An Introduction to Machine Learning Applications in Optical Transport Networks

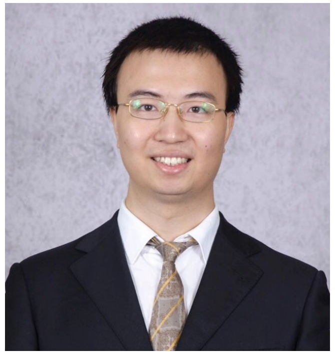
Presented by:

OSA
Optical Communications Technical Group

The OSA Optical Communications Technical Group Welcomes You!

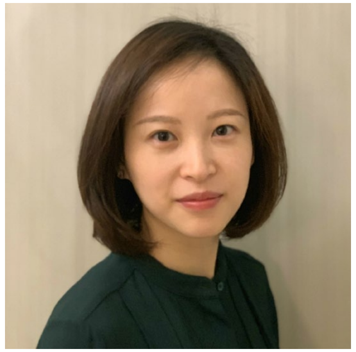
Technical Group Leadership 2019

Chair
Qunbi Zhuge
Shanghai Jiao Tong U., China



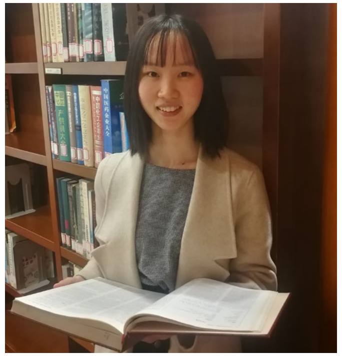
Vice-Chair
Yi Weng
Neophotonics, USA

Webinar Officer
Cosimo Lacava
U. of Southampton, UK



Events Officer
Yanni Ou
Nokia Bell Labs, Germany

Events Officer
Rui Zhang
Georgia Tech., USA



Social Media Officer
Meng Cai
Shanghai Jiao Tong U., China

Technical Group at a Glance

- Focus
  - Our group oversees the development in the field of optical communications. It deals with optical transmission aspects from chip-to-chip, ultra-short haul to long haul range. It deals with optical networking aspects, coding and decoding of information onto photons, optical signal processing and other transmission related aspects.
  - **>3500** members.

- Mission
  - To benefit *YOU*.
  - Webinars, e-Presence, publications, technical events, business events, outreach.
  - Interested in presenting your research? Have ideas for TG events? Contact us at qunbi.zhuge@sjtu.edu.cn.

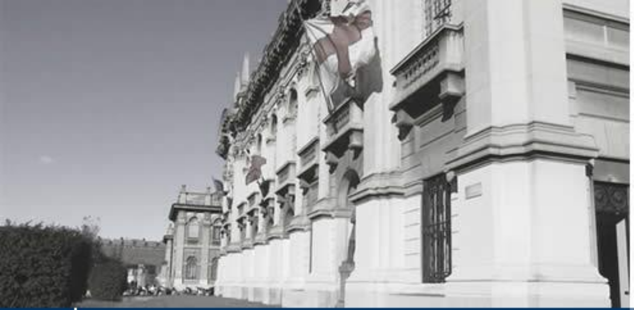
Today's Webinar

An Introduction to Machine Learning Applications in Optical Transport Networks

Prof. Massimo Tornatore

Politecnico di Milano, Italy
massimo.tornatore@polimi.it

Speaker's Short Bio:
Prof. Tornatore is currently an Associate Professor at Politecnico di Milano, Italy. His research interests include performance evaluation and design of communication networks, and machine learning application for network management. In these areas, he co-authored more than 300 peer-reviewed conference and journal papers (with 13 best paper awards). He is an active member of the Editorial Board of, among others, IEEE Communication Surveys and Tutorials, IEEE Communication Letters.



OSA Webinar
November, 20th 2019

POLITECNICO DI MILANO

An Introduction to Machine Learning Applications in Optical Transport Networks

Massimo Tornatore

Politecnico di Milano, Italy & University of California, Davis

@MaxTornatore

Covered topics

- The presentation is organized into two main parts
- Part 1: overview on Machine Learning
  - Basic concepts (supervised/unsupervised learning, neural networks, etc.)
  - Some algorithms
    - Linear regression
    - Neural Networks
    - K-nearest neighbours

Note: this is NOT a "pure" Machine Learning tutorial

The objective is to show **how we applied** ML to **our** research problems

- Part 2: applications of ML to optical-network problems
  - Part 2a): QoT estimation and RSA
  - Part 2b): Failure management
  - Part 2c): Other application at physical and network layer
    - Traffic prediction, virtual topology design,…

What is Machine Learning?

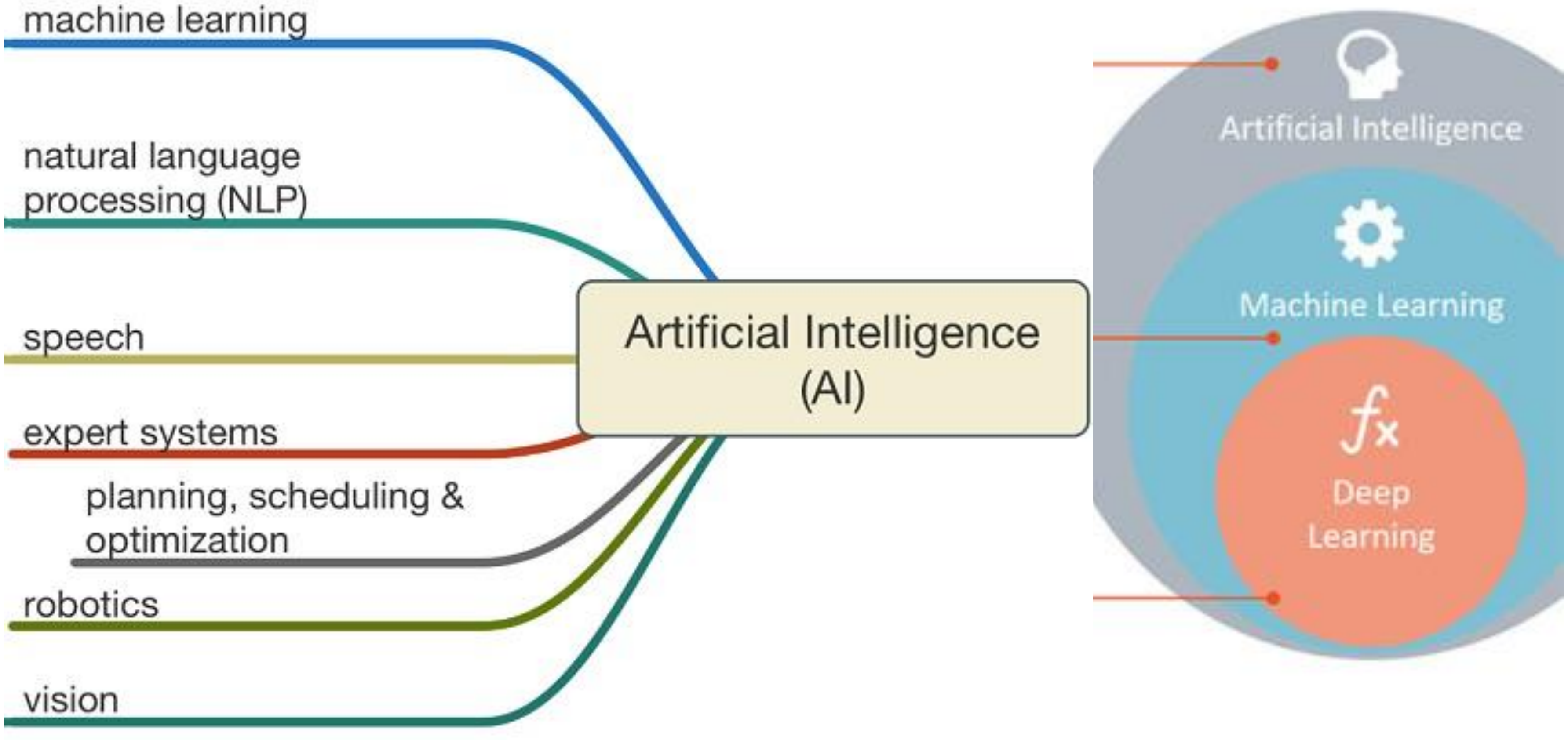
- *"Field of study that gives computers the ability to learn without being explicitly programmed" (A. Samuel, 1959)*
- *"Teaching a computer to automatically learn concepts through data observation"*
- *…*

- For our purposes: An math/statistical instrument to make decisions by inferring statistical properties of monitored data

  …in the context of optical networks

- Sometimes confused with other terms: AI, Deep Learning, Data Analytics, Data Mining, etc.

Many definitions with blurred borders



Why only now in optical networks?

- Dominating complexity
  - Coherent Trasmission /Elastic Networks
    - Several system parameters: channel bandwidth, modulation formats, coding rates, symbol rates..
- New enablers @ *Mngt&Cntr* plane
  - Software Defined Networking
  - Edge computing
  - OPM’s (some of them are for free.. as in coherent receivers..)
- Lack of <u>skilled</u> workforce
  - NTT warning *(OFC 2017)*: aging population, increasing competition for young STEM workforce

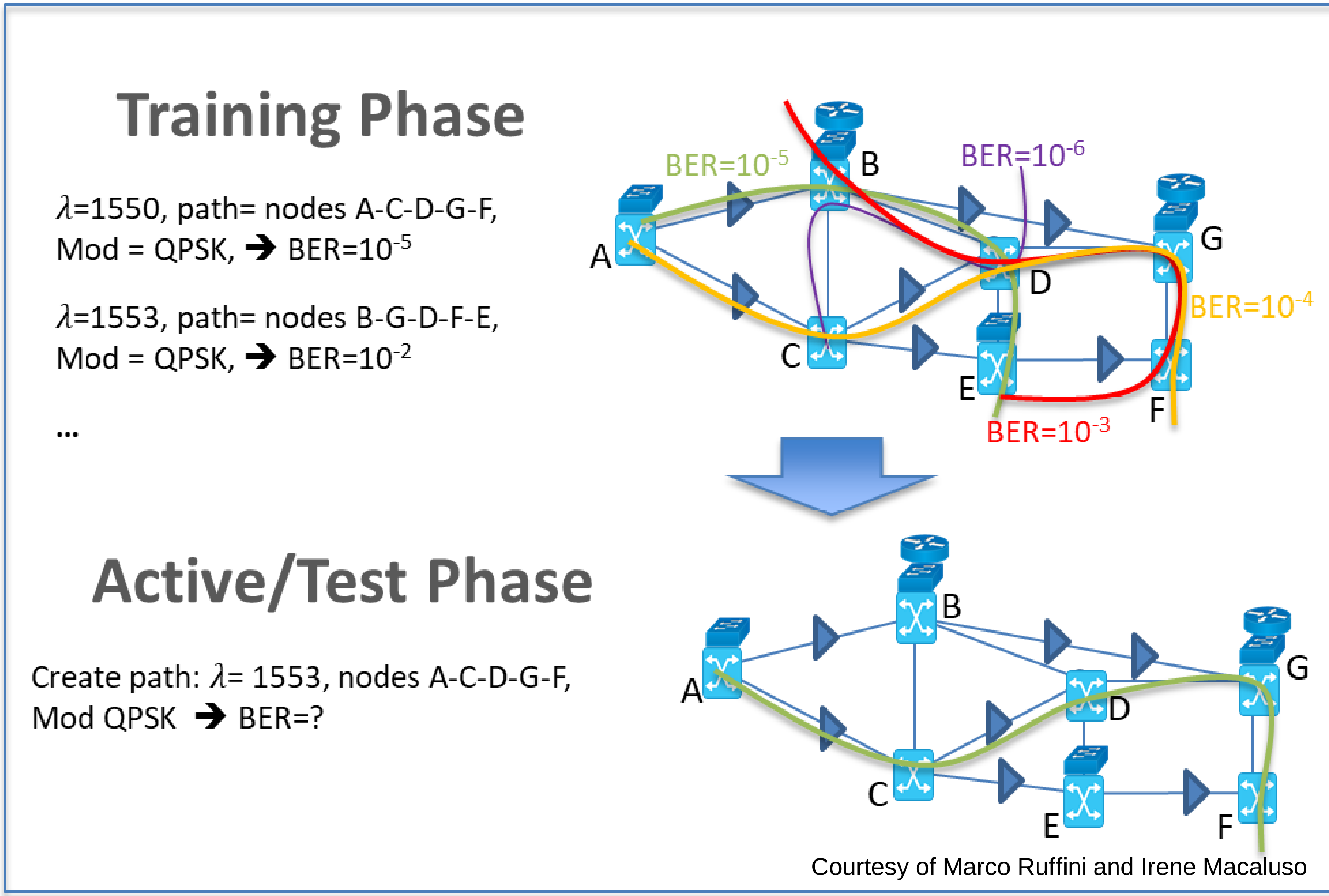
Main categories of ML algorithms (1)

- Supervised-learning algorithms
    - We are given “labeled” data (i.e., “ground truth”)
    - Main objective: given a set of “historical” input(s) predict an output
        - Regression: output value is continuous
        - Classification: output value is discrete or “categorical”

- An example: Traffic forecasts
    - Given traffic during last week/month/year
        - Predict traffic for the next period (regression)
        - Predict if available resources will be sufficient (classification)

- Other examples
    - Speech/image recognition
    - Spam classifier
    - House prices prediction/estimation

Supervised learning: an «optical» example



Supervised Learning: the algorithm is trained on dataset that consists of paths, wavelengths, modulation, and the corresponding BER. Then it extrapolates the BER in correspondence to new inputs.

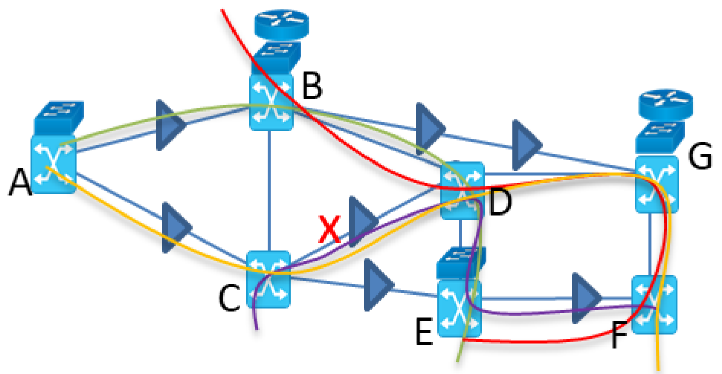
Main categories of ML algorithms (2)

- Unsupervised-learning algorithms
  - Available data is not “labeled”
  - Main objective: derive structures (patterns) from available data
    - Clustering finding “groups” of similar data
    - Anomaly detection

- An example: cell-traffic classification
  - Given traffic traces
  - understand if some cells provide similar patterns
    - Residential, business, close to theatre, cinema, stadium…
    - This information can be used to make network resources planning

- Other example
  - Group people according to their interests to improve advertisement

Unsupervised learning: some examples



Data:

λ=1550, path= nodes A-B-D-E, Mod = QPSK, BER=10^{-6}

λ=1545, path= nodes B-D-G-F-E, Mod = 16-QAM, BER=10^{-7}

λ=1553, path= nodes A-C-D-G-F, Mod = BPSK, BER=10^{-2}

λ=1544, path= nodes C-D-E-F, Mod = DPQPSK, BER=10^{-2}

Anomaly detection for link C-D

Courtesy of Marco Ruffini and Irene Macaluso

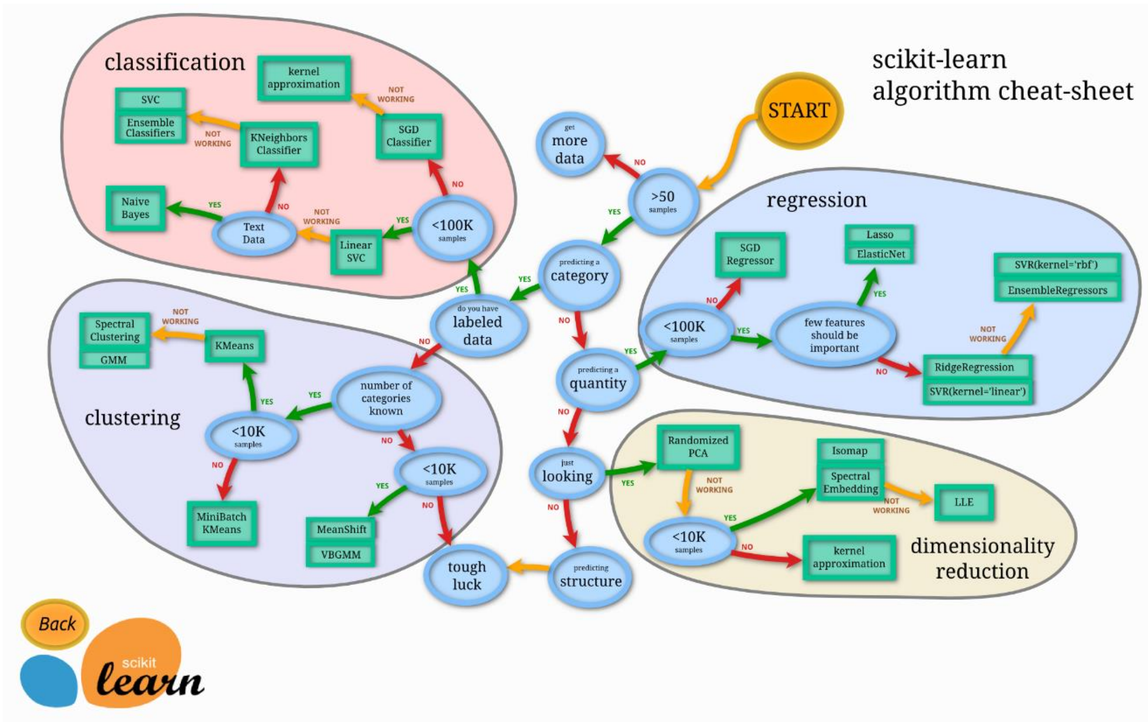
Unsupervised Learning: the algorithm identifies unusual patterns in the data, consisting of wavelengths, paths, BER, and modulation..

Main categories of ML algorithms
Cont'd

- Semi-Supervised learning
  - Hybrid of previous two categories
  - Main objective: most of the training samples are unlabeled, only few are labeled
    - Common when labeled data are scarce or expensive
  - Self-training: start with labeled data, then label unlabeled data based on first phase
- Reinforcement learning
  - Available data is not "labeled"
  - Main objective: learn a policy, i.e., a mapping between in inputs/states and actions. Behavior is refined through rewards
  - Methodologically similar to «optimal control theory» or «dynamic programming»
  - Q-learning

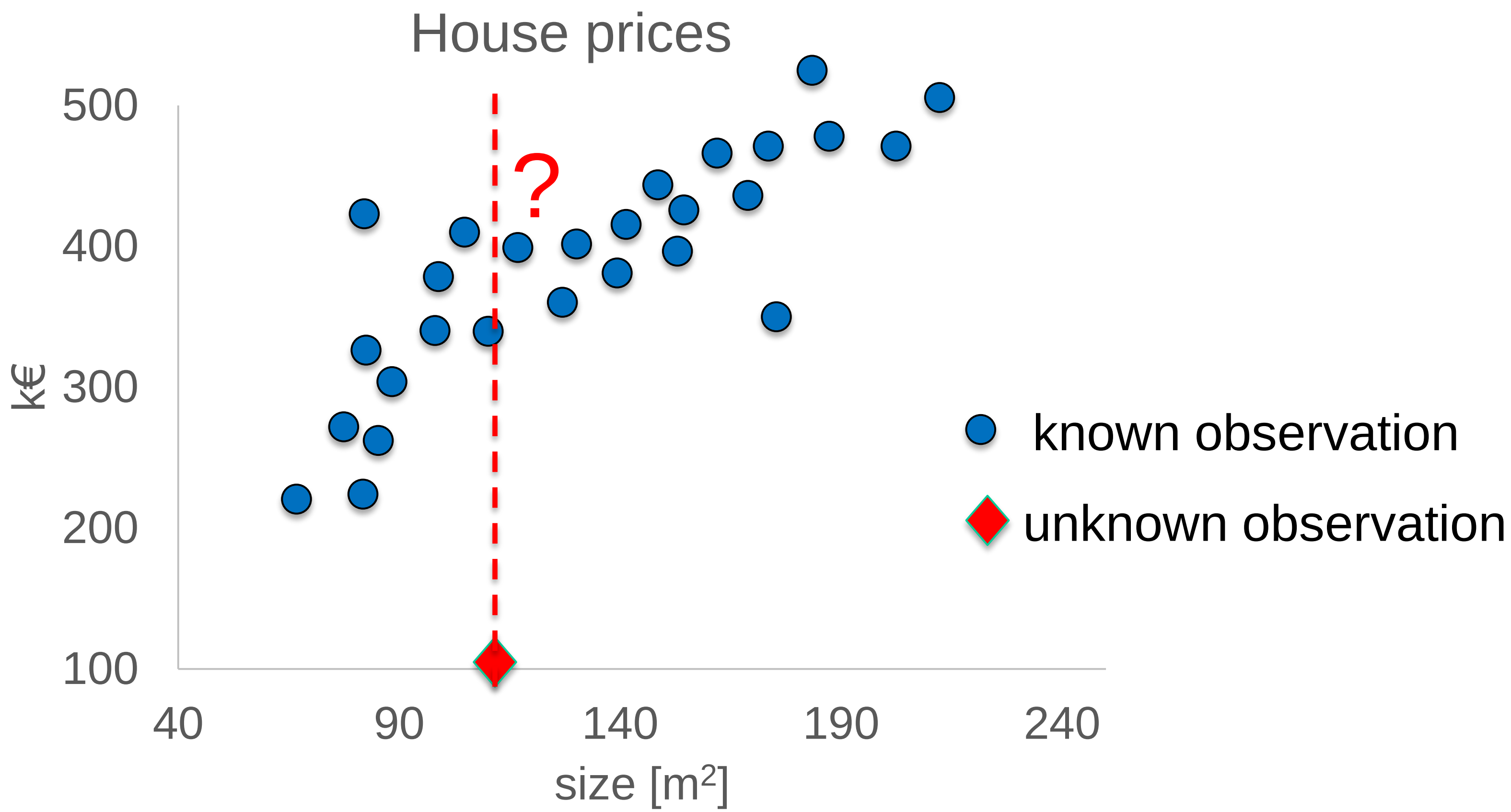
Some algorithms

- Supervised
  - Parametric
    - Linear and logistic regression
    - Neural Networks
    - ..
  - Non parametric
    - K-nearest neighbor
    - Random Forest
    - ...
- Unsupervised
  - Clustering
    - K-means
    - Gaussian Mixture Models
    - …



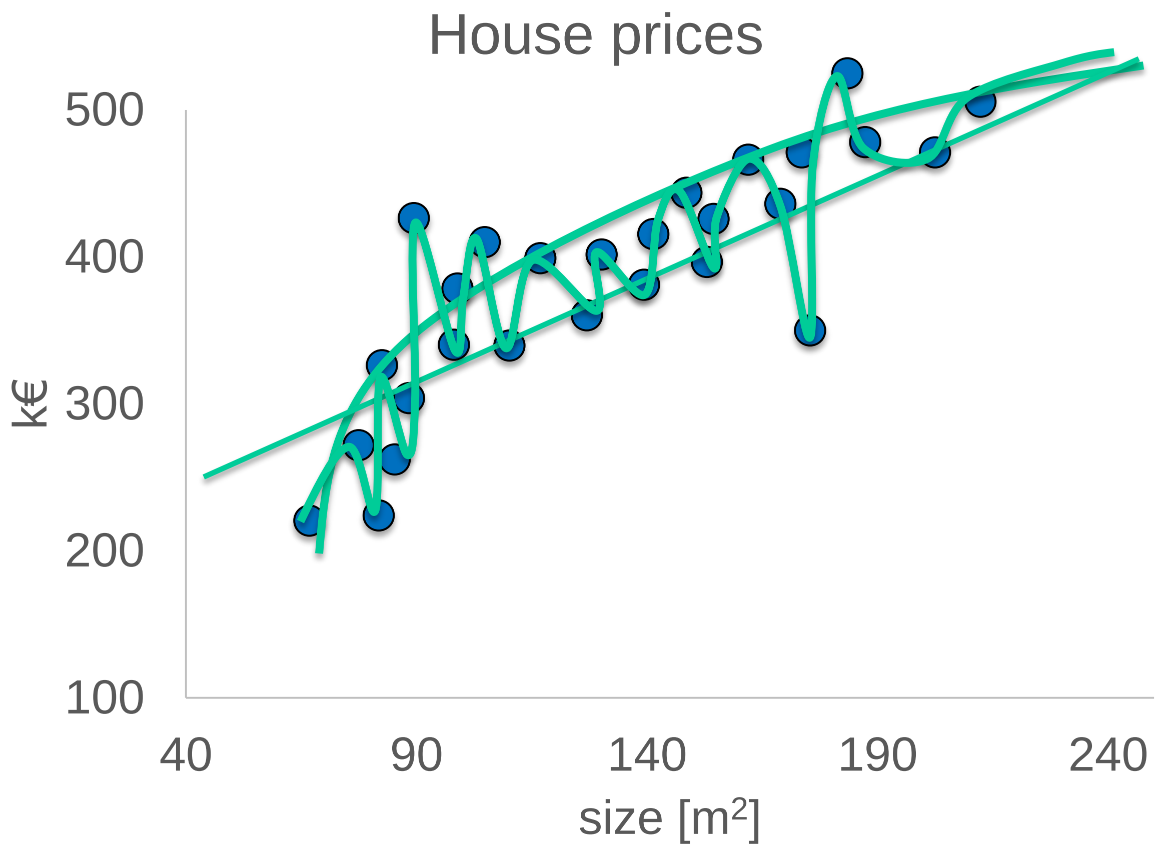
Let's start with a simple problem of regression from data (I)

- Suppose we want to predict a house price given its size



Let’s start with a simple problem of regression from data (II)

- Linear, quadratic, polynomial, non linear

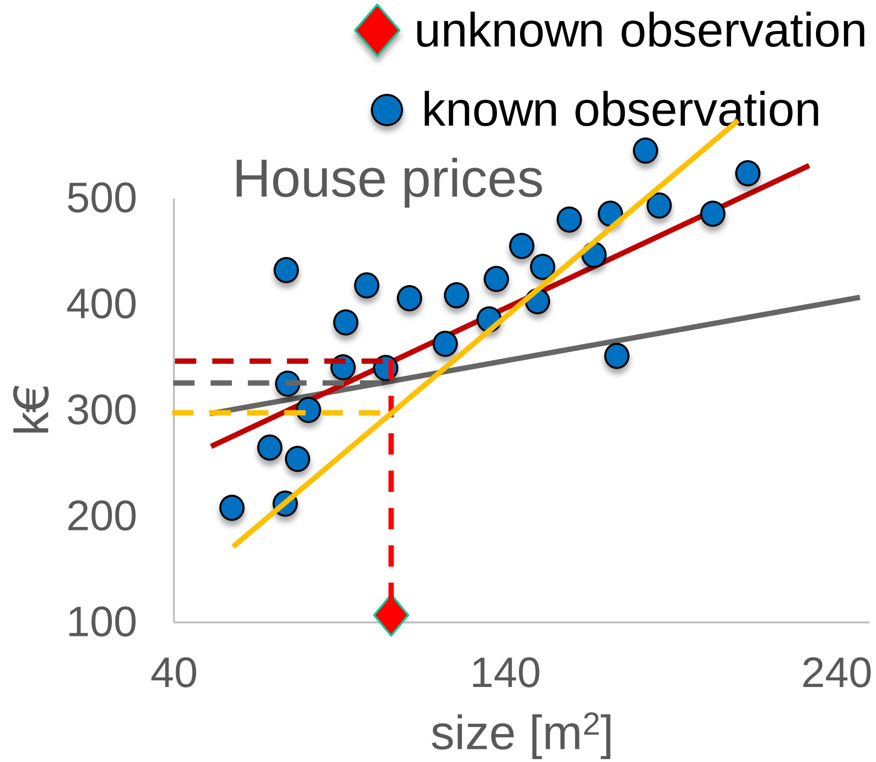


Linear regression
Univariate case

- Simplest model
  - *h(x)* is a ***linear*** function
  - *h(x)* has only **one variable** (univariate), i.e., feature x_1

$h(\underline{x}) = h(x_1) = \theta_0 + \theta_1 x_1$

  - θ_0 and θ_1 are the "weights"
- How to choose θ_0 and θ_1 ?



Minimize the training mean-square error (MSE)

$$\min_{\theta_0,\theta_1}\left\{ MSE = \frac{1}{2m}\sum_{i=1}^{m}\left(h(x^{(i)})-y^{(i)}\right)^2 \right\}$$

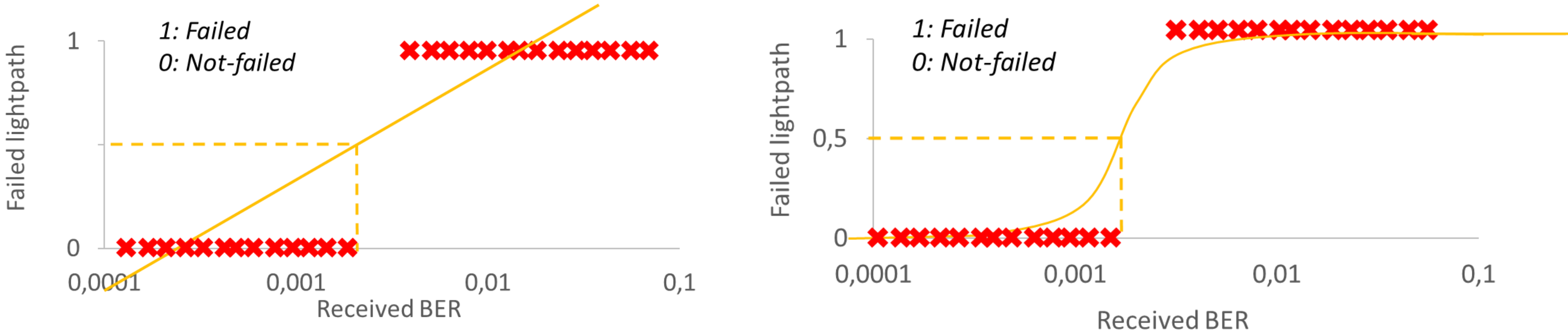
- Multivariate
  - We now have a features **vector** $\underline{x}=(x_1, x_2, \ldots x_n)$
  - $h(\underline{x}) = h(x_1 \ldots x_N) = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \ldots\ldots \theta_N x_N$
    - $\theta_0 \ldots \theta_1$ are the “weights” chosen by the algorithm
- Polynomial
  - We now increasing the order of polynomials in *h(x)*
  - $h(\underline{x}) = \theta_0 + \theta_1 x_1 + \theta_{12} x_1 x_2 + \theta_2 (x_2)^2 \ldots + \theta_n x_n$

What if the relation is completely unknown?

What if I cannot make any assumption regarding input-output relation?

Neural networks!

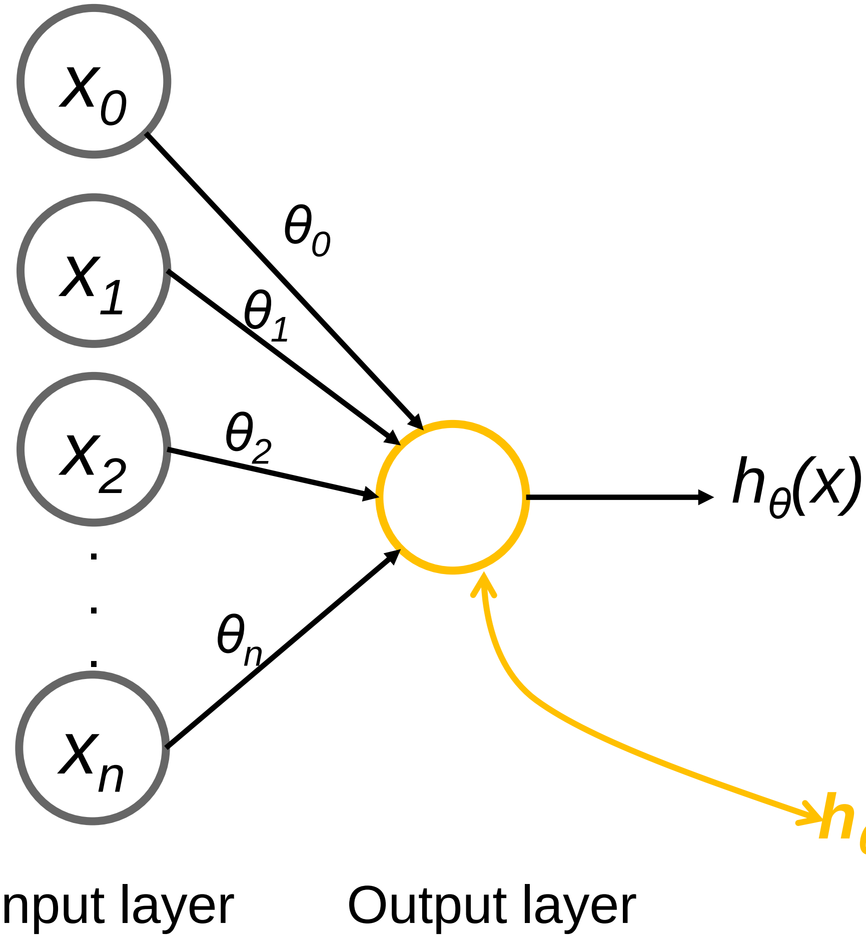
- Output *h(x)* takes only **discrete** values
  - Ex: y={0;1}, e.g., yes/no, good/bad, spam/non-spam…
  - Multiclass classifier: y={A,B,C,…}



- A good candidate function *h* for
  - $h(z) = 1/(1+e^{-z})$ is the "logistic" (or "sigmoid") function
    - for z → -inf: h(z)→0
    - for z → +inf: h(z)→1
    - for z=0: h(z)=0.5

Neural networks representation
Logistic unit or "neuron"

- The simplest neural network



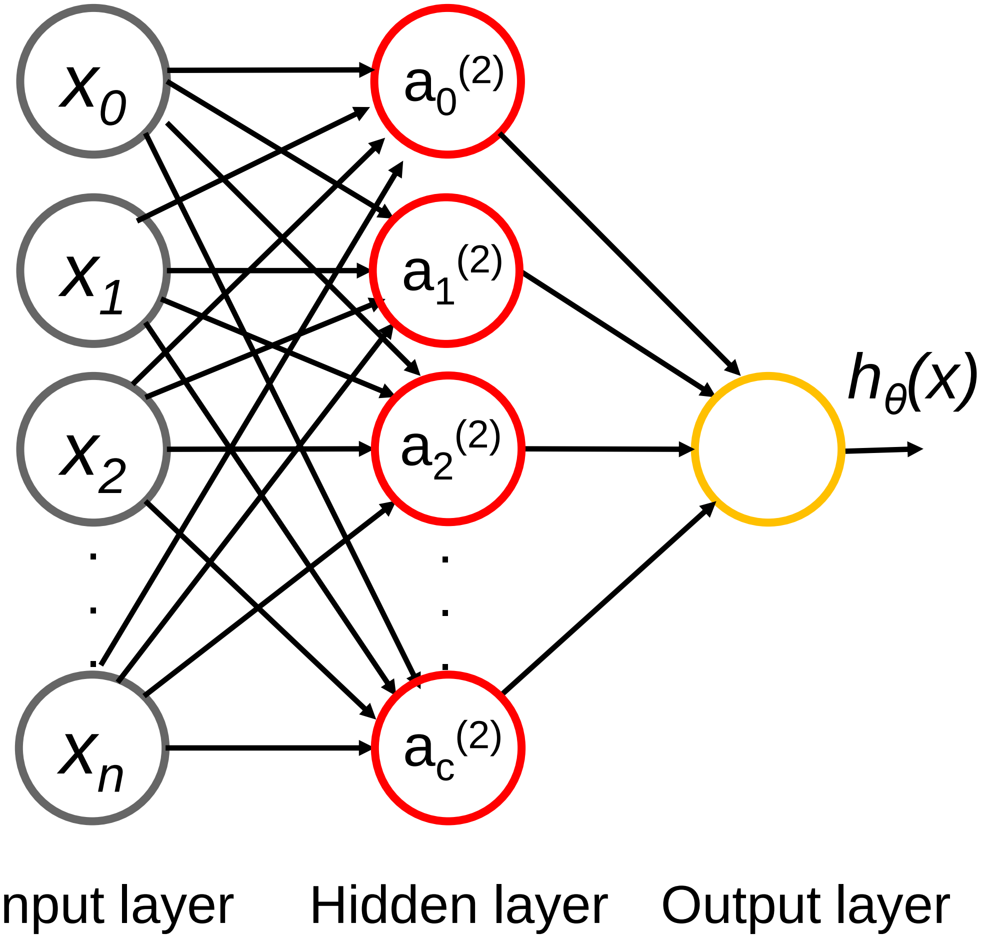
$$x = \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix} \qquad \theta = \begin{bmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \\ \dots \\ \theta_n \end{bmatrix}$$

inputs or features   weights

$$h_\theta(x) = g(\theta^T x) = 1/(1+e^{-(\theta^T x)})$$

Neural Networks (NN) representation
Multiple layers

- A "collection" of interacting neurons



3 observations:

1. NN can capture any relation between x and y
2. Hyperparameters: # of hidden layers , # neurons per hidden layer.
Who decides them?
3. Deep Learning: the more Layer, the less decisions shal be taken by a programmer

1. ML for QoT Estimation for Unestablished Lighpaths

- C. Rottondi, L. Barletta, A. Giusti and M. Tornatore, *A Machine Learning Method for QoT Estimation of Unestablished Lightpaths*, in IEEE/OSA Journal of Optical Comm.& Netw. Vol. 10, No. 2, Feb. 2018
- D. Azzimonti, C. Rottondi, M. Tornatore, “Using Active Learning to Decrease Probes for QoT Estimation in Optical Networks,” in Proceedings of OFC 2019, San Diego, Feb 2019.
- M. Salani, C. Rottondi, M. Tornatore, “Routing and Spectrum Assignment Integrating Machine-Learning-Based QoT Estimation in Elastic Optical Networks,” in Proceedings of INFOCOM 2019, Paris, April 2019.

2. ML for Soft-Failure Identification

- S. Shahkarami, F. Musumeci, F. Cugini, M. Tornatore, “Machine-Learning-Based Soft-Failure Detection and Identification in Optical Networks,“ in Proceedings, OFC 2018, San Diego (CA), Usa, Mar. 11-15, 2018
- A. Vela et al., “BER degradation Detection and Failure Identification in Elastic Optical Networks”, in IEEE/OSA Journal of Lightwave Technology, vol. 35, no. 21, pp. 4595-4604, Nov.1, 1 2017
- Francesco Musumeci ,et al., “A Tutorial on Machine Learning for Failure Management in Optical Networks”, in IEEE/OSA Journal of Lightwave Technology, available online

3. An overview of other applications at network layer

- F. Musumeci et al., “A Survey on Application of Machine Learning Techniques in Optical Networks”, Submitted to IEEE communication surveys and tutorials, available in ArXiv
- Javier Mata, et a., Artificial intelligence (AI) methods in optical networks: A comprehensive survey, Optical Switching and Networking, Volume 28, 2018, pp. 43-57

Outline: Part 2 Applications

1. ML for QoT Estimation for Unestablished Lighpaths

- C. Rottondi, L. Barletta, A. Giusti and M. Tornatore, *A Machine Learning Method for QoT Estimation of Unestablished Lightpaths*, in IEEE/OSA Journal of Optical Comm.& Netw. Vol. 10, No. 2, Feb. 2018
- D. Azzimonti, C. Rottondi, M. Tornatore, “Using Active Learning to Decrease Probes for QoT Estimation in Optical Networks,” in Proceedings of OFC 2019, San Diego, Feb 2019.
- M. Salani, C. Rottondi, M. Tornatore, “Routing and Spectrum Assignment Integrating Machine-Learning-Based QoT Estimation in Elastic Optical Networks,” in Proceedings of INFOCOM 2019, Paris, April 2019.

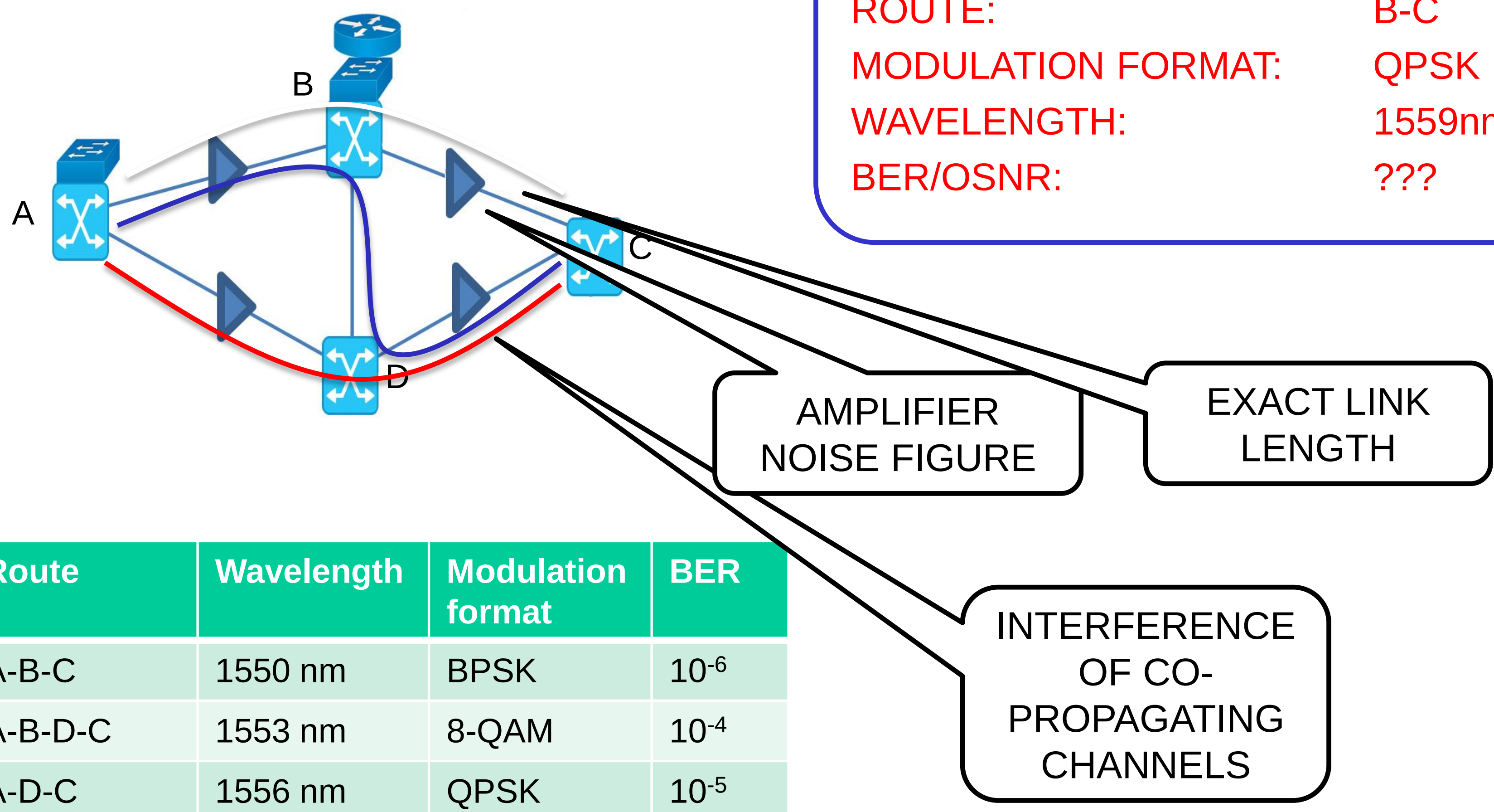
2. ML for Soft-Failure Identification

- S. Shahkarami, F. Musumeci, F. Cugini, M. Tornatore, “Machine-Learning-Based Soft-Failure Detection and Identification in Optical Networks,“ in Proceedings, OFC 2018, San Diego (CA), Usa, Mar. 11-15, 2018
- A. Vela et al., “BER degradation Detection and Failure Identification in Elastic Optical Networks”, in IEEE/OSA Journal of Lightwave Technology, vol. 35, no. 21, pp. 4595-4604, Nov.1, 1 2017
- Francesco Musumeci ,et al., “A Tutorial on Machine Learning for Failure Management in Optical Networks”, in IEEE/OSA Journal of Lightwave Technology, available online

3. An overview of other applications at network layer

- F. Musumeci et al., “A Survey on Application of Machine Learning Techniques in Optical Networks”, Submitted to IEEE communication surveys and tutorials, available in ArXiv
- Javier Mata, et a., Artificial intelligence (AI) methods in optical networks: A comprehensive survey, Optical Switching and Networking, Volume 28, 2018, pp. 43-57

Why QoT estimation?



NEW TRAFFIC REQUEST:

ROUTE: B-C

MODULATION FORMAT: QPSK

WAVELENGTH: 1559nm

BER/OSNR: ???

Route	Wavelength	Modulation format	BER
A-B-C	1550 nm	BPSK	10^{-6}
A-B-D-C	1553 nm	8-QAM	10^{-4}
A-D-C	1556 nm	QPSK	10^{-5}

How (pre-deployment) QoT estimation is done

- **"Exact" analytical models** (e.g., split-step Fourier method)
  - 🙂 Accurate results
  - 🙁 Heavy computational requirements → not scalable / not real time

- **Margined formulas** (e.g., AWGN model…)
  - 🙂 Faster and more scalable

$$\frac{1}{\mathrm{OSNR_{tot}}} = \sum_{k=1}^{N_{\mathrm{span}}} \frac{1}{\mathrm{OSNR}^{(k)}_{\mathrm{ASE,Rx}}} + \sum_{k=1}^{N_{\mathrm{span}}} \frac{1}{\mathrm{OSNR}^{(k)}_{\mathrm{NL}}} \qquad \mathrm{OSNR}^{(k)}_{\mathrm{ASE,Rx}} = \frac{P^{(k)}_{\mathrm{Tx}}}{h\nu B_n G^{(k)} F^{(k)}}$$

  - 🙁 Analitically accurate, but suffers from inaccurate parameter knowledge.
  - 🙁 High margination, underutilization of network resources (up to extra 2 dB for design margins [1])

[1] Y. Pointurier, "Design of low-margin optical networks," in *IEEE/OSA Journal of Optical Communications and Networking*, vol. 9, no. 1, pp. A9-A17, Jan. 2017. doi: 10.1364/JOCN.9.0000A9

Machine Learning as an alternative approach?

- Machine Learning* methods have been proposed to
  - estimate QoT of unestablished lightpaths
  - using field data, e.g., monitored BER/OSNR at the receiver

🙂 No need for complex analytical models

🙂 Fast and scalable

🙁 Requires training phase with historical data

- How long must the training phase be?
- How accurate will the estimation be?
- Objectives of our numerical analysis….

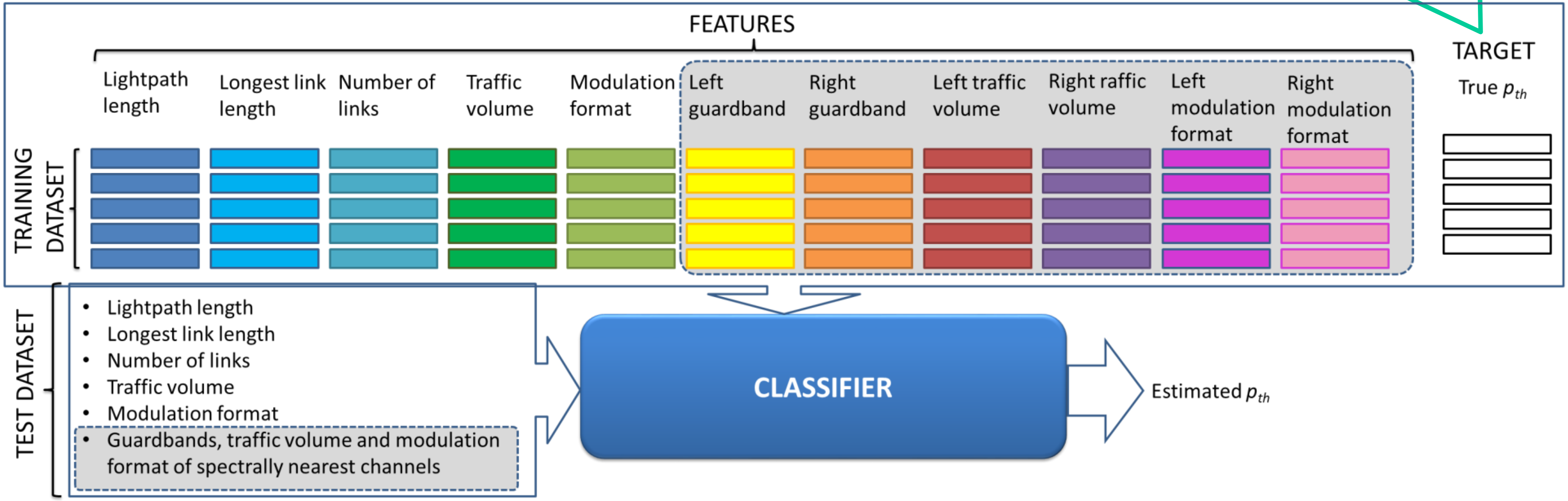
**C. Rottondi, L. Barletta, A. Giusti, M. Tornatore "Machine-learning method for quality of transmission prediction of unestablished lightpaths," IEEE/OSA J. of Optical Comm. and Netw., vol. 10, no. 2, pp. A286–A297, Feb 2018.*

Proposed ML classifier for QoT estimation

Input: *Lightpath features*

Output: *Prob{BER ≤ T*}*

The classifier is trained on a set of *L* experiments to generate ***ground truth***



Which Machine Learning Algorithm?

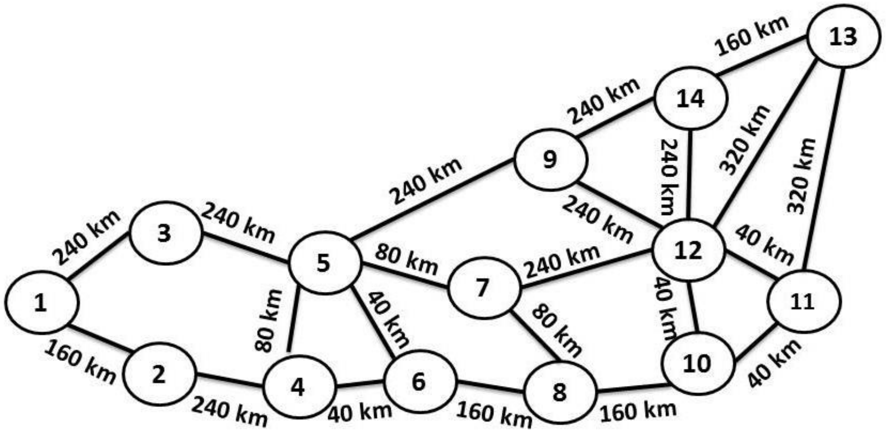
- We used a Random Forest (RF) classifier with 25 estimators
- To take this choice, we compared 5 RFs and 3 kNN classifiers and picked best "accuracy/complexity" tradeoff

Algorithm	Training time (s)	Test time (s)	AUC	Accuracy
Dummy classifier	0.048979	3.83 e-07	0.501	0.539
1 Nearest Neighbor	1.183121	4.83 e-05	0.959	0.957
5 Nearest Neighbor	1.085116	5.05 e-05	0.991	0.965
25 Nearest Neighbor	1.211694	6.91 e-05	0.996	0.965
Random Forest 1 tree	0.076944	3.96 e-07	0.991	0.965
Random Forest 5 trees	0.180835	6.24 e-07	0.995	0.970
Random Forest 25 trees	0.721042	1.56 e-06	0.996	0.968
Random Forest 100 trees	2.830545	5.32 e-06	0.996	0.966
Random Forest 500 trees	14.052182	2.63 e-05	0.996	0.966

- But knowledge is rapidly evolving!
  - Neural Networks… SVMs… (parametric approaches)
  - Gaussian processes (return confidence of classification!)

Training and Testing Scenario

- Some results for a Japanese optical network

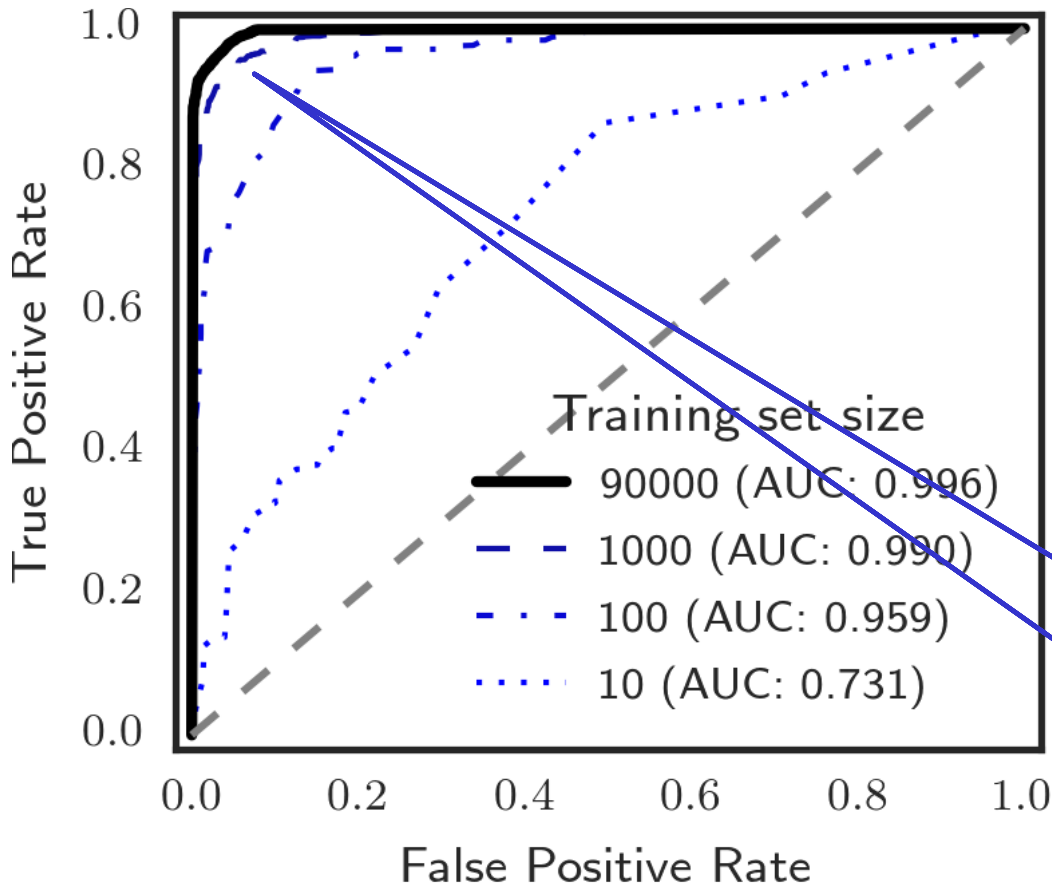


- Flexgrid @ 12.5 GHz
- Transceivers @ 28 GBaud
- 6 Modulation formats
  - (DP) BPSK, QPSK, 8-QAM to 64-QAM,
- Traffic requests: [50;1000] Gbps
- 3 candidate paths per node pair
- BER threshold $T = 4*10^{-3}$

- NB$_1$: We used synthetic data!
- NB$_2$: some data sets are becoming available

- *Monia Ghobadi and Ratul Mahajan. "Optical layer failures in a large backbone." In Proceedings of the 2016 Internet Measurement Conference. ACM, 2016.*
- *Rachee Singh, Monia Ghobadi, Klaus-Tycho Foerster, Mark Filer, and Phillipa Gill. "Run, Walk, Crawl: Towards Dynamic Link Capacities." In Proceedings of the 16th ACM Workshop on Hot Topics in Networks. ACM, 2017.*

How long shall training phase be?
(1) *Accuracy vs training set size*



- «ROC» curve
- Area under the ROC curve (AUC)

Take-Away 1: Training phase has a reasonable duration

How to reduce required data/probes?

- ML requires training phase with historical data

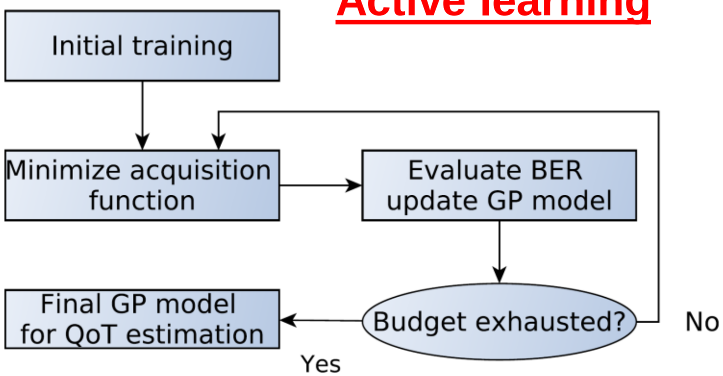
ISSUES

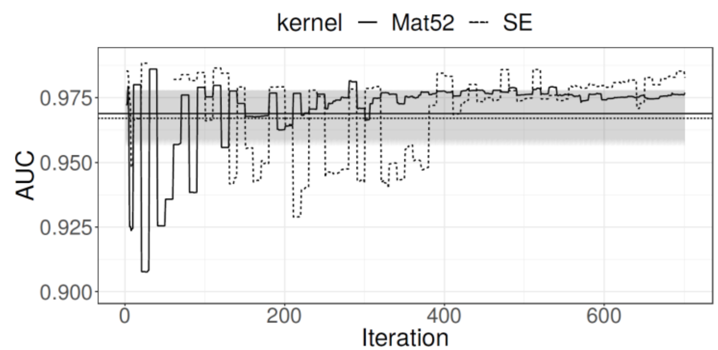
  - Samples from faulty/malfunctioning lightpaths are rare
  - With margined approaches, ligthpaths with risky BER are unlikely deployed (thus never observed)

- Probe ligthtpaths are n... ...gions of feature space not ... data

COSTLY! THE NUMBER OF PROBES MUST BE MINIMIZED!

Active learning

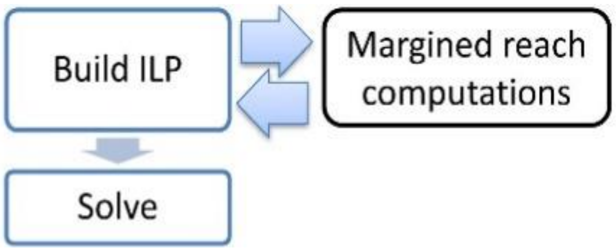




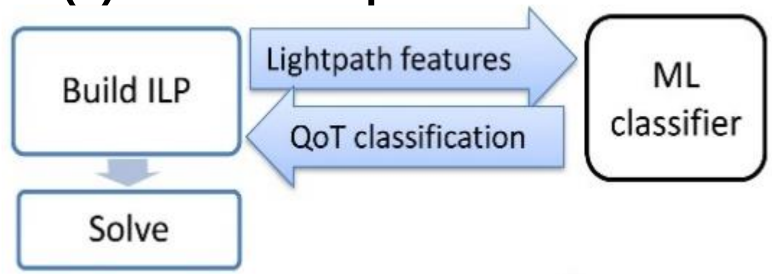
D. Azzimonti, C. Rottondi, M. Tornatore, "Using Active Learning to Decrease Probes for QoT Estimation in Optical Networks," in Proceedings of OFC 2019, San Diego, Feb 2019.

Ok, but, what's the impact on resource saving? Relation between RSA and ML-based QoT estimation

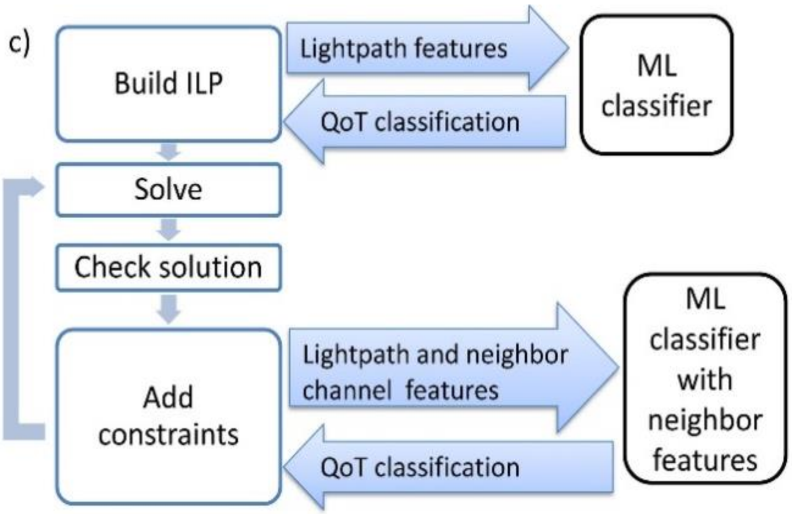
(a) Current mode of operation



(b) ML-based operation



(c) Low-margin design (iterative procedure)



Output of the classifier:
probabilty γ that BER ≤ T*

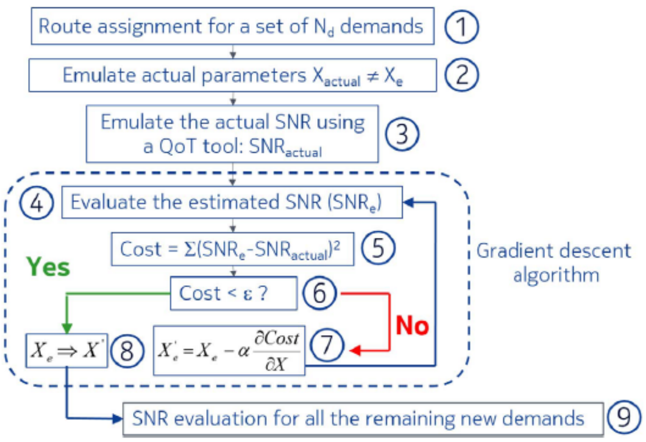
γ: Risk you are willing to accept

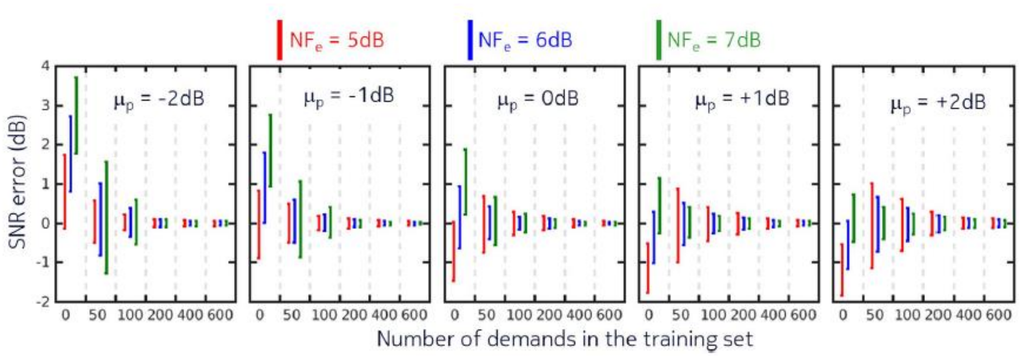
γ	Savings
0.5	35.71%
0.7	32.08%
0.9	27.36%
0.99	26.61%

M. Salani, C. Rottondi, M. Tornatore, "Routing and Spectrum Assignment Integrating Machine-Learning-Based QoT Estimation in Elastic Optical Networks," in Proceedings of INFOCOM 2019, Paris, April 2019.

Another way of looking at this problem.. Estimating unknown parameters in GN model

- General motivation
  - If you have a model, you should use it!
  - No need to reinvent the wheel
- So, for QoT estimation, if we know which parameter is inaccurate (e.g., noise figure), we can use ML to estimate that parameter, and keep the rest of the analytical model





E. Seve, J. Pesic, C. Delezoide, S. Bigo, and Y. Pointurier, "Learning Process for Reducing Uncertainties on Network Parameters and Design Margins," J. Opt. Commun. Netw. 10, A298-A306 (2018)

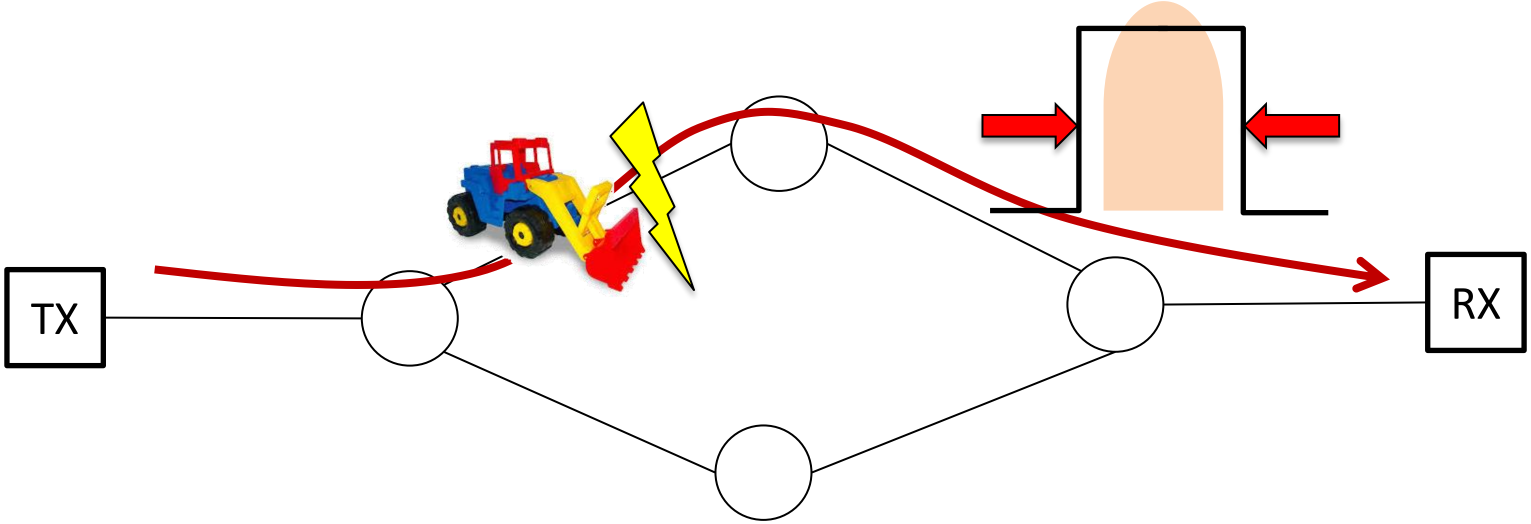
1. ML for QoT Estimation for Unestablished Lighpaths
   - C. Rottondi, L. Barletta, A. Giusti and M. Tornatore, *A Machine Learning Method for QoT Estimation of Unestablished Lightpaths*, in IEEE/OSA Journal of Optical Comm.& Netw. Vol. 10, No. 2, Feb. 2018
   - D. Azzimonti, C. Rottondi, M. Tornatore, “Using Active Learning to Decrease Probes for QoT Estimation in Optical Networks,” in Proceedings of OFC 2019, San Diego, Feb 2019.
   - M. Salani, C. Rottondi, M. Tornatore, “Routing and Spectrum Assignment Integrating Machine-Learning-Based QoT Estimation in Elastic Optical Networks,” in Proceedings of INFOCOM 2019, Paris, April 2019.

2. ML for Soft-Failure Identification
   - S. Shahkarami, F. Musumeci, F. Cugini, M. Tornatore, “Machine-Learning-Based Soft-Failure Detection and Identification in Optical Networks,“ in Proceedings, OFC 2018, San Diego (CA), Usa, Mar. 11-15, 2018
   - A. Vela et al., “BER degradation Detection and Failure Identification in Elastic Optical Networks”, in IEEE/OSA Journal of Lightwave Technology, vol. 35, no. 21, pp. 4595-4604, Nov.1, 1 2017
   - Francesco Musumeci ,et al., “A Tutorial on Machine Learning for Failure Management in Optical Networks”, in IEEE/OSA Journal of Lightwave Technology, available online

3. An overview of other applications at network layer
   - F. Musumeci et al., “A Survey on Application of Machine Learning Techniques in Optical Networks”, Submitted to IEEE communication surveys and tutorials, available in ArXiv
   - Javier Mata, et a., Artificial intelligence (AI) methods in optical networks: A comprehensive survey, Optical Switching and Networking, Volume 28, 2018, pp. 43-57

Two main failure types in optical networks

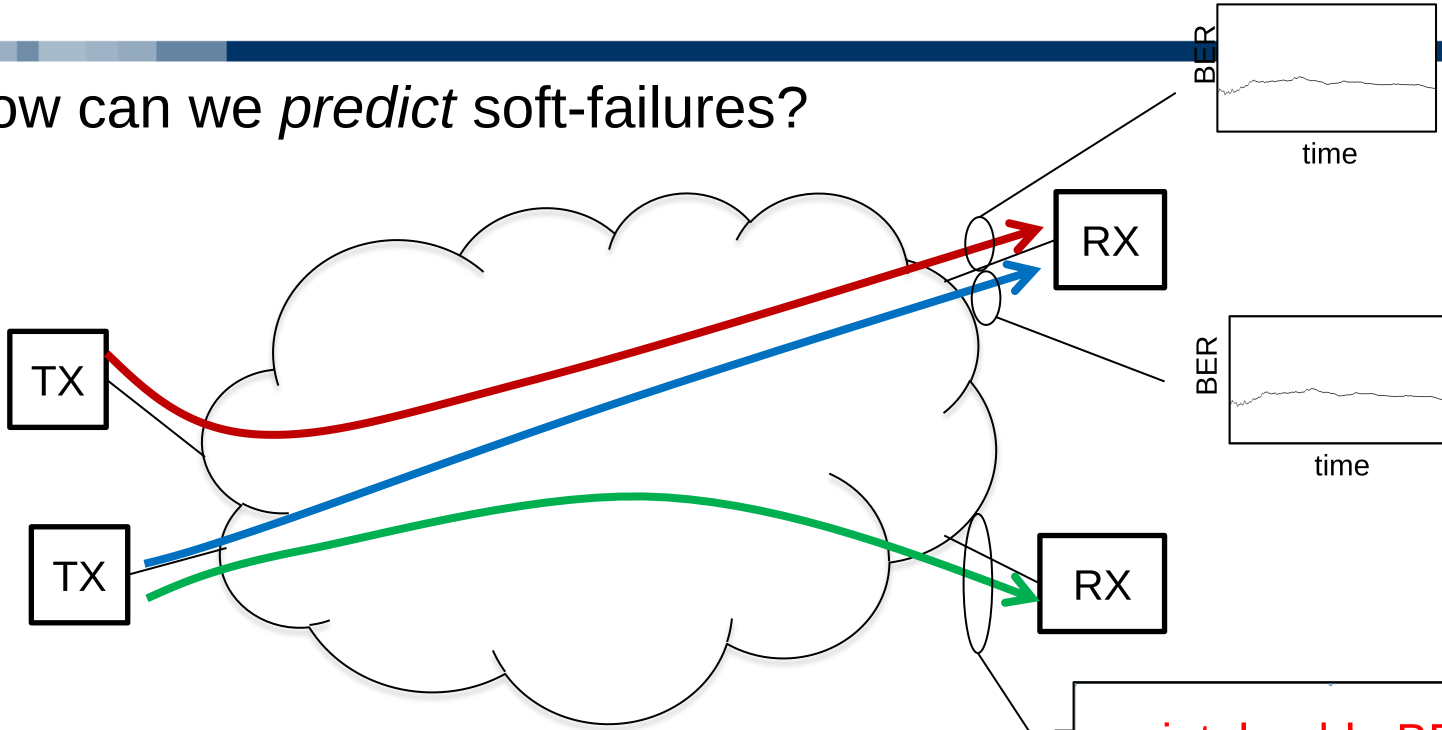
- Hard failures*
  - Sudden events, e.g., fiber cuts, power outages, etc.
  - Require «protection» (*reactive procedures*)
- Soft failures:
  - Gradual transmission degradation due to equipment malfunctioning, filter shrinking/misalignment…
  - Trigger early network reconfiguration (*proactive* procedures)



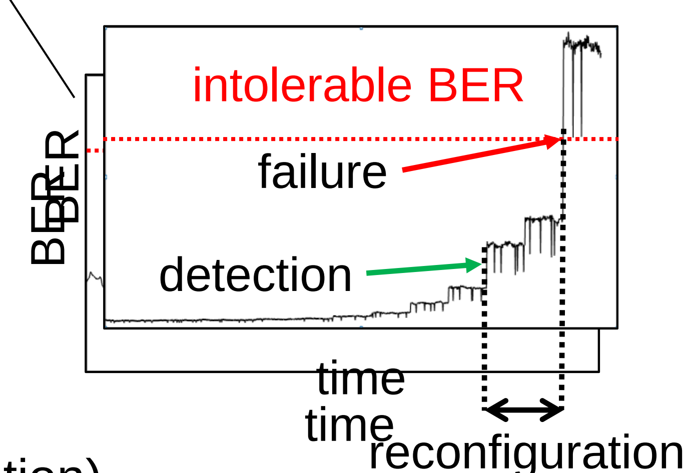
*F. Boitier *et al*., "Proactive Fiber Damage Detection in Real-time Coherent Receiver," *2017 European Conference on Optical Communication (ECOC)*, Gothenburg, 2017, pp. 1-3.

Soft-failure *early detection*

- How can we *predict* soft-failures?



Perform continuous monitoring of BER at the receiver…
… until some “anomalies” are detected

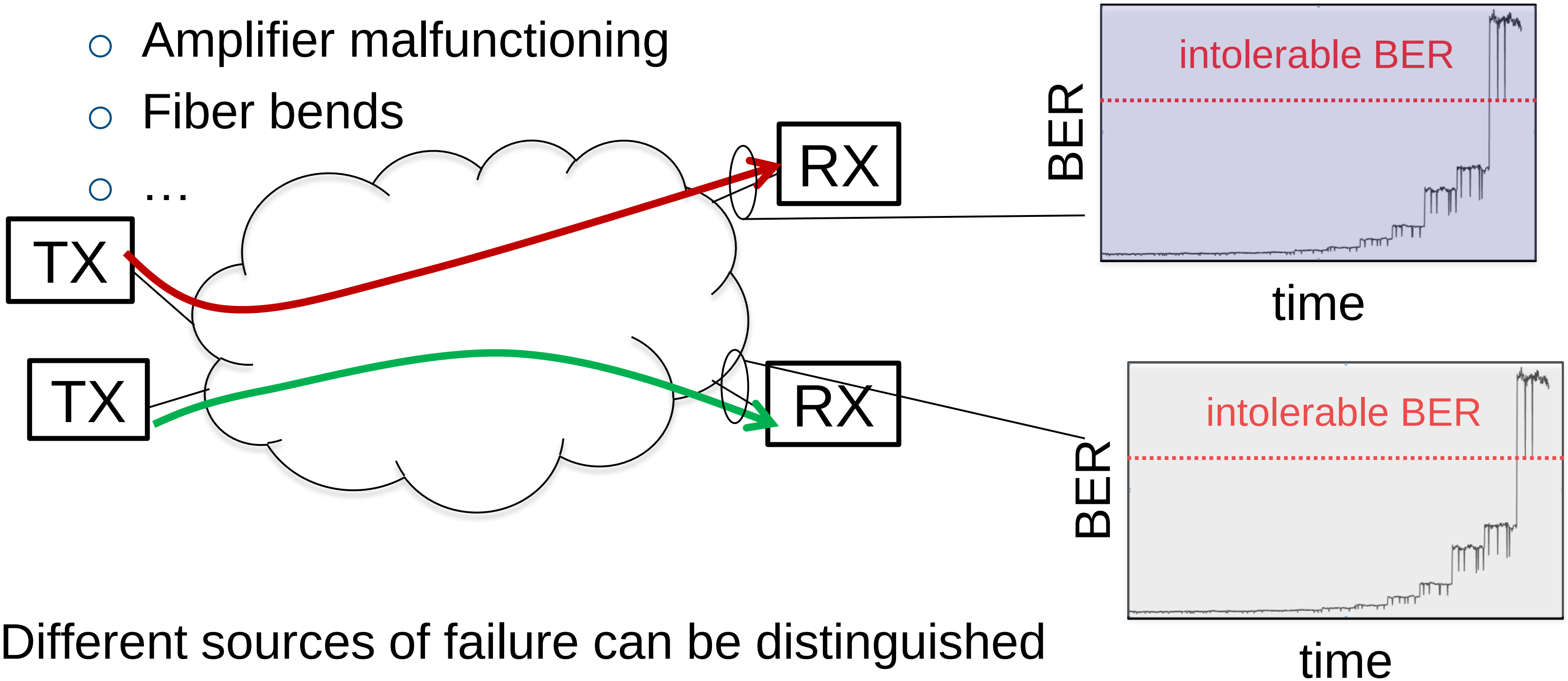


Early-detection helps **preventing** service disruption (e.g., through proactive reconfiguration)

A. Vela et al., “BER degradation Detection and Failure Identification in Elastic Optical Networks”, in IEEE/OSA Journal of Lightwave Technology, vol. 35, no. 21, pp. 4595-4604, Nov.1, 1 2017

Soft-failure *cause identification*

- How can we identify the *cause* of the failure?
  - Failures can be caused by different sources
    - Filters shrinking/misalignment
    - Amplifier malfunctioning
    - Fiber bends
    - …

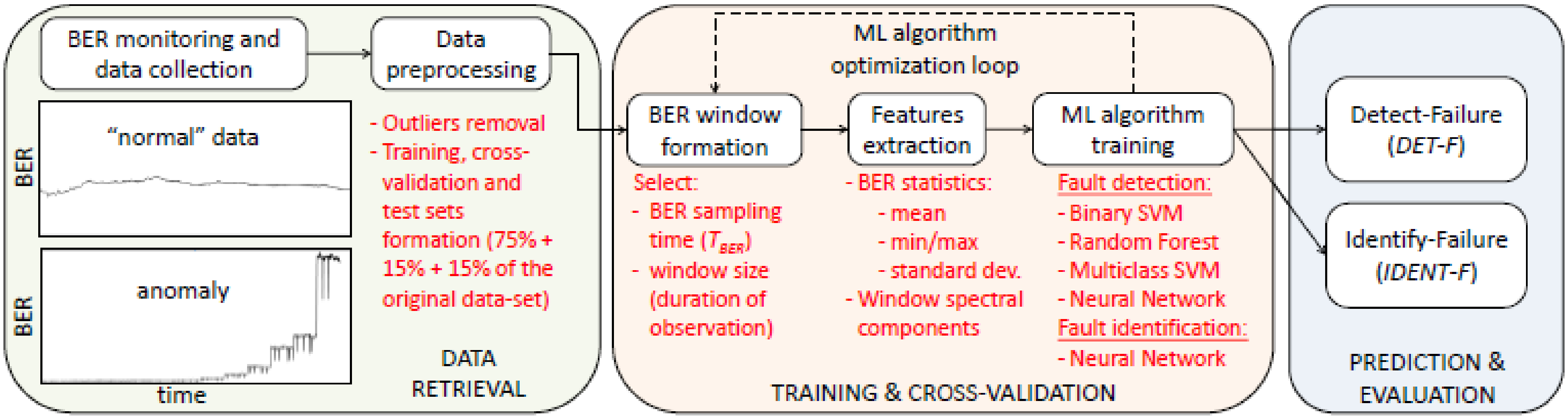


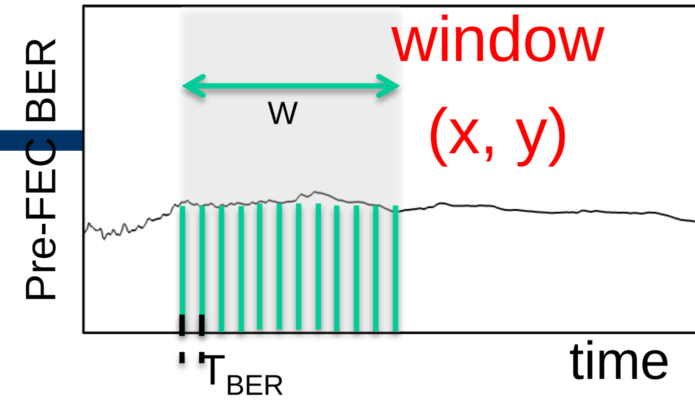
Different sources of failure can be distinguished via the different effects on BER (i.e., via different BER "features")

S. Shahkarami, F. Musumeci, F. Cugini, M. Tornatore, "Machine-Learning-Based Soft-Failure Detection and Identification in Optical Networks,"in Proceedings, OFC 2018, San Diego (CA), Usa, Mar. 11-15, 2018

Phases of our study

Overall view





1. Data Retrieval

3 decisions

Validation (optimization of hyperparameters)

BER window → Features → ML algorithm

BER window

Select:

- BER sampling time (T_{BER})
- window size (duration of observation)

Features

- BER statistics:
  - mean
  - min/max
  - standard dev.
- Window spectral components
- Feature Scaling

ML algorithm

Fault detection:

- Binary SVM
- Random Forest
- Multiclass SVM
- Neural Network

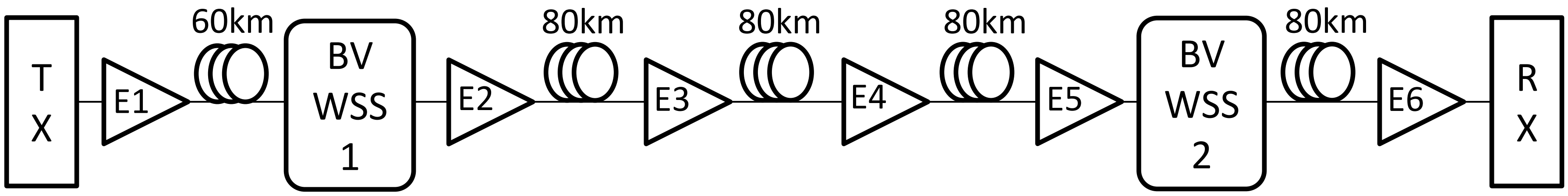
Fault identification:

- Neural Network

3. Prediction and Evaluation

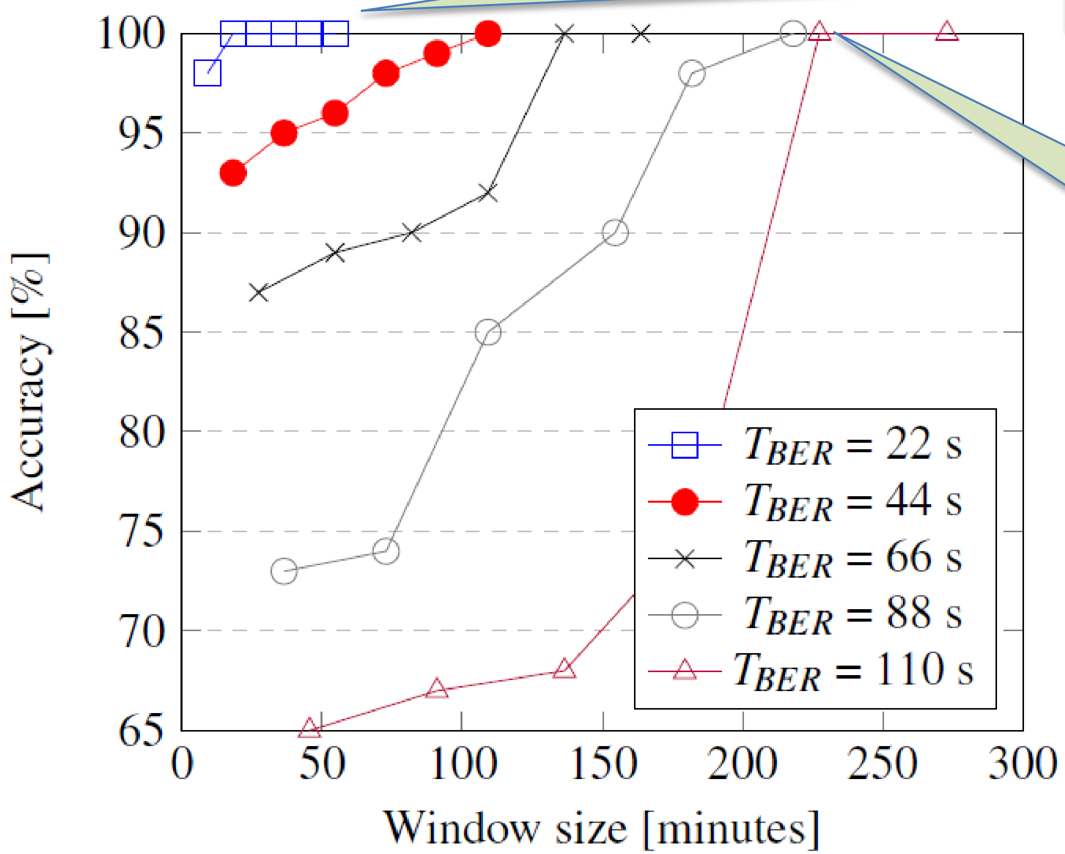
Testbed setup

- Testbed for real BER traces
  - Ericsson 380 km transmission system
    - 24 hours BER monitoring
    - 3 seconds sampling interval
  - PM-QPSK modulation @ 100Gb/s
  - 6 Erbium Doped Fiber Amplifiers (EDFA) followed by Variable Optical Attenuators (VOAs)
  - Bandwidth-Variable Wavelength Selective Switch (BV-WSS) is used to emulate 2 types of BER degradation:
    - **Filter misalignment**
    - Additional attenuation in intermediate span (e.g., due to **EDFA gain-reduction**)



Numerical results: *Detection*
Accuracy vs window features

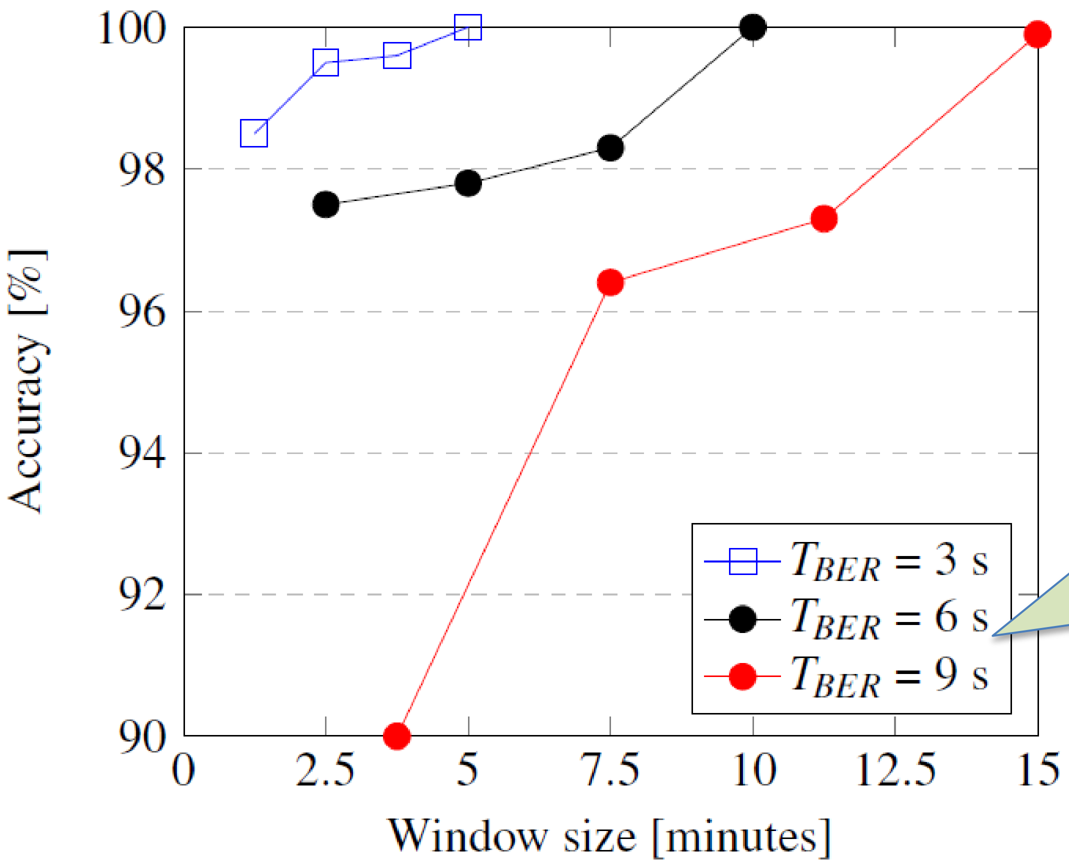
- Binary SVM



Take-away 1: Higher performance for with low sampling time → Fast monitoring equipment is required

Take-away 2: For increasing sampling time, longer "Windows" are needed for high accuracy

- Neural Network

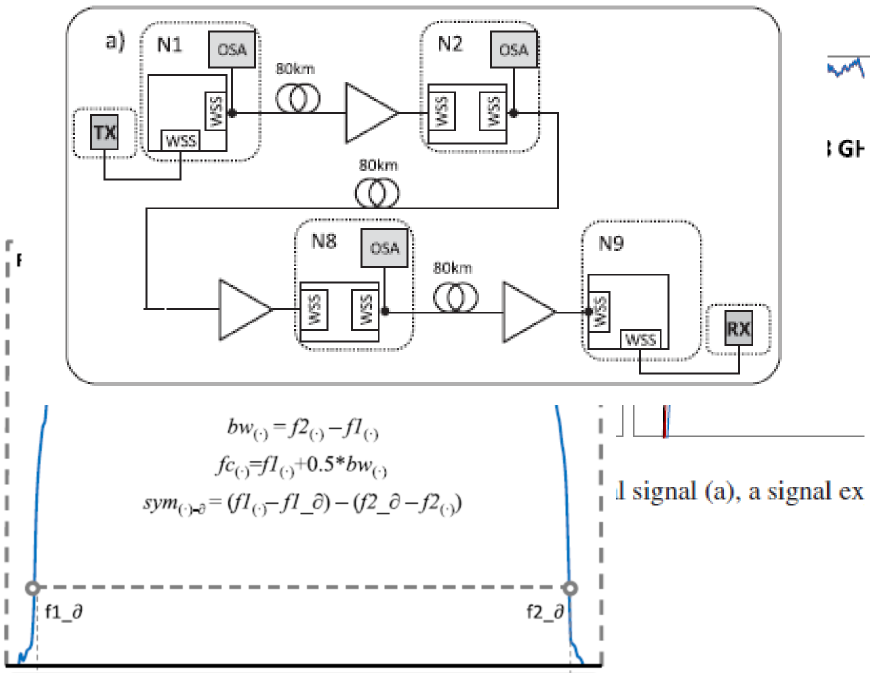


Take-away 3: To perform failure-cause identification, much smaller sampling period is needed wrt failure detection

Detection and Identification using Optical Spectrum Analyzers

- Cost-effective Optical Spectrum Analyzers (OSA) with sub-GHz resolution can be used to monitor spectrum along transmission line

Flexgrid High Resolution Optical Channel Monitor (OCM). [Online]. Available: http://www.finisar.com, Accessed: Jun. 2018.



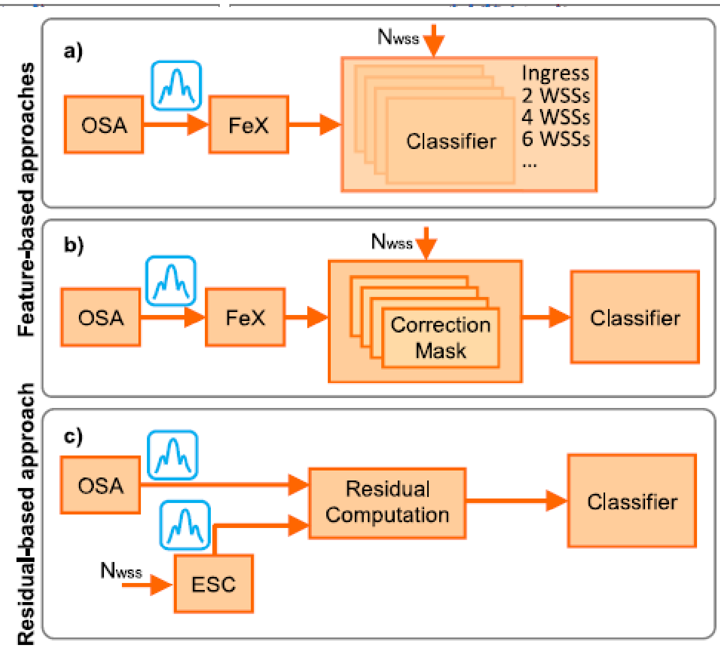


Fig. 4. Approaches to solve the filter cascading problem: (a) multi-classifier, (b) single-classifier, and (c) residual computation.

B. Shariati, M. Ruiz, J. Comellas and L. Velasco, "Learning From the Optical Spectrum: Failure Detection and Identification," in Journal of Lightwave Technology, vol. 37, no. 2, pp. 433-440, 15 Jan.15, 2019

Outline: Part 2 Applications

1. ML for QoT Estimation for Unestablished Lighpaths
   - C. Rottondi, L. Barletta, A. Giusti and M. Tornatore, *A Machine Learning Method for QoT Estimation of Unestablished Lightpaths*, in IEEE/OSA Journal of Optical Comm.& Netw. Vol. 10, No. 2, Feb. 2018
   - D. Azzimonti, C. Rottondi, M. Tornatore, "Using Active Learning to Decrease Probes for QoT Estimation in Optical Networks," in Proceedings of OFC 2019, San Diego, Feb 2019.
   - M. Salani, C. Rottondi, M. Tornatore, "Routing and Spectrum Assignment Integrating Machine-Learning-Based QoT Estimation in Elastic Optical Networks," in Proceedings of INFOCOM 2019, Paris, April 2019.

2. ML for Soft-Failure Identification
   - S. Shahkarami, F. Musumeci, F. Cugini, M. Tornatore, "Machine-Learning-Based Soft-Failure Detection and Identification in Optical Networks," in Proceedings, OFC 2018, San Diego (CA), Usa, Mar. 11-15, 2018
   - A. Vela et al., "BER degradation Detection and Failure Identification in Elastic Optical Networks", in IEEE/OSA Journal of Lightwave Technology, vol. 35, no. 21, pp. 4595-4604, Nov.1, 1 2017
   - Francesco Musumeci ,et al., "A Tutorial on Machine Learning for Failure Management in Optical Networks", in IEEE/OSA Journal of Lightwave Technology, available online

3. An overview of other applications at network layer
   - F. Musumeci et al., "A Survey on Application of Machine Learning Techniques in Optical Networks", Submitted to IEEE communication surveys and tutorials, available in ArXiv
   - Javier Mata, et a., Artificial intelligence (AI) methods in optical networks: A comprehensive survey, Optical Switching and Networking, Volume 28, 2018, pp. 43-57

Overview of other applications

Physical layer

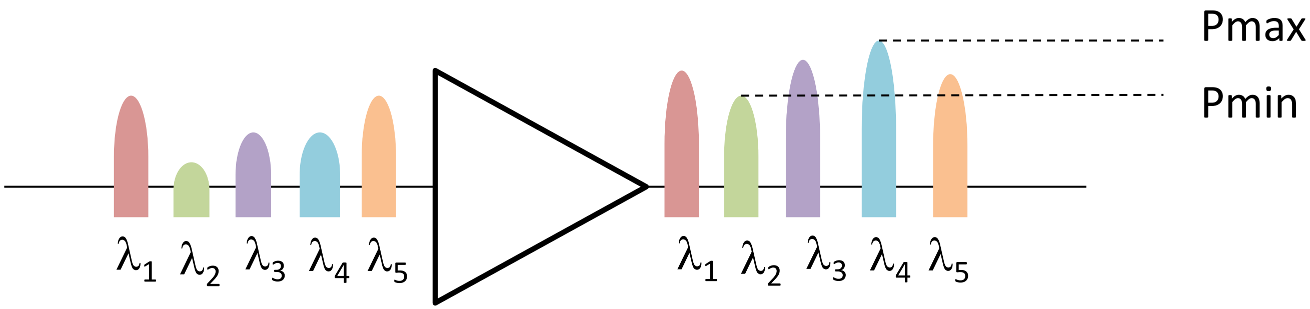
1. Quality of Transmission (QoT) estimation
2. Optical amplifier control
3. Modulation format recognition
4. Nonlinearities mitigation

Network layer

1. Traffic prediction and virtual topology design
2. Failure detection and localization
3. Flow classification

Classification taken from: F. Musumeci et al., “A Survey on Application of Machine Learning Techniques in Optical Networks”, Submitted to IEEE Communication Surveys and Tutorials, soon available in ArXiv

- When adding/dropping channels into/from a WDM system, EDFA gain should be adjusted to re-balance output powers

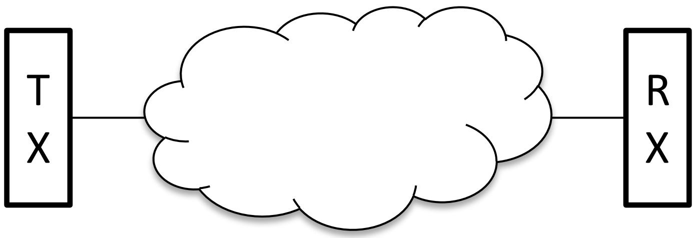


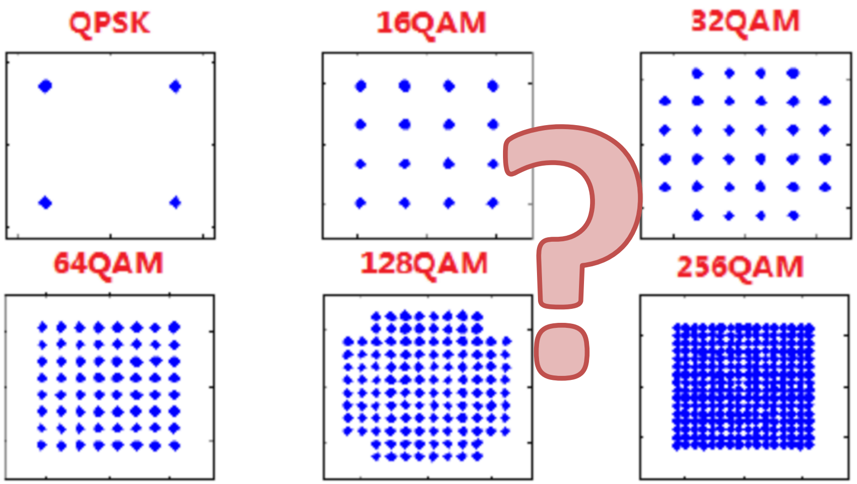
- Analytical models:
  - typically not generalizable
  - depend on the specific system (gain-control mechanism, EDFA gain tilt, nr of EDFAs…) which use to vary during their activity
- ML allows to self-learn typical response patters

Huang et al., "Dynamic mitigation of EDFA power excursions with machine learning", Optics Express, vol. 25 n. 3, Feb. 2017
Bastos et al., "Mapping EDFA Noise Figure and Gain Flatness Over the Power Mask Using Neural Networks", Journal of Microwaves, Optoelectronics and Electromagnetic Applications, vol. 12, n. SI-2, July 2013

Physical layer
Modulation format recognition (MFR)

- Elastic transceiver can to operate with different modulation formats





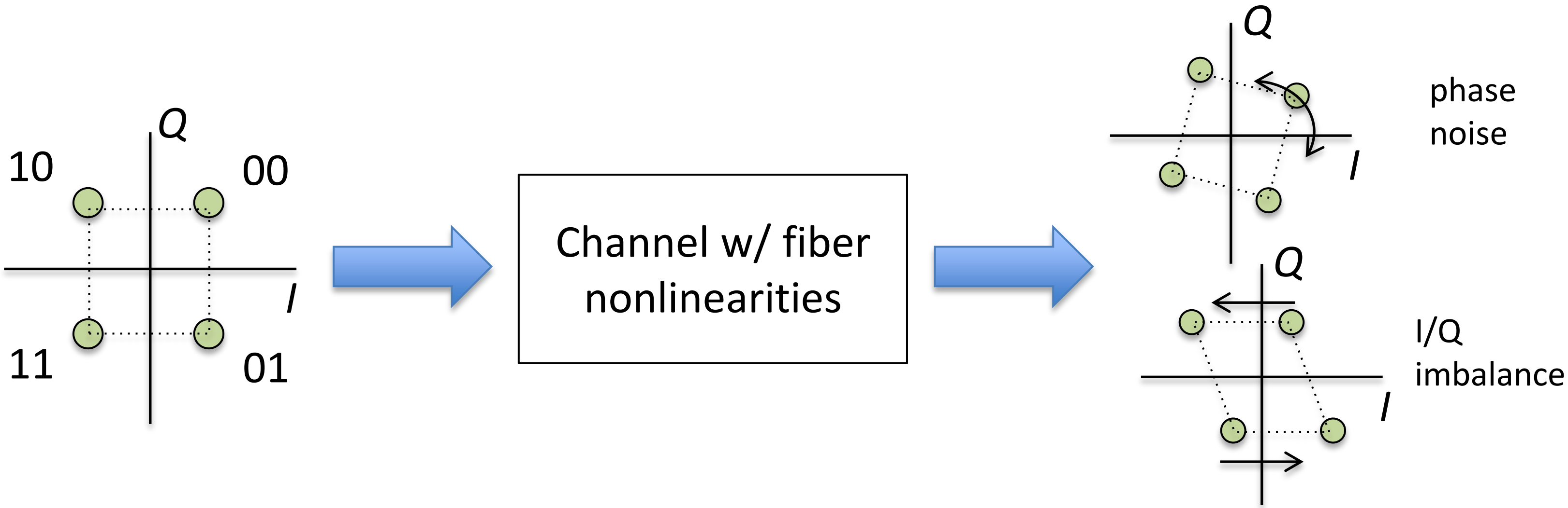
- Traditional MFI requires prior information exchange between end points (from upper layer protocols)
  - additional delay for in signal detection
- ML enables automated MFR from features of the received signal

Khan et al., "Modulation Format Identification in Coherent Receivers Using Deep Machine Learning", Photonics Technology Letters, vol. 28 n. 17, Sep. 2016

Khan et al., "Non-data-aided joint bit-rate and modulation format identification for next-generation heterogeneous optical networks", Optical Fiber Technology, vol. 20 n. 2, Mar. 2014

Tan et al., "Simultaneous Optical Performance Monitoring and Modulation Format/Bit-Rate Identification Using Principal Component Analysis", Journal of Optical Communications and Networking, vol. 6 n. 5, May 2014

- Optical signals are affected by fiber nonlinearities
  - Kerr effect, self-phase modulation (SPM), cross-phase modulation (XPM)…

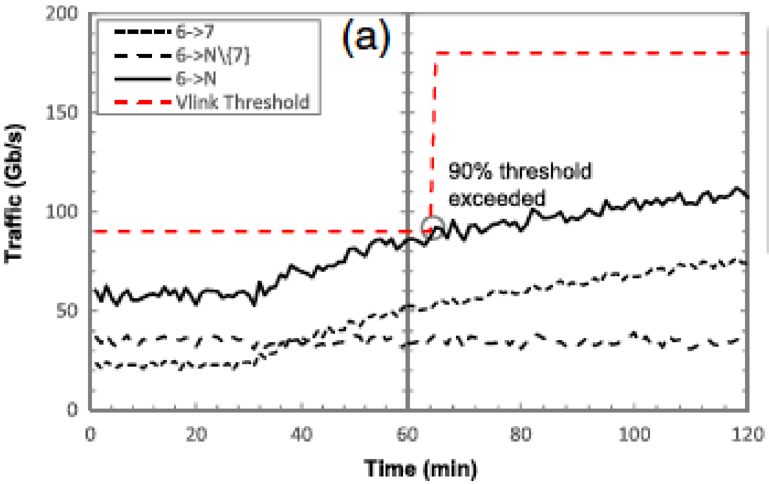


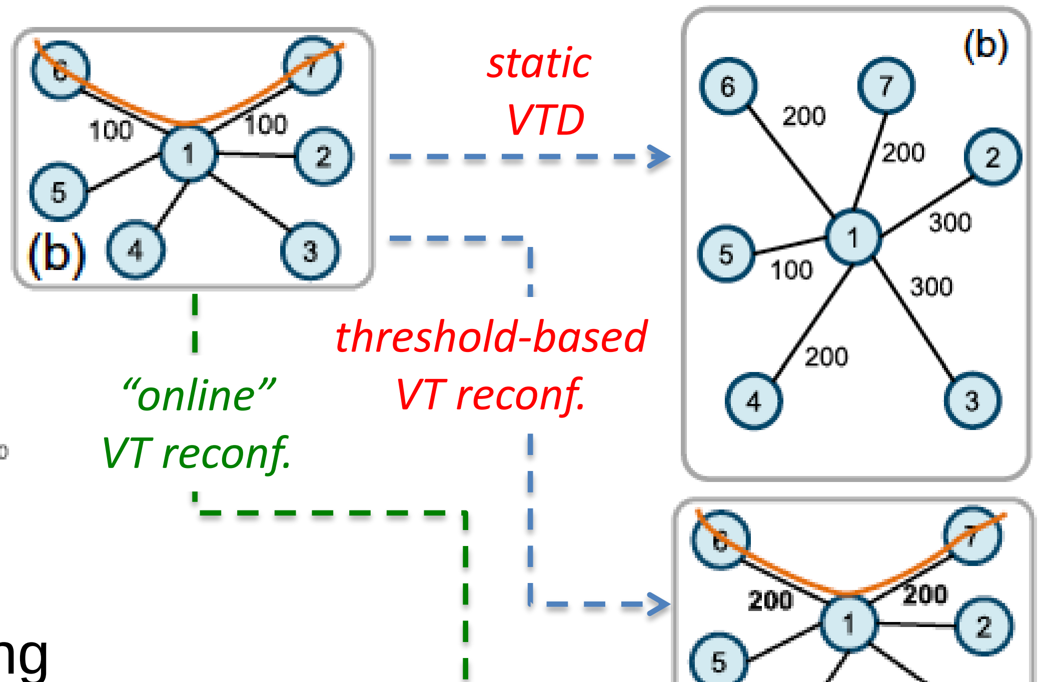
- Traditional methods require complex mathematical models and prior information on the traversed channel
- ML enables "safer" decision by learning from actual channel properties

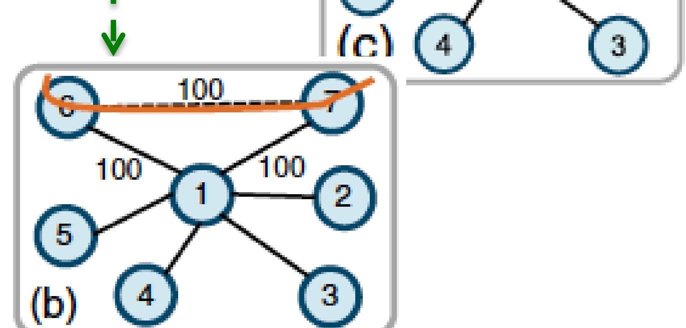
Wang et al., "Nonlinear Decision Boundary Created by a Machine Learning-based Classifier to Mitigate Nonlinear Phase Noise", in ECOC 2015, Sep. 2015

Wang et al., "Nonlinearity Mitigation Using a Machine Learning Detector Based on k-Nearest Neighbors", Photonics Technology Letters, vol. 28 n. 19, Oct. 2016

- New services with high spatio-temporal traffic dynamics







- No reconfiguration → peak-traffic dimensioning
- ML leverages *online* (live) traffic monitoring/prediction to avoid overprovisioning

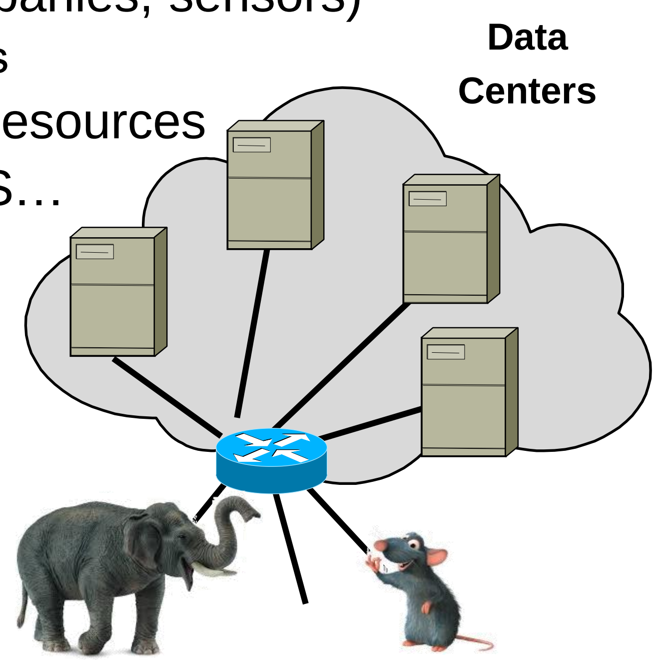
Morales et al., "Virtual Network Topology Adaptability Based on Data Analytics for Traffic Prediction", Journal of Optical Communication and Networking, vol. 9 n. 1, Jan. 2017

Alvizu et al., "Matheuristic with machine learning-based prediction for software-defined mobile metro-core networks", Journal of Optical Communication and Networking, vol. 9 n. 9, Sep. 2017

Network layer
Flow classification

- Traffic flows can be heterogeneous in terms of:
  - protocols (http, ftp, smtp…)
  - services (VoD, data transfer, text messages…)
  - requirements (latency, bandwidth, jitter…)
  - network “customers” (human end-users, companies, sensors)
    - E.g., “mice” vs “elephant” flows in Data Centers
- Distinguish between different flows is crucial for resources (i.e., capacity) allocation, scheduling, SLAs, QoS…

- ML enables traffic classification from direct observation of traffic flows



L. Wang, X. Wang, M. Tornatore, K. Joon Kim, S.-M. Kim, D.-U Kim, K.-E. Han, and B. Mukherjee, “Scheduling With Machine-Learning-Based Flow Detection for Packet-Switched Optical Datacenter Networks, JOCN2018

Viljoen et al., “Machine Learning Based Adaptive Flow Classification for Optically Interconnected Data Centers”, in ICTON 2016, July 2016

Cao et al., “An accurate traffic classification model based on support vector machines”, International Journal on Network Management, 27:e1962, 2017.

Conclusion: My personal opinion

- If you are starting now: not a fast learning curve if you do not simply want to use machine learning as a black box

- Promising directions:
  - QoT estimation
    - Partly. Ok for improving accuracy, or when unknowns are too many
  - Failure management
    - Yes! Root cause analysis (it is a complex semisupervised problem!)
  - Traffic prediction
    - Yes! (Check DC-NN*)
  - Resource allocation (e.g., dynamic traffic allocation)
    - Skeptical
      - Several problems (traffic varies, scalability…)

*D. Andreoletti, S. Troia, F. Musumeci, S. Giordano, G. Maier, M. Tornatore, «Network Traffic Prediction based on Diffusion Convolutional Recurrent Neural Networks», Infocom 2019

Thank You!

..and thanks to them!

POLITECNICO
MILANO 1863

Achille Pattavina
Francesco Musumeci
Luca Barletta

THE UNIVERSITY OF CALIFORNIA DAVIS

Biswanath Mukherjee

Filppo Cugini (CNIT)
Cristina Rottondi (PoliTo)
Dario Azzimonti, Alessandro Giusti, Matteo Salani (Dalle Molle Inst.)

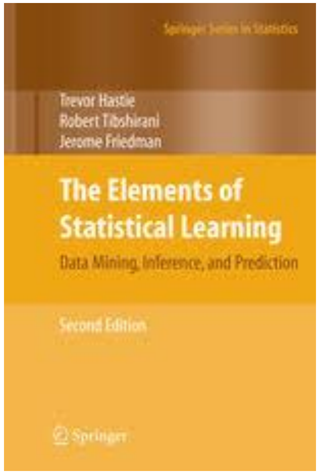
H2020 Metro Haul

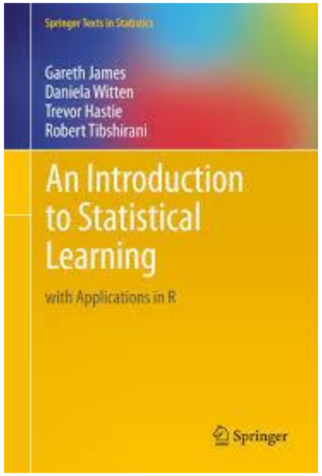
European Commission

Horizon 2020
European Union funding
for Research & Innovation

Some material

- Books (general refs. for ML):
  - T. Hastie, R. Tibshirani, J. Friedman, “The Elements of Statistical Learning”, Ed. Springer
  - G. James, D. Witten, T. Hastie, R. Tibshirani, “An Introduction to Statistical Learning with Applications in R”, Ed. Springer





- Prof. Andrew Ng lectures (Stanford University)
- … Google it!

Some publications (1)

Surveys

- F. Musumeci et al., “A Survey on Application of Machine Learning Techniques in Optical Networks”, Submitted to IEEE communication surveys and tutorials
- Javier Mata, et a., Artificial intelligence (AI) methods in optical networks: A comprehensive survey, Optical Switching and Networking, Volume 28, 2018, pp. 43-57

Some Motivations

- Y. Pointurier, "Design of low-margin optical networks," in *IEEE/OSA Journal of Optical Communications and Networking*, vol. 9, no. 1, pp. A9-A17, Jan. 2017. doi: 10.1364/JOCN.9.0000A9

QoT estimation

- Barletta et al., “QoT Estimation for Unestablished Lighpaths using Machine Learning”, in OFC 2017 Conference, Mar. 2017
- De Miguel *et al.*, “Cognitive Dynamic Optical Networks”, *Journal of Optical Communication and Networking*, vol. 5, n. 10, Oct. 2013
- Thrane *et al.*, “Machine Learning Techniques for Optical Performance Monitoring From Directly Detected PDM-QAM Signals”, *Journal of Lightwave Technology*, vol. 35, n. 4, Feb. 2017
- Caballero *et al.*, “Experimental demonstration of a cognitive quality of transmission estimator for optical communication systems”, *Optics Express*, vol. 20, n. 26, Dec. 2012
- Jimenez *et al.*, “A Cognitive Quality of Transmission Estimator for Core Optical Networks”, *Journal of Lightwave Technology*, vol. 31, n. 6, Mar. 2013
- Angelou *et al.*, “Optimized Monitor Placement for Accurate QoT Assessment in Core Optical Networks”, *Journal of Optical Communication and Networking*, vol. 4, n. 1, Jan. 2012

Some publications (2)

Failure recovery

- S. Shahkarami, F. Musumeci, F. Cugini, M. Tornatore, \Machine-Learning-Based Soft-Failure Detection and Identi cation in Optical Networks,"in Proceedings, OFC 2018, San Diego (CA), Usa, Mar. 11-15, 2017
- A. Vela *et al.*, “Soft Failure Localization during Commissioning Testing and Lightpath Operation”, *Journal of Optical Communication and Networking*, vol. 10 n. 1, Jan. 2018
- A. Vela *et al*., “BER degradation Detection and Failure Identification in Elastic Optical Networks”, in Journal of Lightwave Technology, vol. 35, no. 21, pp. 4595-4604, Nov.1, 1 2017

Others

- E Seve, J Pesic, C Delezoide, A Giorgetti, A Sgambelluri, N Sambo, “Automated Fiber Type Identification in SDN-Enabled Optical Networks, Journal of Lightwave Technology 37 (7), 1724-1731, 2019

Projects

- EU ORCHESTRA and CHRON projects

Some emerging concepts (I)

- **Active Learning**
  - *No explicit separation between training and testing*, continuos training as new data arrives
  - Great in situation where data is scarce expensive

- P. Deisenroth, D. Fox, and C. E. Rasmussen, “Gaussian processesfor data-efficient learning in robotics and control,”IEEE Transactions onPattern Analysis and Machine Intelligence, vol. 37, no. 2, pp. 408–423,Feb 2015
- D. Azzimonti, C. Rottondi, and M. Tornatore, “Using Active Learning to Decrease Probes for QoT Estimation in Optical Networks,” in Optica lFiber Communications Conference (OFC), Mar. 2019

- **Transfer Learning**
  - *Is the training performed over a network/link/failure still valid on a different scenario?*

- ftp://ftp.cs.wisc.edu/machine-learning/shavlik-group/torrey.handbook09.pdf
- W Mo, YK Huang, S Zhang, E Ip, DC Kilper, Y Aono, T Tajima, «ANN-based transfer learning for QoT prediction in real-time mixed line-rate systems», in Optical Fiber Communications Conference (OFC), Mar. 2019

Some emerging concepts (II)

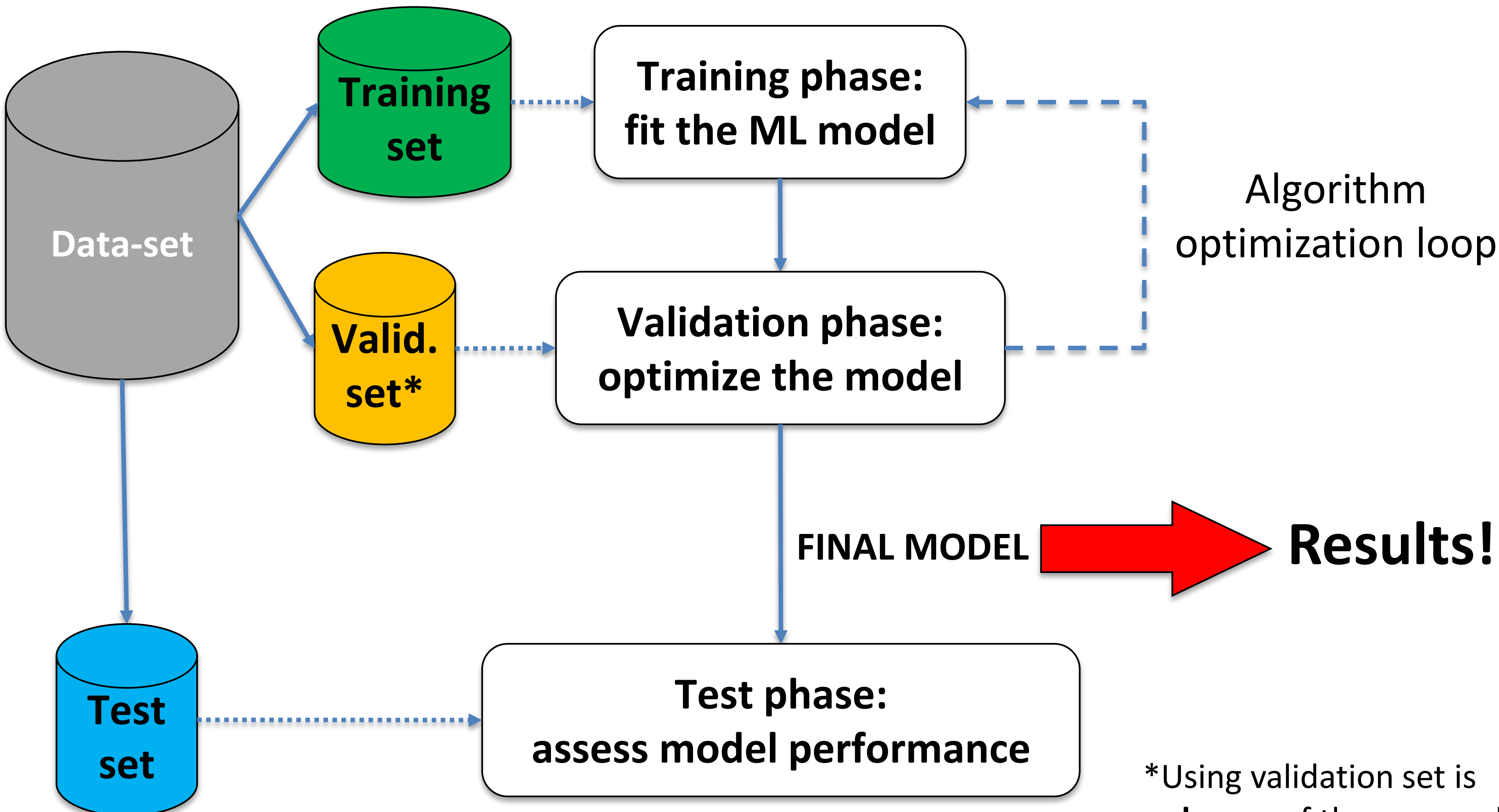
- **Interpretability**
  - Go beyond black-box machine learning outcome!
  - Can we gain insights on our problems

- H.J. Escalante, I. Guyon, S. Escalera X. Baro, Y. Gucluturk, U. Guclu and M. van Gerven, Explainable and Interpretable Models in Computer Vision and Machine Learning, *Springer Series on Challenges in Machine Learning*, 2018.
- F. N. Khan, Q. Fan, C. Lu and A. P. T. Lau, "An Optical Communication's Perspective on Machine Learning and Its Applications," in Journal of Lightwave Technology, vol. 37, no. 2, pp. 493-516, 15 Jan.15, 2019.

- **ML for Optics vs. Optics for ML**
  - See, e.g., in this conference
    - [MG1-1] Image Classification with a 3-Layer SOA-Based Photonic Integrated Neural Network, Bin Shi, et al.

Use of data in Machine Learning

Training, testing, validation



*Using validation set is **only one** of the approaches to optimize a ML algorithm