TELECOM SudParis

Institut Mines-Télécom

Master ROSP Lecture

Machine Learning in Optical Networks

Catherine Lepers

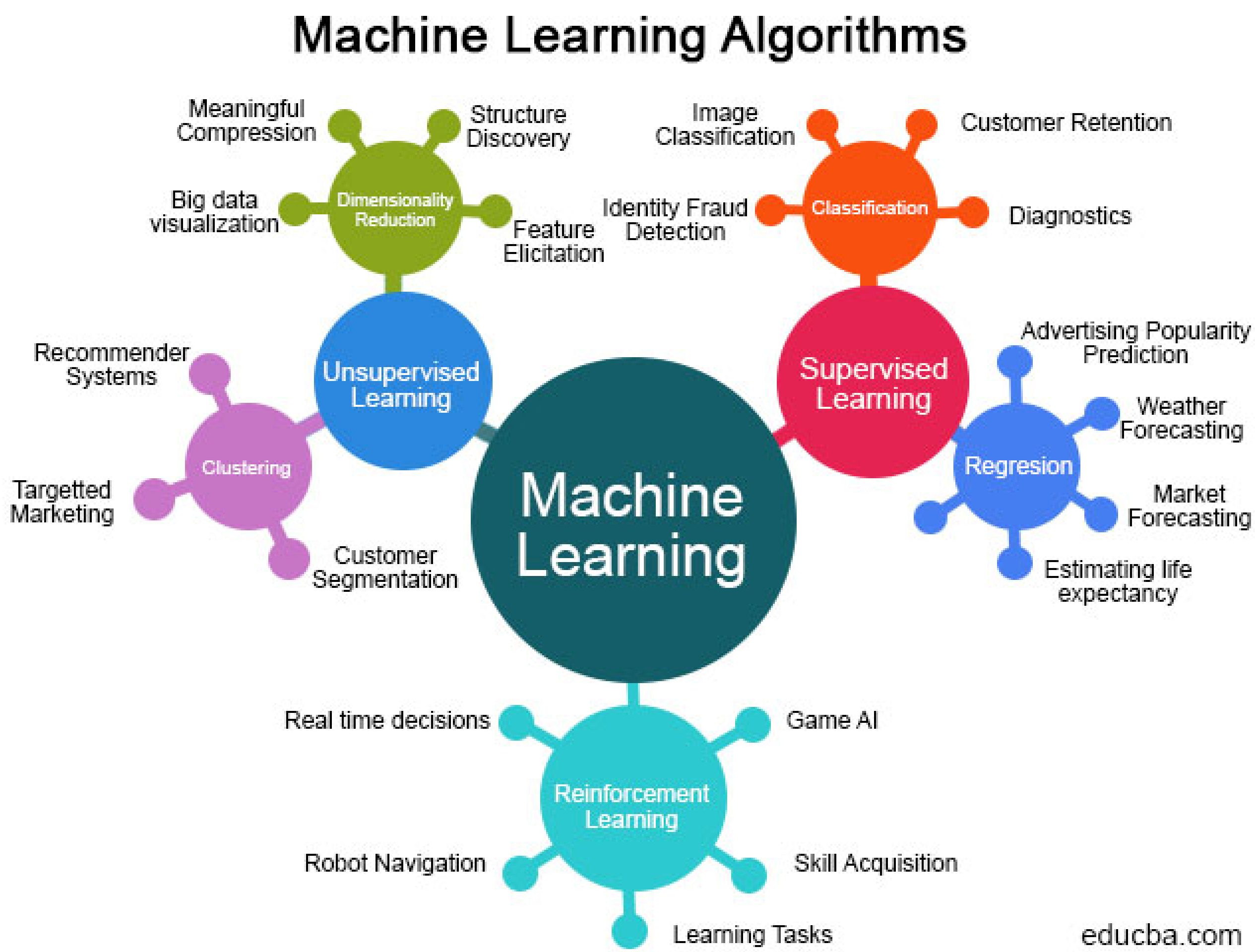
Summary

I. Overview on Machine Learning

II. Applications of ML to optical networks

III. Use Case: Optical amplifier control using ML

Overview on Machine Learning



Overview on Machine Learning

- Many machine learning algorithms

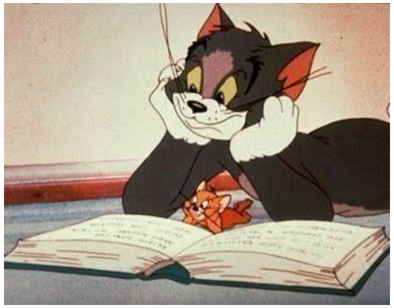
Supervised learning	Unsupervised learning	Reinforcement learning
Train an algorithm on a labeled dataset to predict the correct output value or a class of values for unseen inputs	Train an algorithm to find similarities or abnormalities in a dataset	Learn through trial and error from interaction with an environment

- Data availability
- Performance metrics
- Complexity vs accuracy

Overview on Machine Learning?

Predictive or upervised Learning

- Data are used to train the machine or to let the machine learning
- Two output types:
  - Real-valued scalar; the problem is regression
  - Categorical; the problem is classification or pattern recognition

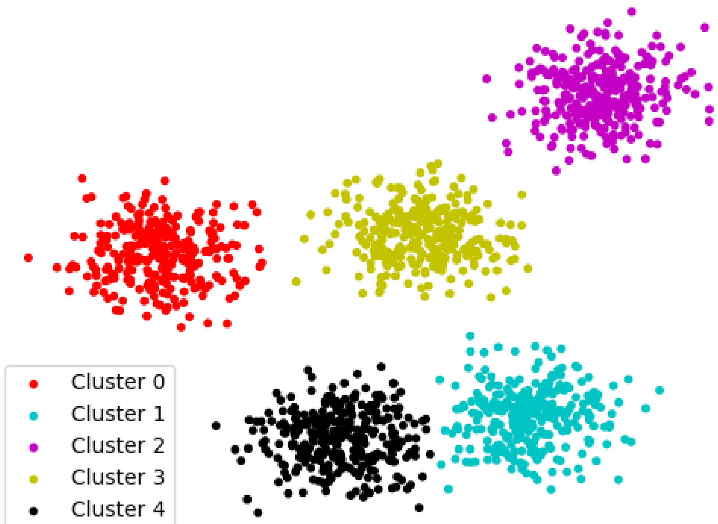


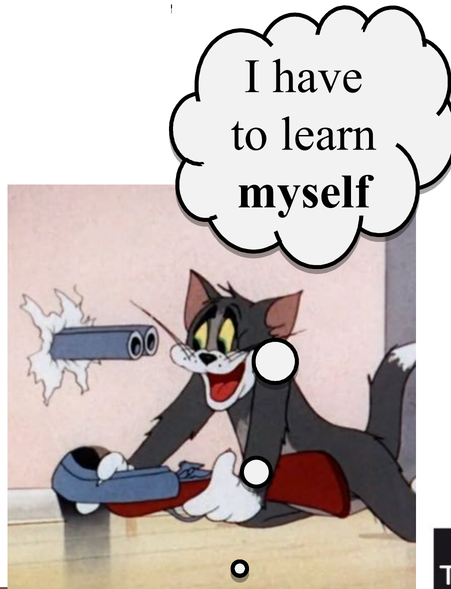
TELECOM SudParis

Overview on Machine Learning?

Descriptive or unsupervised Learning

- **Definition:** Only given inputs; the goal is to find "interesting patterns" in the data
- Mainly used for clustering (**K-Mean, Gaussian clustering**)





Overview on Machine Learning?

Reinforcement Learning

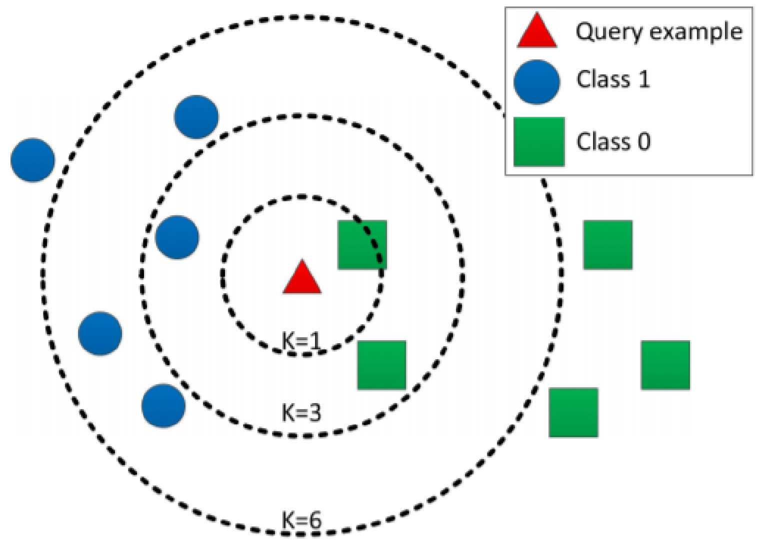
- Useful for learning how to act or behave when given reward or punishment signals
- Data available with feedback (Reward)
- Machine's task is to optimize the policy for more reward

Basic Concepts in ML

- **Parametric versus non-parametric models**
  - Does the model has a fixed number of parameters (parametric model) or does the number of parameters grows with the amount of training data (non-parametric model)?
- **A simple non-parametric classifier : K-nearest neighbors (KNN)**
  - This looks at the K points in the training set that are nearest to the test input x, counts how many members of each class are in the set and returns that empirical fraction as the estimate.
- **Parametric models for classification and regression**
  - Assumptions about the nature of the data distribution
  - Linear regression: the response is a linear function of the inputs
  - Logistic regression: generalization of the linear regression to the binary classification setting (it is a form of classification, not regression) (sigmoid function)

K-nearest neighbors (KNN)

- This looks at the K points in the training set that are nearest to the test input x, counts how many members of each class are in the set and returns that empirical fraction as the estimate.



Logistic Regression

- Generalization of the linear regression to the binary classification setting (it is a form of classification, not regression) (logistic or sigmoid function)
- Sigmoid function means S-shaped: it is also known as a squashing function, since it maps the whole real line to [0,1], necessary for the ouput to be interpreted as a probability

$$\sigma(s_e) = \frac{1}{1 + e^{-(\theta_0 + \theta^T x_e)}}.$$

Artificial Neural Networks

Classification or regression neurons

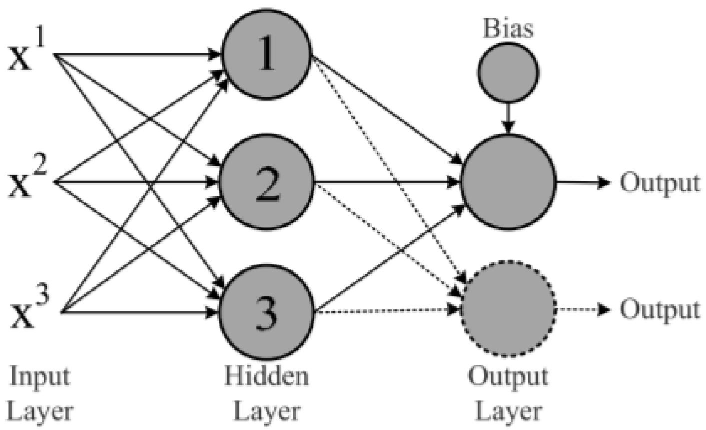
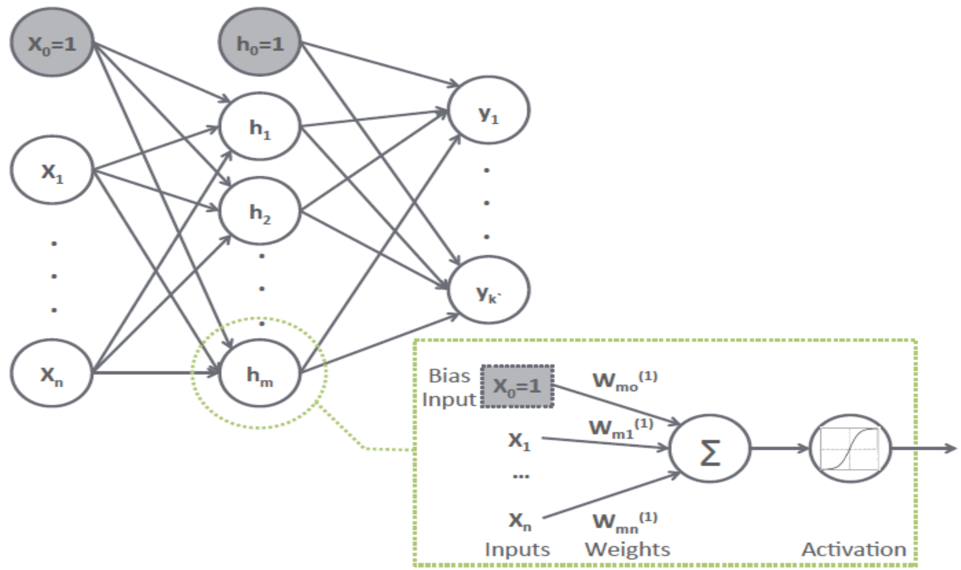


Fig. 5. Example of an ANN with one hidden layer composed of three neurons and an output layer with two neurons.



Machine Learning for Optical Networking

Machine Learning for Optical Networking

Why?

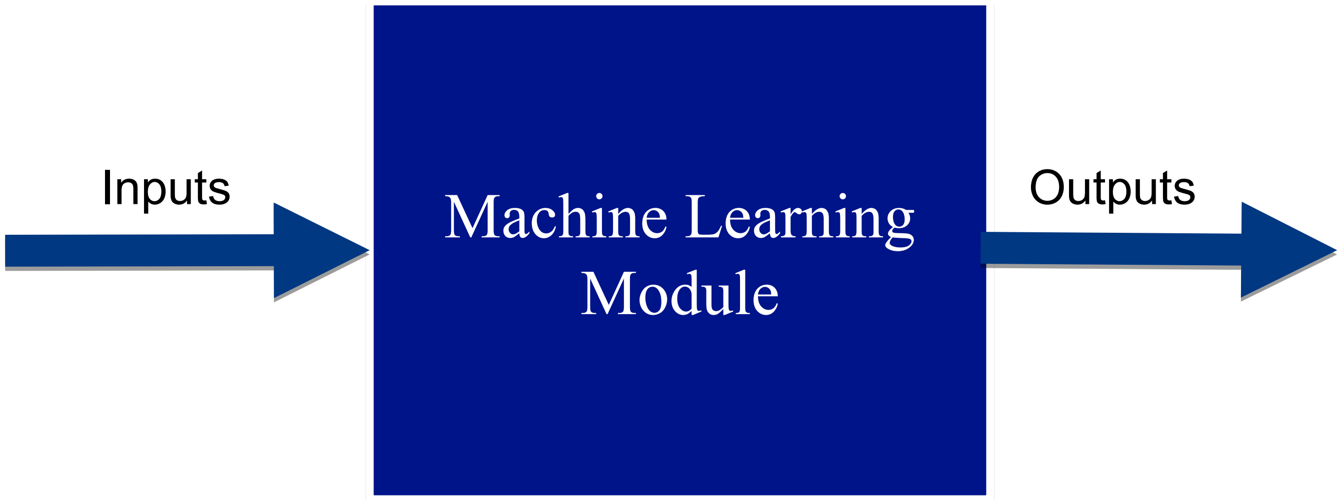
General case	Optical network case
Increased data availability	Large number of monitors
Increased system complexity	Elastic Optical Networks, complex modulation formats…
ML libraries (Keras, TensorFlow…)	

ML provides high accurate solutions to complex problems

TELECOM SudParis

Machine Learning for Optical Networking

- General scheme



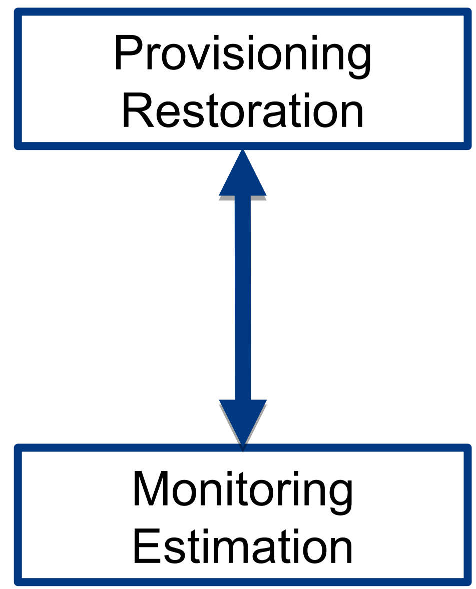
Machine Learning for Optical Networking

Network layer domain

- Traffic prediction
- Virtual topology design
- Failure management
- Path computation

Physical layer domain

- Quality of Transmission (QoT) estimation
- Modulation format recognition
- Nonlinearity mitigation
- Optical performance monitoring
- Optical amplifier control

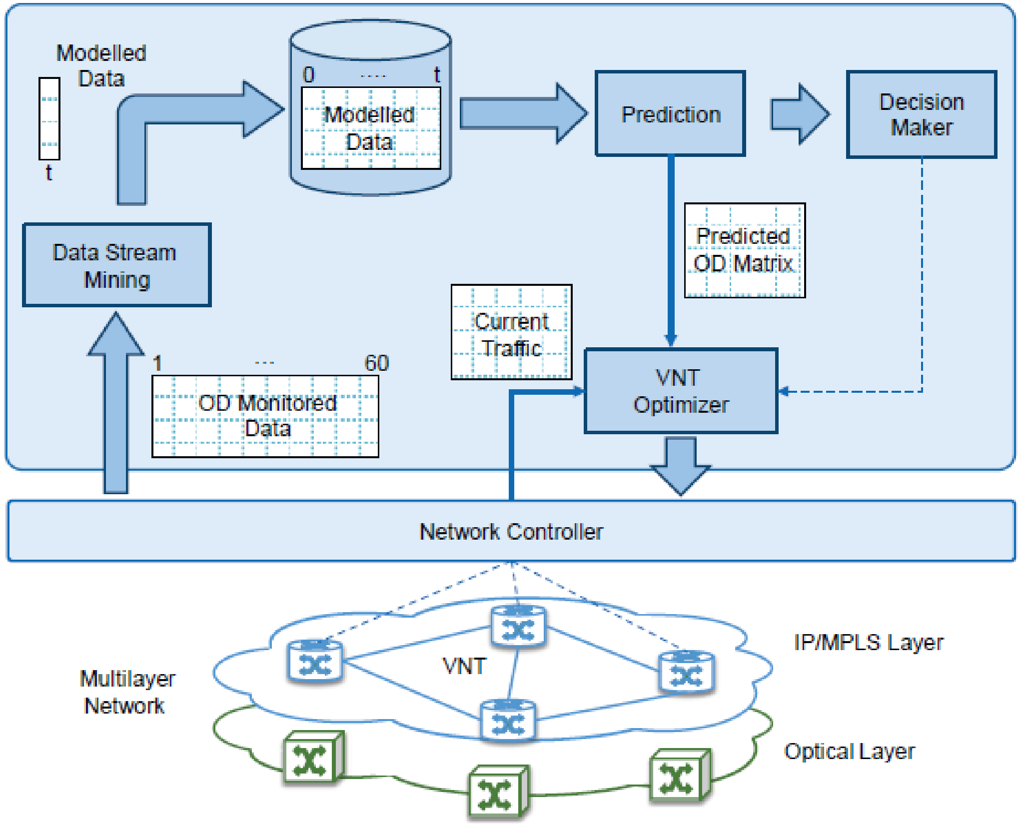


Musumeci, F., Rottondi, C., Nag, A., Macaluso, I., Zibar, D., Ruffini, M., & Tornatore, M. (2018). An overview on application of machine learning techniques in optical networks. *IEEE Communications Surveys & Tutorials*, *21*(2), 1383-1408.

Machine Learning for Optical Networking

Network layer domain

- Traffic prediction
- Virtual topology design



- Input: Historical traffic
- Output: End-to-end traffic
- Algorithm: Neural Network

✓ Design phase: Reduce over-provisioning

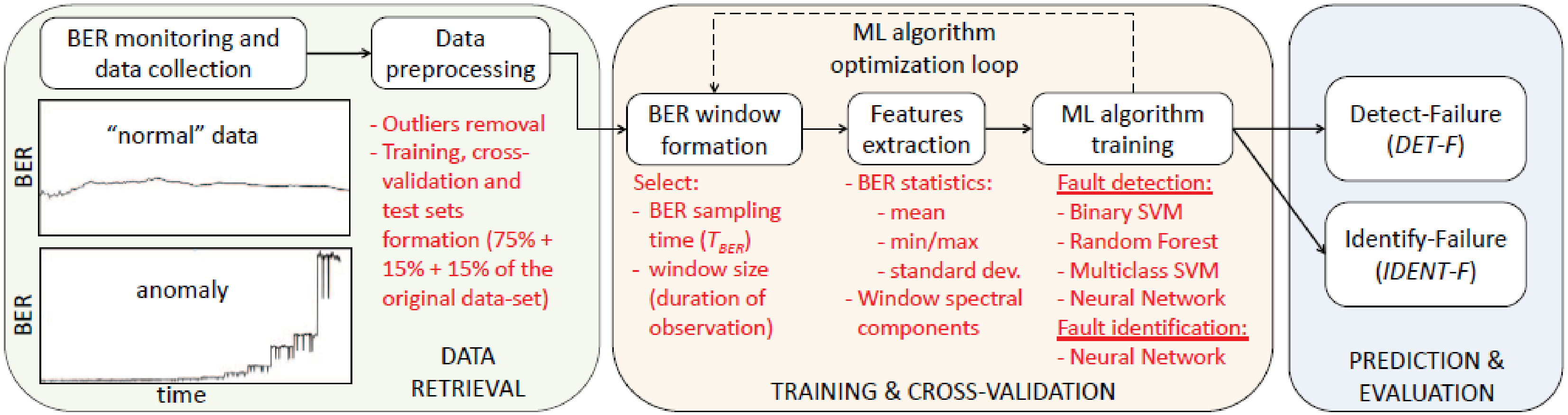
✓ Operation phase: Re-route traffic

F. Morales, M. Ruiz and L. Velasco, "Virtual network topology reconfiguration based on big data analytics for traffic prediction," *2016 Optical Fiber Communications Conference and Exhibition (OFC)*, Anaheim, CA, 2016, pp. 1-3.

Machine Learning for Optical Networking

Network layer domain

- Failure management



S. Shahkarami, F. Musumeci, F. Cugini and M. Tornatore, "Machine-Learning-Based Soft-Failure Detection and Identification in Optical Networks," *2018 Optical Fiber Communications Conference and Exposition (OFC)*, San Diego, CA, 2018, pp. 1-3.

- Input: Bit Error Rate (BER)
- Output: Set of failures
- Algorithm: Support Vector Machine (SVM), Random Forest, Neural Network

Failure:

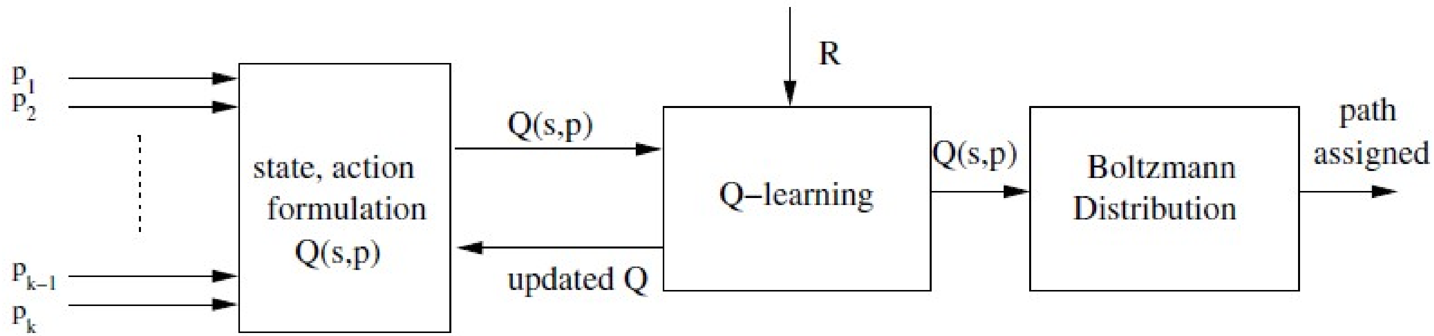
Detection

Identification

Machine Learning for Optical Networking

Network layer domain

- Path computation



Y. V. Kiran, T. Venkatesh and C. S. Ram Murthy, "A Reinforcement Learning Framework for Path Selection and Wavelength Selection in Optical Burst Switched Networks," in *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 9, pp. 18-26, December 2007.

- Input: Traffic requests, candidate paths
- Output: Optimum path for each source-destination path to minimize burst-loss probability
- Algorithm: Q-learning

✔ Select best optical path

Machine Learning for Optical Networking

Physical layer domain

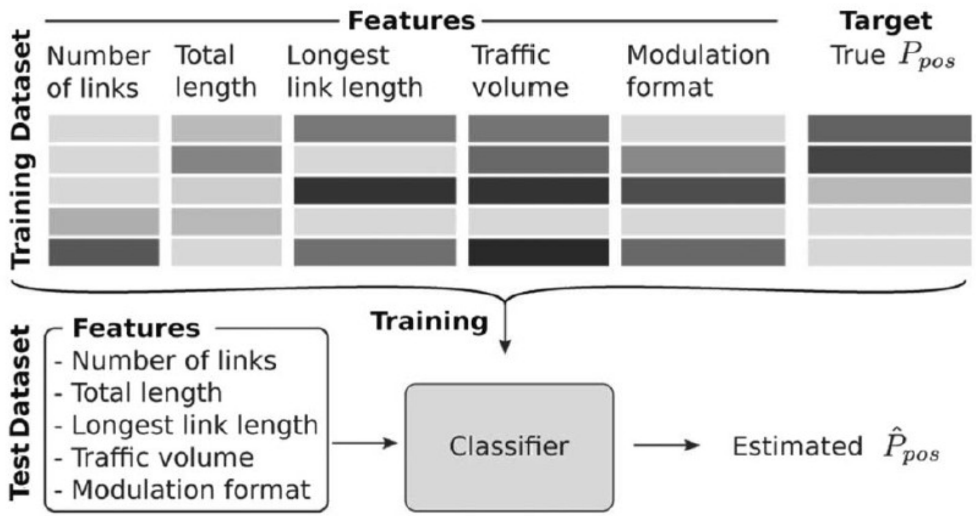
- Quality of Transmission (QoT) estimation

- Analytical models are computationally expensive for real-time requirements and not scalable to large network topologies
- Machine-learning technique using Random Forest Classifier to predict whether the BER of unestablished lightpaths meets a required threshold
- Features: traffic volume to be served, modulation format, total length of light path, length of longest link, number of light path links.
- Training data BER obtained from field by optical performance monitors or by an BER estimation tool (ETool) – input (candidate ligthpath and MF), output (BER, a function of SNR).

Machine Learning for Optical Networking

Physical layer domain

- Quality of Transmission (QoT) estimation



- Input: Lightpath route, length, modulation format, traffic volume
- Output: BER (Bit Error Rate)
- Algorithm: Random Forest

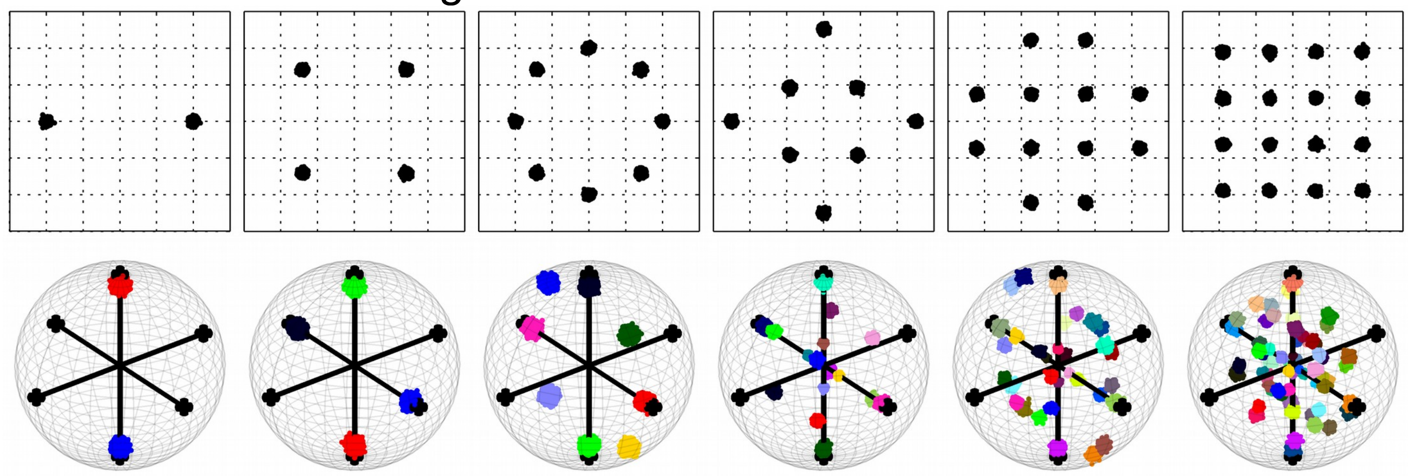
Unestablished lighpaths: Prediction
Established lightpaths: Monitoring

C. Rottondi, L. Barletta, A. Giusti and M. Tornatore, "Machine-learning method for quality of transmission prediction of unestablished lightpaths," in *IEEE/OSA Journal of Optical Communications and Networking*, vol. 10, no. 2, pp. A286-A297, Feb. 2018.

Machine Learning for Optical Networking

Physical layer domain

- Modulation format recognition



R. Borkowski, D. Zibar, A. Caballero, V. Arlunno and I. T. Monroy, "Stokes Space-Based Optical Modulation Format Recognition for Digital Coherent Receivers," in *IEEE Photonics Technology Letters*, vol. 25, no. 21, pp. 2129-2132, Nov.1, 2013.

- Input: Stokes space representation
- Output: Modulation format
- Algorithm: variational Bayesian techniques

✓ Dynamic detection of changes in the modulation format

Machine Learning for Optical Networking

Physical layer domain

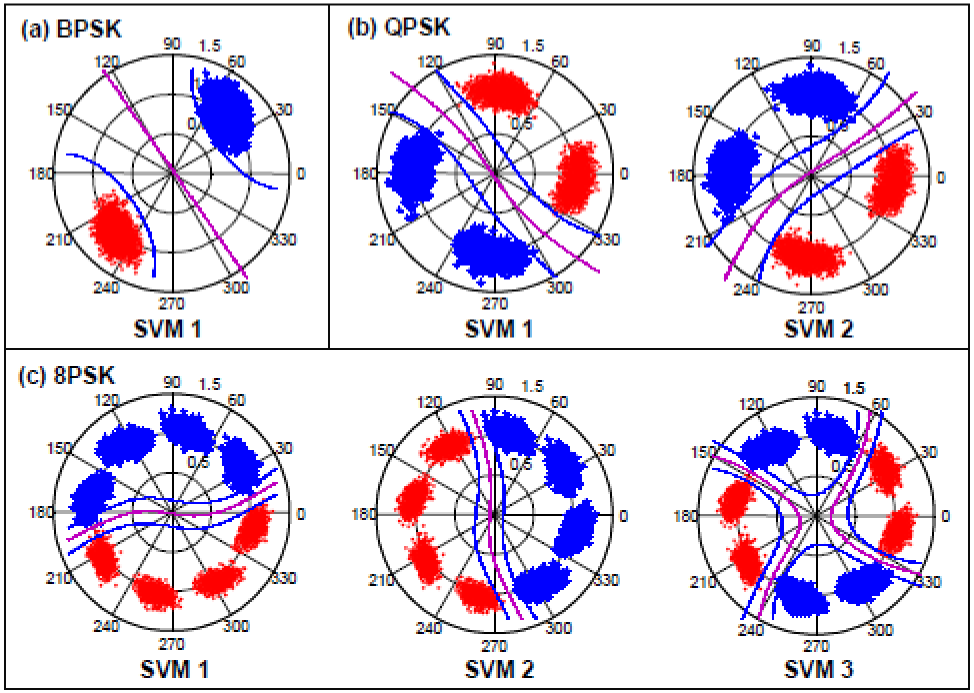
- Nonlinearity mitigation

- Non-linearity due to the Kerr Effect (dependence of fiber refractive index on the optical signal power and channel spacing in multi-channel transmission)
  - Self-phase modulation (SPM)
  - Cross-phase modulation (XPM)
  - Four wave mixing (FWM);
  - Cross-polarization modulation (XPolM)
  - Inelastic processes: Stimulated Brillouin Scattering and Stimulated Raman Scattering (at input powers higher than typical values used in optical communications; can be neglected) [3]
- ML tools for development of adaptive equalizers to deal with nonlinear transmission effects. They can be re-trained, hence suitable for dynamic transmission environment.
- Differs from back-propagation based reception in that signal equalization and demodulation processes are treated jointly as a classification or regression problem by mapping the baseband signal onto a space determined by the training sequence. [5]
- Reduction in computation steps offers potential for real-time application

Machine Learning for Optical Networking

Physical layer domain

- Nonlinearity mitigation



- Input: Received symbols
- Output: BER
- Algorithm: k-nearest neighbours

- ✓ Symbol detection
- ✓ Predistortion

D. Wang *et al.*, "Nonlinear decision boundary created by a machine learning-based classifier to mitigate nonlinear phase noise," *2015 European Conference on Optical Communication (ECOC)*, Valencia, 2015, pp. 1-3.

Nonlinearity Mitigation

K – Nearest Neighbor (KNN) Classifier

- Non-parametric supervised learning
- Large datasets, low dimensions
- Need to investigate a multiclass algorithm that is independent of the link information
- Does not require training process (like ANN)
- Can lead to improvement in launch power dynamic range (2dB) and maximum transmission distance of 240km for a 16QAM coherent optical system

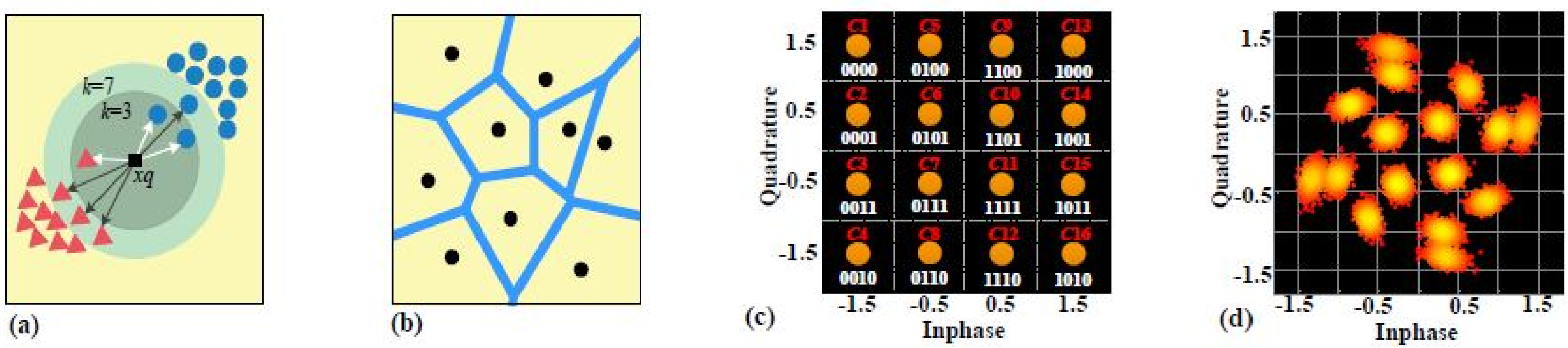
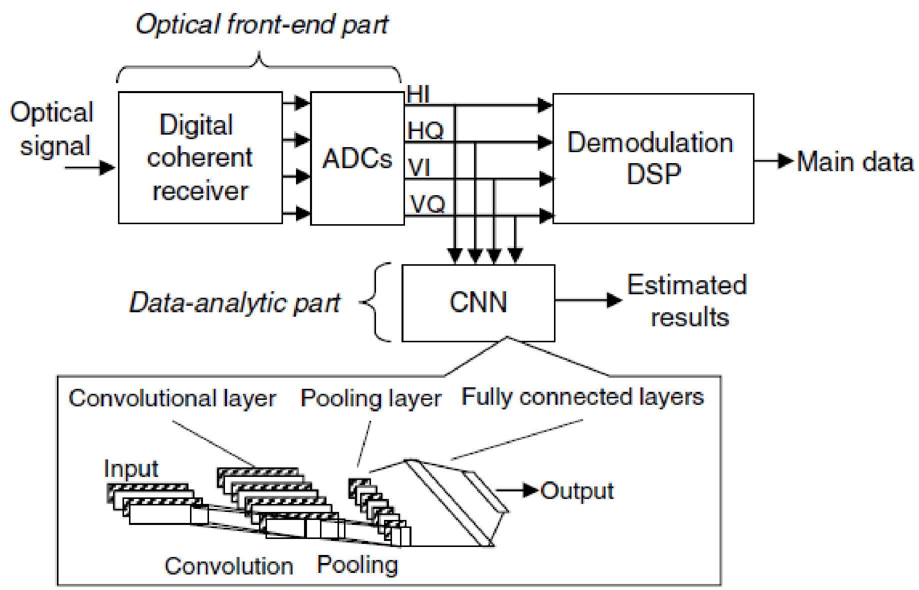


Fig.13 : KNN decision boundaries [7]

Machine Learning for Optical Networking

Physical layer domain

- Optical performance monitoring



T. Tanimura, T. Hoshida, T. Kato, S. Watanabe and H. Morikawa, "Convolutional neural network-based optical performance monitoring for optical transport networks," in *IEEE/OSA Journal of Optical Communications and Networking*, vol. 11, no. 1, pp. A52-A59, Jan. 2019

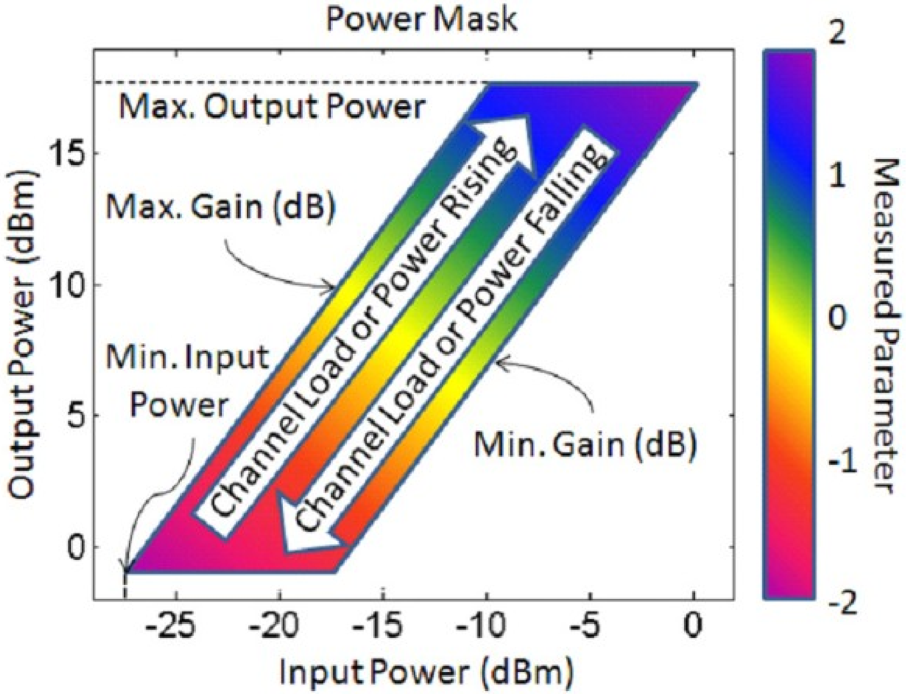
- Input: Received samples
- Output: OSNR
- Algorithm: Convolutional Neural Network

Different parameters can be estimated. Estimations can be used for: re-route lightpaths, adapt modulation format…

Machine Learning for Optical Networking

Physical layer domain

- Optical amplifier control



- Input: EDFA input and output power
- Output: EDFA operating point
- Algorithm: Neural Network, Ridge Regression

✓ Parameter estimation
✓ Spectrum allocation suggestion

Bastos-Filho, Carmelo & Barboza, Erick & Martins Filho, Joaquim & Moura, Uiara & De Oliveira, Juliano. (2013). Mapping EDFA Noise Figure and Gain Flatness Over the Power Mask Using Neural Networks. Journal of Microwaves and Optoelectronics. 12. 128.

Optical Amplifier Control in Optical Networks using Machine Learning

Power Excursion in Optical Networks

- Power excursions occur during the transmission because the wavelength dependent gain and noise figure spectrum of EDFAs can cause wavelength dependent optical power output after the amplification process [5].

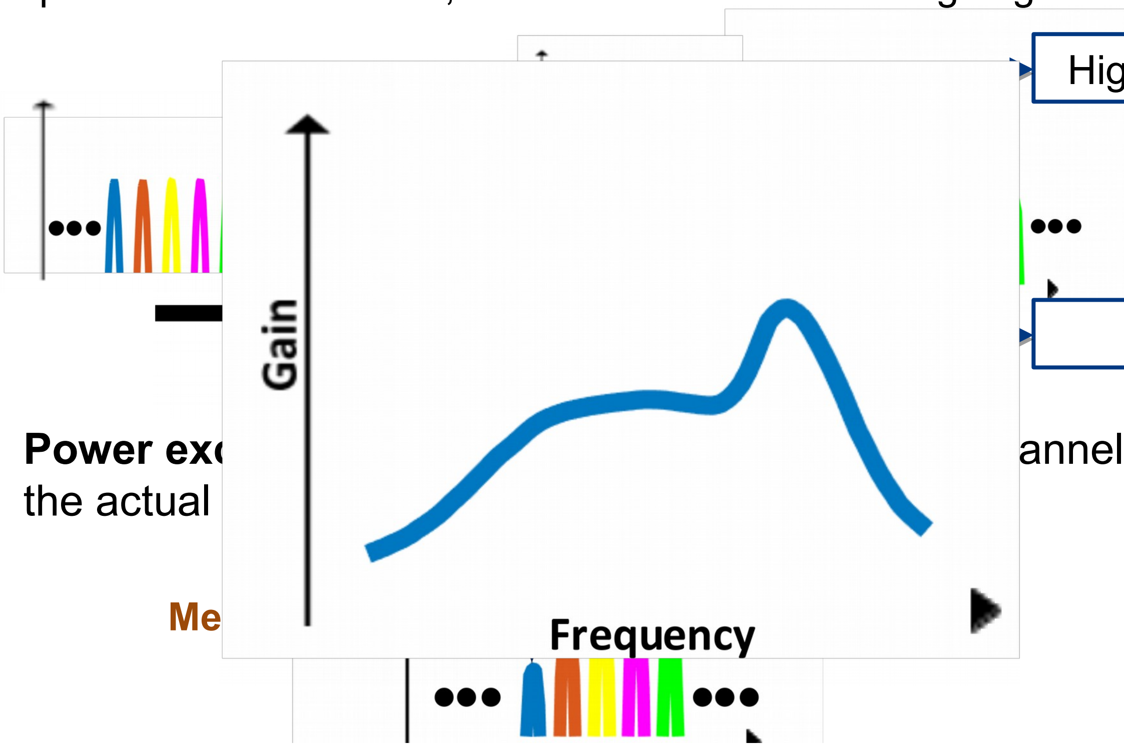
- In metro optical networks, OADMs permit to drop and/or to add WDM channels. If a WDM channel is added with an undesired high gain, Automatic Gain Control module reduces the gain of all channels to maintain constant the global mean gain. High gain channels steal power from low gain channels and the disparity among channel powers increases. Power excursions can reach up to 2 dB through occasional channel additions after 3-cascade EDFAs [6].

5. Ahsan A S et al., Excursion-free dynamic wavelength switching in amplified optical networks, JOCN 2015, 7
6. Lin P J., Reducing optical power variation in amplified optical network, Proceedings of ICCT, 2003.

TELECOM SudParis

Power Excursion in Optical Networks

- **Gain and Noise Figure spectrum of EDFAs are wavelength dependent**
- Remarkable impact in metro networks, with WDM channels undergoing different power excursions



High gain (nonlinearities)

Low gain (noise)

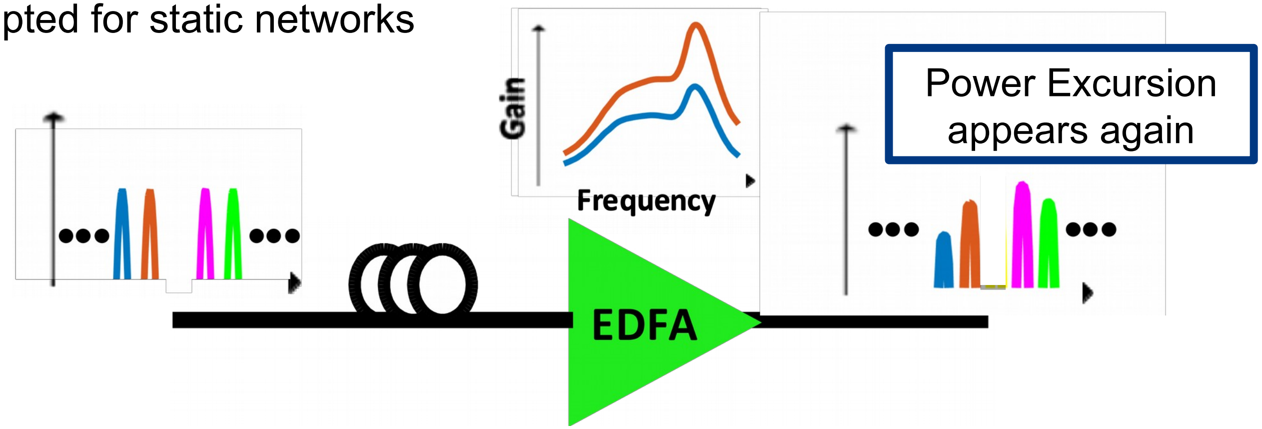
- **Power exc** annel power and the actual

Me

TELECOM SudParis

Power Excursion in Optical Networks

- Output power level of EDFAs can be controlled via **Automatic Gain Control (AGC)** which maintains constant the global mean gain
- **Gain flattening techniques to overcome** disparity among channel powers during add/drop process
  - Well-adapted for static networks



To summarize

- Dynamic gain equalization filters allow operation of EDFAs with constant ACG but does not solve optical power excursion problem
  - Mean global optical gain at each EDFA is maintained constant whatever the number of established wavelengths and the gain of each individual channel is wavelength dependent.

- The wavelength dependence of EDFA channel gain can increase the optical power excursion even stronger with Flex-Grid technology where connections occupy different optical bandwidths depending on their symbol rates.
  - Optical Power Excursion varies from one connection to another one.

- To solve power excursion problem, analytical methods have been implemented based on the prior knowledge of one specific network. What happens in dynamically changing networks?
  - Machine Learning techniques seem well-adapted to deal with the flexibility of the future optical networks [7]

7. Huang Y S et al. A machine learning approach for dynamic optical channel add/drop strategies that minimize EDFA power excursions, Proceedings of ECOC, 2016.

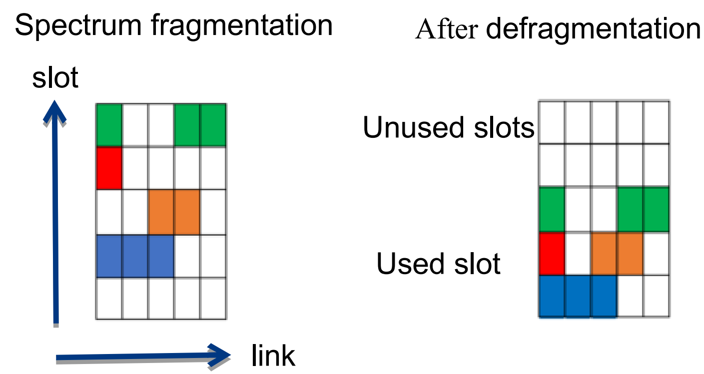
TELECOM SudParis
Institut Mines-Télécom

Power Excursion Mitigation in Optical Networks

Fragmentation and Power Excursion

Fragmentation in Flexible Optical Networks [9]

- Fragmentation of the spectrum due partitioning in frequency slots
- Free fall polynomial time algorithm defragmentation moves connections or demands in order to reduce the used fiber capacity
- This process may be used with Push-Pull technique that requires free intermediate slots before moving connections



Example of 5-slot and 5-link before and after free-fall defragmentation

9. D. Amar et al., Power Excursion Reduction in Flex-Grid Optical Networks with Symbol Rate Adaptation, ACP 2017

Power Excursion Mitigation for Flex-grid Defragmentation

- Objectives
  - Move connections in order to decrease the used fiber capacity
  - Connection symbol rate optimization to reduce power excursion before and after defragmentation process

1- Symbol rate adaptation before spectrum defragmentation

- Impact of connection dynamicity on power excursion during following defragmentation process?

3- Symbol rate adaptation after spectrum defragmentation

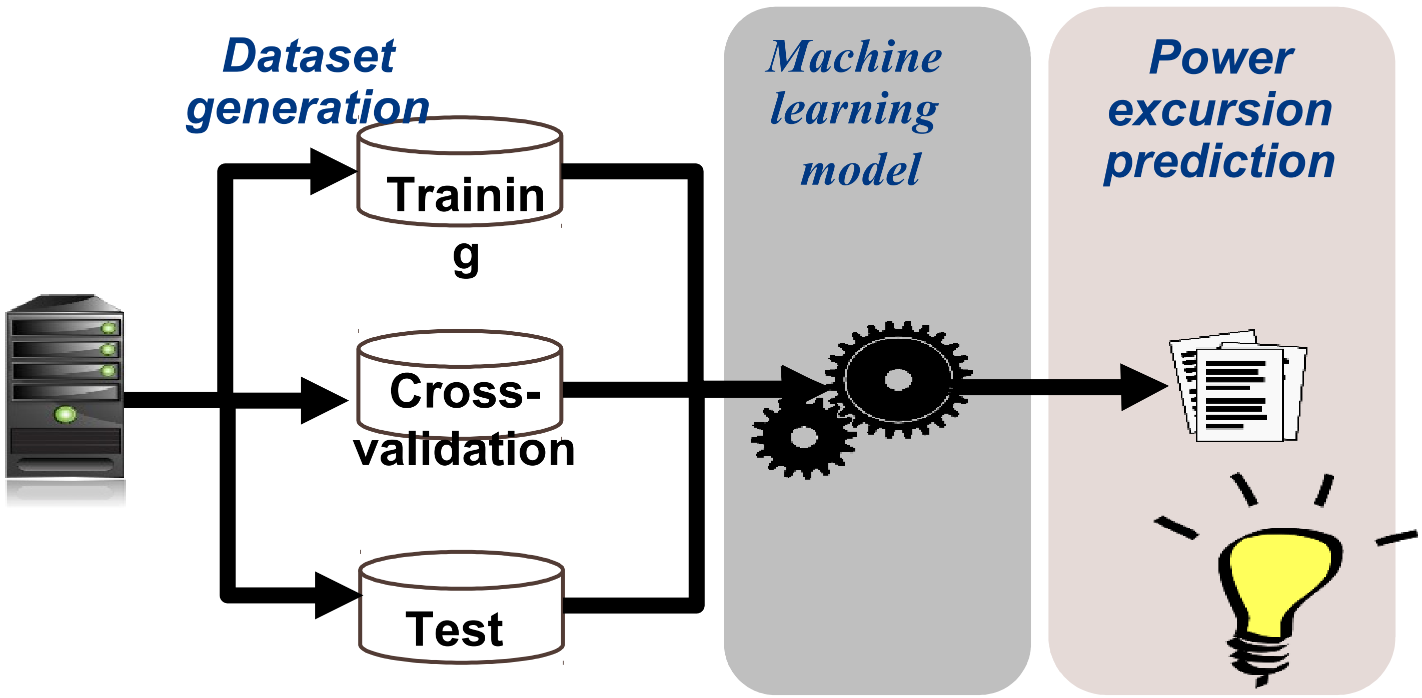
- Symbol rate reconfiguration could be even performed during spectrum defragmentation as most defragmentation processes are not hitless (connection is lost)

TELECOM SudParis
Institut Mines-Télécom

Optical Amplifier Control using Machine Learning

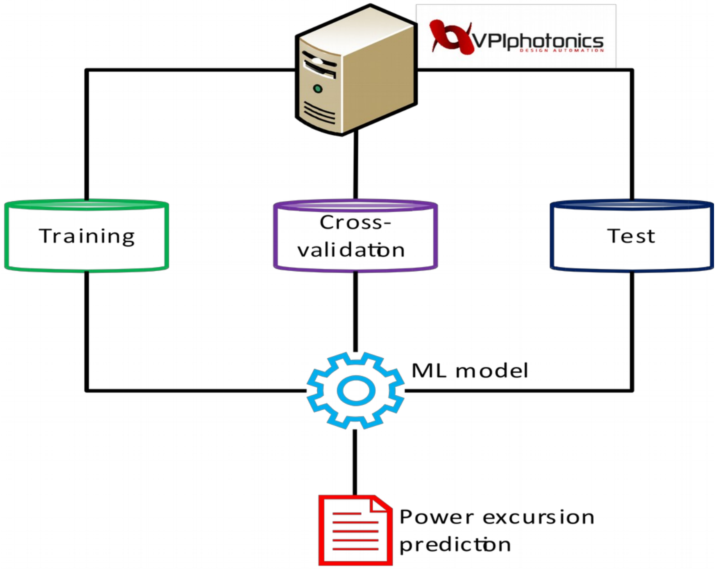
Power Excursion Prediction

Optical Power Excursion Prediction based on Neural Networks



Optical Power Excursion Prediction based on Neural Networks [12]

Power Excursion prediction: output optical power post-compensation or input optical power pre-distortion

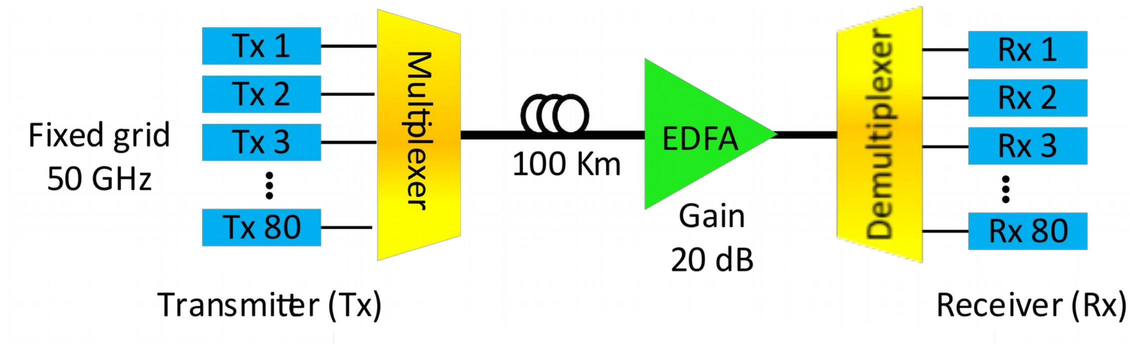


Overview:

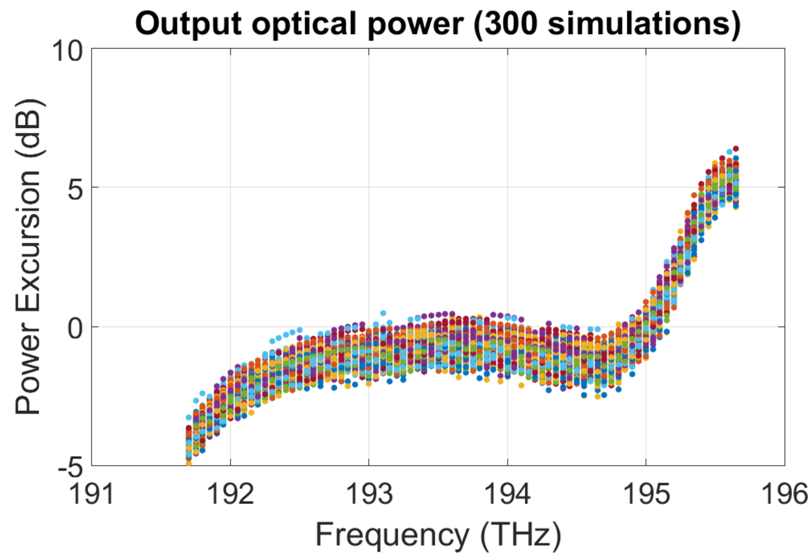
- Training, cross-validation and test dataset using VPI Photonics, an advanced software for simulation of optical communication systems
- Training and cross-validation of the Machine learning module
- Test of the prediction of the optical power excursion

12. M.Freire et al., Predicting Optical Power Excursions in Erbium Doped Fiber Amplifiers using Neural Networks, Proceedings ACP2018

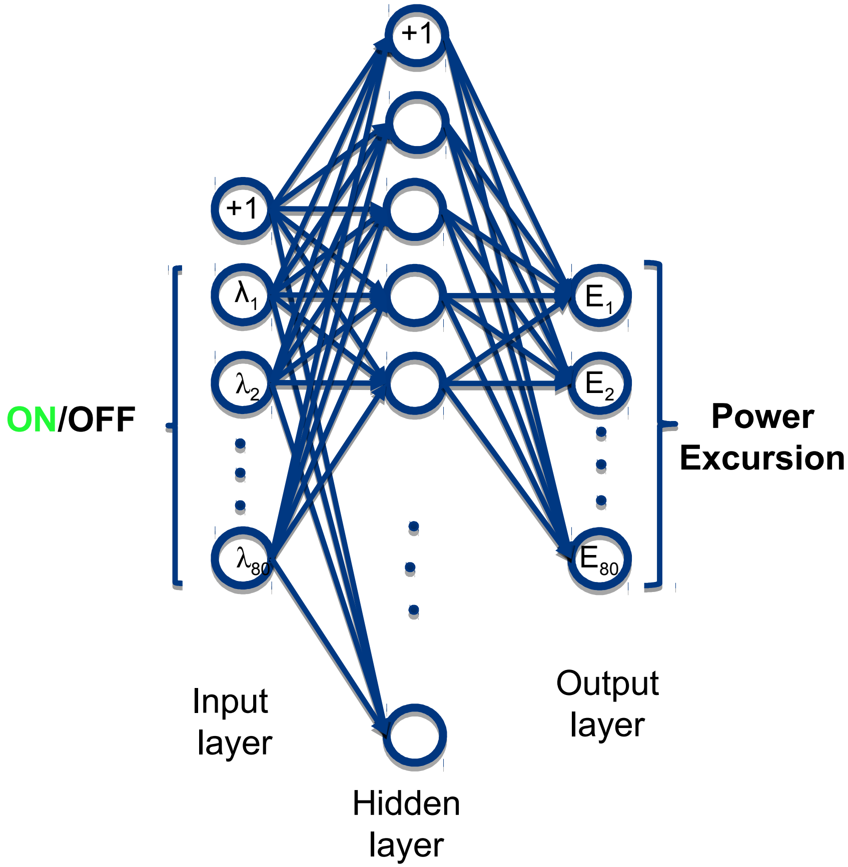
Dataset generation



- Dataset (300 simulations)
  - Channel load from 40% to 87.5%
  - Data stored in every simulation
    - Input feature vector for each channel (active or not)
    - Output vector representing the power excursion for each channel

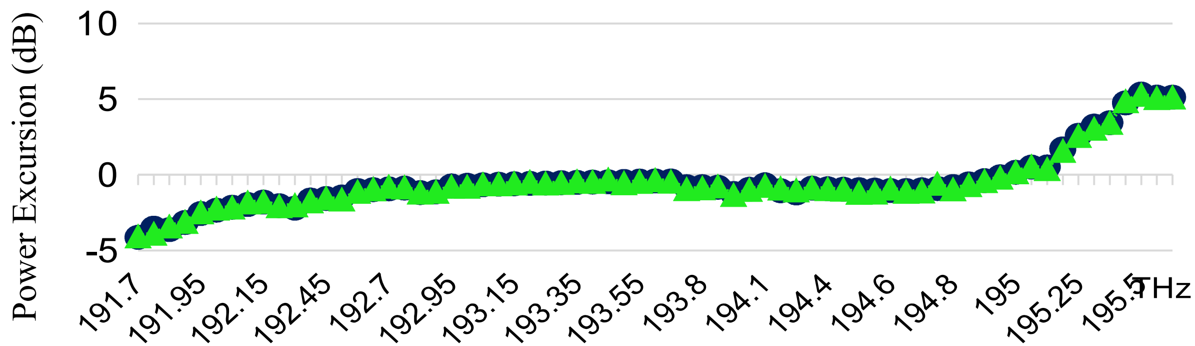


Machine Learning Module

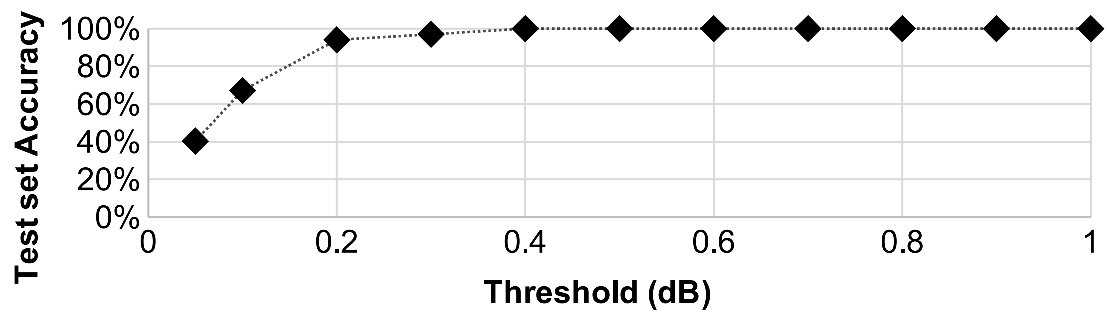


- Machine learning model based on Neural Networks, developed in Keras, to predict power excursion
- Structure: 2 layer Neural Network (160 neurons in the hidden layer)
  - Hidden layer: ReLU (Rectified Linear Unit) activation
  - Output layer: linear activation
- Optimization: Stochastic Gradient Descent, batch size 10 % of the training dataset

Test: Simulated vs predicted power excursion values using machine learning model



- Test set accuracy vs minimum acceptable difference between the simulated values and the predicted values
- 90% accuracy for 0.2 dB threshold



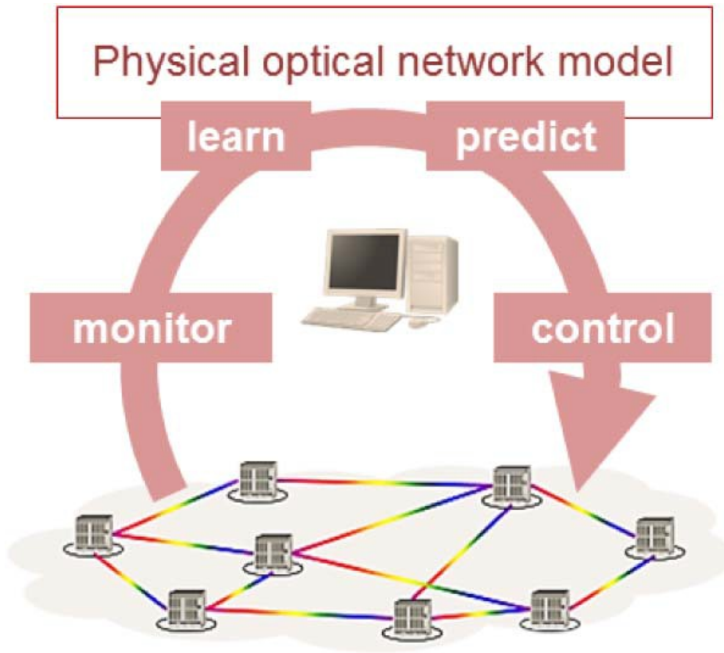
TELECOM SudParis

Institut Mines-Télécom

Optical Amplifier Control using Machine Learning

Power Excursion Precompensation

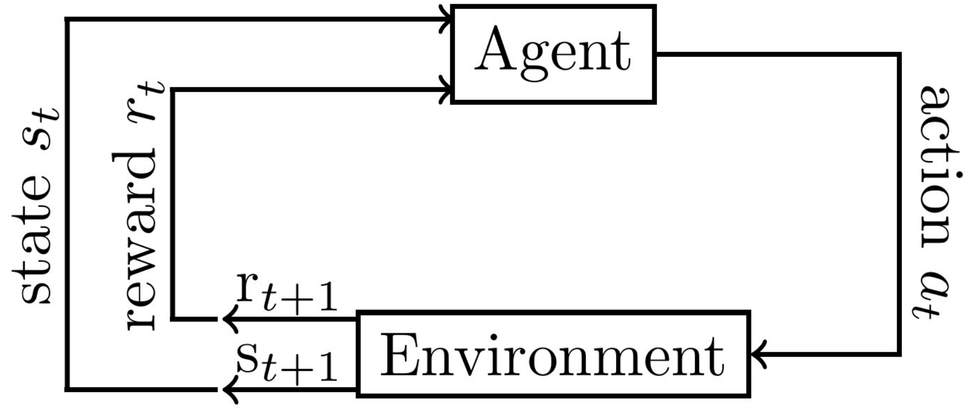
Physical Optical Network Model



13. Bouda et al., Accurate Prediction of Quality of Transmission Based on a Dynamically Configurable Optical Impairment Model, JOCN vol10, n°1, 2018

Power Predistorsion using Reinforcement Learning [14]

- Reinforcement Learning approach



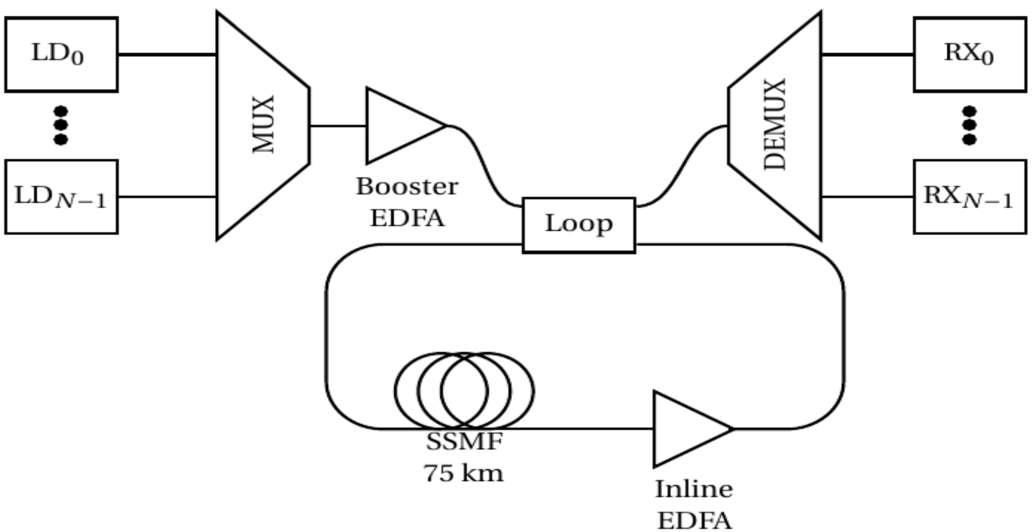
- Environment = Optical link
- Agent = Predistortion Module
- State = Power excursion per channel at the output of the link
- Action = Input power per channel increase (decrease) by 0.1 dB

14. M. Freire et al., Dynamic Power Predistortion Implementation with Reinforcement Learning for Excursion-Free Amplified Optical Systems, ONDM2020

Environment

- Environment based on simulations done with VPItransmissionMakerTM for a cascade of amplifiers (10 spans) experimentally characterized

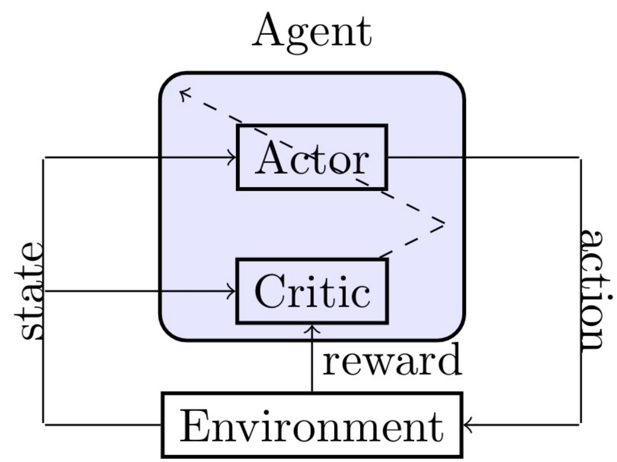
.



- Model based on Neural Networks predicting power deviations.
- Environment's model together with reward function to complete the environment.

Agent [15]

- Actor-critic method:
  - Actor decides which action has to be taken depending on the current state.
  - Critic evaluates the actions of the actor based on the rewards.

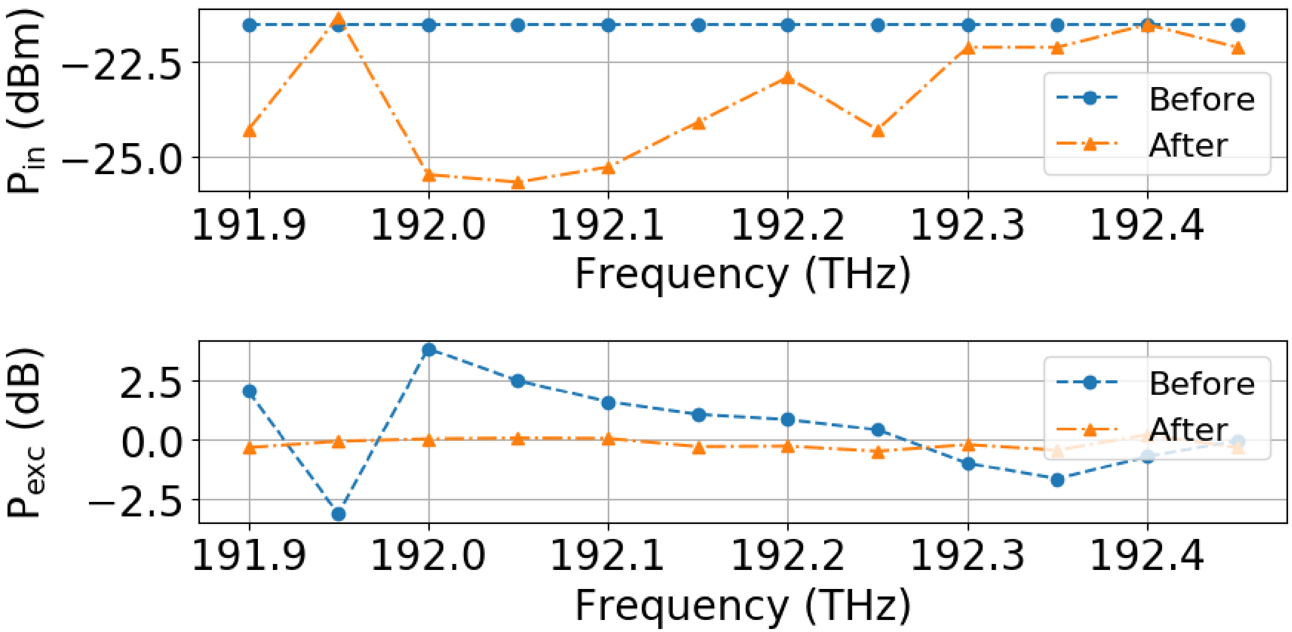


15- Sutton and Barto, Reinforcement learning: An introduction.

TELECOM SudParis

Evaluation

- After training the Reinforcement Learning algorithm, for a 12-channel scenario with 50 GH spacing, we observe that the power predistortion module is able to reduce the power excursion.



- In future optical networks, flexible use of spectrum resources leads to the fragmentation of the spectrum. Due to the dynamicity of these networks, the power excursion at the output of the EDFAs becomes a time varying impairment.

- Machine learning is a very efficient method to predict and to control power excursion at the output of the EDFAs.

- Machine technique permits also to pre-adjust input optical power of the system based on different defragmentation algorithms[16].

In main conclusion, as machine learning is able to solve complex problems efficiently without any ideal assumptions, it will play an important role to design the power optimization strategy in complex flexible heterogeneous optical networks.

16. Huang Y S et al. Power excursion mitigation for flexgrid defragmentation with machine learning, JOCN 2018

Some References

D. Wang, M. Zhang, J. Li, Z. Li, J. Li, C. Song, X. Chen, "Intelligent Constellation Diagram Analyzer Using Convolution Neural Network-based Deep Learning." Optics Express, vol. 25, No. 15. July, 2017.

F. Musumeci, C. Rottondi, A. Nag, I. Macluso, D. Zibar, M. Ruffini, M. Tornatore, "An Overview on Application of Machine Learning Techniques in Optical Networks. IEEE Communications Surveys and Tutorials, vol 21. Issue 2. 2019.

A. Amari, O. Octavia, R. Venkatesan, O. S. Kumar, P. Ciblat, Y. Jaouen, "A Survey on fiber nonlinearity compensation for 400 Gbps and beyond optical communication systems." IEEE Communications Surveys & Tutorials PP(99):1-1. June 2017

L. Barletta, A. Giusti, C. Rottondi, M. Tornatore. "QoT estimation for Unestablished Lightpaths using Machine Learning." Optical Fiber Comminication Conference, OSA Technical Digest. 2017.

O. Sidelnikov, A. Redyuk, S. Sygletos. "Equalization performance and complexity analysis of dynamic deep neural networks in long haul transmission systems." Optics Express, 2018

B. Karanov, M. Chagnon, F. Thouin, T. A. Eriksoon, H. Bulow, D. Lavery, P. Bayvel, L. Schmalen. "End-to-End Deep Learning of Optical Fiber Communications." Journal of Lightwave Technology, vol. 36, No. 20, October 15, 2018

D. Wang, M. Zhang, M. Fu, Z. Cai, Z. Li, Y. Cui, B. Luo. "KNN-based Detector for Coherent Optical System in Presence of Nonlinear Phase Noise. OECC/PS 2016

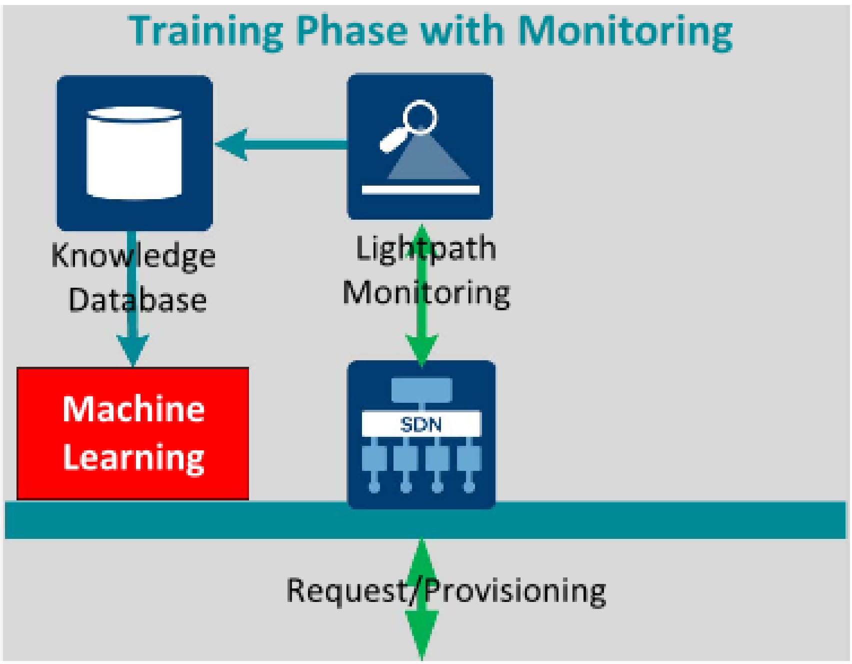
M. Peikari, S. Salama, S. Nofech-Mozes. A. Martel. "A Cluster-then-label Semi-Supervised Learning Approach for Pathology Image Classifiation." Nature, Scientific Reports 8(1), May, 2018.

S. Mogla. "Telecommunication Network Survivability for Improved Reliability in Smart Power Grids. Graduate Theses and Dissertations, 2014. [online] https://scholarcommons.usf.edu/etd/5380

E. Hyytia. "Two Sample applications of MDP and FPI. [Online] https://notendur.hi.is/esa/mdp.html Retrieved Jan 15, 2020.

Other use case

To have a robust and reliable machine learning (ML) model, data from the network itself is mandatory.



1) Request of lightpath
2) Establishment of lightpath
3) Monitoring system addresses the Quality of Transmission (QoT)
4) Monitored measured saved in database and fed the ML algorithm

Scenario and knowledge acquisition

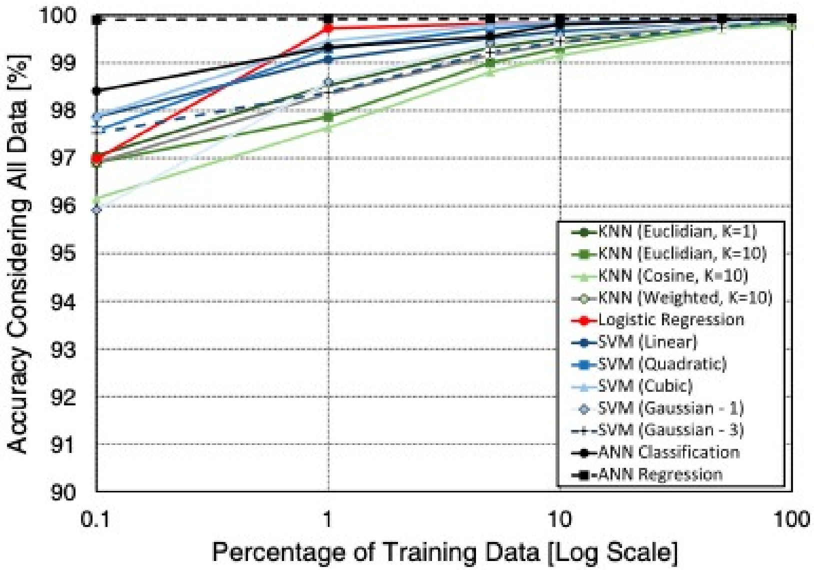
Reference topologies and characteristics:

Network	Fiber Types	Nodes	Links	Av. Link Length (km)	Number of Paths
CORONET	SSMF	75	82	396	105,320
TIM	SSMF LEAF	44	71	174.3	1,134,272

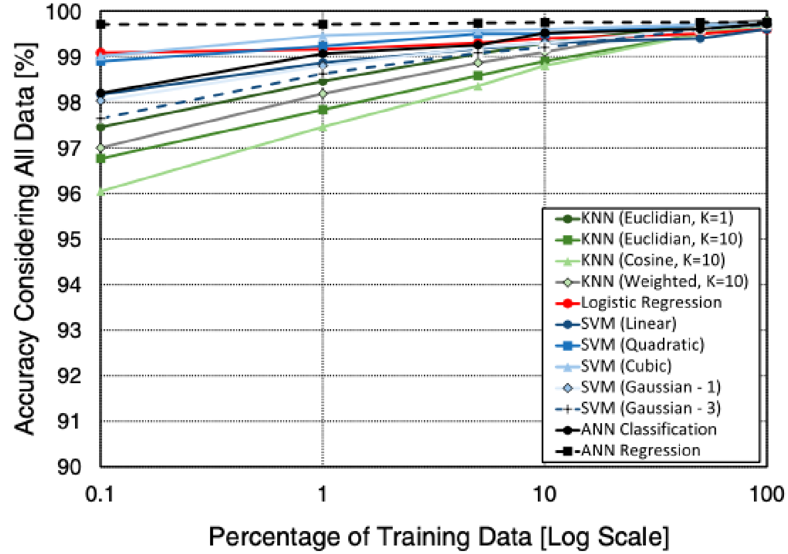
TELECOM SudParis

Comparison and results

CORONET



TIM



- All evaluated models presented an accuracy higher than 90%, the minimum is always higher than 96%
- ANN's tends to have higher accuracy
- KNN worst accuracy
- SVM tends to fall between ANN and KNN
- Logistic regression presents a good trade off between accuracy and training complexity

TELECOM SudParis

Conclusion

- Machine Learning is a valid alternative to estimate the QoT of a lightpath.

- It is noteworthy that ANN proved to be the model achieving the best generalization, with accuracies in the order of 99%.

- Logistic regression is the fastest model among all, showing higher accuracy than KNN and even SVM in some cases.