**PERFORMANCE ANALYSIS OF MACHINE LEARNING ALGORITHMS BASED WDM OPTICAL NETWORKS FOR HIGH SPEED APPLICATIONS**

*A Ph.D. Synopsis*

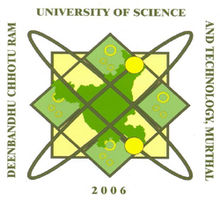
*Submitted in Partial Fulfillment of the*

*Requirement for the Award of the Degree of*

DOCTOR OF PHILOSOPHY

IN

ELECTRONICS AND COMMUNICATION ENGINEERING



Submitted by:

**Saloni Rai**

**Roll no. 19001903012**

Under the supervision of:

**Dr. Amit Kumar Garg**

(Professor, ECE Dept.)

DCRUST, Murthal

Department of ELECTRONICS AND COMMUNICATION ENGINEERING

(faculty of engineering & technology)

DEENBANDHU CHHOTURAM UNIVERSITY OF SCIENCE &TECHNOLOGY, MURTHAL, HARYANA (INDIA)-131039

(A State University Established Under Govt. of Haryana Act. 29 Of 2006)

**September 2020**

**CONTENTS**

1. Introduction
2. Literature Review
3. Problem Statement
4. Research Gaps
5. Research Objectives
6. Research Methodology
7. Tools to be used
8. Workplan
9. References

**1. Introduction**

With the rapid progress of internet services, 5G technology [1], Internet of Things (IoT) , the traffic demand of optical networks has been ever-increasing. Such demand brings forth higher requirements for the capability and efficiency of the next generation optical network. In the past few decades, the capacity of fibre transmission systems has been remarkably increased and is now approaching the theoretical limit. The development of elastic optical network (EON), network resources like spectrum, route, optical power, etc., can be utilised more efficiently to extract more capacity [2-3]. On the other hand, reliability of optical networks is also of great importance. Since a massive amount of data is transmitted in optical networks, even a tiny disruption of a service may cause huge losses for customers [4].

There is also a trend to introduce intelligence into optical networks for its fundamental role in the communication networks and its inherent complexity. Optical networks have been widely used as the major carrier network for traffic for several advantages; include wide bandwidth, low latency, and high anti-interference capability [5]. It connects the upper layer services and the underlying physical resources: on one hand, it needs to provision the bandwidth for different service and on the other hand, optical network involves resources allocation problem in multiple dimensions, such as wavelength, spectrum slot, and time slot. This situation makes optical network operation and maintenance more complicated than other communication networks. If consider, the convergence of optical network and other networks (such as 5G mobile network, and IP networks), it will become more serious.

In general, there are mainly three challenges [6], [7] faced by development and operation of optical networks: (i) Network complexity: The number and complexity of optical network devices increase with the scale of optical networks extend. Additionally, the optical network serves as carrier network in communication systems and may carry multiple heterogeneous networks, such as 5G mobile networks, IoT, vehicle networking, and cloud computing. How to adapt the traffic from these different networks becomes the first challenge for optical network operation and management. (ii) Complexity of services: Various Quality of Service (QoS) agreements enable optical networks to offer differentiated services, the emerging methods and technologies, such as network slicing, allow for the provision of networks to be applied in real time. The optical networks can then have a scalable infrastructure that offers a framework to offer various levels of QoS to different physical domain systems, which is challenging for a large network. (iii) Complexity of resource management: the optical network is a connection between the upper layer traffic and the underlying physical layer infrastructure. It is also responsible for the allocation of physical layer resources for the supply of traffic. There are, however, several dimensional physical resources to be distributed, such as cable, wavelength, bandwidth, modulation format and time slots. Joint allocation of several resources is time-consuming with high computational complexity.

Optical networks are extremely complex systems, posing problems that are difficult to solve using conventional techniques. Recently, machine learning (ML) techniques have been successfully applied to a large number of problems in optical communications. These problems include failure management, optical performance monitoring (OPM), quality of transmission (QoT) estimation, traffic prediction.

**2. Literature Review**

In this section, the state-of-the-arts based on the machine learning techniques are reviewed. For each one of the publications, the concept and highlight of its significant contributions are described.

Letaifa *et.al* [8] proposed a machine learning framework combined with adaptive coding to enhance the Quality of Experience (QoE) in video streaming applications. The proposed approach was developed using a software defined network (SDN). The proposed research implemented a ML-based structural similarity index measure (SSIM) algorithm for developing a centralized architecture to predict the network status. The ML-SSIM algorithm was used for predicting the QoE required for data streaming. The QoE predicted by the ML-SSIM algorithm was combined with the network status for selecting appropriate coding. Experimental analysis was conducted for evaluating the performance of the proposed model. It can be observed from the results that the proposed algorithm effectively enhances the user’s QoE with a success rate of 83%. Also, the ML algorithm increased the throughput of the system and reduced end-to-end delay. As a part of future scope, the study focussed to consider more performance metrics to improve QoE.

Choudhury *et al.* [9] implemented machine learning technique for two applications of internet protocol (IP)/Optical networks. The first application dealt with agile resource management in optical networks by employing ML-based prediction algorithms for predicting short- and long-term traffic prediction in optical networks. The second application was related to migration of networks to open reconfigurable optical add-drop multiplexers (ROADM) networks that allows physical routing without requiring optical network parameters. Various ML algorithms such as multilayer perceptron’s (MLP), ensemble learning (EL), random forest (RF) and regression algorithms were used for evaluating the effectiveness of the proposed approach in terms of prediction error. Results show that random forest regression trees achieved a low error rate of 0.81 (mean squared error -MSE) compared to other algorithms. The study suggested employing deep learning neural networks as an extension of the current research.

Rehman *et al*. [10] presented a QoS-aware, content-aware, and device-aware nonintrusive medical QoE (m-QoE) prediction model using a multilayer perceptron neural network for predicting QoE. MLP is also used as an open source platform for maintaining and optimizing the quality in a device-aware video data streaming process. The proposed framework was trained to enhance the QoS, content features of the output. Experimental validations were performed for evaluating the prediction efficiency obtained through Root Mean-Square-Error (RMSE) and correlation coefficient. The study achieved a RMSE and correlation coefficient of 0.109 and 92.2% respectively. As a part of further research, the study suggested the application of a proposed model for 5G networks.

Wu *et al.* [11] proposed a machine learning based predictive model for predicting link congestion in a SDN data plane. This research aimed to reduce link congestion by effectively predicting the degree of network congestion using four different ML algorithms. The study implemented Support Vector Machine (SVM), Multi-Layer Perceptron (MLP), One Dimensional Convolutional Neural Network (1DCNN) and K-Nearest Neighbour (KNN) algorithms for predicting network congestion. These algorithms were trained to predict the related congestion and the performance of these algorithms was determined through experimental analysis. Experimental evaluation shows that 1DCNN outperforms other ML algorithms by achieving superior prediction accuracy of 98.3%.

Kumar *et al*. [12] presented a novel approach of ML-based models for analysing the traffic flow in SDNs. The study proposed a hybrid approach of K-means clustering algorithm combined with another novel application of the Vector Space Model with cosine similarity. The model was trained by updating the network statistics during multiple instances to extract the knowledge for identifying a least congested path for routing the traffic flow in SDNs. The least congested path was identified from the available data related to probable paths using K-Means and Cosine Similarity approaches. It was observed from the results that, cosine similarity method was more efficient to identify the possible paths in terms of round-trip time (RTT) compared to k-means. However, the speedup ratio of the proposed model was more in K-means algorithms. In future, the proposed model can be used for other network traffic management applications such as multipath routing and traffic load distribution.

Raikar*et al.* [13] discussed the classification of data traffic in SDN systems using supervised machine learning algorithms. Conventional techniques for classifying the data traffic have failed to achieve desired accuracy because of lack of device scalability. This research focussed to overcome this limitation of conventional approaches by employing three advanced supervised learning algorithms such as support vector machine (SVM), nearest centroid and naïve bayes (NB). These algorithms were implemented to categorize the network traffic data based on SDN applications. Experimental evaluations showed that the SVM, NB and nearest centroid algorithms achieved superior accuracy of 92.3%, 96.79% and 91.02% respectively. The study planned to address the challenges related to capturing live network data and classification of network traffic data as a part of further research.

Carvalho *et al.* [14] suggested a device orchestrator that modifies the boundaries to boost the QoE of web-based content. The orchestrator utilized a Device-to-Device (D2D) communications to boost the QoE of the device, thus growing the focus on 4G networks. The usage of D2D was determined by using a machine learning algorithm. The orchestrator utilized ML to determine the video target of the UEs and moved from D2D to 4G based on the results. The study employed a random forest algorithm for improving the QoE. It can be observed from the results that the mean horizontal resolution of the video was significantly enhanced by 150%. For future work, the study planned to employ the predictor in real time traffic networks with background traffic which is one of the main parameters affecting the delivery quality.

Dias *et al.* [15] proposed a novel approach of real time traffic classification using machine learning framework as a solution to package branding, the discovery of operation in the executives' network, the organization of preservation, management forecasts, organization of configuration and security tests. The system depends on three traffic levels, consisting of two distinct video services and one download text, and satisfies the QoS requirements. The proposed approach employs a Naive Bayes algorithm for reducing the hypothesis of independence between the features of the Naive Bayes algorithm. It can be observed from the results that the proposed approach achieved a superior classification accuracy of 98.88% and hence can be applied for real-time traffic conditions. As a part of future work, a thorough analysis of the proposed approach will be performed with respect to accuracy and computational cost using different datasets.

Alawe*et al.* [16] proposed a novel framework for scaling 5G networks by predicting the changes in the traffic using machine learning-based predictive algorithms. The study employed a hybrid approach of long short-term memory cell (LSTM) and deep neural network (DNN) for training the neural network based on a real data set for achieving accurate prediction. The model was simulated for analysing the prediction-based scalability process and it was inferred from the results that the proposed approach exhibited superior performance compared to threshold-based approaches in terms of delay and latency. For future scope, the study suggested training recurrent neural network (RNN) for handling the data plane traffic, in a virtualized environment.

Troia*et al.* [17] proposed a deep learning-based traffic prediction model for predicting the traffic matrices which proactively optimizes the resource allocations in optical networks. The study employed RNN based gated recurrent unit (GRU) for predicting the traffic sequence with respect to speech and handwriting recognition and prediction in the traffic network. GRU was trained to achieve superior accuracy with a mean absolute error of less than 7.4. The predictions were used for allocating the resources in an optimal network proactively and dynamically. For future work, the proposed model will be implemented for optimizing the operator's network with other efficient machine learning techniques.

Fiandrino*et al.* [18] proposed a ML-based approach for optimizing the operation of future networks in 5G mobile networks. The study employed a LSTM and convolutional LSTM algorithms for developing a predictive model that leverages artificial intelligence for predicting the future traffic requirements and to characterize the traffic attributes. This mechanism helped in improving the performance of the network control techniques such as network routing, load balancing and scheduling. The study focussed on orchestrating the ML model for improving two different network systems. For further evaluation, the study intended to validate the implementation of mobile backhaul routing for reducing the packet delay compared to conventional approaches.

Zhang *et al.* [19] proposed an intelligent and application aware network traffic prediction model for enhancing the QoE in smart access gateways. The study employed a deep learning based encrypted packet classifier for identifying different network applications. Deep learning algorithms such as convolution neural networks and recurrent neural networks were employed for obtaining accurate network traffic data prediction. For experimental evaluation, the datasets of the network packets were collected from an open source as well as real time traffic flow data was obtained. Experimental analysis shows that the application awareness was achieved using CNN which was supported by the data packet classifier. It was observed from the results that the proposed model achieved significant prediction accuracy and better QoE using a LSTM recurrent neural network-based model. As a part of future work, the study intended to improve the accuracy of the traffic prediction model considering QoS of each application.

Canovas *et al.* [20] proposed a novel approach of multimedia traffic prediction model for SDN-based management system to predict QoE. The study employs a neural network known as Bayesian regularized neural networks (BRNN) for classifying the traffic pattern. The proposed model is based on estimation of QOE, hence QoE is modelled along with traffic pattern classification. Using the BRNN approach, the system effectively classified various types of traffic data based on the quality perceived by the users. The characteristics of the traffic video was determined based on different quality of service (QoS) parameters such as bandwidth, delay, loss rate and jitter. For QoE estimation, different ML algorithms were used such as Decision trees, SVM, CNN, Naïve Bayes and K-nearest neighbours. It can be observed from the results that the BRNN achieved superior classification results and recall score. For further research, the study intended to employ machine learning based statistical methods of conditional probability for obtaining maximum probability in classification and to optimize the system for improving the execution time.

Song *et al.* [21] evaluated the QoE based on deep learning-based system architecture for web browsing applications. The proposed architecture is divided into two components such as traffic classification sub-system and QoE prediction sub-system. The study employed fully connected neural networks and LSTM for achieving superior classification accuracy. It can be observed from the results that the proposed model achieved a traffic classification accuracy of 96.63% for more than 2000 packets over 6 websites. The achieved QoE prediction results were 0.0975 seconds for 5400 visits. For future work, the study plans to predict the user’s experience directly by conducting crowdsourcing tests.

Bhatia *et al.* [22] proposed a real-time urban traffic analysis in software-defined vehicular networks (SDVNs). The study proposed a data-driven framework for implementing a deep learning based recurrent neural network for predicting the behaviour of vehicular traffic flow. The model was developed by combining various performance metrics such as scalability, flexibility, and adaptability. A long-term short-term memory based neural network (LSTM-NN) architecture was developed to reduce the back-propagated error delay using memory blocks for predicting spatiotemporal traffic. The performance of the LSTM-NN model was evaluated and it can be inferred from the results that the proposed approach achieved an accuracy of 97% in predicting the traffic density. For further research, the study intended to enhance the efficiency of the proposed approach using external parameters such as vehicle density patterns in nearby traffic congestion areas.

Anandet *et al.* [23] suggested a non-linear time series model, generalised self-regressive conditional heteroskedasticity (GARCH), with an invention mechanism generalised to the heavy-tailed distribution class. In this, the model is fitted to the actual data and the results validate the fitness of the model. Author compared to other generic models and this model has increased predictability accuracy. In addition, the estimation of the parameter is less complex than the other models used to date in the modelling of Internet traffic data.

Dong-Chul Park *et al.* [24] have built Bilinear Recurrent Neural Network (BLRNN**)** to further improve the predictability of the BLRNN by incorporating dynamic learning control and layer optimization by layer process. Experiments are carried out on a real-world ethernet network traffic data collection. The results show that the dynamic BLRNN-based prediction scheme outperforms the traditional multi-layer neural-type perceptron network (MLPNN) in terms of normalised mean square error (NMSE).

Chabaa *et al.* [25] developed an artificial neural network (ANN) model based on a multi-layer perceptron (MLP) for the analysis of Internet traffic data over IP networks. ANN is used to evaluate the time series of calculated data for the assessment of the network response. For this purpose, the author used input and output data from internet traffic over IP networks to define the ANN model and compared the performance of some training algorithms used to estimate neuron weights. The comparison between certain training algorithms shows the efficiency and accuracy of the Levenberg-Marquardt (LM) and robust back spread (Rp) algorithms in terms of statistical criteria. As a result, the obtained outcome prove that the established models, using the LM and Rp algorithms, can be used effectively for the analysis of internet traffic over IP networks and can be used as a fundamental method for controlling internet traffic at different times.

Vijayakumar *et al.* [26] built the system to satisfy three primary objectives, i.e., the real-time principle of extracting high-volume data streams, dynamic data flow into a relational hierarchy, and the adaptive reorganisation of the traffic data hierarchy in response to evolving circumstances and time-to-time traffic patterns. As new data traffic is introduced, it is dynamically structured into a hierarchy that has been developed. As a consequence, after this cap has been reached, traffic data must be excreted at a rate equivalent to their congestion.

Junsong *et al.* [27] suggested a model that predicts internet traffic based on the Elman neural network (Elman-NN). Traffic is regarded as a time series, which is non-linear and variable. The Elman-NN is used to model the relationship with sufficient precision, and the Elman-NN based traffic model is used to forecast future traffic. The simulation results show that this approach is feasible and effective for modelling and forecasting traffic.

Peng Wangg *et al.* [28] has induced a system focused on an adaptive learning rate to maximise the speed of convergence of learning. Forecast, first, denoised the traffic time series with wavelet packet transformation to improve predictive accuracy, then compared the capacity of the BP neural network (BPNN) and improved BPWNN (IBPWNN) to predict network traffic. The results of the emulation experiment show that, in the case of one-step prediction, BPNN and IBPWNN have comparable predictive accuracy, however, in the case of multi-step prediction; the BPNN has poor predictive accuracy, while the IBPWNN still has a decent predictive capability.

Samira Chabaa *et al.* [29] used the adaptive neuro-fuzzy inference method (ANFIS) which is generated by an appropriate combination of fuzzy systems and neural networks to predict the input and output data set of the Internet traffic time series. Several statistical parameters are used to ensure the validity of this model. The results show that the ANFIS model provides strong ac curacy in the forecasting phase of internet traffic in terms of statistical indicators. This model suits the actual data well and offers an accurate overview of the network status at various times.

Wei Hu *et al.* [30] induced two data stream mining techniques and applied the problem of anomaly detection. Second, a stream cluster algorithm was used to detect suspicious packets. Using 1-gram and 2-gram tools, this method achieved reasonable success with Generic HTTP and Shell-code attacks but had a higher average false-positive rate. Second, a stream adaptation of the approach to the relative frequency histogram was developed using the Pearson correlation to detect anomalies. While the histogram-based approach obtained marginally better performance, it needed more fine-tuning because of the number of parameters used.

Nafja *et al.* [31] proposed the modular neural network for intrusion detection, which uses the principal component analysis (PCA) as a pre-processing layer to reduce the enormous amount of information provided in the knowledge discovery and data mining (KDD99) data set. PCA greatly reduces the high dimensionality of the data set without loss of information. This pre-process data in the form of the main component is then provided to the batch back propagation neural network for efficient intrusion detection. And these experiments are determined using the root mean square error (RMSE) modular neural network on the KDD 99 data set. These experimental results show an increase in the learning time due to a reduction in the high data dimensions.

Gupta *et al.* [32] proposed an intrusion detection system that uses the rough set theory for the selection of features, which extracts the relevant attributes from the whole set of attributes describing the data packet and uses the same theory to classify the packet if it is normal or an attack. Select or reduce the features after the simplification of the discernibility matrix. The Rosetta method for obtaining the rules on reductions and classification. The NSL KDD dataset is used as a training set and is given to Rosetta for the purpose of obtaining the classification rules.

Zhou Mingqiang *et al.* [32] proposed a graph-based intrusion detection algorithm using an outlier detection method based on a local deviation coefficient (LDCGB). Compared to other clustering intrusion detection algorithms, this algorithm is not required for the initial cluster number. In the meantime, it is resilient in the outlier 's affection and capable of detecting the form of a cluster rather than a single circle. In addition, it still has a good detection rate for unknown or muted attacks. LDCGB uses a graph-based cluster algorithm ( to get an initial partition of the data set that depends on the cluster precision parameter rather than the initial cluster number. On the other hand, because this intrusion detection model is based on a mixed training data set, it must have high label accuracy to ensure its performance. Therefore, in the label word, the algorithm imposes a local deviation coefficient external detection algorithm to re-mark the output of the GB algorithm. This test is capable of enhancing the consistency of labelling.

Mohammad Mahdi Tajiki *et al.* [34] induced QoS-aware resource reallocation algorithm that exploits software-defined networking advantages to proactively avoid congestion by using traffic prediction techniques. The author proposed two schemes: I an effective solution; and (ii) a fast, sub-optimal one. The proposed schemes are contrasted with the viewpoint of accuracy. In addition, the effect of the forecast on the reallocation of capital is addressed. The proposed scheme reduces the loss of packets and greatly increases the throughput.

Yan Zheng *et al.* [35] proposed advanced prediction and navigation models on a complex traffic network. In comparison to the conventional shortest path algorithms, which concentrate on the static network, the first approach considered possible traffic jams and was developed to provide optimal driving advice for different periods of the day. Consequently, by dividing the real-time global taxi positioning system data in Shenzhen city into 50 regions, the Markov balance chain model was designed to dispatch vehicles and applied to ease city congestion. With the development of field tests, traffic congestion on city traffic networks can be minimised easily and efficiently, and the efficiency of the system can also be improved.

Yongli Zhao *et al*. [36] suggested a novel SDON-based control architecture that could support both control layer AI and on-board AI at the same time. In particular, edge computing on-board AI is proposed to support various ML applications. To evaluate the proposed architecture, the author built an experimental test bed and demonstrated a typical use case, i.e. the prediction of alarm information. Experimental findings show that synchronisation and cross-layer optimisation between the control layer AI and the on-board AI can be accomplished.

In [37], the authors suggested a data-driven system for the allocation of bandwidth in EON and considered the QoS specifications of the network. Strengthening learning (RL) is used to solve the problem of bandwidth allocation. This structure is sensitive to increasing the network load to meet the quality-of - service requirements of the network. The Bandwidth Allocation (BA) problem is laid out as a partially observable Markov Decision Process (POMDP). The design is scalable using POMDP and RL; the central controller is scalable.

Suchao Xiao et al. [38] concentrated on the issue of dynamic resource allocation of bandwidth in transport network parts. And implemented a new LSTM-based traffic-predicting dynamic transport network slicing architecture. The solution consists of two phases: the phase of traffic prediction and the phase of bandwidth configuration. Long and short memory models are used in the first step to forecast traffic. In the second step, the model was modelled as a fractional knapsack problem, and the greedy algorithms were used to find approximate solutions. The decentralised distribution of resources to programmes with various goals can be accomplished, thereby enhancing the quality of service and the user experience of the entire system.

Yang *et al.* [39] proposed a hybrid multi-fail location approach with Hopfield neural network (HNN) in radio and optical wireless networks. The actual failure nodes and links can be localised via alarm processing, network topology and service information. The relationship between detected failures and alarms is then modelled to create a bipartite graph. The bipartite graph is computed using HNN 's quick processing to solve combination optimization problems. The HNN-based method can be used in parallel computing training to minimise computing time. The simulation results show the time-efficiency of the proposed system.

A case-based justification (CBR) methodology is proposed for lightpath QoT estimation, which classifies lightpaths into two groups, namely high or low-quality impairment-conscious resource allocation in optical networks [40]. In this analysis, the CBR method stores a information database with each sample consisting of a set of features related to the lightpath and the corresponding Q-factor value. When determining the QoT of a new sample, the characteristics of the new lightpath will be compared with the samples in the CBR database and the weighted Euclidean distance is measured to determine the similarity between the two lightpaths. The corresponding Q-factor of the light path in the knowledge base with the greatest similarity is assigned to the current light path as its Q-factor. The lightpath Q-factor is contrasted with the Q-factor threshold (Q-threshold) for the lightpath efficiency classification. The authors have addressed the CBR with learning and forgetting strategies to refine the knowledge base in order to reduce its difficulty.

Deep neural networks (DNNs) are also used for traffic prediction within data centre optical networks. Predicted traffic information is used for resource allocation to increase bandwidth efficiency and reduce the risk of blocking [41]. The hybrid electro-optical system data centre network uses both optical switching (OCS) and electrical packet switching (EPS) circuits. OCS is used for the transport of large-capacity networks and EPS is used for the transmission of short-lived and burst data streams. This hybrid structure is capable of transporting traffic of different granularities at low congestion rates. Prediction of traffic is crucial to early decision whether traffic can pass via OCS or EPS transmission. The nonlinear auto-regressive neural network (NARNN) is used in a busy service scenario for traffic prediction in a hybrid electro-optical DCN scenario. Predicted traffic information helps managers to target the worst traffic flows for optical switching and electrical switching in the future.

Nicolas Jara *et al.* [42] has suggested a fault tolerance technique that is robust to any collection of simultaneous link failures, as long as the network topology allows re-connection through links that remain in service. This scheme is executed before a network service, normally taking just a few seconds of execution time. This fast execution also makes it possible to quickly overcome any link failure scenario during network operation. In addition, the network operation based on this method is fast and straightforward, as the routes (primary and secondary) are stored in routing tables and checked only on requests.

In [43], the authors proposed optical chaos and hybrid wavelength division multiplexing / time division multiplexing (WDM / TDM) based on large-capacity quasi-distributed sensing networks. With WDM / TDM technology, hundreds of sensing units could be multiplexed in multiple fibre sensing lines. This sensing network could be used to track fibre faults in real time.

In [44], the authors used a distributed fibre optic sensor to enhance the optical fibre cable conditioning system. A variety of surviving technologies designed to ensure optical sensing and optical communication are already in place to avoid link / node failures. For example, the classic p-cycle scheme ensures that there is at least one available light path between any node pair located in this cycle after single failure [45]. Traditional survival methods can be classified into two categories: security schemes and rehabilitation schemes. The security scheme reserves backup resources for workflows and only needs safety switching at source and destination nodes when failures occur [46]. Restoration schemes, such as link-based restoration and path-based restoration, attempt to create recovery channels for broken workflows using the remaining available network resources after failure [47].

In [48], ANN was adopted with the same input as in [49], and the output is the power excursion of each channel and the accuracy can be up to 90%. However, the experiment environment [49] did not consider channel interactions in WDM systems and wavelength switching operations. To solve this problem, authors in [50] have expanded the application to multiple reconfigurable optical add-drop multiplexers (ROADM) and complete C-band WDM channels. Under the same setup, in [50] the use of the neural network in [51] can be further enhanced with an RMSE of less than 0.15 dB compared to the RR model. And the precision of the wavelength assignment can be up to 99 percent. Similarly, the DNN was used in [52]. It shows that DNN obtains 0.1 dB of RMSE for 8400 randomised samples and has better performance than the regression method and the random forest.

**3. Problem Statement**

The problem statement is inferred pertaining to design an optimized form machine learning technique which efficiently predicts traffic, optimally allocates resource, controls congestion and guarantees fault tolerance by improving QoE, QoT and QoS parameters.

**4. Research Gaps**

In recent times, the increase in the data volume has affected the performance of the traffic prediction models in optical networks. In order to enhance the performance of the traffic prediction models in a cost-efficient way, it is essential to overcome dynamic allocation problems such as optimal resource allocation and different prediction problems associated with network performance such as QoE, QoS and QoT. Researchers have proposed various machine learning-based intelligent traffic prediction models to analyze the behaviour of traffic prediction in optical network systems in real-time scenarios. However, there are certain issues which are left unanswered by the existing predictive models. The prominent challenges which need to be addressed are:

1. With the increase in functional complexity of the networks, it is essential to develop a system which is capable of self-learning and self-optimizing.
2. Most of the conventional approaches have used supervised and reinforcement learning for dealing with network traffic prediction. However, very little focus is given for unsupervised machine learning approaches.
3. Conventional traffic prediction models are incorporated with multiple mechanisms and different resources which increase the complexity of the predictive models.
4. Most of the existing approaches do not deal with high-dimensional problems such as prediction of network performance in terms of end-use quality of experience while maintaining the accuracy and reliability of the system.
5. It is essential to develop an effective network traffic management tool which is capable of handling exponential network traffic growth with optimal utilization of the network resources.

**5. Research Objectives**

Based upon Comprehensive literature survey, the following are the research objectives that have been framed for deployment of WDM optical networks.

1. To propose an efficient traffic prediction and pattern extraction technique using supervised machine learning for optimal routing and wavelength assignment in optical networks.
2. To investigate and to propose novel machine learning based dynamic bandwidth allocation approach for bandwidth demand of end users to increase the bandwidth utilization on consideration with reduction in control overheads.
3. To overcome congestion control issues in optical networking by using efficient supervised machine learning technique thereby meeting both QoS and QoE parameters.
4. To propose machine learning based QoT impairments modeling and monitoring schemes for WDM optical networks supporting 5g application.

**6. Research Methodology**

The research methodology for executing the research in the direction of timely completion of the target objectives can be divided in the following phases:

**Phase I**

**Literature Review:** During this phase, various machine learning techniques for traffic prediction in optical networks are reviewed and various parameters influencing the traffic prediction are analysed. The review will be thoroughly conducted for a period of around 30 months.

**Phase II**

**Research Planning:** In this phase, evaluation of required research materials and design planning are carried out. This phase will be conducted for a period of 6 months along with the next phases till the documentation phase initiates.

**Phase III**

**Experimentation and Simulation:** In this phase, the proposed network traffic prediction model will be developed by implementing the proposed machine learning techniques and the model will be subjected for performance validation. All the performance evaluation will be made with respect to designated performance metrics such as QoE, QoS, QoT, BER and OSNR analysis. To theoretically validate the equivalent circuit models suggested experimentally in the initial course of this phase, simulation of the traffic prediction model will be done using required software tools.

**Phase IV**

**Analysis of Results:** In this phase, the obtained results will be analysed so that it can be known that the proposed machine learning algorithm is effective in achieving the target research objectives by comparing it with other effective network prediction techniques.

**Phase V**

**Publishing the work:** This is the phase which will continue throughout the research work. When some findings will come, the same will be published in the form of research papers in journals of repute.

**Phase VI**

**Documentation:** The final phase will be the documentation in the form of a thesis. This phase will enable me to compile my experience throughout my research period in the form of various chapters of my thesis starting right from literature review to final result compilation and conclusion.

**7. Tools to be used**

**NS3 / MATLAB:** It is used to create a pin to point connection between the nodes in a wireless sensor network. It is used to simulate the network of communicating nodes and the traffic between the nodes.

**8. Work Plan**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **S No.** | **Activity** | |  |  | | --- | --- | | **SEM 1** | **SEM 2** | | | | | |  |  | | --- | --- | | **SEM 3** | **SEM 4** | | | | | |  |  | | --- | --- | | **SEM 5** | **SEM 6** | | | | | **Timeline** |
| 1 | Course Work | **✓** | **✓** |  |  |  |  |  |  |  |  |  |  | Completed |
| 2 | Literature Review | **✓** | **✓** | **✓** | **✓** | **✓** | **✓** | **✓** | **✓** | **✓** | **✓** |  |  | On-going |
| 3 | Presentation of Proposed Ph.D. Synopsis |  |  |  |  | **✓** |  |  |  |  |  |  |  | On 25.09.2020 |
| 4 | Simulation Analysis |  |  |  |  | **✓** | **✓** | **✓** |  |  |  |  |  |  |
| 5 | Analysis & Interpretation of Results |  |  |  |  |  |  | **✓** | **✓** | **✓** |  |  |  |  |
| 6 | Validation of Results |  |  |  |  |  |  |  | **✓** | **✓** |  |  |  |  |
| 7 | Publications |  |  |  |  |  |  |  |  | **✓** | **✓** |  |  | Parallel with Objectives |
| 8 | Thesis Writing |  |  |  |  |  |  |  |  |  | **✓** | **✓** | **✓** |  |

**References**

|  |  |
| --- | --- |
| [1] | Mata, J., de Miguel, I., Duran, R. J., Merayo, N., Singh, S. K., Jukan, A., & Chamania, M., “Artificial Intelligence (AI) Methods in Optical Networks: A Comprehensive Survey” *Optical switching and networking,* vol. 28, pp. 43-57, 2018. |
| [2] | Michal Aibin,“Traffic prediction based on machine learning for elastic optical networks,”*Optical Switching and Networking*, Vol. 30, pp. 33-39,2018. |
| [3] | S. K. Singh and A. Jukan, “Machine-learning-based prediction for resource (Re)allocation in optical data center networks,”*in IEEE/OSA Journal of Optical Communications and Networking*, vol. 10, no. 10, pp. D12-D28, Oct. 2018. |
| [4] | Xiong, Y., Yang, Y., Ye, Y., & Rouskas, G. N., “A machine learning approach to mitigating fragmentation and crosstalk in space division multiplexing elastic optical networks, ”*Optical Fiber Technology,*vol. 50, pp. 99-107, 2019. |
| [5] | Gu, R., Yang, Z., & Ji, Y., “Machine learning for intelligent optical networks: A comprehensive survey,” *Journal of Network and Computer Applications*, vol. 157, pp. 1- 36, 2020. |
| [6] | J. Wang, C. Jiang, H. Zhang, Y. Ren, K. Chen and L. Hanzo, “Thirty Years of Machine Learning: The Road to Pareto-Optimal Wireless Networks,”*in IEEE Communications Surveys & Tutorials*, vol. 22, no. 3, pp. 1472-1514,2020. |
| [7] | Ruan, L., Dias, M. P., & Wong, E., “Enhancing latency performance through intelligent bandwidth allocation decisions: a survey and comparative study of machine learning techniques,”*IEEE/OSA Journal of Optical Communications and Networking*, Vol. 12, Issue 4, pp. B20-B32, 2020. |
| [8] | A. Ben Letaifa, G. Maher and S. Mouna, “ML based QoE enhancement in SDN context: Video streaming case,”*2017 13th International Wireless Communications and Mobile Computing Conference (IWCMC), Valencia,* pp. 103-108, 2017. |
| [9] | Choudhury, G., Lynch, D., Thakur, G., & Tse, S., “Two use cases of machine learning for SDN-enabled IP/optical networks: Traffic matrix prediction and optical path performance prediction,” *IEEE/OSA Journal of Optical Communications and Networking*, vol. 10(10), pp. D52-D62, 2018. |
| [10] | I. Rehman, M. Nasralla, and N. Philip, “Multilayer Perceptron Neural Network-Based QoS-Aware, Content-Aware and Device-Aware QoE Prediction Model: A Proposed Prediction Model for Medical Ultrasound Streaming Over Small Cell Networks,” *Electronics,* vol. 8, no. 2, p. 194, Feb. 2019. |
| [11] | J. Wu, Y. Peng, M. Song, M. Cui and L. Zhang, “Link Congestion Prediction using Machine Learning for Software-Defined-Network Data Plane,”*2019 International Conference on Computer, Information and Telecommunication Systems (CITS), Beijing, China,*pp. 1-5, 2019. |
| [12] | S. Kumar, G. Bansal and V. S. Shekhawat, “A Machine Learning Approach for Traffic Flow Provisioning in Software Defined Networks,”*2020 International Conference on Information Networking (ICOIN), Barcelona, Spain,*pp. 602-607, 2020. |
| [13] | Raikar, M. M., Meena, S. M., Mulla, M. M., Shetti, N. S., &Karanandi, M., “Data Traffic Classification in Software Defined Networks (SDN) using supervised-learning,” *Procedia Computer Science*, vol.171, pp. 2750-2759, 2020. |
| [14] | Carvalho, M., Silva, V. F., e Silva, E. D. B., Macedo, D. F., de Resende, H. C., Marquez-Barja, J. M., & Wickboldt, J., “QoE-Based Video Orchestration for 4G Networks,” *In 2019 IEEE 30th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, pp. 1-6, 2019. |
| [15] | Dias, K. L., Pongelupe, M. A., Caminhas, W. M., & de Errico, L., “An innovative approach for real-time network traffic classification,” *Computer Networks*, vol. 158, pp. 143-157, 2019. |
| [16] | Alawe, I., Ksentini, A., Hadjadj-Aoul, Y., & Bertin, P., “Improving traffic forecasting for 5G core network scalability: A Machine Learning approach, “*IEEE Network*, vol. 32(6), pp. 42-49, 2018. |
| [17] | Troia, S., Alvizu, R., Zhou, Y., Maier, G., & Pattavina, A., “Deep learning-based traffic prediction for network optimization,” *In 2018 20th International Conference on Transparent Optical Networks (ICTON)*, pp. 1-4, 2018. |
| [18] | C. Fiandrino, C. Zhang, P. Patras, A. Banchs and J. Widmer, “A Machine-Learning-Based Framework for Optimizing the Operation of Future Networks,”*in IEEE Communications Magazine,* vol. 58, no. 6, pp. 20-25, 2020. |
| [19] | J. Zhang, F. Ye and Y. Qian, “Intelligent and Application-Aware Network Traffic Prediction in Smart Access Gateways,”*in IEEE Network*, vol. 34, no. 3, pp. 264-269,2020 |
| [20] | Canovas, A., Rego, A., Romero, O., & Lloret, J., “A robust multimedia traffic SDN-Based management system using patterns and models of QoE estimation with BRNN’’, *Journal of Network and Computer Applications*, vol. 150, 102498, 2020. |
| [21] | E. Song et al., “Threshold-oblivious on-line web QoE assessment using neural network-based regression model,”*in IET Communications,* vol. 14, no. 12, pp. 2018-2026,2020 |
| [22] | Bhatia, J., Dave, R., Bhayani, H., Tanwar, S., & Nayyar, A., “Sdn-based real-time urban traffic analysis in vanet environment,”*Computer Communications*, vol. 149, pp. 162-175, 2020 |
| [23] | Anand, N. C., Scoglio, C., & Natarajan, B.,“GARCH—Non-linear time series model for traffic modeling and prediction”, *In NOMS 2008-2008 IEEE Network Operations and Management Symposium,*pp. 694-697, 2008. |
| [24] | Park, D. C., & Woo, D. M.,“Prediction of network traffic using dynamic bilinear recurrent neural network”,*In 2009 Fifth International Conference on Natural Computation,* vol. 2, pp. 419 – 423, 2009. |
| [25] | Chabaa, S., Zeroual, A., & Antari, J., “Identification and prediction of internet traffic using artificial neural networks,” *Journal of Intelligent Learning Systems and Applications*, vol. 2(03), pp. 147, 2010. |
| [26] | Vijayakumar, M., & Parvathi, R. M. S., “Concept mining of highvolume data streams in network traffic using hierarchical clustering,” *European Journal of Scientific Research*, vol. 39(2), pp. 234-242, 2010. |
| [27] | Junsong, W., Jiukun, W., Maohua, Z., & Junjie, W., “Prediction of internet traffic based on Elman neural network” *In 2009 Chinese Control and Decision Conference*, pp. 1248-1252, 2009. |
| [28] | Wang, P.,Liu, Y., “Network traffic prediction based on improved BP wavelet neural network,” *In 2008 4th International Conference on Wireless Communications, Networking and Mobile Computing*, pp. 1-5, 2008. |
| [29] | Chabaa, S., Zeroual, A.,Antari, J., “ANFIS method for forecasting internet traffic time series,”*In 2009 Mediterrannean Microwave Symposium (MMS),* pp. 1-4, 2009. |
| [30] | Miller, Z., Deitrick, W., Hu, W., “Anomalous network packet detection using data stream mining,” *Journal of Information Security*, vol. 2(04), pp. 158, 2011. |
| [31] | Al-Nafjan, K., Al-Hussein, M. A., Alghamdi, A. S., Haque, M. A., & Ahmad, I., “Intrusion detection using PCA based modular neural network,”*International Journal of Machine Learning and Computing*, vol. 2(5), pp.583, 2012. |
| [32] | N.Gupta, N.Singh, V. Sharma, T. Sharama, A.S. Bhandra, “Feature Selection and Classification of intrusion detection using rough set,” *International Journal of Communication Network Security,*vol. 2(2), 2013. |
| [33] | Z. Mingqiang, H. Hui and W. Qian, “A graph-based clustering algorithm for anomaly intrusion detection,”*2012 7th International Conference on Computer Science & Education (ICCSE), Melbourne, VIC,* pp. 1311-1314, 2012. |
| [34] | M. Tajiki, B. Akbari, M. Shojafar, and N. Mokari, “Joint QoS and Congestion Control Based on Traffic Prediction in SDN,” *Applied Sciences*, vol. 7, no. 12, pp. 1265, 2017. |
| [35] | Yan Zheng, Yanran Li, Chung-Ming Own, Zhaopeng Meng, Mengya Gao, “Real-Time Predication and Navigation on Traffic Congestion Model with Equilibrium Markov Chain.” *International Journal of Distributed Sensor Networks*, vol. 14(4), pp. 1 – 13, 2018, |
| [36] | Y. Zhao, B. Yan, Z. Li, W. Wang, Y. Wang and J. Zhang, "Coordination between control layer AI and on-board AI in optical transport networks [Invited]," *in IEEE/OSA Journal of Optical Communications and Networking*, vol. 12, no. 1, pp. A49-A57, 2020. |
| [37] | T. Panayiotou et al., “A data-driven bandwidth allocation framework with QoS considerations for EONs,”*Journal of Lightwave Technology*, 2019. |
| [38] | S. Xiao and W. Chen, “Dynamic Allocation of 5G Transport Network Slice Bandwidth Based on LSTM Traffic Prediction,”*2018 IEEE 9th International Conference on Software Engineering and Service Science (ICSESS), Beijing, China,*pp. 735-739, 2018 |
| [39] | Yang H, Wang B, Yao Q, “Efficient Hybrid Multi-Faults Location Based on Hopfield Neural Network in 5G Coexisting Radio and Optical Wireless Networks,”*IEEE Transactions on Cognitive Communications and Networking*, vol. 5, pp. 1218-1228, 2019. |
| [40] | J. Li, D. Wang and M. Zhang, “Low-Complexity Adaptive Chromatic Dispersion Estimation Scheme Using Machine Learning for Coherent Long-Reach Passive Optical Networks,”*in IEEE Photonics Journal*, vol. 11, no. 5, pp. 1-11,2019. |
| [41] | Q. Guo et al., “Proactive Dynamic Network Slicing with Deep Learning Based Short-Term Traffic Prediction for 5G Transport Network,”*2019 Optical Fiber Communications Conference and Exhibition (OFC), San Diego, CA, USA*, pp. 1-3, 2019. |
| [42] | N. Jara, H. Pempelfort, G. Rubino and R. Vallejos, “A Fault-Tolerance Solution to Any Set of Failure Scenarios on Dynamic WDM Optical Networks with Wavelength Continuity Constraints,”*in IEEE Access*, vol. 8, pp. 21291-21301, 2020. |
| [43] | Luo, Y.; Xia, L.; Xu, Z.; Yu, C.; Sun, Q.; Li, W.; Huang, D.; Liu, D.,“Optical chaos and hybrid WDM/TDM based large capacity quasi-distributed sensing network with real-time fiber fault monitoring,”*Opt. Express*, vol. 23, pp. 2416–2423, 2015. |
| [44] | Tang, B.; Zhou, Z., “The design of communication network optical fiber cable condition monitoring system based on distributed optical fiber sensor,”*In Proceedings of the 2018 International Conference on Electronics Technology (ICET), Chengdu, China,* pp. 97–101, 2018. |
| [45] | Stamatelakis, D.; Grover, W.D.,“Theoretical underpinnings for the efficiency of restorable networks using preconfigured cycles (“p-cycles”),”*IEEE Trans. Commun*, vol. 48, pp. 1262–1265, 2000. |
| [46] | Kamal, A.E. “1 + N Network Protection for Mesh Networks: Network Coding-Based Protection Using p-Cycles,”*IEEE/ACM Trans. Netw.,* vol. 18, pp. 67–80, 2009. |
| [47] | Azim, M.; Jiang, X.; Ho, P.H.; Horiguchi, S.; Guo, M.,“Restoration Probability Modelling for Active Restoration-Based Optical Networks with Correlation Among Backup Routes,”*IEEE Trans. Parall. Distrib,* vol. 18, pp. 1592–1606, 2007. |
| [48] | You Y R, Jiang Z P, Janz C., “Machine learning-based EDFA gain model,”*In: Proceedings of European Conference on Optical Communication*, pp. 1–3, 2018. |
| [49] | Huang Y S, Gutterman C L, Samadi P, et al.,“Dynamic mitigation of EDFA power excursions with machine learning,” *Opt Express,*vol. 25, pp. 2245, 2017. |
| [50] | Freire M, Mansfeld S, Amar D, et al., “Predicting optical power excursions in erbium doped fiber amplifiers using neuralnetworks,” *In: Proceedings of Asia Communications and Photonics Conference*, pp.1–3, 2018 |
| [51] | Gutterman C L, Mo W Y, Zhu S X, et al., “Neural networkbased wavelength assignment in optical switching,”*In: Proceedings of the Workshop on Big Data Analytics and Machine Learning for Data Communication Networks,*pp. 37–42, 2017. |
| [52] | Delezoide C, Christo doulopoulos K, Kretsis A, et al., “Pre-emptive detection and localization of failures towards marginless operations of optical networks,” *In: Proceedings of International Conference on Transparent Optical Networks,* pp. 1 – 4, 2018. |