FOREX Publication
Open Access | Rapid and quality publishing

International Journal of
Electrical and Electronics Research (IJEER)
Research Article | Volume 13, Issue 1 | Pages 89-100 | e-ISSN: 2347-470X

Advanced Artificial Intelligence Techniques for Fault Distance Prediction in Optical Fibres

Omar W. Albawab[1*]

[1]*College of Education for Human Sciences, University of Mosul, Mosul – Iraq*

**Correspondence:* *omar.albawab@uomosul.edu.iq*

ABSTRACT- It is necessary to estimate the value of distances to faults in the optical fibres for proper functioning and fault diagnosis of the fibre optic systems. This research proposes a comparison of result outcomes within numerous categories of machine learning algorithms such as Decision Tree, Random Forest, Gradient Boosting, XGBoost, LightGBM, Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM) on a new and emerging Fibre Optic Fault Distance Dataset. Given dataset contains sequenced OTDR signatures corresponding to different types as well as positions of the faults. The data then went through data pre-processing and was separated into training and test sets where models were trained from 80% of the data and tested on 20%. Model accuracy was also determined by the commonly used performance parameters including R-Square and Mean Squared Error (MSE). They found out that the Random Forest and LSTM models represented the highest R-Square values and the lowest MSE, thus proving the algorithm's capability to forecast the fault distances. The greatest results were obtained by the proposed hybrid LSTM–RF model which combines the sequence processing and ensemble learning. These studies show that effective diagnosis of optical fibre faults using higher level and more complex techniques in machine learning is possible and specify directions on further study and application of this subject.

Keywords: Optical Fibre Fault Detection, Machine Learning, Fault Distance Prediction, OTDR Signatures, Hybrid Models.

ARTICLE INFORMATION

Author(s): Omar W. Albawab;

Received: 24/10/2024; **Accepted**: 05/01/2025; **Published**: 30/03/2025;
e-ISSN: 2347-470X;
Paper Id: IJEER 2410-21;
Citation: 10.37391/ijeer.130113
Webpage-link:
https://ijeer.forexjournal.co.in/archive/volume-13/ijeer-130113.html

crossref member CROSSREF.ORG THE CITATION LINKING BACKBONE

Publisher's Note: FOREX Publication stays neutral with regard to Jurisdictional claims in Published maps and institutional affiliations.

1. INTRODUCTION

Low-loss (ULL) optical fibres can transmit signals in extremely low loss, various working conditions and features with high reliability, and have excellent long-distance transmission performance, is widely used in tough transmission condition such as sparse population area, no man's land and high-altitude area and so on [1]. The research review by our team has revealed that high altitude influences the loss as well as the strength of fibres [2] and has given serious consequences on the splicing points in optical fibre communication system and the safety of the entire system. Literature coming from established authors that gives information on the use of optical fibres for different functions is freely accessible [3].

A new technology that deals with the issues of fibre optic performance monitoring is targeting at efficiently introducing an effective space and time localization and estimation of faults, and at the same time, it does not interrupt the data traffic; therefore, enhancing the reliability of the fibre optic network, especially in real-time remote faulting detection [4]. ToF also adds that these sensors are also being applied to develop multiple use of the monitoring systems of various sorts such as long distance, hostile environments, structural and engine interiors and health assessment [5]. But most of these implementations have taken place in plain areas. The dependable performance of optical fibre communication links at altitude is not thoroughly investigated and suffers from the campaign's extreme conditions and high maintenance costs. These systems should take a very minimal amount of maintenance if needed to be installed in a facility for its proper functioning [6].

Besides, manufacturing defects or various external conditions can negatively impact the fibre optic lines, causing system breakdowns and possible signal interception, which threatens the security of the communication. Sufficient early detection and repair of these faults are therefore inevitable in ensuring the reliability of the communication networks. OTDR is applied for the inspection of the optical cable lines with the help of which fibre-optic monitoring systems can easily locate a fault or a splice loss [7].

The combination and utilization of ML has become a powerful tool in the prediction of fibre optical faults with high accuracy and quick diagnosis of the problems affecting fibre optic networks. The use of ML algorithms can be accomplished for big sets of data from OTDR devices where patterns and abnormalities are detected that may show the presence of faults [8].

For example, the supervised learning algorithms can be applied, in which an actionable model is trained using data acquired from the history and then precise fault locations can be estimated using the specific characteristics of the signals. Clustering and anomaly detection, which are the approach of unsupervised learning, can discover the new type of faults even without the labelling. CNN approaches have been explored for processing OTDR signal patterns to identify fault features that may be difficult for conventional methods to capture [9].

Scientists have described many approaches and methods used for fault identification on the basis of OTDR signals. That is why J. Z. Christopher used the algorithm based on digital signal processing with Gabor series and adopted Rosanne's minimum description length criterion; nevertheless, the problem of noise persists in real-world applications [10].

2. BACKGROUND THEORY

To detect faults in optical fibre frameworks, it's essential to perceive the likely sorts of deficiencies that might happen. In optical fibre correspondence frameworks, two fundamental sorts of shortcomings are regularly experienced: fibre link property flaws and fibre cuts [11]. Fibre link property shortcomings allude to issues with the qualities or properties of the fibre link itself, like imperfections in the material or assembling process. Then again, fibre cuts happen when the actual progression of the fibre is disturbed, frequently because of outside factors like unplanned harm or conscious damage. Distinguishing and tending to these flaws are fundamental for keeping up with the respectability and usefulness of optical fibre frameworks [12].

2.1 Fibre Cable Attribute Faults

While assessing fibre optic reasonableness for correspondence frameworks, fibre link trait deficiencies come first. The key transmission qualities to consider incorporate data transfer capacity, which is impacted by scattering and weakening levels [12]. Scattering alludes to the spreading of signs after some time or distance, while lessening alludes to the deficiency of sign strength. These attributes are constrained by different components, including radiation, assimilation, and dispersing. Guaranteeing ideal scattering and constriction levels is vital for keeping up with dependable and productive correspondence over fibre optic organizations [13].

2.2 Attenuation

Attenuation, or sign misfortune, is a basic calculate deciding the reasonableness of fibre optic transmission for correspondence frameworks. It straightforwardly impacts the greatest distance over which signs can be sent prior to requiring recovery. The allure of optical fibre associations expanded essentially when constriction levels diminished to not exactly those of customary metal connectors, normally under 5 dB/km [14]. This decrease in lessening was viewed as a significant achievement in the seriousness of fibre optic frameworks contrasted with existing correspondence innovations. Prominently, critical headways in decreasing constriction were accomplished during the 1970s, with the presentation of strands displaying weakening levels under 20 dB/km, denoting a limit for the reasonability of fibre optic frameworks in viable applications [15].

2.3 Dispersion

In digital communication systems using optical strands, data is encoded in beats of light communicated from the transmitter to the beneficiary. Be that as it may, during proliferation through the fibre, these heartbeats experience scattering, prompting different types of sign debasement [16]. Beat scattering causes the spreading of heartbeats over the long haul or distance, bringing about peculiarities, for example, between image impedance, where covering beats become vague at the recipient. Scattering in optical filaments can be arranged into two principal types: multi-purpose scattering, which happens in multi-mode strands because of contrasts in mode lengths and speeds, and intramural scattering, predominant in single-mode strands at high information rates, causing beat expanding. Overseeing scattering is significant for keeping up with the respectability and execution of optical correspondence frameworks, guaranteeing solid information transmission overstretched distances [11].

2.4 Fibre Cable Cut

The event of a functioning fibre-optic link interference because of work directed at the link's area is named a fibre cut peculiarity. The degree of administration interruption relies upon the area and number of dynamic fibre optic links impacted by the cut. This stances critical dangers to the broadcast communications industry, influencing network availability, activity, support, and income edges [17]. fibre optics, with their better highlights analysed than conventional transmission mediums, are progressively supplanting microwave transmission frameworks in correspondence organizations. Be that as it may, guaranteeing the unwavering quality and consistent activity of optical fibre frameworks, which normally convey huge volumes of information traffic, stays a critical test [11].

Persistent fibre cutting represents a significant test for telecom administrators, as shown by measurements from a neighbourhood optical fibre organization in 2018. Deficiencies are arranged in light of their effect on framework parcels, administrations, and main drivers. In spine organizations, where fibre link lengths are significantly longer than in metro organizations and hub numbers are higher, safeguarding the fibre length boundary is vital because of the greater disappointment rate [17].

2.5 OTDR

The Optical Time Domain Reflectometer (OTDR) is a pivotal device for following flaws in optical links. Its functional rule depends on the usage of Rayleigh dissipating and Fresnel reflection methods to quantify issue separates precisely. Furthermore, OTDRs are utilized to check join misfortune, measure link length, and distinguish flaws in optical links, particularly during beginning establishment [18].

At the point when an OTDR sends a high-power optical heartbeat through the fibre, Rayleigh dispersing happens, creating a backscattered signal that reflects shortcomings in the link back to the gadget. This backscattered light is distinguished by a delicate optical collector, changed over completely to computerized structure, and found the middle value of to upgrade the sign to-commotion proportion. The subsequent information is shown as a follow chart, giving a visual portrayal of backscattering action, including cuts, join misfortune, curves, constriction, and shortcoming distances in the optical organization [19].

Fresnel reflection, one more procedure utilized by the OTDR, recognizes discrete reflections brought about by changes in refractive file components, for example, air holes or particles discouraging light stream. These reflections show issue areas, and by breaking down Fresnel reflection information, OTDRs can foresee both delicate and hard disappointments in the organization foundation [20].

Notwithstanding Rayleigh dissipating and Fresnel reflection, OTDRs can use other logical standards like Raman dispersing, Mie dispersing, and Brillouin dispersing to follow shortcomings in optical organizations. These standards permit OTDRs to precisely quantify the distance of underground fibre links, improving shortcoming identification abilities across different situations [11].

2.5.1. Application of machine learning in optical-network failures

Machine learning (ML) is progressively being applied to address difficulties connected with optical organization disappointments. Here are a few key applications:

1. *Fault Location and Classification:* ML calculations can investigate information gathered from optical organizations, including OTDR follows, to recognize and arrange various sorts of deficiencies, for example, fibre cuts, twists, and sign corruption. Via preparing models on verifiable information, ML can distinguish designs demonstrative of explicit kinds of disappointments, empowering proactive support and quicker issue goal [21].
2. *Predictive Maintenance:* ML models can foresee expected disappointments in optical organizations by breaking down different boundaries like sign strength, lessening, and natural circumstances. By checking these elements continuously and contrasting them and authentic information, ML calculations can gauge when and where disappointments are probably going to happen, permitting administrators to make preventive moves before issues heighten [22].
3. *Anomaly Detection:* ML procedures, for example, unaided learning can recognize oddities in optical organization conduct that might demonstrate looming disappointments or strange circumstances. By ceaselessly checking network execution measurements, ML calculations can recognize deviations from typical activity and trigger cautions for additional examination [23].
4. *Optical Signal Quality Optimization:* ML calculations can upgrade optical sign quality by changing boundaries, for example, power levels, regulation arrangements, and scattering pay settings because of changing organization conditions. By gaining from past execution information, ML models can powerfully adjust network setups to amplify signal quality and limit the gamble of disappointments [24].
5. *Dynamic Steering and Asset Allocation:* ML-based traffic designing calculations can upgrade directing choices and asset allotment in optical organizations to moderate the effect of disappointments and guarantee productive utilization of organization assets. By dissecting traffic examples and organization geography, ML models can powerfully reroute traffic around bombed connections or hubs to keep up with administration progression and limit clog [25].
6. *Performance Forecast and Limit Planning:* ML models can anticipate future organization execution and limit prerequisites in light of verifiable information and projected development patterns. By estimating traffic interest, transmission capacity usage, and asset accessibility, ML calculations can assist administrators with arranging network overhauls and extensions to forestall bottlenecks and oblige expanding request [26].

In general, the utilization of ML in optical-network disappointments holds extraordinary potential to improve network dependability, effectiveness, and execution by empowering proactive shortcoming discovery, prescient upkeep, and canny asset the executives.

3. RELATED WORKS

Automated image analysis and ML were used by Lindström et al. [27], to enhance pulp testing approach in the pulping and papering industry. Four supervised ML algorithms were employed on the fibre data collected from the fibre suspension micrographs; Lasso regression, SVM, FFNN, and RNN. The author of this study concluded that the best of all the combinations examined in this study was the FFNN algorithm with Yeo's Johnson's pre-processing, with a record passing rate of 81% of a commercial fibre analyser. This approach thereby provides a consequent, fast and cheap option especially to the labour- and time-intensive pulp testing. However, the study points weaknesses in terms of generality because of the used techniques and software; thus, suggesting the replication of the study in realistic industrial environments for validation and fine-tuning purposes.

Singh et al. [28] put forward a brand new approach named Sea Lion fine-tuned Long Short-Term Memory (SL-FLSTM) for forecasting distributed denial-of-service (DDoS) attacks of fibre-optical networks. The SL-FLSTM strategy based on the Sea Lion behaviour increases the effectiveness in the processing of sequential data and involves bio-inspired amendments into the LSTM structure for increasing abilities to model long-term dependencies. These goals were achieved during the performance of the suggested approach with the recall being 98.1%, precision of 98.2% while the F1 score was noted to be approximately 98.8% Through this technique, it will be possible to identify the areas to focus on for improvement and hence increase efficiency by about 3%, coupled with increased accuracy of 98.4% which increase the performance in predicting the attack of DDoS in contrast with the conventional methods. However, its usefulness is given if used for the prediction of DDoS attacks in the fibre-optical networks, and it is not effective if used in other types of cyber-attacks.

Manzoni et al. [29] discussed fault detection in Continuous Glucose Monitoring (CGM) sensors in an artificial pancreas system and its model-based fault detection using a Kalan predictor. They also noted that large differences of measured and predicted values are signs of faults. This model-based approach does help in considering the dynamics of a system but this approach necessitates the establishment of a model and depends on prognosis. This research implies the need to incorporate novel accurate mathematical models that make the identification of faults in crucial well-being monitoring instruments authentic and unerring.

Fault detection and isolation was examined by Jihani et al. [30] as it relates to WSN with a parity space consideration involving mathematical models. It is a method that looks at the variance between the actual and modelled values in order to highlight faults, using the fact of measured redundancy in sensor readings. However, it has a drawback that prior knowledge is employed in dealing with model construction and this can be an issue in instances where environments or systems involve dynamic changes or occurrence of events and there is not sufficient information at one's disposal. In this respect, the study proves the usefulness of mathematical modelling in increasing certainty of WSNs.

Hashimoto et al. [31] study is on fault detection and diagnosis of the internal sensor in mobile robots utilizing a multimodal approach using Kalan filters. The fault decision is made from mode probabilities determined from the sensor gain, to accommodate the various failure modes. This approach hinges on the efficient Kalan estimation and this is essential to decision by filters. In particular, the study shows the necessity for the use of the multiple mode fault diagnosis in the enhancement of the performance and dependability of mobile robotic networks.

Based on the fault detection of optical fibre sensors in aero-engine systems, He et al. [32] proposed a model-based method when there are disturbances and uncertainties. The authors illustrated the use of the proposed technique using a simple model of a gas turbine and the method's capability of addressing system uncertainty. The key objective of this approach is to increase the reliability and safe operation of aero-engine through the accurate modelling of the system dynamics particularly for detection of sensor faults. The study underlines the importance of the accurate and liberal modelling to implement the fault detection systems particularly for the applications in engineering.

Yan et al. [33] describe a method to detect minor soft faults of air conditioning sensors based on Kernel Principal Component Analysis (KPCA), Deep Learning (DL) and Bidirectional Long Short-Term Memory (BiLSTM). Their method having higher detection rate than individual methods due to employ advanced machine learning to improve the accuracy of fault detection. Nevertheless, the performance can be influenced by the type of the fault and therefore flexible models proficient in many types of faults must be employed. The paper illustrates how various approaches of ML can be combined in order to enhance the identification of faults in HVAC systems.

The long-segmental faults detected by Alwan et al. [34] time-series clustering technique involved sensor nodes that showed continuous changes with time. Thus, this method effectively identifies long-segmental outliers, which allows using an alternative to the predictive analysis. The effectiveness of this technique, as noticed before, is contingent upon the quality and sample representativeness of the data, which in turn underlines the significance of acquiring high-quality data and their pre-processing. The results of the paper depict time-series clustering as a viable solution to improve the dependability of sensor networks concerning the prolonged types of anomalies.

In their work, Zhao [35] aligned the sliding window approach with control limits to monitor fault symptoms of process industries. This method looks for a continual drift and a decline in precision by using control limits that are based on historical data of various types of faults and processes. Thus, the proposed approach may be used for early detection of faults and actually offers a practical solution of this problem, though its usage can be effective only for some types of faults and industries. The study thus highlights the usefulness of empirical process in keeping up the working of functions in the industrial chain.

The early fault predictions in IoT setting were examined by Uppal et al. [36] employing Random Forest along with other machine learning techniques. According to the strategy they applied, the classification accuracy achieved a high percentage of 94%. 25% as to prove the efficiency of the developed ML algorithm in fault prediction of IoT systems. However, the performance of the proposed method depends upon the data and the algorithm used and may need constant fine-tuning and evaluation. It explains how the field of ML can be used to increase the dependability and proactive maintenance of IoT gadgets.

Liu et al [37] proposed a deep learning method of both prediction and fault detectors for the self-diagnosis of sensor fault in wind turbine blade. Their method provided good results in prediction and fault diagnosis; they used deep learning approaches to diagnose the faults. This method entails a massive volume of data for training the model and large computational power, making it suitable for large corporations. This research shows how deep learning can be used for the betterment of the maintenance and use of renewable energy systems.

Wahid et al. [38] introduced the CNN-LSTM model for the diagnosis of Machine Faults Magnitude (MFM) and for time series analysis to achieve predictive maintenance. Their model is replete with a set of effective and accurate predictions based on the offer of the state-of-the-art neural network architectures to process time-series data. The usage of this approach might depend on the quality and amount of the data which is why it is crucial to gather and pre-process the data properly. This work also affirms that the integration of CNN and LSTM

Research Article | Volume 13, Issue 1 | Pages 89-100 | e-ISSN: 2347-470X

improves the condition of predictive maintenance in industries.

Uppal et al. [39], described the classification of faults in office appliances integrated with IoT using the machine learning approach. What the authors offer is the process of monitoring and categorizing faults using data from IoT for the purpose of fault identification. This confirms that applying such a method might be sensitive to the complexity of the appliances, hence the need to develop and use versatile and resilient models. The study also demonstrates how ML is useful in increasing the dependability and effectiveness of smart offices via sound fault categorization

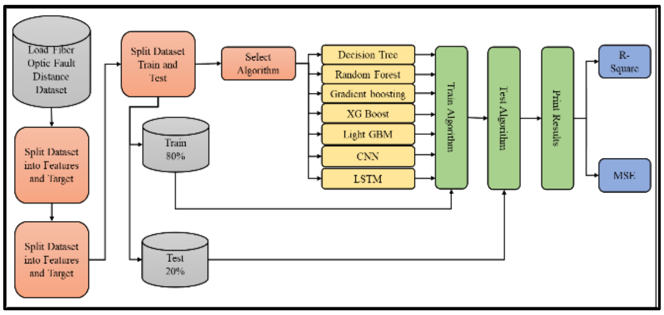

Figure 1. Overall Model Block Diagram

Safavi et al. [40] discussed about the electronic sensors of self-driving cars through the assessment of its health status which was based on the feature extraction and multi class DNN. Their approach determines impaired sensors and types of the impairments; the paper employs sophisticated ML algorithms for fault detection. The assignment of the features related to the faults and the labelling of these features is very important to the success of this method that seeks to improve the safety and reliability of the self-driving automobiles. The study explains about the perfect ML models for sustaining functional correctness of multi-system vehicles.

4. METHODOLOGY

The fibre Optic Fault Distance Dataset will be uploaded consisting of features like Fault Circuits, Fault Locations, Fault Times and a class attribute namely Fault Distance. Data Pre-processing: The given dataset will be divided initially into Feature(X) and Target (y) sets, primarily, this dataset will be further divided into 80% Training data and 20% Testing data for evaluating the performance of the model. While deciding on the number of algorithms to be used, some of them include Decision Tree, Random Forest, Gradient Boosting, XGBoost, LightGBM, Convolutional Neural Network (CNN), and Long Short-Term Memory (LSTM). These algorithms will be trained using the training data and will be tested on the testing data respectively. These two metrics will be used to assess the performance of each model and the results will then be printed to contrast. *Figure (1)* shows overall model block diagram.

4.1 Dataset Description

The dataset consists of sequenced OTDR signatures that consist of common types of fibre faults like; fibre cut, fibre eavesdropping, dirty connector, and bad splice. Each sample is made up of the SNR, a 30-length normed sequence of the OTDR trace [P1….. P30], the nature of event (normal, fibre tapping, splice that is bad, bending event, dirty connector, cut fibre, PC connector, and reflector), location of the event and its reflectance and loss both of which are normed. An exception is made for the identification number of the event location, which is obtained by dividing the index of the event location within the sequence by 100.

This dataset is significant for the development of machine learning techniques regarding the identification of the nature and the locations of the faults in optical fibres. Additional raw OTDR traces are also stored in the data folder with various kinds of optical faults introduced by manipulating the experimental settings, for instance, changing the bending radius and adjusting the amount of the variable optical attenuation to change the height of the reflector. Figure (2) shows dataset collection circuit.

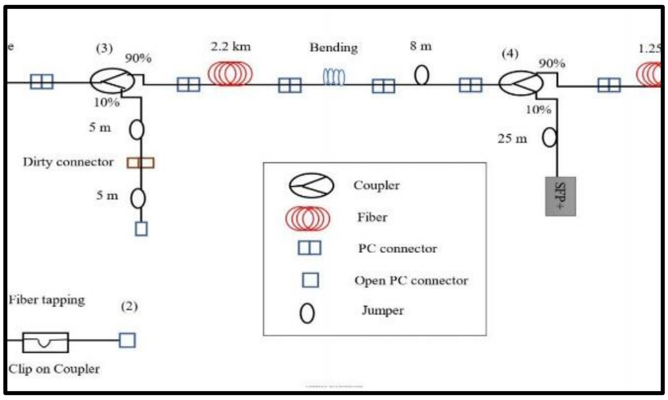

Figure 2. Dataset Collection Circuit

4.2 Dataset Pre-processing

The fibre Optic Fault Distance Dataset will be pre-processed initially and the pre-processed dataset will be split into the feature set (X) and target set (y). Among the features that will be incorporated as variables in our study are: Fault Circuits, Fault Locations, and Fault Times and the parameter of study being the Feature Distance study. After this first split the dataset will then be split into 80/20 to train the models and later on use to test their performance.

4.3 Used Algorithms

4.3.1. Decision Tree: Quinlan. A decision tree comprises of decision nodes and the leaf nodes. Every decision node checks only one factor of the input characteristics and has several outcomes, each of which corresponds to the decision node. Each of the two leaves is a classification outcome for a case where the set of potential outcomes is given by the elements of the class.

The problem in building such trees is the determination of which attribute is most suitable for specification at the decision nodes. C4. 5 It uses the Gain Ratio Criterion that is an improvement of the Gini Index in that it chooses an attribute that offers the greatest information gain but unlike the Gini index which tends to favour attributes with numerous

FOREX Publication
Open Access | Rapid and quality publishing

International Journal of
Electrical and Electronics Research (IJEER)
Research Article | Volume 13, Issue 1 | Pages 89-100 | e-ISSN: 2347-470X

outcomes, has its values normalized. After that decision tree is constructed, this means that one can be able to classify the testing data which has similar characteristics of the training data. The test commences at the root node and then moves along the branches relative to the test case's attribute values until a leaf node is attained. The class of the identified leaf node is the output of the classification of the test case. (*Figure 3*).

4.3.2. Random Forest: Random Forest is a method of ensemble learning to be used in classification, regression and numerous tasks, integrating multiple decision trees into one. In this model, many decision trees are induced within different data randomly selected, and the last decision is made by combining all the decision trees. This reduces the problem of overfitting and increases the ability to generalize on new unseen data. Also, Random Forest incorporates feature bagging where a certain number of features are randomly chosen for a specific tree to consider and hence minimizing variance and concentrating on the appropriate features. It offers several advantages over other machine learning algorithms: It creates decision trees with high interpretability; it can deal with a massive number of input features, be it categorical or numerical; It has a good approach to the missing data and outliers; Overfitting occurs less frequently than in the case of individual decision trees or models resulting from the use of a single algorithm; and the output of the function includes information regarding variable importance. Because of its high accuracy, it is strong and can be employed across different fields some of which are image recognition, text categorization, and biological informatics among others hence the reason the machine learning engineers prefer it [41][42].

4.3.3. Gradient Boosting Algorithm: XGBoost uses different algorithms of regularization when it is iteratively building the best approximation to a target function using the best-chosen loss function. Needless to say, in each cycle of the XGBoost algorithm mentioned in Equation (1), a weak learner is created. This weak learner is trained using an objective function which integrates the regularization techniques into the loss function. At each iteration, only this objective function is minimized, and the fit of the model is incrementally improved each time. Thus, iteratively, XGBoost forms a strong learner from pieces or weak learners that fit the various characteristics of the data to improve the predictive model. [43][44].

$$L_{(t)(x_i,y_i)=\sum_{i=1}^{N}\zeta\left(y_i,y_i^{t-1}+f_1(x_1)+\Omega(f_1)\right)} \quad \ldots\ldots\ldots\ldots\ldots\ (1)$$

4.3.4. LightGBM Algorithm: Using decision tree techniques, LightGBM is a distributed, fast, and effective gradient-boosting system. It is widely used in machine learning applications like classification, regression, and ranking. LightGBM, as a Boosting method, combines multiple weak models into a strong predictive model. Boosting algorithms' fundamental idea is to reduce the weights of correctly identified instances and increase the weights of poorly categorized ones. This modification guarantees that later training cycles concentrate more on the data that was first incorrectly identified. In the end, the combined model—which is made up of all the separate models—is linearly integrated using weights that have been adjusted based on the error rate of each model. Equation (2) can be used to describe the fundamental idea [45]:

$$f(x) = \sum_{q=1}^{Q} \alpha_q T(x, \theta_q) \quad (2)$$

In this instance, the goal value connected to a training sample is denoted by $f(x)$. The number of base learners is indicated by (Q), and the weight coefficient for the q-th base learner is indicated by (α_q). The training sample is represented by the variable (x), and the parameter unique to the learner's categorization is (θ_q). During training, the q-th base learner is represented by $T(x, \theta_q)$. The training procedure of the Boosting algorithm turns into an optimization problem with the goal of minimizing the loss function after the training data and loss function are defined. This process's objective function is stated as follows [45]:

$$\arg\min \sum_{h=1}^{H} L(y_h, f(x_h)) \quad (3)$$

The total number of samples in this framework is indicated by (*H*), and the index for each sample is shown by (*h*). While $f(x_h)$ represents the goal value connected to the h-th sample, (y_h) represents the actual value of the data. As for the h-th sample, $L(y_h, f(x_h))$ indicates the loss function value. LightGBM, a GBM variation, successfully addresses the issues that standard GBM has with big datasets. LightGBM has two main characteristics [46]:

1. ***Leaf-wise Tree Growth Method:*** LightGBM uses the concept of a growing tree from the bottom up also known as a leaf-wise growth strategy instead of a level-wise growth strategy. Leaving a minimal number of data points per leave and the depth of a tree to construct also reduces the computational cost and prevents the model from overfitting.
2. ***Decision Tree Algorithm Based on Histograms:*** LightGBM uses a decision tree algorithm based on histograms of splitting criteria instead of using gradient tree. This does less of the computation and searching when it is time to select features to split on because it only looks up values of the best split in a discrete histogram. This feature will enable LightGBM to learn and make more accurate predictions with the large-scale datasets.

4.3.4. Convolutional Neural Network: Standard deep CNNs are commonly referred to as "2D CNNs" because they are expected to process 2D data such as images and videos. Nonetheless, there is a version called 1 Dimensional Convolutional Neural Networks (1D CNNs) which has been developed as there is strength in 1D signals. Studies [47][48] highlight several benefits of 1D CNNs: they use for forward and backward propagation simple array operations which make the computational complexity to be less as compared to 2D CNNs.

Furthermore, the 1D CNNs can cope with complex tasks with relatively simple structure, therefore, it is easy to train and apply. Also, in contrast to 2D CNNs, which, as rule, need

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of
Electrical and Electronics Research (IJEER)
Research Article | Volume 13, Issue 1 | Pages 89-100 | e-ISSN: 2347-470X

additional helps like GPU or cloud, compact 1D CNNs are successfully trained on standard CPUs. This makes them suitable for low delay, low-cost applications such as real-time and or portable devices like handheld or mobile devices. Later work reveals that compact 1D CNNs perform well when there is a small amount of labelled data and large fluctuations in signals, for example, patient ECGs and mechanical and aerospace structures.

These networks consist of two main layer types: CNN layers for which 1D convolution and pooling take place and MLP layers similar to what is found in Multi-Layer Perceptron. The hyper parameters concerning the configuration of a 1D CNN are the number of hidden CNN and MLP layers, filter size, subsampling factor, and the type of pooling and the activation function to be used [49].

Like normal 2D CNNs, the input layer of any 1D CNN is an input layer that only takes the original one-dimensional signal passively and the final output layer is an MLP layer with neurons equal to the number of classes [49].

4.3.5. Long-short Term Memory: Several RNN models have been developed based on parameter configurations, including considerations related to input-output data types and layer weights [50]. During the training of time series data, upgrades were made to the RNN structure to enable the incorporation of time-series characteristics over extended durations [51].

Simple RNN, a fundamental form of a recurrent model, involves the sequential delivery of input data to adjacent layer cells within a chain of general neural networks. Notably, the cells that transmit data do not retain information about the conveyed data, rendering them incapable of preserving long-term time-series information. When the hyperbolic tangent is employed as the activation function, the memory (ht) utilized for predicting the output vector is defined as follows [51]:

$$h_t = \tanh(U_t x_t + W h_{t-1} + b) \quad \ldots\ldots\ldots\ldots\ldots\ldots (4)$$

In this context, (x) represents an input vector, (U) and (W) denote adaptable matrices, and (b) signifies the bias vector. The model's training process entails the adjustment of weights to align (xt) with the correct value (yt), which is used for predicting (ht) generated by the combination of (Ut), (W) and (ht−1). Given that (ht−1) from the previous time step plays a crucial role in forecasting 'ht' for the current time step, it establishes a recurrent structure within the neural network. This structure incorporates feedback from prior steps into the computation of the current step, empowering the RNN to effectively process time-series data [51].

LSTM represent models that have undergone alterations in the cell structure connecting input and output vectors to address the limited long-term dependency issue observed in a simple RNN. The distinguishing characteristic of LSTM is the additional sharing of the cell state parameter, (zt) among neighbouring layers. This preservation of the existing input vector enhances long-term memory storage capabilities. Within a single layer, various parameters, contingent upon the dimensions of input and output vectors, exist [50]. These parameters collectively form the output gate, forget gate, and input gate, serving to safeguard and govern the cell state. Notably, the forget gate directly influences the determination of how much information from the previous vector should be retained or discarded [52].

4.4 Training Parameters

The training process of the models will be carried out with specific parameters tailored for each algorithm to be used. Concerning the Decision Tree, its max_depth will be set to 10 while min_samples_split will be set at 2. An important decision related to tuning is the number of estimators, which shall be set to 100, while the minimum number of samples to split shall be set to 2. For Gradient Boosting, 100 estimators will also be used with the minimum sample split set to 2.

XGBoost will be set with parameters maximum depth 6 and the learn rate shall be set to 0. 3. LightGBM is now set to use 100 estimators in a tree and a learning rate of 0.1. Neural network models comprise of the CNN model and the LSTM model for the CNN model, it shall be trained using 32 batch size and a learning rate of 0.001. These parameters are selected to achieve the best results for each of the models on the defined dataset. *Table (1)* summarize training parameters.

Table 1. Training Parameters

Algorithm	Parameters			
Decision Tree	Max Depth	10	Min Sample Split	2
Random Forest	N- Estimation	100	Min Sample Split	2
Gradient Boosting	N- Estimation	100	Min Sample Split	2
XG Boost	Max Depth	6	Learning Rate	0.3
Light GBM	N- Estimation	100	Learning Rate	0.1
CNN	Batch Size	32	Learning Rate	0.001
LSTM	Batch Size	32	Learning Rate	0.001

4.5 Build Deep Algorithms

The CNN next starts with a Conv1D layer that uses 64 filters and a kernel size of 2, with ReLU as the activation function. This layer takes the input data which has a shape defined by the number of features in the training set (X_train. shape [1]) and a single channel, if the input is grayscale or single feature. This has taken the form of dropout regularization with a rate of 0. The result of this convolutional layer is passed through 5, which random ½ neurons during the training period to reduce overfitting.

After the dropout layer, one more Conv1D layer is included with 32 filters and length=2, activation ReLU is used. This layer yet again distils new features from the learnt data by the previous convolutional layer. Before the final prediction, the model flattens the output of the convolutional layers to a final vector of one dimensionality using the Flatten layer. This step makes the multiple values output as a format that is suitable in feeding into the densely connected layers.

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of
Electrical and Electronics Research (IJEER)
Research Article | Volume 13, Issue 1 | Pages 89-100 | e-ISSN: 2347-470X

With the densities flattened, the model then adds a Dense layer which has 50 neurons, with ReLU activation function being used. This dense layer enables the network to learn representations of higher order after flattening them from the convolutional layers. Last but not the least, the model is concluded with another Dense layer with 1 neuron for simple regression if the output layer is meant to produce a continuous value. This last layer does not have an activation function following it which suggests that it produces a numerical output.

In general, this CNN architecture builds the CNN model that iteratively performs the input data by the convolutional layer to get and learn hierarchical representations of the given data, flattening layer and dense layer to perform final prediction. This design is especially suitable for a task of processing the sequential data such as time-series or other signal data. Figure (3) shows the build CNN.

The LSTM (Long Short-Term Memory) model described in this paper is designed to handle sequences as input. It is composed of a sequential model with long short-term memory units followed by a dropout layer of 0. 2, to avoid overfitting, and rectified linear unit activation of 64. This layer is return sequences True since its outputs are to be passed to the next layer as sequences. A Dropout layer with a 0. 5 dropout rate follows as a method of preventing the model from overfitting through a process that involves the temporary disconnection of fifty percent of its neurons during the training stage.

The next LSTM layer has 32 units and ReLU activation; however, it does not outputs sequences to combine the learned features into a single output vector. Adding and normalizing this vector to the next Dense layer with 50 neurons and ReLU activation allows the model to learn the next set of features or higher abstraction.

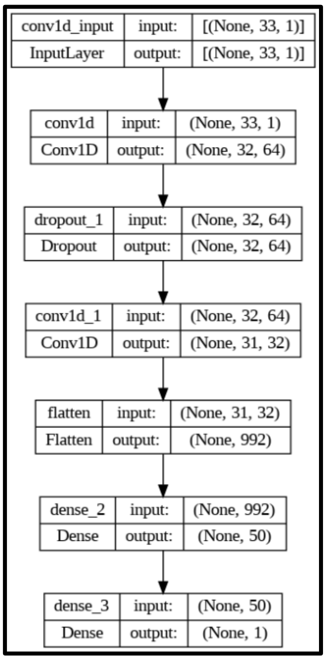

Figure 3. Builded CNN Model

The last dense layer has a single neuron and no activation function; thus, it provides continuous predictions. The LSTM structure can be applied to the problems involves sequential data processing such as time series forecasting, natural language processing by bringing out the temporal dependencies and yielding precise predictions. *Figure (4)* shows the build LSTM model.

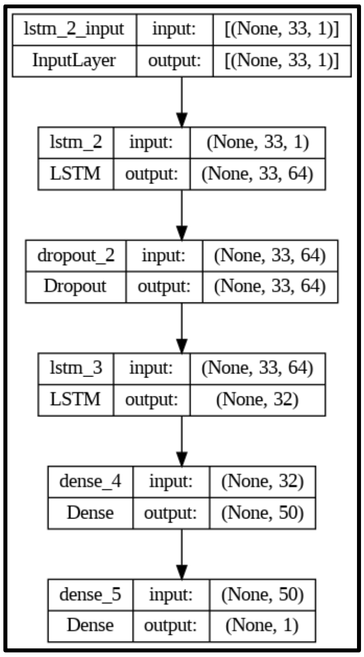

Figure 4. Builded LSTM Model

4.6. Evaluation Metrics

R-Square (Coefficient of Determination) is defined as extent to which change in target variable can be explained by change in features. It varies from 0 to 1, where 1 here means the target is perfectly predicted. Mathematically, it is defined as:

$$R^2 = 1 - \frac{\Sigma(y_i - \hat{y}_i)^2}{\Sigma(y_i - \bar{y})^{2}{}^{2}} \qquad (5)$$

where y_i are the actual values, $\hat{y}_i$ are the predicted values, and $\bar{y}$ is the mean of the actual values

Mean Squared Error (MSE) measures the level of variability for the potential differences between the actual value of a data point and a predicted value. It gives a measure of how good the model is, or how well it fits the data with lower values being optimal.

$$MSE = \frac{1}{n}\Sigma_{i=1}^{n}(y_i - \hat{y}_i)^6 \qquad (6)$$

where y_i are the actual values, $\hat{y}_i$ are the predicted values, and n is the number of samples. These metrics are essential for assessing the quality and predictive power of machine learning models.

MAE (Mean Absolute Error): Calculates the mean of the observation space deviations from the prediction space (the lower the value the better).

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \check{y}_i| \quad (7)$$

Where:
y_i= actual value of the dependent variable
$\hat{y}_i$= predicted value
n = number of data points

RMSE (Root Mean Squared Error): Squares the differences between the values and then takes the mean of those squared differences and takes square root of that (smaller is better).

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \check{y}_i)^2} \quad (8)$$

y_i= actual value of the dependent variable
$\hat{y}_i$= predicted value
n = number of data points

5. RESULTS AND DISCUSSION

That point shall be further discussed to show the importance of the dataset in developing new approaches and or improving existing ones that are used in the machine learning process of diagnosing optical fibres and precisely pinpointing the location of the defects in it. The given dataset is quite versatile because it contains a rather diverse set of fault types and possesses accurate descriptions of the particular events that occurred. In the case of multiple raw OTDR traces with originally altered experimental parameters, the set becomes even more versatile to better study the various types of faults.

From the comparison of several methods in the framework of this dataset, the effectiveness of each is determined. Comparing the performance, Decision Tree model is marked as a moderate baseline where its accuracy is less and MSE is comparatively higher that does not seem to capture the same degree of the fault patterns.

Random forest and LSTM models on the other end reveal outstanding results shown by high R-Square and low MSE. This indicates that these models have great potential of learning from the sequence data and can be used to identify type of faults and their locations. The outperformance of LSTM is specifically worth it due to its suitability to solve problems involving some sequence of successive data.

XG Boost and Light GBM are also on top followed by Gradient Boosting though the performance differentials are minute. The fact that shown models possess a high value of R-Square as well as low MSEs proves decent performance of these models to describe the data and relations within it. It can be seen that CNN, a deep learning model, has achieved a fairly good result, which proves that there is great potential for using neural networks in fault diagnosis fields.

Regarding the combination methods, models such as CNN-RF and LSTM-RF provide the feature extraction workflow of CNNs/LSTMs and the ensemble learning of random forests. In details, being most accurate, the LSTM-RF model is again the most outstanding one as suggested by the best R-Square rate and the lowest MSE rate noted in the study. This type of architecture combines LSTM that processes numerical sequences with RFC which makes a decision on the basis of the processed data; this combination makes the hybrid approach very efficient for this application.

Table 2. Models Results

Algorithm	R-Square	MSE	MAE	RMSE
Decision Tree	0.740581	0.00048	0.0169	0.0219
Random Forest	0.885504	0.00021	0.013	0.0145
Gradient Boosting	0.773956	0.00042	0.0153	0.0205
XG Boost	0.872512	0.00023	0.0136	0.0152
Light GBM	0.866684	0.00024	0.014	0.0155
CNN	0.870759	0.00024	0.0142	0.0155
LSTM	0.887053	0.00021	0.0129	0.0145
CNN-RF	0.864591	0.00025	0.0143	0.0157
LSTM-RF	0.910105	0.00016	0.012	0.0126

6. RESULTS DISCUSSION

The provides the performance comparison of many machine learning algorithms and some latest advanced AI models for the fault distance prediction of optical fibres. These models include basic models such as Decision Tree, Random Forests, Gradient Boosting, XGBoost, LightGBM and CNN, LSTM and some combined model such as CNN-RF and LSTM-RF. The evaluation criteria which will be used are R-Square (R^2), Mean Squared Error (MSE), Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE).

6.1 Model Performance Overview

- LSTM-RF (R-Square = 0.910105, MSE = 0.00016, MAE = 0.012, RMSE = 0.0126): In all the evaluation criteria, LSTM-RF model has emerged as the most efficient one. The models provide the greatest R-Square of 0.910, thus signifying that the design accounts for nearly 91% of the variability in the fault distance prediction. It also boasts the lowest value of MSE 0.00016 and MAE 0.012 meaning minimal prediction errors and clearly shows the model is accurate in its predictions.

- Random Forest (R-Square = 0.885504, MSE = 0.00021, MAE = 0.013, RMSE = 0.0145): Random Forest also give good results with the value of adjusted R square equals to 0.885. Despite a slight decrease in the values of R-Square and other parameters compared to LSTM-RF, it provides a very high accuracy and can be recommended for diagnosing optical fibres faults distance. Compared to the models like LSTM, the Random Forest model is in general more interpretable, and computationally less demanding.
- LSTM (R-Square = 0.887053, MSE = 0.00021, MAE = 0.0129, RMSE = 0.0145): In this case, the mean value of MSE, MAE and RMSE of the LSTM model is close to that of the Random Forest classification. Though, slightly higher R-Square shows that LSTM, which has a slightly higher value than the Random Forest, may capture higher order features of the data, which may be important for time series data or sequential data in fault predictions. However, LSTM models used in the experiments is relatively less accurate and computationally intensive, hence the need to fine tune.
- XGBoost (R-Square = 0.872512, MSE = 0.00023, MAE = 0.0136, RMSE = 0.0152): XGBoost which is generally recommended for large data set as well as complex non-linear relationship, yields high R-Square and low error. Although it is a little slower than LSTM and LSTM-RF, the authors assert that it is a candidate of choice because of its efficiency and adaptability.
- Gradient Boosting (R-Square = 0.773956, MSE = 0.00042, MAE = 0.0153, RMSE = 0.0205): Nonetheless, Gradient Boosting takes relatively a longer time for the execution and is not as good as Random Forest or XGBoost models. They have slightly higher MSE and MAE indicating that it can be outperformed by the others on fault distance prediction because of the model's complexity. Still, it could still be useful in cases where the inter pretableness of the model as well as the training time matters.
- LightGBM (R-Square = 0.866684, MSE = 0.00024, MAE = 0.014, RMSE = 0.0155): It also provides satisfactory performance, especially when analysing big data as a fraction of a second. The algorithm is somewhat better than Gradient Boosting in R-Square and MAE but does not beat Random Forest or LSTM. Since AdaBoost can train quickly and accommodating to large sets of data may be useful in industries where model updates are regularly required.
- CNN (R-Square = 0.870759, MSE = 0.00024, MAE = 0.0142, RMSE = 0.0155): As expected, the model achieves comparable results to LightGBM and XGBoost in this case, which it outperforms in tasks involving spatial data and pattern recognition. Despite the fact that CNNs are highly flexible they could be less efficient compared to other models in terms of accuracy and computational complexity of the fault distance prediction in optical fibres.
- CNN-RF (R-Square = 0.864591, MSE = 0.00025, MAE = 0.0143, RMSE = 0.0157): The hybrid CNN-RF model gives a low accuracy compared with other models like LSTM and Random Forest. CNN for feature extraction might not be optimally exploited here and thus the ensemble with Random Forest does not show much improvement. It implies that for this particular application, other rudimentary methods such as Random Forest can work well.
- Decision Tree (R-Square = 0.740581, MSE = 0.00048, MAE = 0.0169, RMSE = 0.0219): The Decision Tree model has the worst performance compared to all other four models based on all the four metrics used. Although the squared term is easy to interpret, it has a problem with over-fitting and less suitable in terms of capturing relationship in the data. While this model may be useful for the case of exploration, or when interpretability is paramount, they are suboptimal for high accuracy

6.2. Analysis of Model Performance

- *Best Overall Model: LSTM-RF:* The LSTM-RF model should be noticed because of its high R-Square and low MSE, MAE, and RMSE rates. It is a blend of merits from LSTM for sequential dependence and Random Forest for regression stand-point. This makes it the best suited for this application.
- *Scalability and Efficienc*y: It is known that the Random Forest, XGBoost and LightGBM are fast and scalable for big data. They also don't demand as much computational power as deep learning models such as LSTM and CNN. In most practical cases these models would seem to be quite well balanced between their accuracy and the number of resources that they can or will require.
- *Interpretability vs. Complexity:* However, there exist other algorithms with higher interpretability, but lower complexity, which do not allow for accurate determination of fault distance, for example, Decision Tree. Better models like LSTM-RF, XGBoost & Random Forest have better precision but have a higher value of computational overhead and model complexity than the linear model.

6.3. Recommendations

- When the importance of accuracy is paramount, and computational resources are available, LSTM-RF should be preferred.
- Random Forest and XGBoost are equally good trade-offs between accuracy and computational costs and could be great options for real time or low computational power use cases.
- Perhaps, LightGBM could be beneficial for those tasks where fast model training and its scalability is valuable.
- Decision Tree can still be used in simple problems or when one is testing models for the first time.

7. CONCLUSION

This paper provides proof of different machine learning models when it comes to determining fault distances in optical fibres and Random Forest, LSTM models are deemed to be the best models that have the highest R-Square and lowest MSE score. The most optimized solution suggested is the LSTM-RF approach, which integrates the ability of sequence handling in LSTMs and ensemble learning in Random Forest, and shows the highest accuracy and lowest errors in fault distance prediction. The Fibre Optic Fault Distance Dataset indeed helped, as it is rather diversified in terms of fault types and offers detailed descriptions of the events, which is useful for training and testing the ml algorithms. Consequently, the insights provided in this study have practical application on how to maintain and troubleshoot with efficiency and cost-effectiveness the optical fibre network's fault distance therefore improving the effectiveness of fibre optic communication. This research should investigate in more detail the ways of enhancing the proposed hybrid LSTM-RF model, as well as examine the potential of using other forms of sophisticated ML algorithms. Further, the data set could be enriched with more fault scenarios and conditions that should make the models more effective. These conclusions put forward the prospects of machine learning in improving optical fibre fault diagnosis and propose further study on the application of machine learning in this field and the development of relevant technologies.

AUTHOR CONTRIBUTION

This piece of research brings a major contribution to the area of optical fibre fault diagnosis since it compares and evaluates the performance of different machine learning approaches as far as the estimation of the distances to faults in optical fibres is concerned. Using Decision Tree, Random Forest, Gradient Boosting, XGBoost, LightGBM, CNN, LSTM, the research contributes to the understanding of how these algorithms perform in terms of inputting, transforming and analysing OTDR logos. The data pre-processing is methodical and precise whereas the division of the dataset into training (80%) and testing (20%) actually help in proper model assessment. Employment of the R-Square and Mean Squared Error (MSE) performances increase the validity of the results. Of these models, Random Forest and LSTM are found to be the most accurate, having higher R-Square Values and the lowest MSE in the research findings which goes to support their efficacy in terms of predictive capability in fault distance prediction. Additionally, the development of the new hybrid LSTM–RF model is considered as an enormous breakthrough and gives the highest performance among all the developed models. This work stresses the importance of higher-level computational methods in improving accuracy and reliability of diagnosis of optical fibres faults, which can be used as directions for further researches and developments in the field.

REFERENCES

1. K. Xu, X. Li, Y. Tang, and C. Yuan, "Serial-parallel combined all optical sequence matching system using highly nonlinear fibres for photonic firewall," *Optik (Stuttg).*, vol. 244, no. July, p. 167571, 2021, doi: 10.1016/j.ijleo.2021.167571.
2. Z. Cui, C. Yuan, K. Xu, Y. Sun, and S. Yin, "Modeling the Splice Loss of Single-Mode Optical Fibres Affected by Altitude," *IEEE Access*, vol. 7, pp. 99283–99289, 2019, doi: 10.1109/ACCESS.2019.2927395.
3. B. Gelkop, L. Aichnboim, and D. Malka, "RGB wavelength multiplexer based on polycarbonate multicore polymer optical fibre," *Opt. Fibre Technol.*, vol. 61, no. July 2020, p. 102441, 2021, doi: 10.1016/j.yofte.2020.102441.
4. A. A. Ibrahim, M. M. Fouad, and A. A. Hamdi, "A Design Fibre Performance Monitoring Tool (FPMT) for Online Remote Fibre Line Performance Detection," *Electron.*, vol. 11, no. 21, 2022, doi: 10.3390/electronics11213627.
5. C. Pendão and I. Silva, "Optical Fibre Sensors and Sensing Networks: Overview of the Main Principles and Applications," *Sensors*, vol. 22, no. 19, 2022, doi: 10.3390/s22197554.
6. A. A. Ibrahim, M. M. Fouad, and A. A. Hamdi, "Remote Real-Time Optical Layers Performance Monitoring Using a Modern FPMT Technique Integrated with an EDFA Optical Amplifier," *Electron.*, vol. 12, no. 3, pp. 1–19, 2023, doi: 10.3390/electronics12030601.
7. M. K. Barnoski and S. M. Jensen, "Fibre waveguides: a novel technique for investigating attenuation characteristics," *Appl. Opt.*, vol. 15, no. 9, p. 2112, 1976, doi: 10.1364/ao.15.002112.
8. S. P. Abdula, M. J. Llagas, A. M. Fernandez, and E. Arboleda, "Machine Learning Applications for Fault Tracing and Localization in Optical Fibre Communication Networks: A Review," 2024, doi: 10.20944/preprints202405.1285.v1.
9. E. Quatrini, F. Costantino, G. Di Gravio, and R. Patriarca, "Machine learning for anomaly detection and process phase classification to improve safety and maintenance activities," *J. Manuf. Syst.*, vol. 56, pp. 117–132, Jul. 2020, doi: 10.1016/j.jmsy.2020.05.013.
10. H. Zhang, J. Gao, and B. Hong, "Φ-OTDR Signal Identification Method Based on Multimodal Fusion," *Sensors*, vol. 22, no. 22, pp. 1–12, 2022, doi: 10.3390/s22228795.
11. T. A. Ali and J. J. H. Ameen, "Study of Fault Detection Techniques for Optical Fibres," *Zanco J. Pure Appl. Sci.*, vol. 31, no. s3, 2019, doi: 10.21271/zjpas.31.s3.20.
12. K. Xu and C. Yuan, "A Fault Location Analysis of Optical Fibre Communication Links in High Altitude Areas," *Electron.*, vol. 12, no. 17, pp. 1–15, 2023, doi: 10.3390/electronics12173728.
13. O. Nyarko-Boateng, A. F. Adekoya, and B. A. Weyori, "Predicting the actual location of faults in underground optical networks using linear regression," *Eng. Reports*, vol. 3, no. 3, pp. 1–13, 2021, doi: 10.1002/eng2.12304.
14. S. D. S. A. N. Ahmed, "Attenuation and Dispersion through Single Mode fibre Optic Simulation," 2016.
15. A. H. Al-Fatlawi, M. M. Abdul Zahra, and H. A. Rassool, "Simulation of optical fibre cable regarding bandwidth limitations," *Int. J. Nonlinear Anal. Appl.*, vol. 12, no. Special Issue, pp. 1159–1174, 2021, doi: 10.22075/IJNAA.2021.5607.
16. G. A. Roth *et al.*, "Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980–2017 : a systematic analysis for the Global Burden of Disease Study 2017," *Lancet*, vol. 392, no. 10159, pp. 1736–1788, 2018, doi: 10.1016/S0140-6736(18)32203-7.
17. [17] T. Hayford-Acquah and B. Asante, "Causes of Fibre Cut and the Recommendation to Solve the Problem," *IOSR J. Electron. Commun. Eng.*, vol. 12, no. 01, pp. 46–64, 2017, doi: 10.9790/2834-1201014664.
18. R. Tariq *et al.*, "An Optimized Solution for Fault Detection and Location in Underground Cables Based on Traveling Waves," *Energies*, vol. 15, no. 17, 2022, doi: 10.3390/en15176468.
19. S. Aldi, K. Thallapally, D. Bussa, S. T. Rayini, and Y. Shekar, "Detection and Location of Faults in Underground Cables," *Int. J. Sci. Res. Eng. Dev.*, vol. 3, no. 2, pp. 0–4, 2020.
20. O. Nyarko-Boateng, A. F. Adekoya, and B. A. Weyori, "Adopting Intelligent Modelling to Trace Fault in Underground Optical Network : A Comprehensive Survey," *J. Comput. Sci.*, vol. 16, no. 10, pp. 1355–1366, 2020, doi: 10.3844/jcssp.2020.1355.1366.
21. V. Kumar and D. Rajouria, "Fault Detection Technique by using OTDR : Limitations and drawbacks on practical approach of measurement," *Int. J. Emerg. Technol. Adv. Eng.*, vol. 2, no. 6, 2012.
22. A. L. Paul and J. Oluwaseyi, "Predictive Maintenance : Leveraging Machine Learning for Equipment Health Monitoring," no. January, 2024, [Online]. Available: https://www.researchgate.net/publication/377411657
23. A. B. Nassif, M. A. Talib, Q. Nasir, and F. M. Dakalbab, "Machine Learning for Anomaly Detection: A Systematic Review," *IEEE Access*, vol. 9, pp. 78658–78700, 2021, doi: 10.1109/ACCESS.2021.3083060.
24. M. Moshawrab, M. Adda, A. Bouzouane, H. Ibrahim, and A. Raad, "Reviewing Federated Learning Aggregation Algorithms; Strategies, Contributions, Limitations and Future Perspectives," *Electron.*, vol. 12, no. 10, pp. 1–35, 2023, doi: 10.3390/electronics12102287.

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 13, Issue 1 | Pages 89-100 | e-ISSN: 2347-470X

25. N. Xiong *et al.*, "Lecture Notes on Data Engineering and Communications Technologies 153," vol. 53, no. 0, p. 6221, 2022.
26. L. Skyttner, "Artificial Intelligence and Life," *Gen. Syst. Theory*, pp. 319–351, 2006, doi: 10.1142/9789812774750_0007.
27. S. B. Lindström *et al.*, "Pulp Particle Classification Based on Optical Fibre Analysis and Machine Learning Techniques," *Fibres*, pp. 1–21, 2024.
28. A. Singh, P. Grover, A. K. Gautam, B. Nagappan, and N. Sharma, "Intelligent Systems And Applications In Engineering Novel prediction mechanism for Attack Prevention in Fibre-Optical Networks using AI-based SDN," *Int. J. Intell. Syst. Appl. Eng.*, vol. 12, pp. 1408–1414, 2024.
29. E. Manzoni, M. Rampazzo, and S. Del Favero, "Detection of glucose sensor faults in an artificial pancreas via whiteness test on kalman filter residuals," *IFAC-PapersOnLine*, vol. 54, no. 7, pp. 274–279, 2021, doi: 10.1016/j.ifacol.2021.08.371.
30. N. Jihani, M. N. Kabbaj, and M. Benbrahim, "Sensor fault detection and isolation for smart irrigation wireless sensor network based on parity space," *Int. J. Electr. Comput. Eng.*, vol. 13, no. 2, pp. 1463–1471, 2023, doi: 10.11591/ijece.v13i2.pp1463-1471.
31. M. Hashimoto, H. Kawashima, and F. Oba, "A multi-model based fault detection and diagnosis of internal sensors for mobile robot," in *Proceedings 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2003) (Cat. No.03CH37453)*, 2003, pp. 3787–3792 vol.3. doi: 10.1109/IROS.2003.1249744.
32. C. He, X. Zhang, and B. Jia, "UIO based robust fault diagnosis approach for aero-engine fibre-optic sensor," in *2013 IEEE International Conference on Automation Science and Engineering (CASE)*, 2013, pp. 550–553. doi: 10.1109/CoASE.2013.6653947.
33. X. Yan, T. Guan, K. Fan, and Q. Sun, "Novel double layer BiLSTM minor soft fault detection for sensors in air-conditioning system with KPCA reducing dimensions," *J. Build. Eng.*, vol. 44, no. July, p. 102950, 2021, doi: 10.1016/j.jobe.2021.102950.
34. A. A. Alwan, A. J. Brimicombe, M. A. Ciupala, S. A. Ghorashi, A. Baravalle, and P. Falcarin, "Time-series clustering for sensor fault detection in large-scale Cyber–Physical Systems," *Comput. Networks*, vol. 218, p. 109384, 2022, doi: https://doi.org/10.1016/j.comnet.2022.109384.
35. W. Zhao, H. Luo, Q. Liu, H. Ji, and N. Sheng, "Incipient Sensor Fault Detection by Directly Monitoring Sliding Window Based Singular Values∗," *IFAC-PapersOnLine*, vol. 55, no. 6, pp. 637–642, 2022, doi: 10.1016/j.ifacol.2022.07.199.
36. M. Uppal *et al.*, "Cloud-Based Fault Prediction for Real-Time Monitoring of Sensor Data in Hospital Environment Using Machine Learning," *Sustain.*, vol. 14, no. 18, 2022, doi: 10.3390/su141811667.
37. W.-X. Liu, R.-P. Yin, and P.-Y. Zhu, "Deep Learning Approach for Sensor Data Prediction and Sensor Fault Diagnosis in Wind Turbine Blade," *IEEE Access*, vol. 10, pp. 117225–117234, 2022, doi: 10.1109/ACCESS.2022.3219480.
38. A. Wahid, J. G. Breslin, and M. A. Intizar, "Prediction of Machine Failure in Industry 4.0: A Hybrid CNN-LSTM Framework," *Appl. Sci.*, vol. 12, no. 9, 2022, doi: 10.3390/app12094221.
39. M. Uppal, D. Gupta, S. Juneja, G. Dhiman, and S. Kautish, "Cloud-Based Fault Prediction Using IoT in Office Automation for Improvisation of Health of Employees," *J. Healthc. Eng.*, vol. 2021, 2021, doi: 10.1155/2021/8106467.
40. S. Safavi, M. A. Safavi, and H. Hamid, "Health Forecasting for Autonomous Vehicles," *Sensors*, pp. 1–23, 2021.
41. L. Hellerstein, "Learning with Maximum-Entropy Distributions ∗," pp. 123–145, 2001.
42. J. Ali, R. Khan, N. Ahmad, and I. Maqsood, "Random forests and decision trees," *IJCSI Int. J. Comput. Sci. Issues*, vol. 9, no. 5, pp. 272–278, 2012.
43. A. Natekin and A. Knoll, "Gradient boosting machines, a tutorial," *Front. Neurorobot.*, vol. 7, no. DEC, 2013, doi: 10.3389/fnbot.2013.00021.
44. Abdullah-All-Tanvir, I. Ali Khandokar, A. K. M. Muzahidul Islam, S. Islam, and S. Shatabda, "A gradient boosting classifier for purchase intention prediction of online shoppers," *Heliyon*, vol. 9, no. 4, p. e15163, 2023, doi: 10.1016/j.heliyon.2023.e15163.
45. S. Li, N. Jin, A. Dogani, Y. Yang, M. Zhang, and X. Gu, "Enhancing LightGBM for Industrial Fault Warning : An Innovative Hybrid Algorithm," *Processes*, vol. 12, no. 1, 2024, doi: 10.3390/pr12010221.
46. M. R. Machado, S. Karray, and I. T. De Sousa, "LightGBM: An effective decision tree gradient boosting method to predict customer loyalty in the finance industry," *14th Int. Conf. Comput. Sci. Educ. ICCSE 2019*, no. Nips, pp. 1111–1116, 2019, doi: 10.1109/ICCSE.2019.8845529.
47. O. Avci, O. Abdeljaber, S. Kiranyaz, M. Hussein, and D. J. Inman, "Wireless and real-time structural damage detection: A novel decentralized method for wireless sensor networks," *J. Sound Vib.*, vol. 424, pp. 158–172, 2018, doi: 10.1016/j.jsv.2018.03.008.
48. O. Avci, O. Abdeljaber, S. Kiranyaz, and D. Inman, "Structural damage detection in real time : Implementation of 1D convolutional neural networks for SHM applications," *Conf. Proc. Soc. Exp. Mech. Ser.*, vol. 7, pp. 49–54, 2017, doi: 10.1007/978-3-319-54109-9_6.
49. S. Kiranyaz, O. Avci, O. Abdeljaber, T. Ince, M. Gabbouj, and D. J. Inman, "1D convolutional neural networks and applications : A survey," *Mech. Syst. Signal Process.*, vol. 151, pp. 1–20, 2021, doi: 10.1016/j.ymssp.2020.107398.
50. K. M.Tarwani and S. Edem, "Survey on Recurrent Neural Network in Natural Language Processing," *Int. J. Eng. Trends Technol.*, vol. 48, no. 6, pp. 301–304, 2017, doi: 10.14445/22315381/ijett-v48p253.
51. Y. S. Shin and J. Kim, "Sensor Data Reconstruction for Dynamic Responses of Structures Using External Feedback of Recurrent Neural Network," *Sensors*, vol. 23, no. 5, 2023, doi: 10.3390/s23052737.
52. A. Sherstinsky, "Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) network," *Phys. D Nonlinear Phenom.*, vol. 404, no. March, 2020, doi: 10.1016/j.physd.2019.132306.