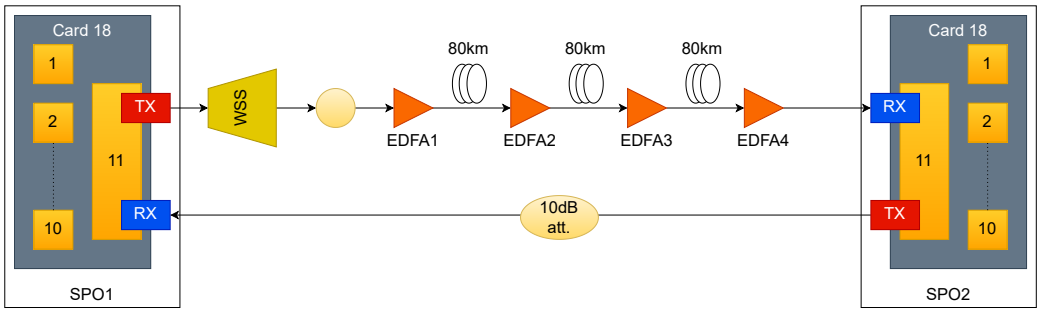


Figure 11 – Transparent optical network testbed.

The testbed includes two Ericcson SPO 1400 devices (i.e., SPO-TX/SPO1 and SPO-RX/SPO2) commercial transport nodes with Optical Transport Network (OTN) functionalities. Each SPO device is equipped with a 100Gb/s OTN muxponder (installed at slot 18) presenting one Dense Wavelength Division Multiplexing (DWDM) optical line (port 11) and ten 10Gb/s Ethernet tributaries. The muxponder performs coherent detection, collecting the main coherent metrics (i.e., pre-FEC BER, OSNR). The testbed includes one WSS that receives as input the line port of the 100G muxponder of SPO-TX and as output the multi-span link. At the exit of the WSS, a 10 dB attenuator is installed. The link going from SPO-TX to SPO-RX consists of 3 spans. A single-mode-fiber spool with a length of 80 km is included in each span, resulting in a 240 km total link length.

4 EDFAs compensate for the power attenuation experienced along each traversed span. All the EDFAs are controlled via SPO devices (SPO-TX controls EDFA1 and EDFA2, while SPO-RX controls EDFA3 and EDFA4) and present a configurable gain in the range of 15-25dB, with output mute power of 0.4 dBm. All the amplifiers are configured in constant gain mode, with a gain value that allows one to enter each span with 0 dBm of optical power. The reverse link (i.e., from SPO-RX to SPO-TX) is in a back-to-back configuration, presenting only an attenuator of 10 dB.

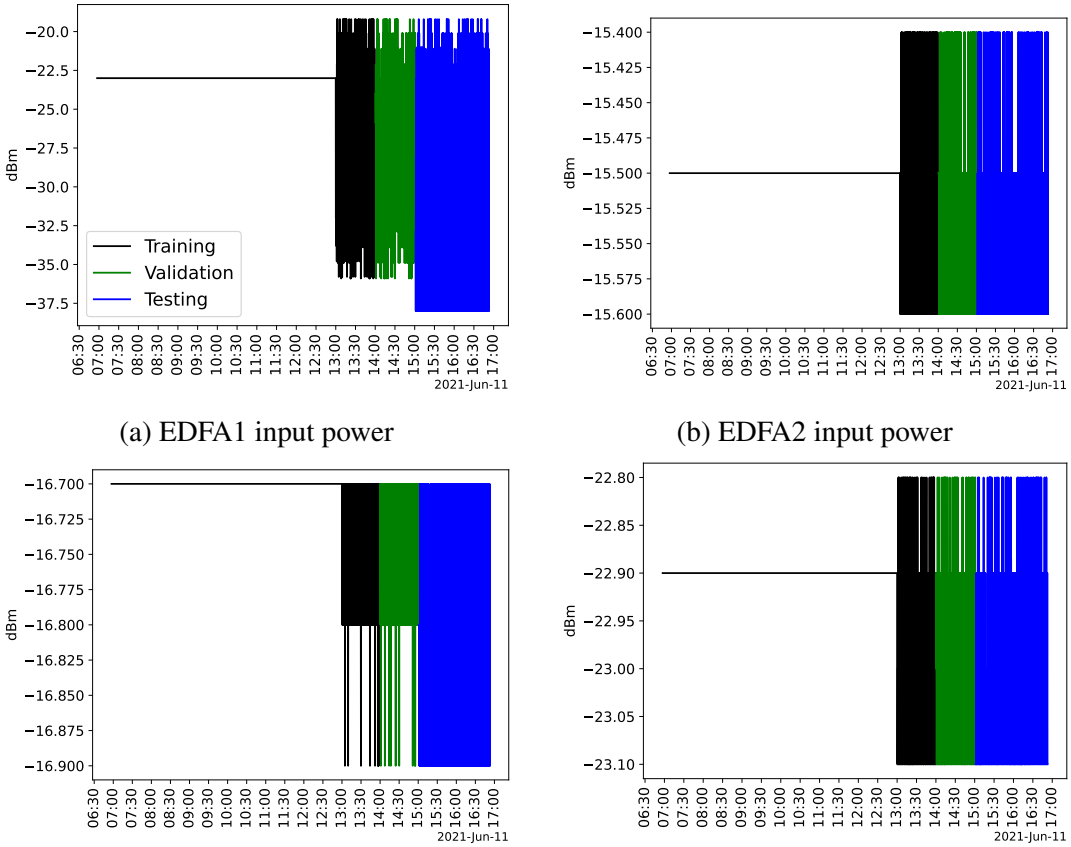
The monitoring system proposed by Sgambelluri et al. (SGAMBELLURI et al., 2021) has been used to collect the data. By selecting the appropriate metric at each traversed element (i.e., the input/output optical power levels at each amplifier and the OSNR and the BER at the transponders), we relied on the Kafka-based telemetry system to distribute the selected metrics. In contrast, a new collector (i.e., Kafka consumer, receiving the telemetry) has been designed to export the data in a CSV format. The dataset has been uploaded into a public GitHub repository (INRETE, 2021).

5.2 Dataset pre-processing and evaluation metrics

The dataset comprises samples collected with a sampling frequency of approximately 3.5 seconds, limited by the minimum time required by the commercial equipment to provide coherent transmission data (i.e., BER). The dataset considered was collected around 10 hours. Before the models are trained, the dataset undergoes exploratory data analysis (EDA).

Some missed values are found in the data, representing a considerable loss of information. This loss might affect the performance of the proposed techniques; thus, an interpolation technique is used to replace these missing data. Several interpolation techniques were evaluated, and a spline linear interpolation technique (JUNNINEN et al., 2004) presented the best fitting, totaling 13.948 samples in the final dataset. For localization purposes, only the EDFA's input power data were used. Among them, the first 80% of the data is used for training and the last 20% for testing, as shown in Figure 12. It is possible to notice that only the EDFA1 input power data show clear and noticeable drops in magnitude, as expected due to the experimental configuration. Therefore, detected outliers should be expected only for the data arising from this equipment. Furthermore, only in the last two hours are the simulated failure conditions corresponding to the 20% used for testing. For every 50 seconds of data collection, 10 seconds of random failures in the optical system are simulated, corresponding to 711 samples related to failures.

Moreover, the scenario is based on the injection of soft failures into the system (by setting 25 dB of attenuation, the received signal goes below the receptor sensitivity). Although the generation of the link failures has been designed to make the detection process more complicated, in the training and validation data, only samples from the normal system conditions are available, allowing the novelty detection approaches to detect deviations from the normal conditions as



(a) EDFA1 input power

(b) EDFA2 input power

(c) EDFA3 input power

(d) EDFA4 input power

Figure 12 – Optical dataset along the training, validation, and testing sets.

possible outliers. This data partitioning can be seen in Figure 12.

The outlier detection performance of the compared techniques is evaluated in terms of Type I (false-positive) and Type II (false-negative) indications of faults. Type I and Type II errors can be calculated using the following equations:

$$Type\,I\ = \frac{FP}{FP + TN}, \tag{5.1}$$

$$Type\,II\ = \frac{FN}{FN + TP}, \tag{5.2}$$

where TP is the total number of normal samples that are correctly classified as normal samples, FP is the total number of normal samples that are wrongly classified as fault samples, TN is the total number of fault samples that are correctly classified as fault samples and FN is the total number of fault samples that are wrongly classified as normal samples.

As previously highlighted in Chapter 4, thresholds are required to realize the fault detection/localization by the models. Hence, a linear threshold is defined for 99% confidence over the training data. This threshold value is based on the 99% percentile from the FIs derived

from the training data. Figure 13 presents boxplots of FIs for each EDFA parameter for illustrative purposes.

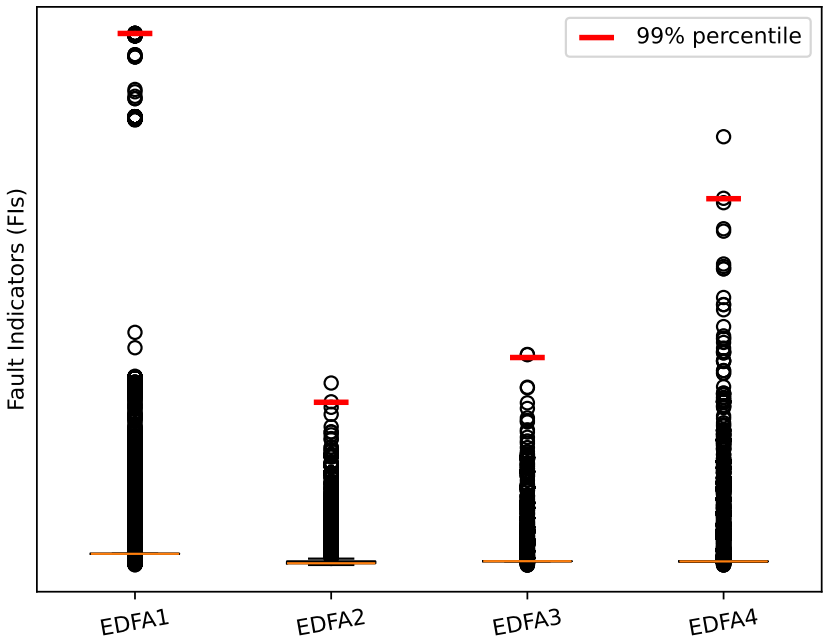


Figure 13 – Illustrative figure presenting the percentile indication through boxplot of FIs derived from the training data.

5.3 Hyperparameter tuning

A critical aspect of working with ML models is the fine-tuning process. From the experimental setup, we conducted intensive experiments with the previously described techniques to find the best choice of hyperparameters for each technique.

For DBSCAN, several values for its two main parameters were evaluated. The values for ϵ and $MinPts$ were: $[0.1, 0.001, 0.0001, 0.00001, 0.000001]$ and $[5, 50, 100, 800, 1500, 2000]$, respectively. Based on the training accuracy, the best values correspond to $\epsilon = 0.00001$ and $MinPts = 1500$. For the PCA model, the optimal number of principal components (PCs) is required as PCA reduces the dimensionality of the data using these parameters. There are several ways to find that value, but this work defines them based on the retaining variance. More specifically, different numbers of PCs ranging from 1 to 3 were tested to find which value could retain the maximum variance of the data when reduced dimensionality (Figure 14). From these tests, one can note in Figure 14 that when reducing the dimensionality of the data (initially having four features corresponding to the four EDFAs) to a 1-dimensional space (using PCs = 1), approximately 99% of the variance of the data is retained. Therefore, the final PCA MODEL has only one dimension in the latent space.

Regarding AE model optimization, defining the optimal number of neurons and layers that compose an AE architecture is essential to finding the most appropriate model in terms of

performance and efficiency. In this work, the AIC is used to estimate the quality of an AE model compared to others.

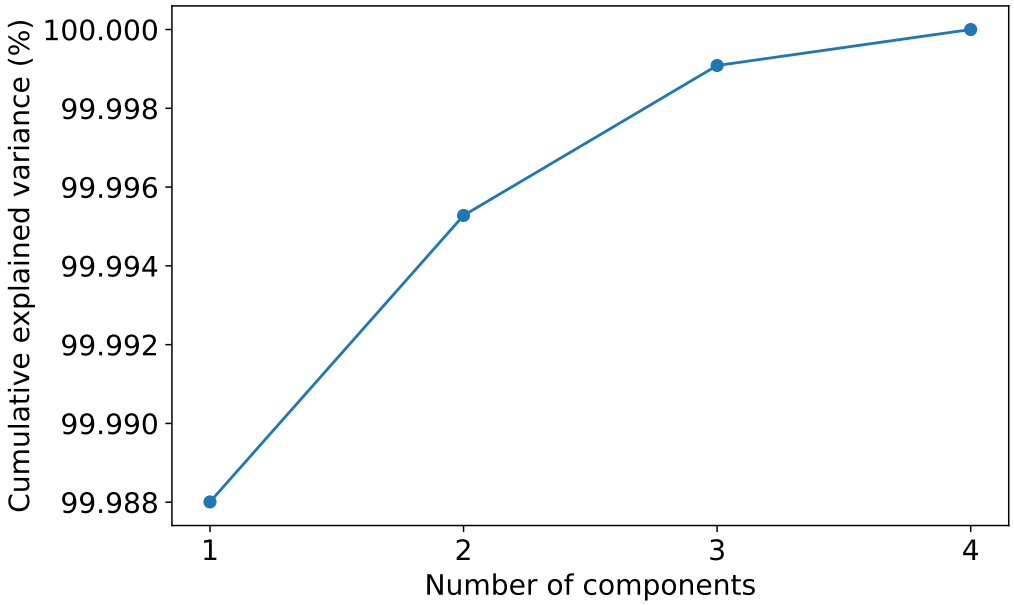


Figure 14 – Explained variance retained by PCA per number of components.

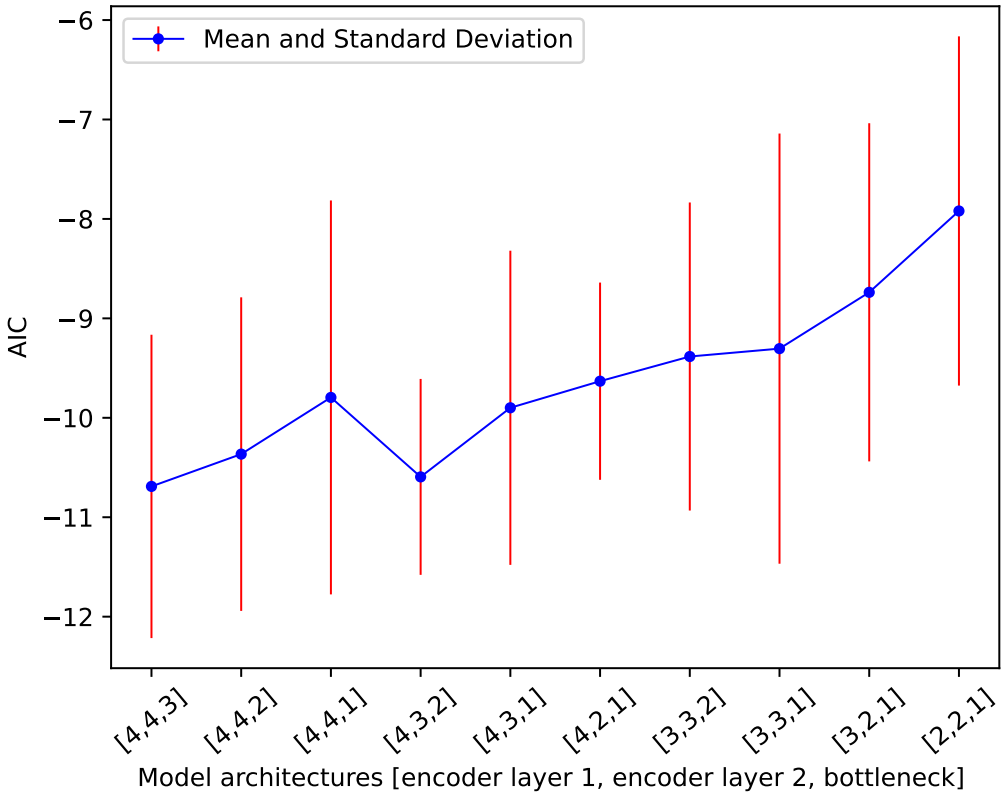


Figure 15 – Average AIC values of different AE node combinations.

Ten node combinations with different numbers of neurons for five layers (two encoders, one bottleneck, and two decoders) were tested on 20 training trials of 200 epochs each. At the

end of each training, the mean squared error (MSE) is used in Equation 4.6 to calculate the AIC. The final AIC value for each combination of mapping/demapping nodes is the average AIC value from the 20 training trials. Fig. 15 presents the respective AIC values for each model architecture. Note that each architecture is represented in the figure by only the encoder and bottleneck layers, as the decoder layers are the same as the encoder, following a symmetric model architecture.

By analyzing the location of the minima in AIC presented in Fig. 15, the two models with the architectures [4,4,3] and [4,3,2] can be chosen as the most adequate models for our specific scenario. In that sense, the model with fewer neurons ([4,3,2]) was selected to be further trained during 20 trials based on the minimum variation of training loss throughout a pre-defined maximum number of epochs. More specifically, the model starts the training and keeps it until the training loss variation between the actual epochs is smaller than a specific pre-defined value. It prevents the model from suffering *overfitting* by stopping the training whether the model does not present a satisfactory decrease in the training loss. Fig. 16 shows the average training loss and the standard deviation from the chosen AE model.

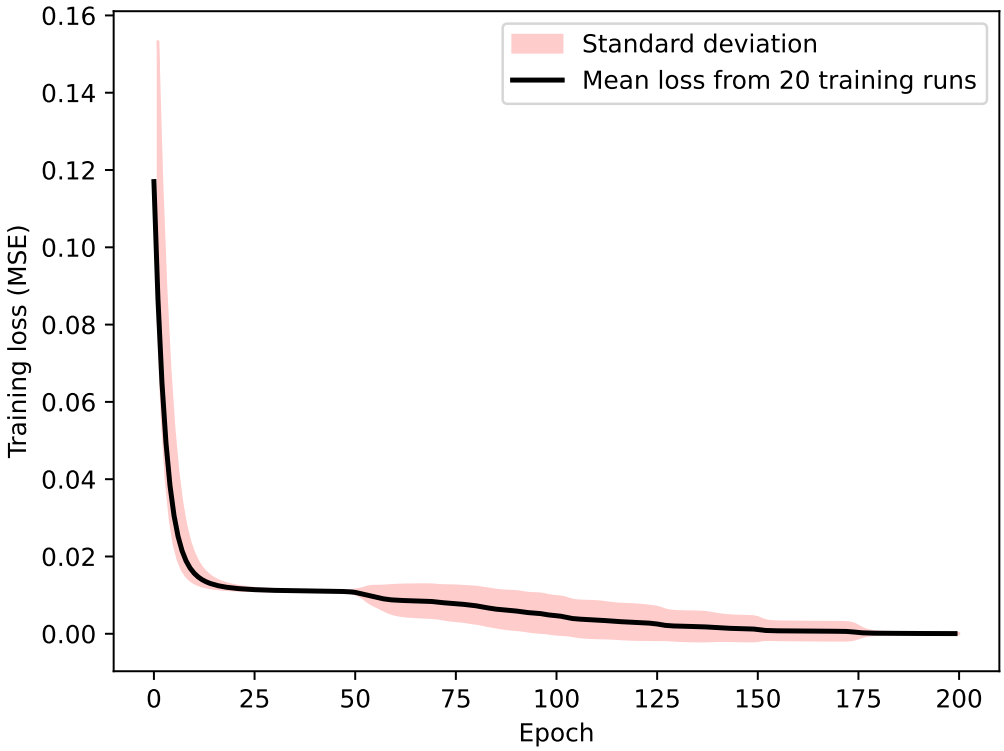
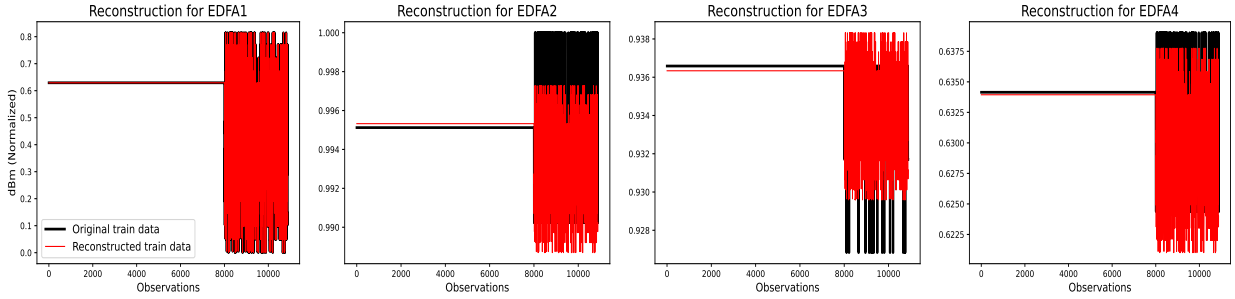


Figure 16 – Average training loss and standard deviation from 20 trials of the best AE model.

It is noted that the model quickly achieved a proper convergence. In that case, the AE model does not learn the stochastic variations in the training data; rather, its underlying functions can be used for model generalization purposes. Additionally, Figure 17 shows the reconstruction ability performance over the training data for the optimal models of AE and PCA. Note that both models presented a poor reconstruction for the data portion under variations in EDFA2 and EDFA3. However, it can still be considered a satisfactory reconstruction result as, in a general

manner, both models could learn the main characteristics of the normal condition in every EDFA.

(a) PCA



(b) AE

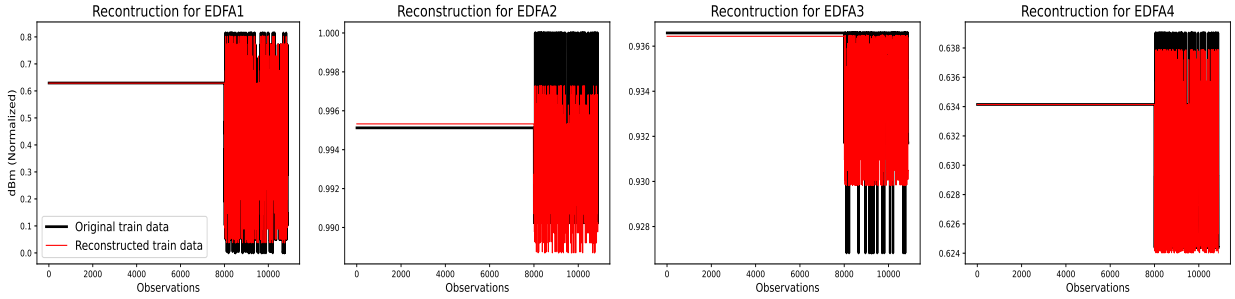


Figure 17 – Comparison between original and reconstructed input data for PCA and AE for each EDFA.

For the cluster-based approaches, the AIC equation described in Equation 4.5 was used to find the optimal number of clusters for each K-means, FCM, and GMM. Finding the optimal value for this parameter is crucial for model generalization purposes, as a poor number of components may not be enough to represent the different distributions of the data properly. Hence, numbers of components (clusters) ranging from 2 to 10 were evaluated for 20 trials for each model. Figure 18 presents the average AIC results and the standard deviation.

As shown in Figure 18, K-means and FCM presented almost the same AIC values (a few tenths of a difference) for each number of components. Moreover, one can note that the K-means, FCM, and GMM models presented the same optimal number of components (2 components), as the minimum AIC values were presented for such value. In fact, the training dataset fed into the models is composed of two different distributions, which justify the resulting AIC-based best number of components. Therefore, two clusters are sufficient to represent the training data adequately.

5.4 Fault management results

5.4.1 Cluster-based approaches

For evaluation of the cluster-based approaches, the number of Type I and Type II errors for each EDFA device in the testing set are presented in Table 1. Type II indications are not

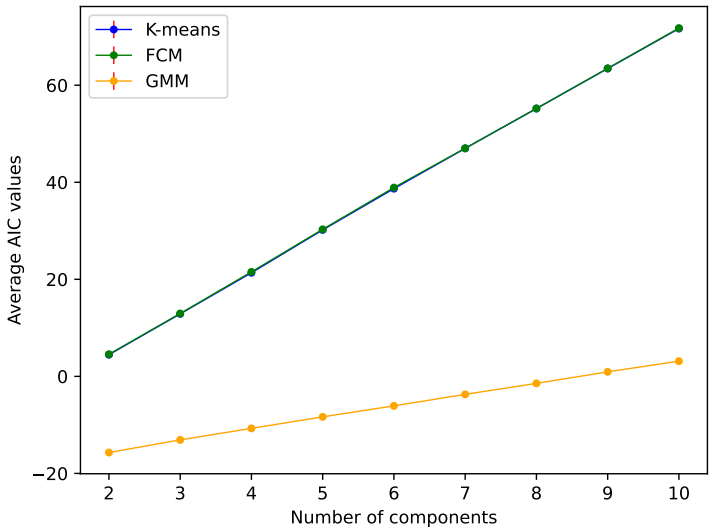


Figure 18 – Average AIC values of each cluster-based algorithm for various number of components.

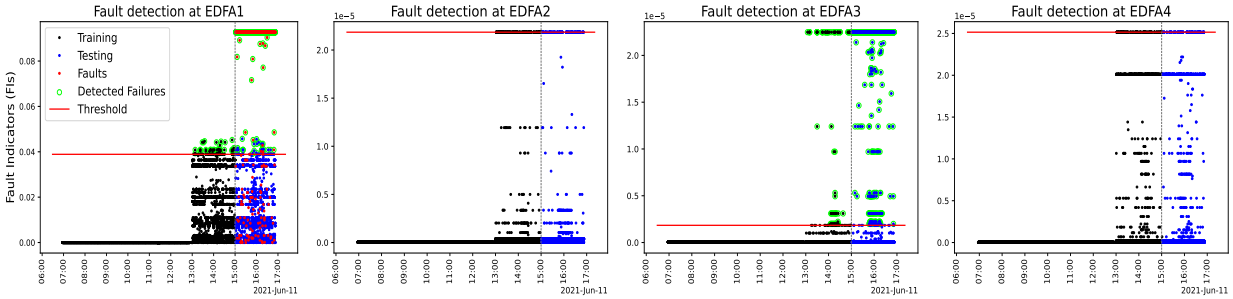
presented for EDFA2, EDFA3, and EDFA4, as no failures were simulated at these devices. The table shows that all models present marginal differences in results over the two comparison metrics. Overall, it can be said that GMM presented a marginal superiority. Furthermore, one may note in Table 1 that all the models misclassified some failure samples presented in EDFA1, corresponding to the respective Type II errors. Part of this is due to the simulated process of failures in the dataset. When the failure condition was simulated, although the label of the respective sample was set to failure, the system gradually changed from one state to the other, creating a few samples labeled as "failure". Thus, those failure samples that were not detected have characteristics related to normal conditions, making them quite hard to identify.

Table 1 – Fault detection results for cluster-based approaches per EDFA.

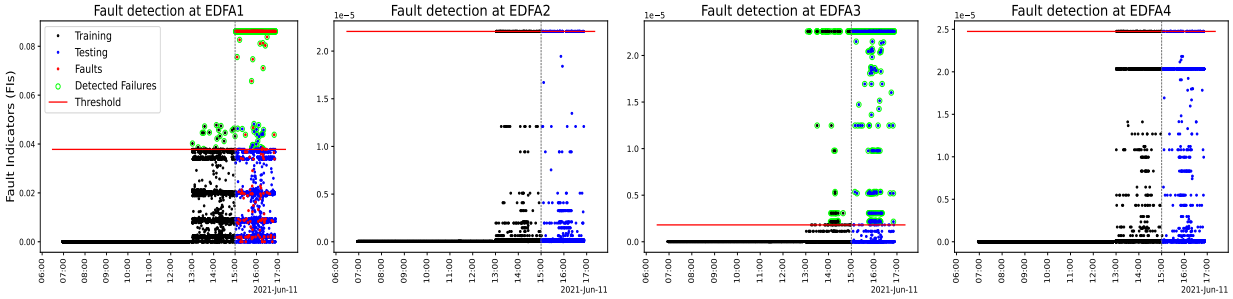
Device	EDFA1		EDFA2		EDFA3		EDFA4	
Model	Type I (%)	Type II (%)	Type I (%)	Type II (%)	Type I (%)	Type II (%)	Type I (%)	Type II (%)
K-means	0.88	3.36	0	-	27.41	-	0	-
FCM	0.68	3.45	0	-	27.41	-	0	-
MSD	0.42	3.39	0	-	26.01	-	0.26	-
GMM	0	3.49	0	-	26.53	-	0	-
DBSCAN	0.42	3.39	0	-	26.01	-	0.26	-

Moreover, one can note the relatively high rate of Type I errors in EDFA3. These results can be explained by observing Figure 12. As shown in the figure, unlike other EDFAs, the

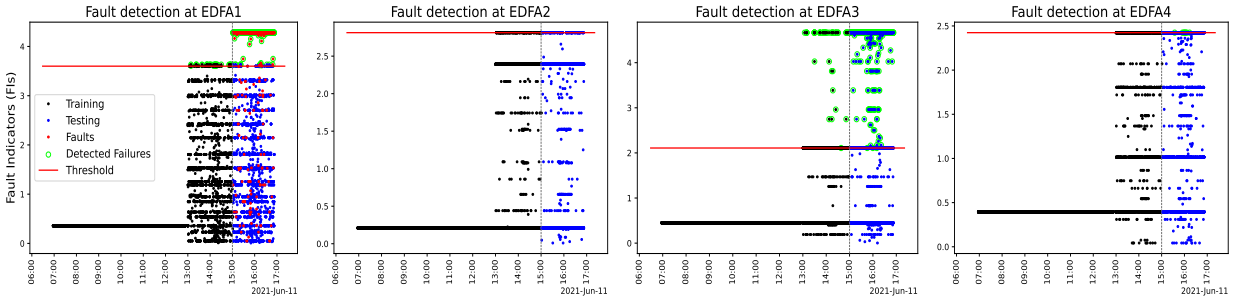
(a) K-means



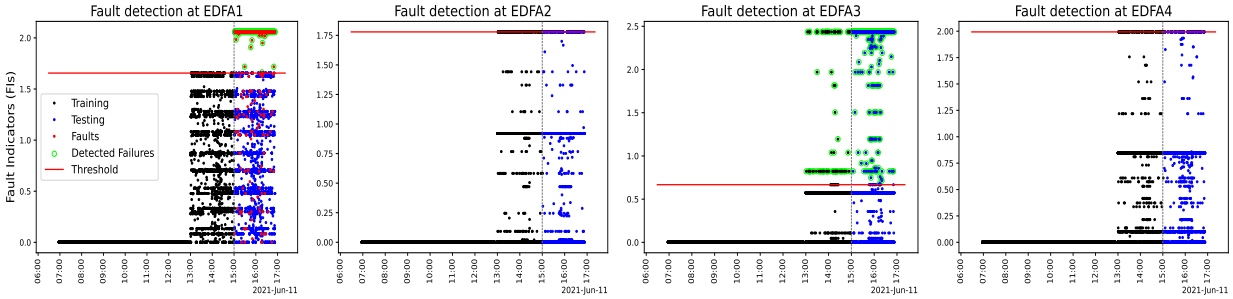
(b) FCM



(c) MSD



(d) GMM



(e) DBSCAN

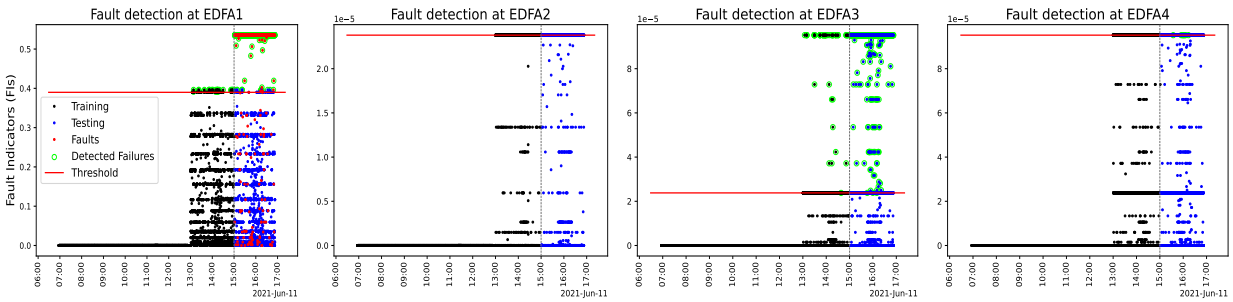


Figure 19 – Fault indicators along with a threshold defined over the training data for each cluster-based approach.

EDFA3 input power values do not present a significant amount of data with variation from the normal condition during the period used for training. Therefore, this fact affects the fault detection performance of the clustering models by reducing model generalization when submitted for fault scenarios in the test phase.

For further evaluation, the FIs for the entire dataset from all cluster-based approaches are shown along the threshold and over time in Figure 19. From a general perspective, all the models successfully detect/locate the faults presented in the EDFA1 equipment, achieving less than 3.50% of Type II errors. Red dots in the EDFA1 figure of each model represent the faults. One can note these dots are mostly arranged above the threshold line, corresponding to around 85% of the 711 fault samples. The red dots below the threshold line represent the amount of Type II errors previously discussed.

In fact, all previous analyses show that all cluster-based approaches achieved marginal differences for both evaluation metrics. In that sense, it is required to closely look at Type I and II errors to find the most appropriate clustering model. For Type I errors, the model accuses the network of being in a fault condition, even though the network is in a normal state. From an economic point of view, the maintenance costs associated with a recurring Type I error can be a problem, increasing operational expenses (OPEX). As shown in Table 1, GMM is the most appropriate option in this scenario since it presented a marginally better result. However, the situation changes when we look at the Type II error. In this scenario, the model indicates normality while the network is in a fault condition. This error affects the availability of network services and directly affects SLA requirements. In that regard, K-means may be the most suitable for a scenario where network availability is critical, as it presents marginally better results than other algorithms for Type II errors.

5.4.2 Dimensionality reduction approaches

The results of the PCA and AE models over the testing set using the comparison metrics are presented in Table 2. Initially, one can note that the models achieve different results for each EDFA. For EDFA1, AE overcomes PCA for Type I errors, as PCA duplicated the number of this type of error. On the other hand, PCA achieved a marginally better result for Type II errors compared to AE. Similar results are shown for EDFA2 and EDFA3, with AE presenting better results for Type I errors. In that case, for those EDFAs, AE could properly learn the variations of the data under normal conditions, being able to correctly detect almost all the normal samples in the testing set.

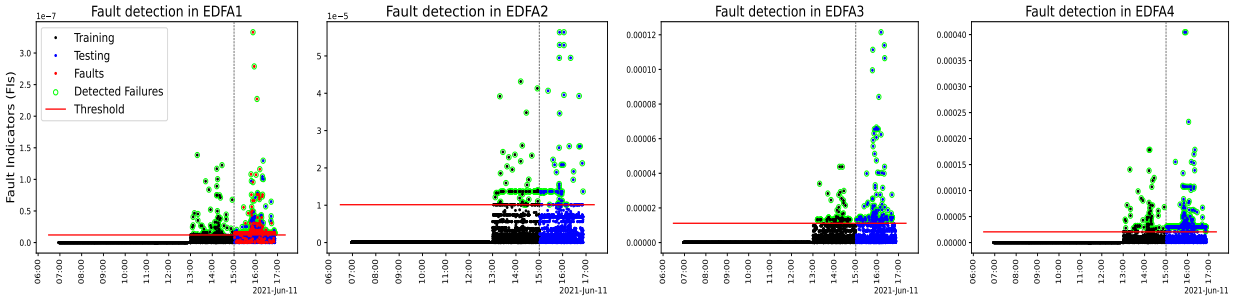
For EDFA4, AE overcomes PCA by far. PCA presented 22.45% of Type I errors. Quantitatively, from 2,353 samples under normal condition in the testing set, PCA misclassified approximately 528 samples. It means that PCA would trigger 528 false alarms. Conversely, AE presented 3.98% of Type I errors in that EDFA. For further analysis, similarly made for cluster-based approaches, Figure 20 presents the fault indicators generated by each model for

Table 2 – Fault detection results for PCA and AE for each EDFA.

Device	EDFA1		EDFA2		EDFA3		EDFA4	
Model	Type I (%)	Type II (%)	Type I (%)	Type II (%)	Type I (%)	Type II (%)	Type I (%)	Type II (%)
PCA	3.39	3.28	2.41	-	7.60	-	22.45	-
AE	1.40	3.36	0.39	-	5.05	-	3.98	-

each EDFA parameter along the entire dataset.

(a) PCA



(b) AE

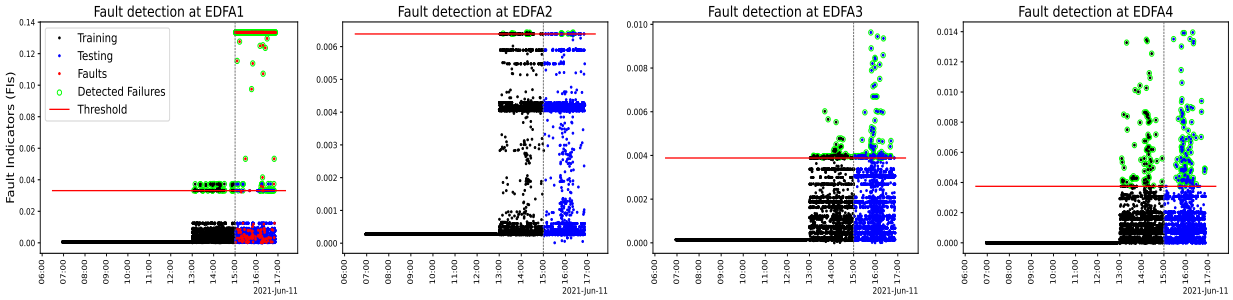


Figure 20 – Fault indicators along with a threshold defined over the training data for PCA and AE models.

As shown for EDFA1 in Figure 20, both models generated large FIs for samples under fault conditions (red dots) compared to the normal condition (blue dots) in the testing set. In that case, PCA and AE could properly reconstruct the input data from normal condition at the output, thereby generating small FIs for these data. In contrast, large FIs are generated for fault samples as the models generate large reconstruction errors for this kind of data. However, one can note in the figure that PCA generated fault sample FIs with values with slight differences from the ones from normal samples, as the red dots are concentrated almost in the same region as the blue dots.

Furthermore, note that for EDFA4 in Figure 20, although in a visual manner, the amount of Type I errors (blue dots above the threshold line) seems the same for both PCA and AE, the difference between these numbers is quite different, as shown in Table 2. Moreover, one can say that if the threshold line were set to a higher value, the number of Type I errors achieved by PCA would be much smaller, as the large amount of blue dots in Figure 20 would appear below the threshold line.

5.4.3 Overall analysis

In general, for the evaluated testbed, one can verify that all the proposed unsupervised learning-based approaches presented reasonable results regarding the minimization of misclassification. In this sense, the models could properly detect samples from fault condition only training with data from normal condition. Among the cluster-based techniques, it is quite difficult to define the best model, as Table 1 shows that the models have very similar results. However, one can say that GMM performed marginally better, as the model correctly classified the samples under normal conditions in the testing set. The GMM modeling of the different distributions presented in the training data can explain these results. On the other hand, among the dimensionality reduction-based approaches, AE outperformed PCA, presenting a better balance of Type I and Type II errors.

Furthermore, comparing the two categories of UL-based techniques proposed in this work, all the cluster-based approaches presented poor fault detection at EDFA3, as seen in Table 1. The 27% of Type I errors result in approximately 600 false alarms in the evaluated testing dataset. In this case, the limitations imposed by their clustering properties compromise the proper learning of normal conditions distributions presented in the dataset for this parameter. Conversely, the two dimensionality-reduction proposed techniques overcome this limitation. In this case, PCA and AE reached less than 8% of Type I errors for EDFA3, as seen in Table 2. AE, specifically, reached 5% of Type I errors in this parameter, the best performance over all the techniques for this EDFA.

Moreover, AE can be chosen as the most appropriate technique for the evaluated dataset, as this model presented stable and consistent results for fault localization. As AE works based on neural networks, it can potentially learn non-linear distributions in the data, taking into account hidden relevant information from the data. Compared to PCA, the AE model can use non-linear mapping functions to better represent the data in the latent space, thereby better learning the variations presented in data from normal conditions. Furthermore, although neural network-based models are commonly more expensive in terms of computational cost than cluster-based techniques (depending on the number of neurons and hidden layers), they do not represent a limitation for the particular fault detection task in optical networks. The computational cost is irrelevant in the training phase as the models are trained offline. Therefore, once trained, all the proposed techniques have similar computational costs in the testing phase, as the only procedure required in this phase is to calculate the FI for every sample from the testing dataset.

6 Conclusion, Future Research and Published Works

6.1 Main conclusions

It is essential to promote accurate approaches to fault management that can correctly distinguish telemetry data with variations caused by fault conditions from those caused by operational conditions. Although several ML-based approaches have been used to address this challenge, most rely on supervised learning algorithms that require large amounts of telemetry data derived from fault conditions. In this regard, requiring data from these conditions represents a significant limitation as these data are absent in real-world scenarios. Herein, to cope with this limitation, this dissertation exploited the use of unsupervised learning methods to properly detect and localize faults in optical networks without the explicit acquisition of fault data.

Specifically, this work leveraged the ability of UL methods to train with only data from normal conditions. In this sense, the fault detection and localization performances of different cluster-based algorithms and dimensionality reduction techniques were compared. As centroid-based clustering algorithms, K-means and FCM could accurately distinguish fault data from normal data by generating centroids that properly represent the operational variations of the data from normal conditions. On the other hand, the distribution-based clustering algorithms MSD and GMM demonstrated proper learning of the different distributions related to normal conditions. As the only density-based algorithm used in this work, the DBSCAN model was also demonstrated to be suitable as it adequately learned the normal operational variations by analyzing the different densities within the dataset. However, although all the cluster-based approaches posed as excellent UL models to distinguish fault samples from normal ones, these approaches demonstrated inconsistency in the case of fault localization, triggering significant amounts of false-alarms by misclassifying data from normal conditions.

Conversely, the dimensionality reduction techniques PCA and AE demonstrated better results in terms of fault localization. By leveraging their particular model properties, these techniques could learn hidden patterns from the network's normal conditions, extracting the most relevant information necessary to perform consistent fault localization. In this case, AE can be pointed out as the best model overall, as unlike PCA, which can only learn linear distributions of the data, AE leverages a neural network with non-linear mapping functions to extract intrinsic information from the data. Hence, AE presented the most consistent and stable results for fault localization, correctly detecting both normal and fault samples.

The cluster-based approaches and the dimensionality reduction techniques were evaluated and compared through their applications on a dataset generated from an optical testbed. This

testbed comprises several optical parameters, such as the input power of the four EDFA that compose the total link. An attenuator inserted at the beginning of the link was designed to simulate normal and fault variations that cause changes in the optical parameters of the telemetry data. Specifically, only the first EDFA of the network is affected by the fault variations, which means that the fault is expected to be detected through the variations in the optical input power of this parameter.

Ultimately, it can be concluded that the AE model had the best fault management performance in terms of minimization of Type I/II errors, providing, on average, a Type I rate of 2.7% and a Type II rate of 3.36%. These results demonstrate that AE can be the chosen UL approach to deploy in real monitoring optical systems if the goal is the minimization of misclassifications. However, cluster-based algorithms can be the most appropriate approach when the goal is to provide fault management with model interpretation concerns. In terms of the overall analysis, based on the fault management performance results, the proposed approaches demonstrated: (i) be effective in extracting intrinsic characteristics related to operational variability of the data, being able to model the normal conditions of the data properly and thereby distinguishing them from the fault ones, and (ii) have the potential to be easily deployed in practical optical networks fault management scenarios as they can properly be trained with only data derived from network's normal conditions.

6.2 Future research topics

For future research, there are several promising directions for extending the work. First, further studies evaluating deep learning-based approaches that, besides fault detection and localization, can promote fault forecasting to trigger an alarm to the network operator before the occurrence of an eventual fault, reducing OPEX. Second, additional tests leveraging a more extended dataset with different numbers of optical devices and telemetry parameters must be evaluated regarding fault management scalability, planning to evaluate the performance of the proposed techniques on large-scale optical networks commonly seen in real-world environments. Third, the evaluation of the use of physics-informed neural network models that integrate physical properties of the optical systems to improve the ability to manage faults in optical networks. Finally, the evaluation of large language model (LLM)-based approaches integrated with retrieval augmented generation (RAG) techniques to develop an AI agent that can provide optical network services, such as fault management and resource allocation, in a fully automated manner, which requires minimum human intervention.

6.3 Published works

The main original works that serve as foundations of this dissertation are listed below. The works are organized by the publication period and the type of conference (national or

international).

Original works published in international conferences:

1. ANDREI RIBEIRO, RAFAEL SALES, FABRÍCIO R. LOBATO, JOÃO C.W.A. COSTA, MOISÉS SILVA, ANDREA SGAMBELLURI, LUCA VALCARENGHI, LENA WOSINSKA, “PCA-Assisted Fuzzy Clustering Approach for Soft-Failure Detection in Optical Networks", 2024 International Conference on Optical Network Design and Modeling (ONDM), Madrid, Spain, 2024, pp. 1-5, DOI: 10.23919/ONDM61578.2024.10582591

2. RAFAEL SALES, ANDREI RIBEIRO, MOISÉS SILVA, FABRÍCIO R. LOBATO, ANDREA SGAMBELLURI, LUCA VALCARENGHI, JOÃO C.W.A. COSTA, “Disaggregated Confidentiality-Preserving Scheme for Fault Detection in Optical Networks", 2024 Optical Fiber Communications Conference and Exhibition (OFC), San Diego, CA, USA, 2024, pp. 1-3.

Original works published in national conferences:

1. INGRID RAMOS, ADRYELE OLIVEIRA, ANDREI RIBEIRO, RAFAEL SALES, FABRÍCIO R. LOBATO, JOÃO C.W.A. COSTA, MOISÉS SILVA, “A Comparative Study on Principal Components Analysis for Failure Detection in Optical Networks", 2024 Brazilian XLII Symposium on Telecommunications and Signal Processing (SBrT), Belém, Brazil, 2024, DOI: 10.14209/sbrt.2024.1571036264

2. RAFAEL SALES, ANDREI RIBEIRO, FABRÍCIO R. LOBATO, MOISÉS SILVA, JOÃO C.W.A. COSTA, “Comparative Analysis on Cluster-based Algorithms for Failure Management in Optical Networks", 2024 Brazilian XLII Symposium on Telecommunications and Signal Processing (SBrT), Belém, Brazil, 2024, DOI: 10.14209/sbrt.2024.1571035792

3. RAFAEL SALES, ANDREI RIBEIRO, FABRÍCIO R. LOBATO, MOISÉS SILVA, JOÃO C.W.A. COSTA, “A Cluster-based Approach for Single Failure Management in Transparent Optical Networks", 2023 XIII National Conference in Communications, Networks, and Information Security (ENCOM), Belém, Brazil, 2023.

4. ANDREI RIBEIRO, FABRÍCIO R. LOBATO, VICTOR CARDOSO, JOÃO C.W.A. COSTA, “An Unsupervised ML-driven Approach for Water Leak Detection in Localized Irrigation System", 2023 XIII National Conference in Communications, Networks, and Information Security (ENCOM), Belém, Brazil, 2023.

Bibliography

ABDI, H.; WILLIAMS, L. J. Principal component analysis. *Wiley interdisciplinary reviews: computational statistics*, Wiley Online Library, v. 2, n. 4, p. 433–459, 2010. Cited on page 24.

AGRAWAL, G. P. *Fiber-optic communication systems*. [S.l.]: John Wiley & Sons, 2012. Cited on page 10.

AHMED, M.; SERAJ, R.; ISLAM, S. M. S. The k-means algorithm: A comprehensive survey and performance evaluation. *Electronics*, MDPI, v. 9, n. 8, p. 1295, 2020. Cited on page 17.

ALHAWARAT, M.; HEGAZI, M. Revisiting k-means and topic modeling, a comparison study to cluster arabic documents. *IEEE Access*, IEEE, v. 6, p. 42740–42749, 2018. Cited on page 18.

ASYALI, M. H. et al. Gene expression profile classification: a review. *Current Bioinformatics*, Bentham Science Publishers, v. 1, n. 1, p. 55–73, 2006. Cited on page 20.

BANK, D.; KOENIGSTEIN, N.; GIRYES, R. Autoencoders. *Machine learning for data science handbook: data mining and knowledge discovery handbook*, Springer, p. 353–374, 2023. Cited 3 times on pages 25, 26, and 27.

BARZEGAR, S. et al. Soft-failure localization and device working parameters estimation in disaggregated scenarios. In: OPTICA PUBLISHING GROUP. *Optical Fiber Communication Conference*. [S.l.], 2020. p. Th1F–2. Cited on page 1.

BEZDEK, J. C. *Pattern recognition with fuzzy objective function algorithms*. [S.l.]: Springer Science & Business Media, 2013. Cited on page 19.

BORLAND, J. *Analyzing the Internet collapse*. [S.l.]: MIT Technology Review, 2008. Cited on page 16.

CARROZZO, G. et al. Ai-driven zero-touch operations, security and trust in multi-operator 5g networks: A conceptual architecture. In: IEEE. *2020 European conference on networks and communications (EuCNC)*. [S.l.], 2020. p. 254–258. Cited on page 2.

CHEN, X. et al. Leveraging deep learning to achieve knowledge-based autonomous service provisioning in broker-based multi-domain sd-eons with proactive and intelligent predictions of multi-domain traffic. In: IEEE. *2017 European Conference on Optical Communication (ECOC)*. [S.l.], 2017. p. 1–3. Cited on page 2.

CHEN, X. et al. On cooperative fault management in multi-domain optical networks using hybrid learning. *IEEE Journal of Selected Topics in Quantum Electronics*, IEEE, v. 28, n. 4: Mach. Learn. in Photon. Commun. and Meas. Syst., p. 1–9, 2022. Cited on page 4.

CHEN, X. et al. Knowledge-based autonomous service provisioning in multi-domain elastic optical networks. *IEEE Communications Magazine*, IEEE, v. 56, n. 8, p. 152–158, 2018. Cited 2 times on pages 1 and 2.

DEMPSTER, A. P.; LAIRD, N. M.; RUBIN, D. B. Maximum likelihood from incomplete data via the em algorithm. *Journal of the royal statistical society: series B (methodological)*, Wiley Online Library, v. 39, n. 1, p. 1–22, 1977. Cited 2 times on pages 21 and 22.

DENG, D. Dbscan clustering algorithm based on density. In: IEEE. *2020 7th international forum on electrical engineering and automation (IFEEA)*. [S.l.], 2020. p. 949–953. Cited 2 times on pages 22 and 23.

ESTER, M. et al. A density-based algorithm for discovering clusters in large spatial databases with noise. In: *kdd*. [S.l.: s.n.], 1996. v. 96, n. 34, p. 226–231. Cited on page 23.

FERREIRA, M. R.; CARVALHO, F. D. A. D. Kernel fuzzy c-means with automatic variable weighting. *Fuzzy Sets and Systems*, Elsevier, v. 237, p. 1–46, 2014. Cited on page 19.

FOERSTER, K.-T.; SCHMID, S.; VISSICCHIO, S. Survey of consistent software-defined network updates. *IEEE Communications Surveys & Tutorials*, IEEE, v. 21, n. 2, p. 1435–1461, 2018. Cited on page 13.

FOSTER, J. S. et al. Report of the commission to assess the threat to the united states from electromagnetic pulse (emp) attack. *Critical National Infrastructures Report*, v. 1, 2004. Cited on page 16.

FURDEK, M. et al. An overview of security challenges in communication networks. In: IEEE. *2016 8th International Workshop on Resilient Networks Design and Modeling (RNDM)*. [S.l.], 2016. p. 43–50. Cited on page 16.

GOODFELLOW, I. *Deep learning*. [S.l.]: MIT press, 2016. Cited 2 times on pages 17 and 25.

GU, R.; YANG, Z.; JI, Y. Machine learning for intelligent optical networks: A comprehensive survey. *Journal of Network and Computer Applications*, Elsevier, v. 157, p. 102576, 2020. Cited 2 times on pages 1 and 12.

GUO, J.; ZHU, Z. When deep learning meets inter-datacenter optical network management: Advantages and vulnerabilities. *Journal of Lightwave Technology*, IEEE, v. 36, n. 20, p. 4761–4773, 2018. Cited on page 2.

HENRY, P. Introduction to lightwave transmission. *IEEE Communications Magazine*, IEEE, v. 23, n. 5, p. 12–16, 1985. Cited on page 10.

HINTON, G. E.; SALAKHUTDINOV, R. R. Reducing the dimensionality of data with neural networks. *science*, American Association for the Advancement of Science, v. 313, n. 5786, p. 504–507, 2006. Cited on page 24.

HO, K.-P.; KAHN, J. M. Mode-dependent loss and gain: statistics and effect on mode-division multiplexing. *Optics express*, Optica Publishing Group, v. 19, n. 17, p. 16612–16635, 2011. Cited on page 11.

HUI, R. *Introduction to fiber-optic communications*. [S.l.]: Academic Press, 2019. Cited on page 11.

INRETE, L. *Optical Failure Dataset*. [S.l.], 2021. Available at https://github.com/Network-AndServices/optical-failure-dataset. Cited 2 times on pages 32 and 33.

IWAHASHI, E. Trends in long-wavelength single-mode transmission systems and demonstrations in japan. *IEEE Journal of Quantum Electronics*, IEEE, v. 17, n. 6, p. 890–896, 1981. Cited on page 10.

JAIN, A. K.; MURTY, M. N.; FLYNN, P. J. Data clustering: a review. *ACM computing surveys (CSUR)*, Acm New York, NY, USA, v. 31, n. 3, p. 264–323, 1999. Cited on page 17.

JUNNINEN, H. et al. Methods for imputation of missing values in air quality data sets. *Atmospheric environment*, Elsevier, v. 38, n. 18, p. 2895–2907, 2004. Cited on page 33.

KAO, K. C.; HOCKHAM, G. A. Dielectric-fibre surface waveguides for optical frequencies. In: IET. *Proceedings of the Institution of Electrical Engineers*. [S.l.], 1966. v. 113, n. 7, p. 1151–1158. Cited on page 9.

KODINARIYA, T. M.; MAKWANA, P. R. et al. Review on determining number of cluster in k-means clustering. *International Journal*, v. 1, n. 6, p. 90–95, 2013. Cited 2 times on pages 19 and 31.

KRAMER, M. A. Nonlinear principal component analysis using autoassociative neural networks. *AIChE journal*, Wiley Online Library, v. 37, n. 2, p. 233–243, 1991. Cited 4 times on pages 24, 26, 27, and 31.

KRUSE, L. E. et al. Experimental investigation of machine-learning-based soft-failure management using the optical spectrum. *Journal of Optical Communications and Networking*, Optica Publishing Group, v. 16, n. 2, p. 94–103, 2024. Cited 2 times on pages 1 and 5.

KURITA, T. Principal component analysis (pca). *Computer vision: a reference guide*, Springer, p. 1–4, 2019. Cited on page 24.

LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. *nature*, Nature Publishing Group UK London, v. 521, n. 7553, p. 436–444, 2015. Cited on page 17.

LI, B. et al. Deep-learning-assisted network orchestration for on-demand and cost-effective vnf service chaining in inter-dc elastic optical networks. *Journal of Optical Communications and Networking*, Optica Publishing Group, v. 10, n. 10, p. D29–D41, 2018. Cited on page 2.

LI, T. Optical fiber communication-the state of the art. *IEEE Transactions on Communications*, IEEE, v. 26, n. 7, p. 946–955, 1978. Cited on page 10.

LI, T. Advances in optical fiber communications: An historical perspective. *IEEE Journal on Selected Areas in Communications*, IEEE, v. 1, n. 3, p. 356–372, 1983. Cited on page 10.

LIU, S. et al. Semi-supervised anomaly detection with imbalanced data for failure detection in optical networks. In: OPTICA PUBLISHING GROUP. *Optical Fiber Communication Conference*. [S.l.], 2021. p. Th1A–24. Cited on page 5.

LUN, H. et al. Gan based soft failure detection and identification for long-haul coherent transmission systems. In: IEEE. *2021 Optical Fiber Communications Conference and Exhibition (OFC)*. [S.l.], 2021. p. 1–3. Cited on page 4.

MACQUEEN, J. Some methods for classification and analysis of multivariate observations. In: *Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability/University of California Press*. [S.l.: s.n.], 1967. Cited on page 18.

MAIONE, T.; SELL, D.; WOLAVER, D. Atlanta fiber system experiment: Practical 45-mb/s regenerator for lightwave transmission. *Bell System Technical Journal*, Wiley Online Library, v. 57, n. 6, p. 1837–1856, 1978. Cited on page 10.

MAROM, D. M. et al. Survey of photonic switching architectures and technologies in support of spatially and spectrally flexible optical networking. *Journal of Optical Communications and Networking*, IEEE, v. 9, n. 1, p. 1–26, 2017. Cited on page 11.

MAS, C.; TOMKOS, I.; TONGUZ, O. K. Failure location algorithm for transparent optical networks. *IEEE Journal on Selected Areas in Communications*, IEEE, v. 23, n. 8, p. 1508–1519, 2005. Cited 2 times on pages 1 and 15.

MAYER, K. S. et al. Soft failure localization using machine learning with sdn-based network-wide telemetry. In: IEEE. *2020 European Conference on Optical Communications (ECOC)*. [S.l.], 2020. p. 1–4. Cited on page 4.

MAYER, K. S. et al. Machine-learning-based soft-failure localization with partial software-defined networking telemetry. *Journal of Optical Communications and Networking*, Optica Publishing Group, v. 13, n. 10, p. E122–E131, 2021. Cited 2 times on pages 1 and 2.

MCLACHLAN, G. Finite mixture models. *A wiley-interscience publication*, 2000. Cited on page 22.

MELLO, D. A. de A.; BARBOSA, F. A. *Digital Coherent Optical Systems*. [S.l.]: Springer, 2021. Cited 3 times on pages 9, 10, and 11.

MIYA, T. et al. Ultimate low-loss single-mode fibre at 1.55 μm. *Electronics Letters*, IET, v. 15, n. 4, p. 106–108, 1979. Cited on page 9.

MUKHERJEE, B.; HABIB, M. F.; DIKBIYIK, F. Network adaptability from disaster disruptions and cascading failures. *IEEE Communications Magazine*, IEEE, v. 52, n. 5, p. 230–238, 2014. Cited on page 16.

MUSUMECI, F. Machine learning for failure management in optical networks. In: OPTICA PUBLISHING GROUP. *Optical Fiber Communication Conference*. [S.l.], 2021. p. Th4J–1. Cited 2 times on pages 1 and 2.

MUSUMECI, F. et al. An overview on application of machine learning techniques in optical networks. *IEEE Communications Surveys & Tutorials*, IEEE, v. 21, n. 2, p. 1383–1408, 2018. Cited 2 times on pages 1 and 2.

NAYAK, J.; NAIK, B.; BEHERA, H. Fuzzy c-means (fcm) clustering algorithm: a decade review from 2000 to 2014. In: SPRINGER. *Computational Intelligence in Data Mining-Volume 2: Proceedings of the International Conference on CIDM, 20-21 December 2014*. [S.l.], 2015. p. 133–149. Cited on page 19.

NSAFOA-YEBOAH, K. et al. Software-defined networks for optical networks using flexible orchestration: Advances, challenges, and opportunities. *Journal of Computer Networks and Communications*, Wiley Online Library, v. 2022, n. 1, p. 5037702, 2022. Cited on page 12.

ODA, S. et al. A learning living network with open roadms. *Journal of Lightwave Technology*, IEEE, v. 35, n. 8, p. 1350–1356, 2017. Cited on page 2.

PATEL, E.; KUSHWAHA, D. S. Clustering cloud workloads: K-means vs gaussian mixture model. *Procedia computer science*, Elsevier, v. 171, p. 158–167, 2020. Cited on page 21.

POINTURIER, Y. Design of low-margin optical networks. *Journal of Optical Communications and Networking*, Optica Publishing Group, v. 9, n. 1, p. A9–A17, 2017. Cited on page 1.

PROIETTI, R. et al. Experimental demonstration of machine-learning-aided qot estimation in multi-domain elastic optical networks with alien wavelengths. *Journal of Optical Communications and Networking*, Optica Publishing Group, v. 11, n. 1, p. A1–A10, 2019. Cited on page 2.

RAFIQUE, D. et al. Cognitive assurance architecture for optical network fault management. *Journal of Lightwave Technology*, IEEE, v. 36, n. 7, p. 1443–1450, 2018. Cited 2 times on pages 2 and 3.

RAK, J. et al. Disaster resilience of optical networks: State of the art, challenges, and opportunities. *Optical Switching and Networking*, Elsevier, v. 42, p. 100619, 2021. Cited on page 1.

RAK, J. et al. Recodis: Resilient communication services protecting end-user applications from disaster-based failures. In: IEEE. *2016 18th International Conference on Transparent Optical Networks (ICTON)*. [S.l.], 2016. p. 1–4. Cited on page 16.

RAMASWAMI, R.; SIVARAJAN, K.; SASAKI, G. *Optical networks: a practical perspective*. [S.l.]: Morgan Kaufmann, 2009. Cited on page 9.

RANGISETTI, A. K.; TAMMA, B. R. Software defined wireless networks: A survey of issues and solutions. *Wireless Personal Communications*, Springer, v. 97, p. 6019–6053, 2017. Cited on page 12.

REDDY, G. T. et al. Analysis of dimensionality reduction techniques on big data. *Ieee Access*, IEEE, v. 8, p. 54776–54788, 2020. Cited on page 24.

REJEB, R.; LEESON, M. S.; GREEN, R. J. Fault and attack management in all-optical networks. *IEEE Communications Magazine*, IEEE, v. 44, n. 11, p. 79–86, 2006. Cited on page 1.

RIBEIRO, A. et al. Pca-assisted fuzzy clustering approach for soft-failure detection in optical networks. In: IEEE. *2024 International Conference on Optical Network Design and Modeling (ONDM)*. [S.l.], 2024. p. 1–5. Cited on page 5.

RUSSELL, S. J.; NORVIG, P. *Artificial intelligence: a modern approach*. [S.l.]: Pearson, 2016. Cited on page 17.

SALES, R. et al. Disaggregated confidentiality-preserving scheme for fault detection in optical networks. In: IEEE. *2024 Optical Fiber Communications Conference and Exhibition (OFC)*. [S.l.], 2024. p. 1–3. Cited on page 5.

SARMADI, H. et al. Ensemble learning-based structural health monitoring by mahalanobis distance metrics. *Structural Control and Health Monitoring*, Wiley Online Library, v. 28, n. 2, p. e2663, 2021. Cited on page 21.

SGAMBELLURI, A. et al. Reliable and scalable kafka-based framework for optical network telemetry. *Journal of Optical Communications and Networking*, Optica Publishing Group, v. 13, n. 10, p. E42–E52, 2021. Cited on page 33.

SHAHKARAMI, S. et al. Machine-learning-based soft-failure detection and identification in optical networks. In: IEEE. *2018 Optical Fiber Communications Conference and Exposition (OFC)*. [S.l.], 2018. p. 1–3. Cited 2 times on pages 2 and 3.

SHIMIZU, D. Y. et al. A deep neural network model for link failure identification in multi-path roadm based networks. In: IEEE. *2020 Photonics North (PN)*. [S.l.], 2020. p. 1–1. Cited 2 times on pages 2 and 4.

SILVA, M. F. et al. Learning long-and short-term temporal patterns for ml-driven fault management in optical communication networks. *IEEE Transactions on Network and Service Management*, IEEE, v. 19, n. 3, p. 2195–2206, 2022. Cited on page 5.

SILVA, M. F. et al. Confidentiality-preserving machine learning algorithms for soft-failure detection in optical communication networks. *Journal of Optical Communications and Networking*, Optica Publishing Group, v. 15, n. 8, p. C212–C222, 2023. Cited on page 5.

SKORIN-KAPOV, N. et al. Physical-layer security in evolving optical networks. *IEEE Communications Magazine*, IEEE, v. 54, n. 8, p. 110–117, 2016. Cited on page 15.

SORZANO, C. O. S.; VARGAS, J.; MONTANO, A. P. A survey of dimensionality reduction techniques. *arXiv preprint arXiv:1403.2877*, 2014. Cited on page 23.

SRIVASTAVA, V.; PANDEY, R. S. A dominance of the channel capacity in load balancing of software defined network. *Wireless Personal Communications*, Springer, v. 112, p. 1859–1873, 2020. Cited 2 times on pages 9 and 13.

SU, M.-C.; CHOU, C.-H. A modified version of the k-means algorithm with a distance based on cluster symmetry. *IEEE Transactions on pattern analysis and machine intelligence*, IEEE, v. 23, n. 6, p. 674–680, 2001. Cited on page 18.

TAYLOR, M. G. Coherent detection method using dsp for demodulation of signal and subsequent equalization of propagation impairments. *IEEE Photonics Technology Letters*, IEEE, v. 16, n. 2, p. 674–676, 2004. Cited on page 11.

THYAGATURU, A. et al. Software defined optical networks (sdons): a comprehensive survey. *IEEE Communications Surveys Tutorials*, vol. 18, n. no. 4, 2016. Cited 2 times on pages 12 and 13.

VELA, A. et al. Ber degradation detection and failure identification in elastic optical networks. *IEEE/OSA Journal of Lightwave Technology*, vol. 35, n. no. 21, 2017. Cited 2 times on pages 2 and 3.

VELA, A. et al. Soft failure localization during commissioning testing and lightpath operation. *IEEE/OSA Journal of Optical Communications and Networking*, vol. 10, n. no. 1, 2018. Cited 2 times on pages 1 and 3.

VELASCO, L. et al. Learning from the optical spectrum: Soft-failure identification and localization. In: IEEE. *2018 Optical Fiber Communications Conference and Exposition (OFC)*. [S.l.], 2018. p. 1–3. Cited on page 4.

VILLA, G. et al. Machine learning techniques in optical networks: a systematic mapping study. *IEEE Access*, IEEE, v. 11, p. 98714–98750, 2023. Cited on page 14.

WANG, D. et al. A review of machine learning-based failure management in optical networks. *Science China Information Sciences*, Springer, v. 65, n. 11, p. 211302, 2022. Cited 4 times on pages 9, 1, 15, and 16.

WINZER, P. J.; ESSIAMBRE, R.-J. Advanced modulation formats for high-capacity optical transport networks. *Journal of Lightwave Technology*, IEEE, v. 24, n. 12, p. 4711–4728, 2006. Cited on page 11.

WINZER, P. J.; FOSCHINI, G. J. Mimo capacities and outage probabilities in spatially multiplexed optical transport systems. *Optics express*, Optica Publishing Group, v. 19, n. 17, p. 16680–16696, 2011. Cited on page 11.

WORDEN, K.; MANSON, G.; ALLMAN, D. Experimental validation of a structural health monitoring methodology: Part i. novelty detection on a laboratory structure. *Journal of sound and vibration*, Elsevier, v. 259, n. 2, p. 323–343, 2003. Cited on page 21.

XIE, M. et al. Ai-driven closed-loop service assurance with service exposures. In: IEEE. *2020 European Conference on Networks and Communications (EuCNC)*. [S.l.], 2020. p. 265–270. Cited on page 2.

YAMADA, J.; MACHIDA, S.; KIMURA, T. 2 gbit/s optical transmission experiments at 1.3 μm with 44 km single-mode fibre. *Electronics Letters*, IET, v. 17, n. 13, p. 479–480, 1981. Cited on page 10.

ZAHMATKESH, A.; KUNZ, T. Software defined multihop wireless networks: Promises and challenges. *Journal of Communications and Networks*, KICS, v. 19, n. 6, p. 546–554, 2017. Cited on page 12.

ZHANG, M. et al. Dynamic and adaptive bandwidth defragmentation in spectrum-sliced elastic optical networks with time-varying traffic. *Journal of Lightwave Technology*, IEEE, v. 32, n. 5, p. 1014–1023, 2014. Cited on page 1.