ITU

ML APPLICATIONS FOR 5G NETWORK ENHANCEMENT

Prof. Ridha Hamila

Electrical Engineering Department

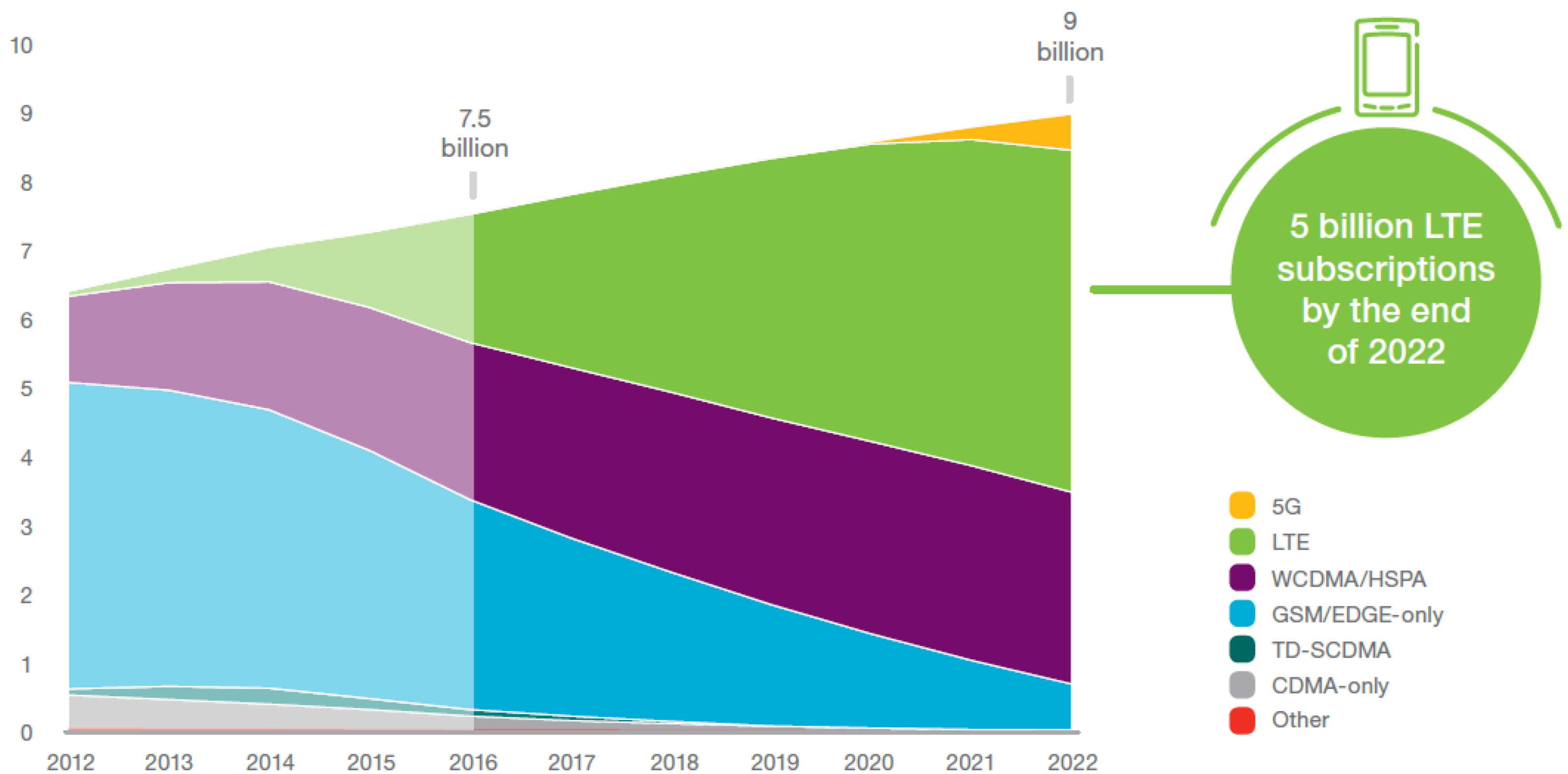
كلية الهندسة
College of Engineering
QATAR UNIVERSITY جامعة قطر

OUTLINE

1. Where we are heading?
2. What are the Requirements of 5G generation?
3. 5G Open Issues and Trends
4. Machine Learning (ML) Overview
5. ML for 5G Challenges
6. ML Applications for Wireless Networks
7. Summary: Challenges and Research Directions

Mobile subscriptions by technology (billion)

- Globally, traffic in mobile networks increased by 70% between from Q1 2016 – Q1 2017



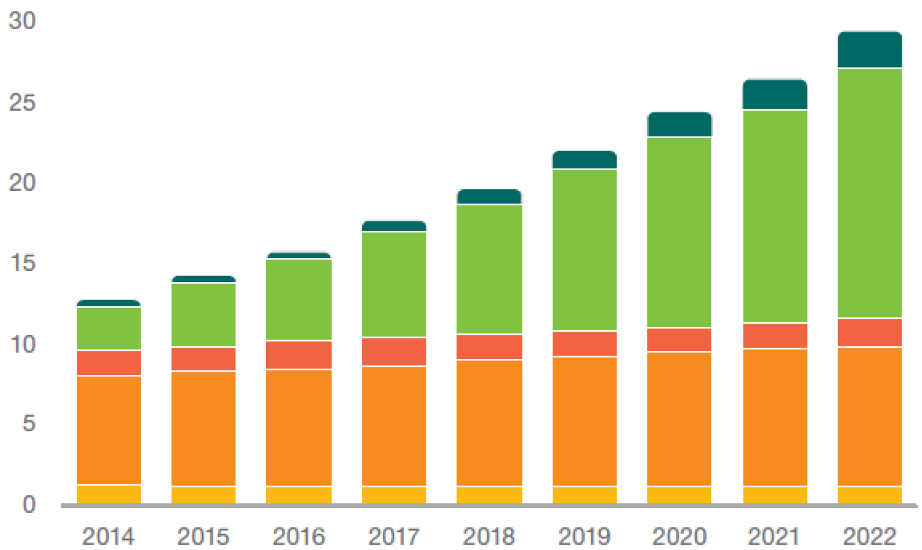
[1] Non-Standalone 5G NR will utilize the existing LTE radio and Evolved Packet Core network as an anchor for mobility management and coverage, while adding a new 5G radio access carrier to enable certain 5G use cases starting in 2019

Figure note: IoT connections and Fixed Wireless Access (FWA) subscriptions are not included in the above graph

ERICSSON MOBILITY REPORT JUNE 2017

IoT OUTLOOK

- IoT devices with cellular connections end 2016 is ~0.4 billion
- Forecast: IoT device connections by 2022 is 18 billions !

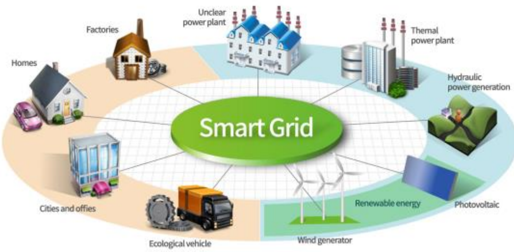


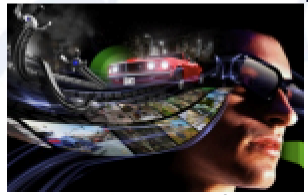
	2016	2022	CAGR
Wide-area IoT	0.4	2.1	30%
Short-range IoT	5.2	15.5	20%
PC/laptop/tablet	1.6	1.7	0%
Mobile phones	7.3	8.6	3%
Fixed phones	1.4	1.3	0%
	16 billion	29 billion	

ERICSSON MOBILITY REPORT JUNE 2017

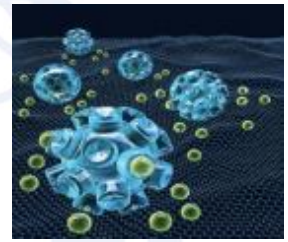
WHERE WE ARE HEADING?

Autonomous driving





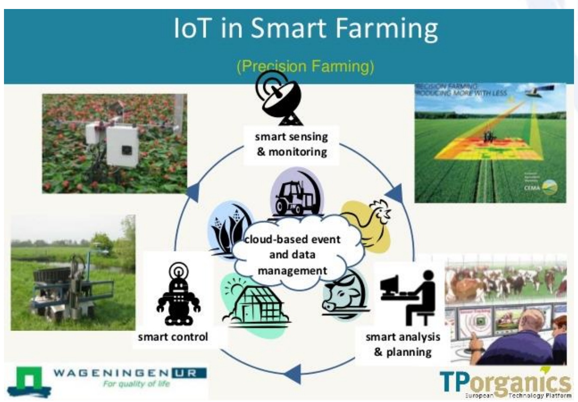
Extreme HD video steaming

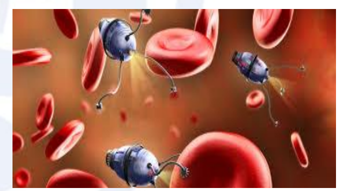


Nano IoT

Holographic watch

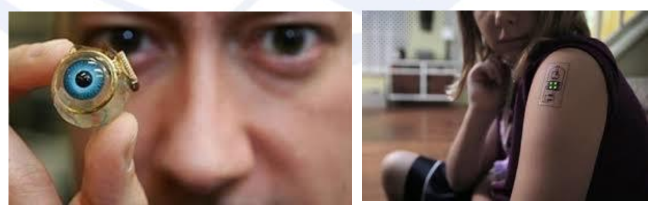
Platforms for Telemedicine





Nano Swarms

Haptic holography



Implantable wearables

In body networks

2018 2020 2024 2028 2030 2034 2050

How can we handle this terrific amount of Network Traffic?

➢ **Real Revolution to the current Mobile broadband wireless network infrast. !!**

SOLUTION:

- Novel intelligent and standalone Radios and Network equipment's having cognitive functions capable to mimic and learn from their environments and act on their own without humans intervention.
- Intelligent and self-sustained Radios and network elements !!

EXPECTATIONS FROM 5G

- Ability to handle 10,000 times more data/call traffic than 3G/4G
- Data download speeds to be 100’s times more than 4G
    - Pages will upload almost instantly
- Connects 100’s of thousands of simultaneous wireless nodes
- Improve spectrum utilization (bits/Hertz/unit area)
- Extreme data rates (more than 1 Gbps)
- Extreme capacity
- Finest Quality Of Service(QOS)
- Support interactive multimedia, voice, streaming video, Internet, and other broadband services etc.

NEXT GENERATION 5G REQUIREMENTS

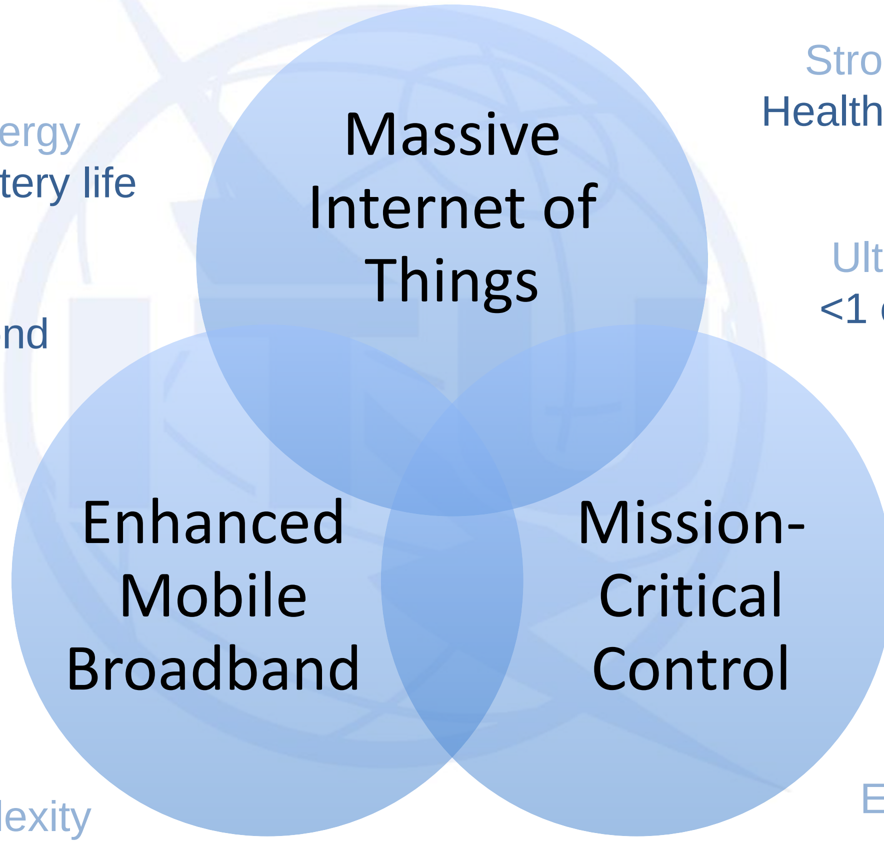
Deep coverage
To reach challenging locations

Strong security
Health/ government

Ultra low energy
10 years of battery life

Extreme data rates
Multi-gigabits per second

Ultra high reliability
<1 out of 100 million packets lost



Ultra high density
1 million nodes per km^2

Ultra low latency
As low as millisecond

Extreme capacity
10 Tbps per km^2

Extreme user mobility

Ultra low complexity
10s of bits per second

Deep awareness
Discovery and optimization

Courtesy Qualcomm Technologies

5G – OPEN ISSUES AND TRENDS

Heterogenous multi-tier networks:

- Interference management, dynamic mode selection, unified MAC.

Full-duplex communication:

- Resource management, dynamic mode selection, MAC protocol design.

Energy harvesting:

- Multiuser scheduling, advanced channel acquisition, energy beamforming.

Network virtualization:

- Software-defined networking, resource allocation, mobility management.

Machine Learning Applications

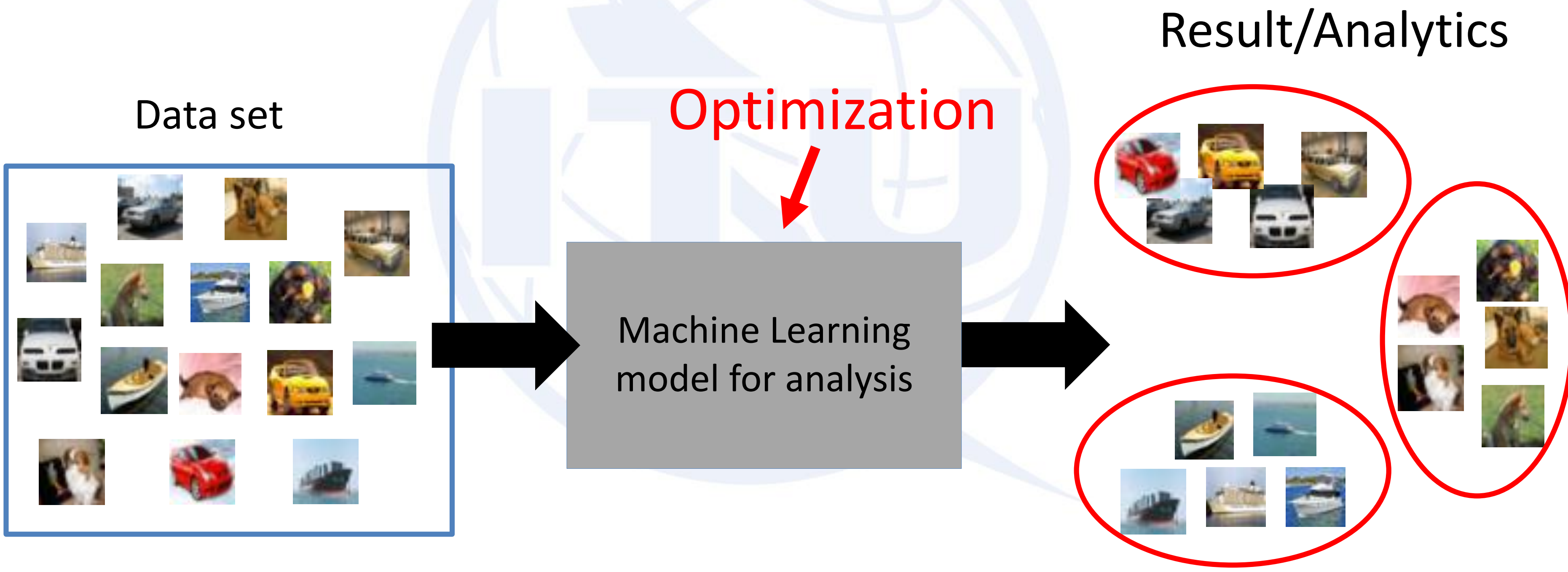
ML field has proved track in areas of:

- Social behavior analysis
- Economics, project management
- Computer vision
- Image recognition
- Speech recognition
- Natural language processing
- Biomedical engineering (MIT ECG benchmark dataset)

➢ Increasing popularity of ML algorithms are gaining attention from several other fields !

Connection of ML to Optimization

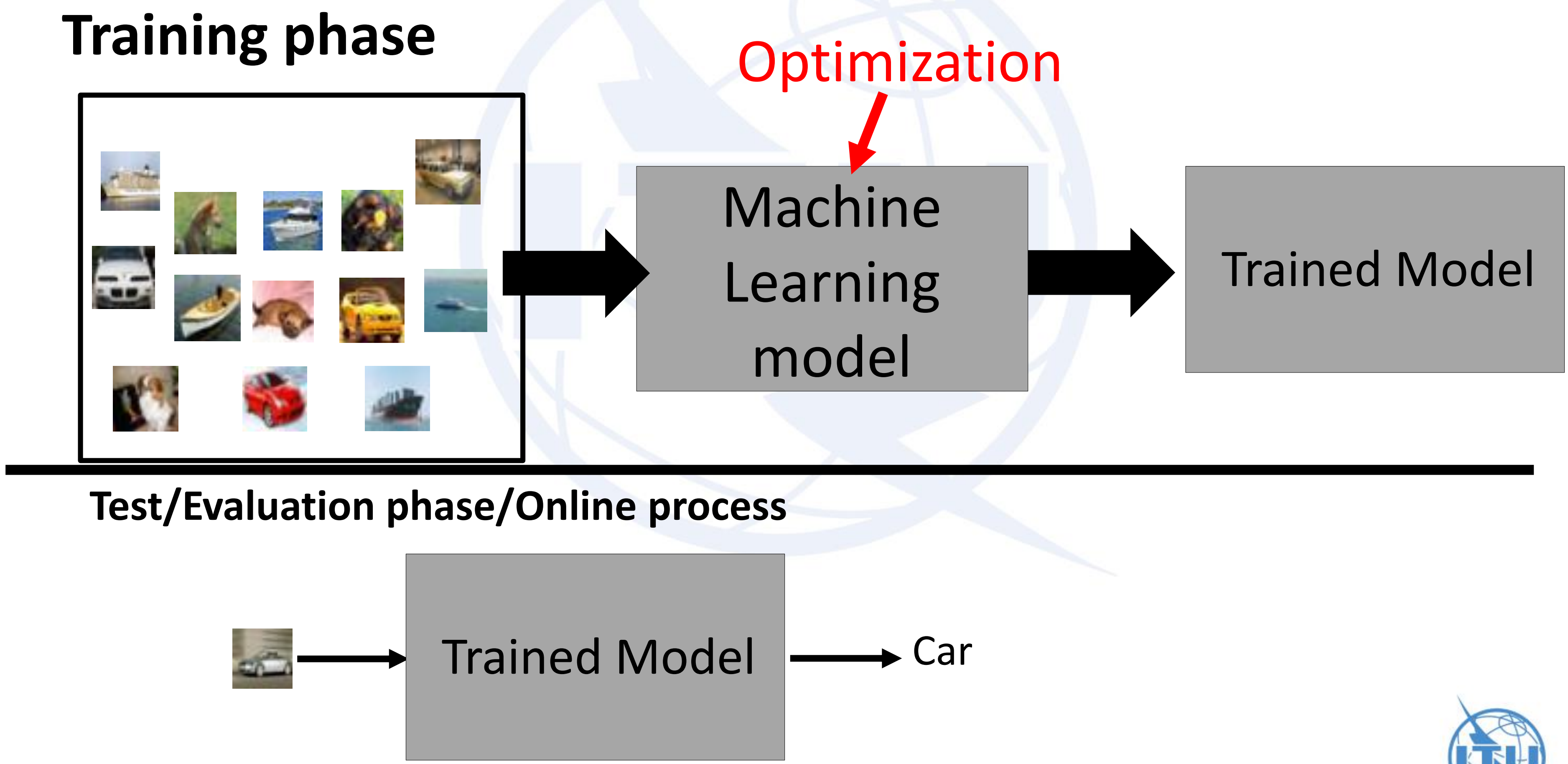
- Data analysis (in-sample analysis)



***Analytics**: discovery, interpretation and communication of meaningful patterns*

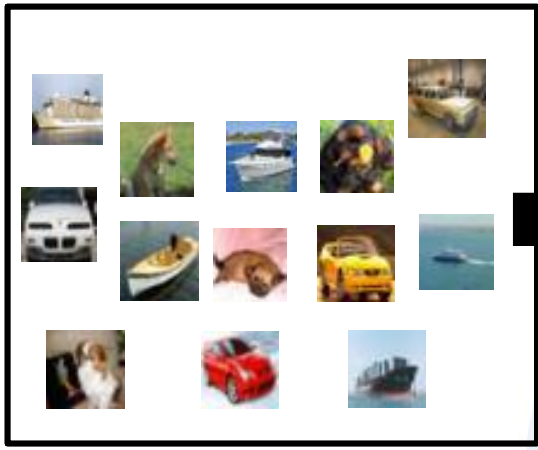
Connection of ML to Optimization

- Learning (out-of-sample generalization)



***Optimization:** selection of Best element from a set with regard to some criterion.*

Bottom-up view of classical ML models



Machine Learning model → Trained Model

Data pre-processing → Data representation → Repr. pre-processing → Model selection/ training

For example:
UE: Throughput, delay, jitter, call set-up,..
Cell Data: RSSI, RSRP, CQI, RSRQ, user/cell..
Network Data: Link utilization, call drop ratio

For example:
- Vector per sample
- Vectors per sample

For example:
- Centering
- Standardization

For example:
- Classification
- Clustering

MACHINE LEARNING FOR 5G CHALLENGES

Exploding traffic volume

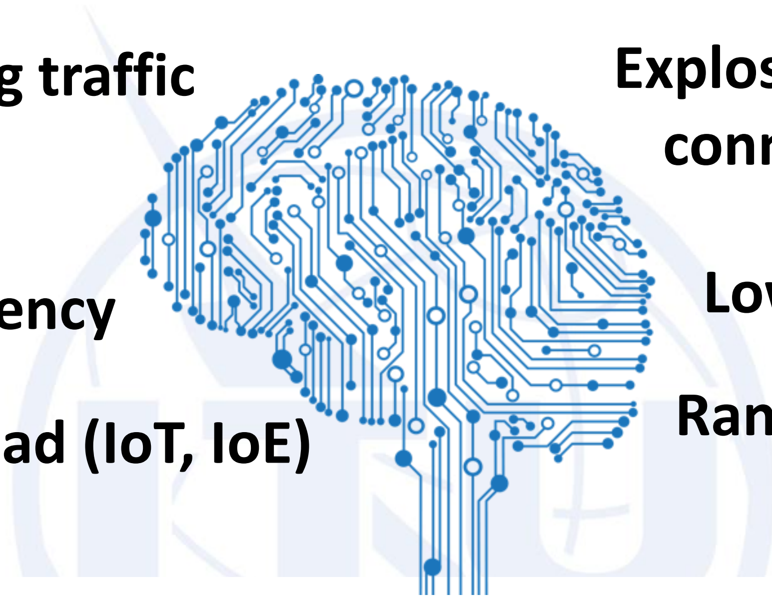
Explosive growth of connected devices

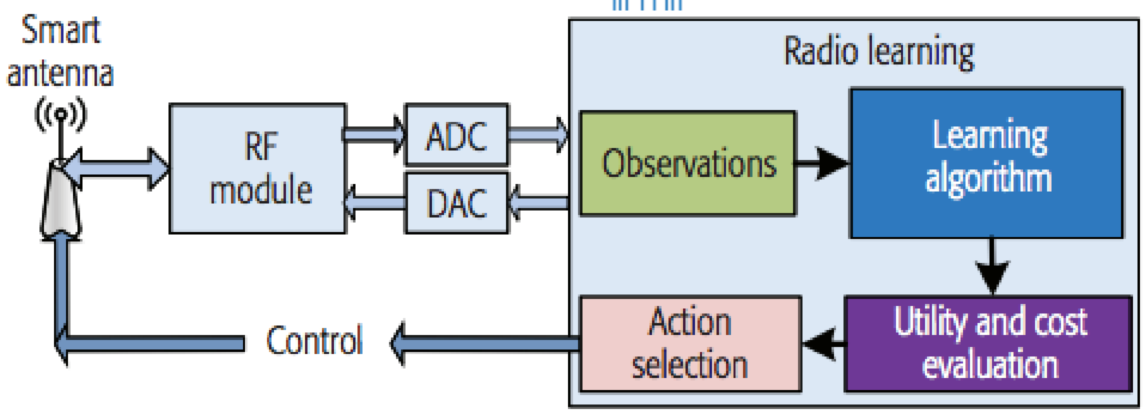
Energy efficiency

Low cost

Control plane load (IoT, IoE)

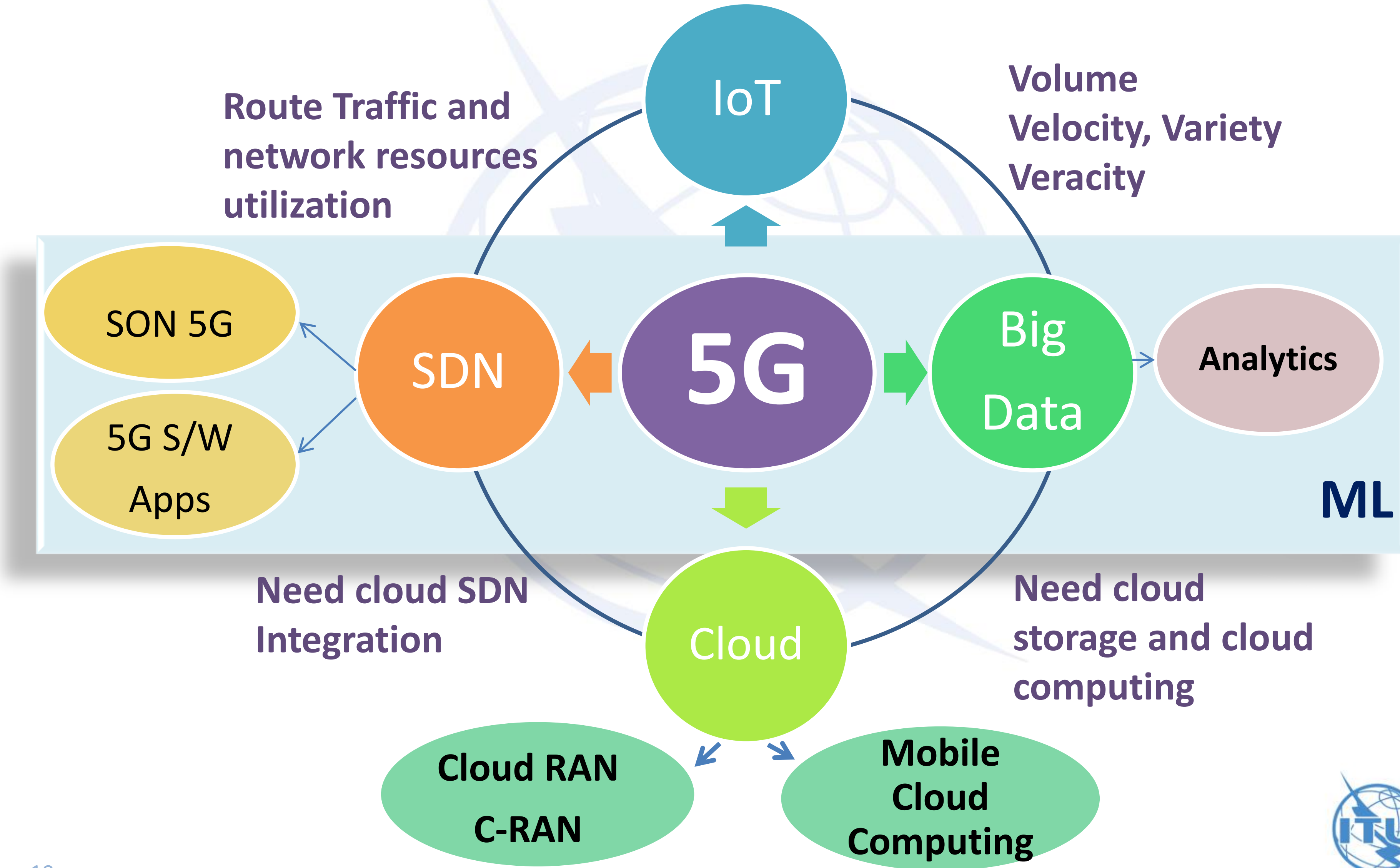
Random and diverse traffic





➢ **ML decisions to solve 5G challenges**

5G Integration to IoT, Big Data, Cloud, and SDN Augmented by ML



MACHINE LEARNING IN 5G

Machine learning in 5G

Supervised learning

Regression model, KNN, SVM app in 5G:
Massive MIMO channel estimation and detection.
User location and behavior learning and classification.

Bayesian learning apps in 5G:
- Massive MIMO channel estimation;
- spectrum sensing, detection and learning in CR.

Deep learning apps:
- Anomaly detection;
- flow aggregation
-application identification.

Unsupervised learning

k-means clustering apps in 5G:
- Small cell clustering.
-WiFi association.
-D2D user clustering.
-HetNet clustering.

PCA and ICA apps in 5G:
- spectrum sensing.
- anomaly, fault and intrusion detection.
- signal dimension reduction.

Reinforcement learning

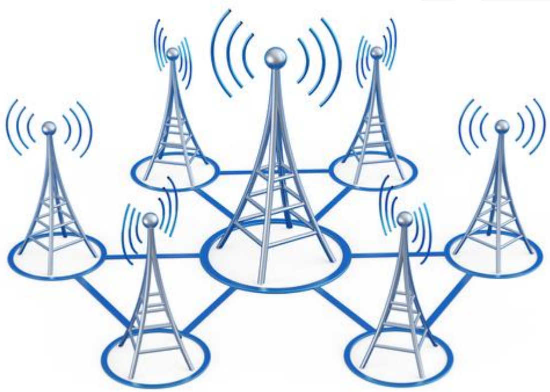
MDP, POMDP, Q-learning multi-armed bandit apps in 5G:
- Decision making under unknown network conditions.
- Resource competition in femto/small cell **channel selection and spectrum sharing for D2D network.**
- Energy modeling in energy harvesting.
- HetNet selection/ association.

PCA (Principal Component Analysis) ICA (Ind. Comp. Analysis)
MDP (Marckov Decision Process)/POMDP (Partially Observable Marckov Decision Process)

1- SUPERVISED LEARNING: **MIMO CHANNEL AND ENERGY LEARNING**

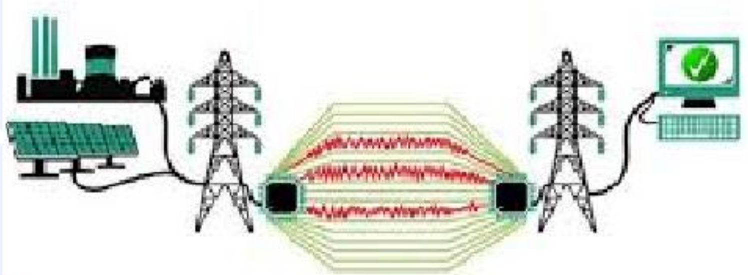
REGRESSION MODELS; KNN AND SVM

Challenge: In massive MIMO systems associated with hundreds of antennas, both **detection and channel estimation lead to high-dimensional search-problems**, which can be addressed by the learning models: KNN, SVM, Regression models.



ML solution: Estimating and predicting radio parameters.

Challenge: Calculating the Gaussian channel's noise level in a MIMO-aided wireless network having "t" transmit antennas and "r" receive antennas.

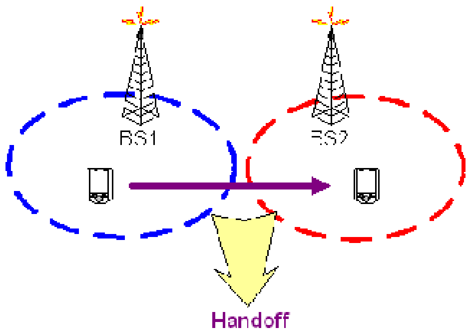


ML solution: By **exploiting the training data**, the **SVM** model was trained **for the estimation of the channel noise statistics**.

1- SUPERVISED LEARNING: **MIMO CHANNEL AND ENERGY LEARNING**

REGRESSION MODELS; KNN AND SVM

Challenge: heterogeneous networks constituted by diverse cells, handovers may be frequent.



ML solution: KNN and SVM can be **applied to find the optimal handover solutions**.

Challenge: Calculating the configuration to be used in specific locations and times.

ML solution: At the application layer, these models can also be used for **learning the mobile terminal's specific usage pattern in diverse spatio-temporal and device contexts**. This may then be exploited for prediction of the configuration to be used in the location-specific interface.

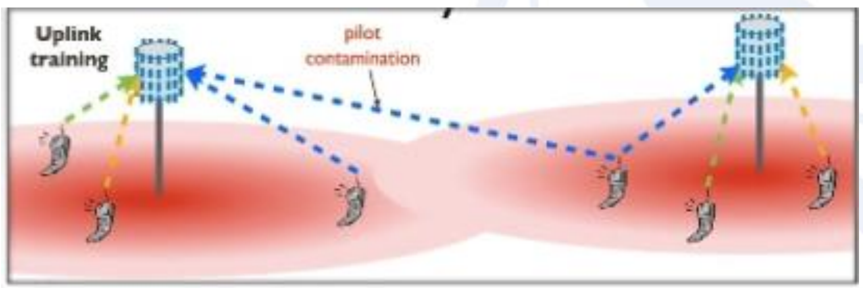
Challenge: Energy efficiency.

ML solution: Machine learning algorithms are capable to learn the user context and preferences. Hence, by dynamically classifying these preferences into a system of state, energy can be saved.

1- SUPERVISED LEARNING: **MASSIVE MIMO AND COGNITIVE RADIO**

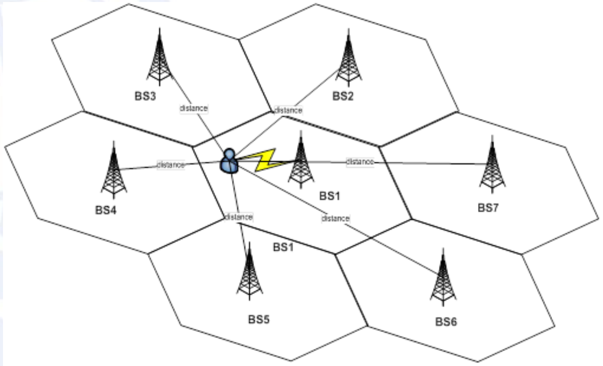
BAYESIAN LEARNING

Challenge: Pilot contamination problem encountered in massive MIMO systems.



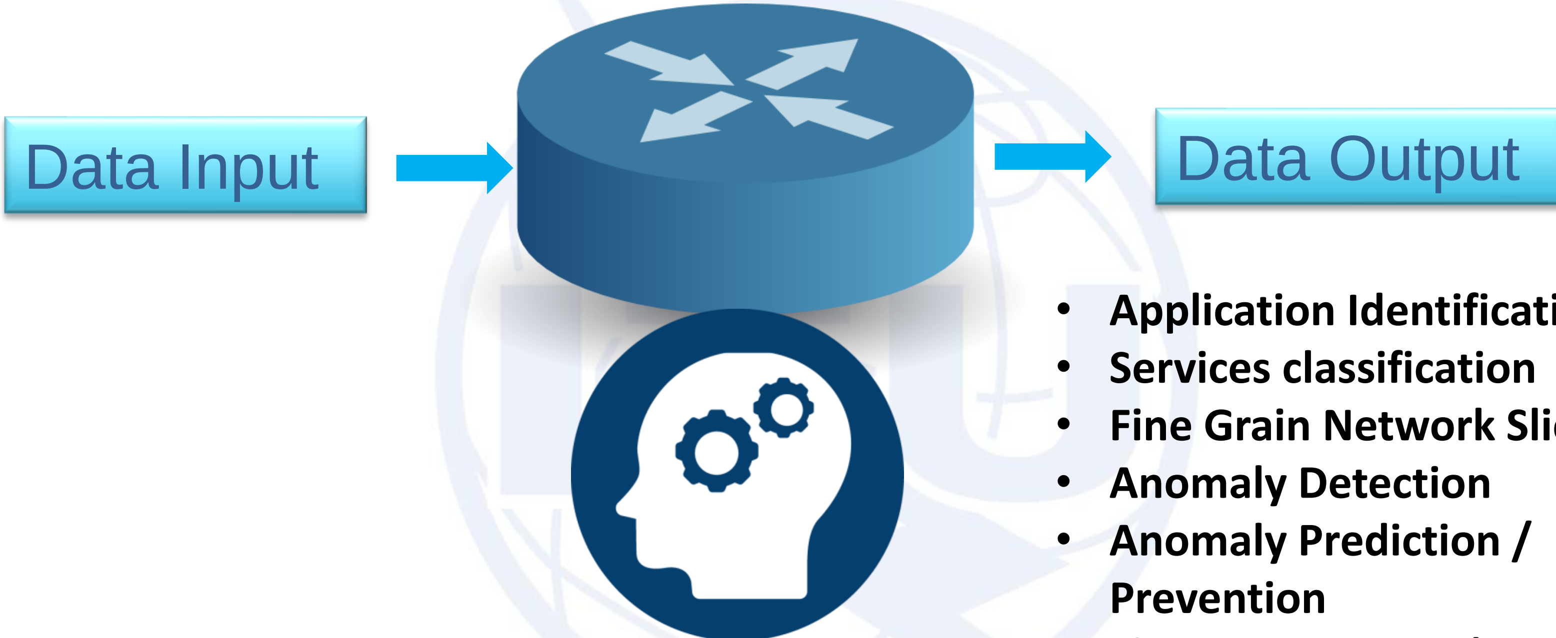
ML solution: we can estimate both the <u>channel parameters of the desired links in a target cell as well as those of the interfering links of the adjacent cells</u>.

Challenge: **Detection of a primary user (PU)** supported by a multi-antenna assisted cognitive radio network.



ML solution: Bayesian learning can be carried out to jointly detecting both the PU signal as well as estimating the channel's unknown frequency response and the noise variance of multiple sub-bands.

1- SUPERVISED LEARNING: IN-NETWORK DEEP LEARNING FOR **TRAFFIC CONTROL AND OPTIMIZATION**



In-network Deep learning

- **Application Identification**
- **Services classification**
- **Fine Grain Network Slicing**
- **Anomaly Detection**
- **Anomaly Prediction / Prevention**
- **Flow Aggregation (Heavy Hitter)**
- **Optimal assignment of resources.**

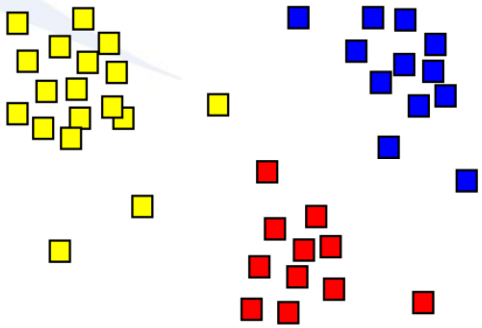
2- UNSUPERVISED LEARNING IN HETEROGENEOUS NETWORKS

K-MEANS CLUSTERING:

Challenge: Clustering is a common problem in 5G networks, especially in heterogeneous scenarios associated with diverse cell sizes as well as WiFi and D2D networks:

- Small cells are clustered to avoid interference using coordinated multi-point transmission (CoMP).
- Mobile users are clustered to obey an optimal offloading policy.
- Devices are clustered in D2D networks to achieve high energy efficiency.
- WiFi users are clustered to maintain an optimal access point association.

ML solution: k-means clustering.



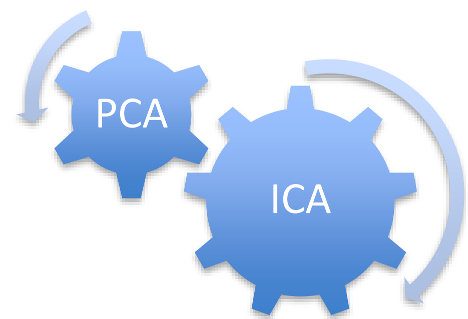
2- UNSUPERVISED LEARNING IN COGNITIVE RADIO:

PRINCIPAL AND INDEPENDENT COMPONENT ANALYSIS (PCA/ ICA)

Challenges and solutions: Both the PCA and ICA constitute powerful **statistical signal processing techniques devised to recover statistically independent source signals from their linear mixtures.**

Application in wireless, sensor and mesh networks:

- Anomaly-detection.
- Fault-detection.
- Intrusion-detection.



Application in cognitive radio networks:

- Distinguishing and characterizing the activities of PUs in the context of collaborative spectrum sensing.
- Observations of the secondary users (SUs).

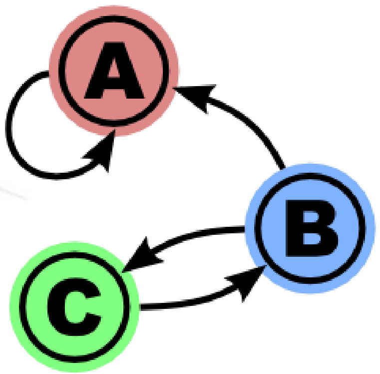
3- REINFORCEMENT LEARNING IN WIRELESS COMMUNICATIONS

MDP/POMDP MODELS

Challenges and solutions: The family of MDP (Marckov Decision Process)/POMDP (Partially Observable Marckov Decision Process) **constitutes ideal tools for supporting decision making in 5G networks**, where the **users** may be regarded as **agents** and the network constitutes the environment.

Classical applications are:

- Network selection/association problems of heterogeneous networks (HetNets).
- Channel sensing.
- User access in cognitive radio networks.
- Energy harvesting (EH).



CHANNEL LEARNING

- Channel learning at device level aids to select the best channel for transmission/reception.
  - **Traditional approach:** use of “pilot symbols”.
  - Waste of bandwidth resources, hardware cost.

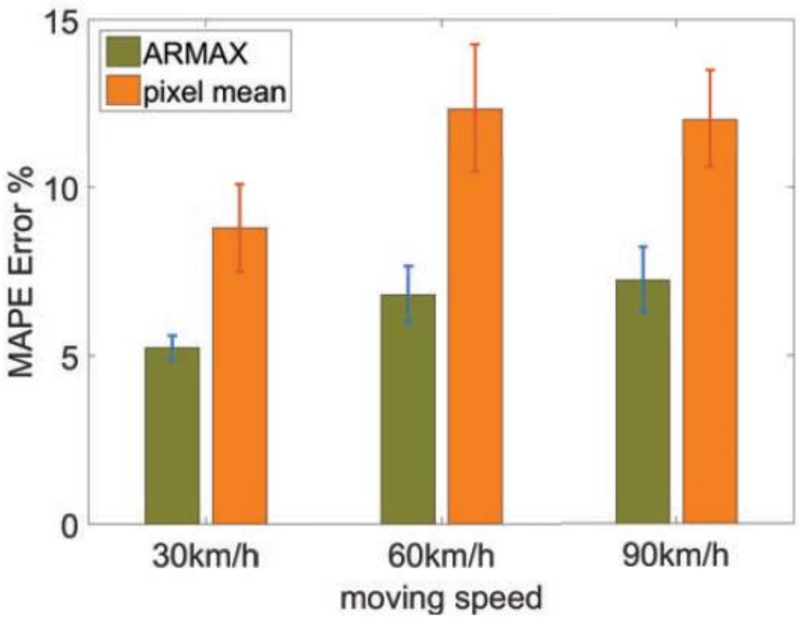
- **Network wide channel maps** in cellular networks are used for **network optimization**.
  - **Traditional approach:** require extensive driving tests.
  - High cost of data acquisition.

Channel learning using ML

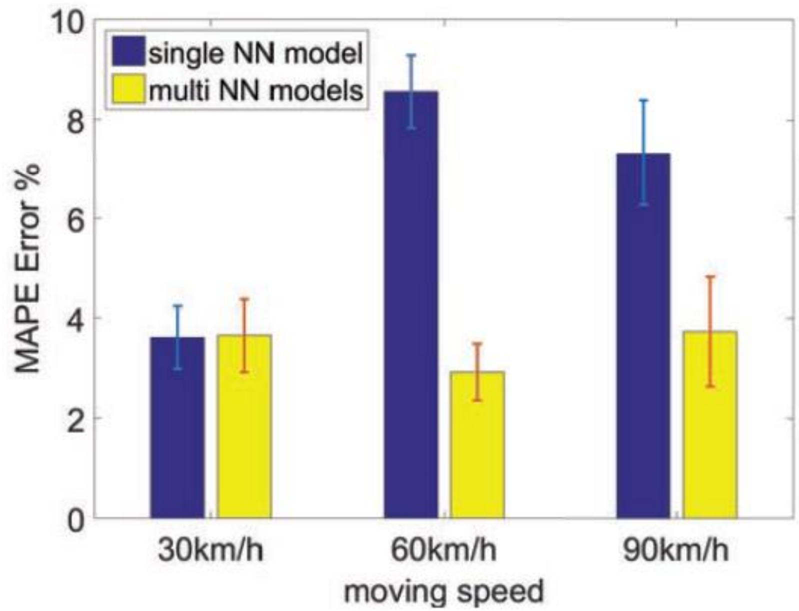
From past CSI and location information [3]

- **Medium term ~** 1 second prediction (moving average versus ARMAX)
- **Long term ~ 5, 10** seconds prediction (Single NN versus Multi NN model)

➢ NN model is more accurate, but more complex



(a) medium term



(b) long term

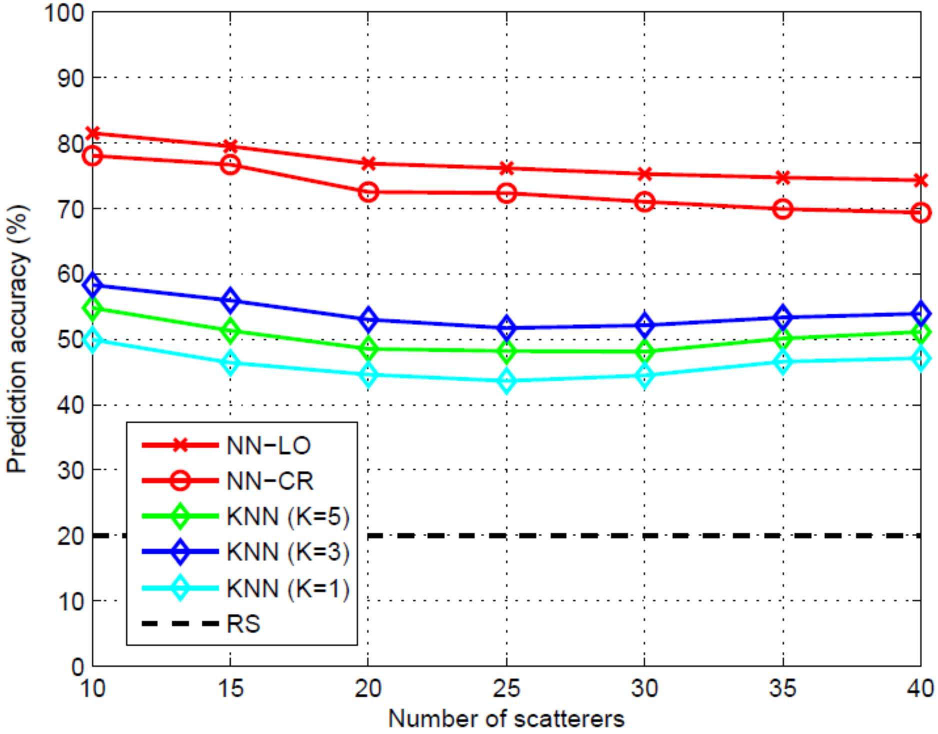
Mean Absolute Percentage Error (MAPE)

Channel learning using ML

From observable CSI and location information [5]

- Predict the unobservable CSI from the observable CSI and location information.

1. RS (Random Selection)
   ~20% Accuracy
2. KNN (K nearest neighbors)
   ~ 50% accuracy
3. NN (neural networks)
   ~73% accuracy



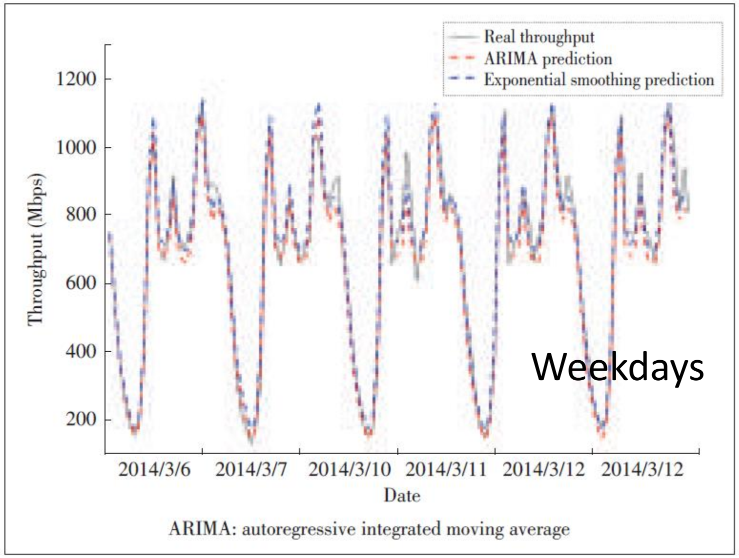
QoS Prediction

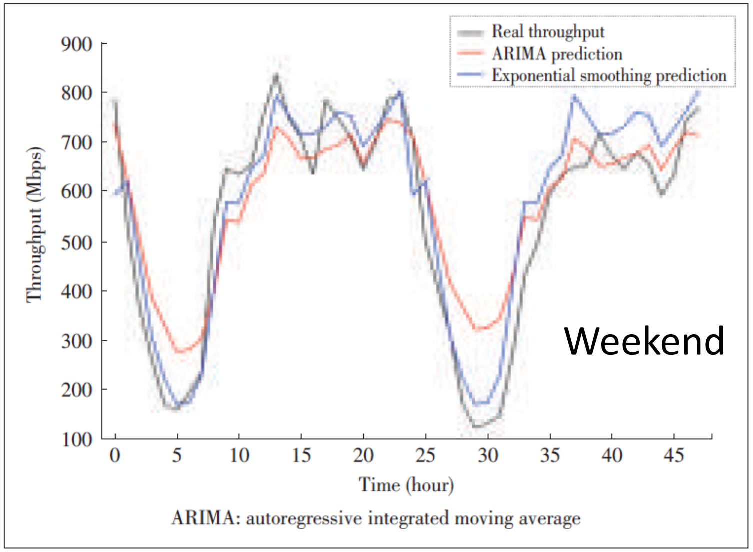
Throughput prediction in cellular networks

- **Traditional methods:** Based on counters and call traces of a live network [6].
- **ML based approaches:**
  1. ARIMA (autoregressive integrated moving average) model [7].
  2. Generalized Linear Models (GLM), Random Forest (RF), Neural Networks (NN) [8].
  3. Multi-layer Perceptrons (MLPs) [9].

QoS Prediction

Network Throughput prediction using ARIMA ML Model [7]





Throughput prediction using RF, GLM and NN Models [8]

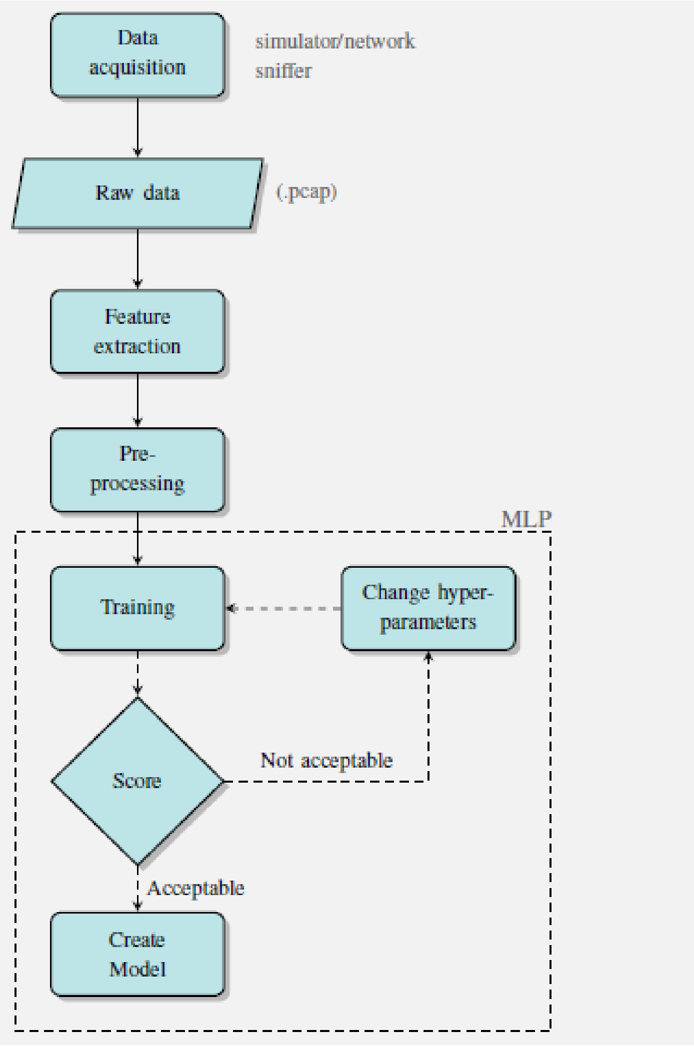
	GLM	GLM 10-fold	NNET	NNET 10-fold	RF	RF 10-fold
Context	0.82	0.80	0.84	0.80	0.84	0.84
Context + RAN data	0.88	0.85	0.89	0.86	0.89	0.88
Context + RAN data + other E2E measurements	0.95	0.93	0.95	0.93	0.95	0.94

TABLE 1: Coefficient of determination (R^2) of models.

QoS Prediction [Contributed]

Throughput prediction using MLPs [9]

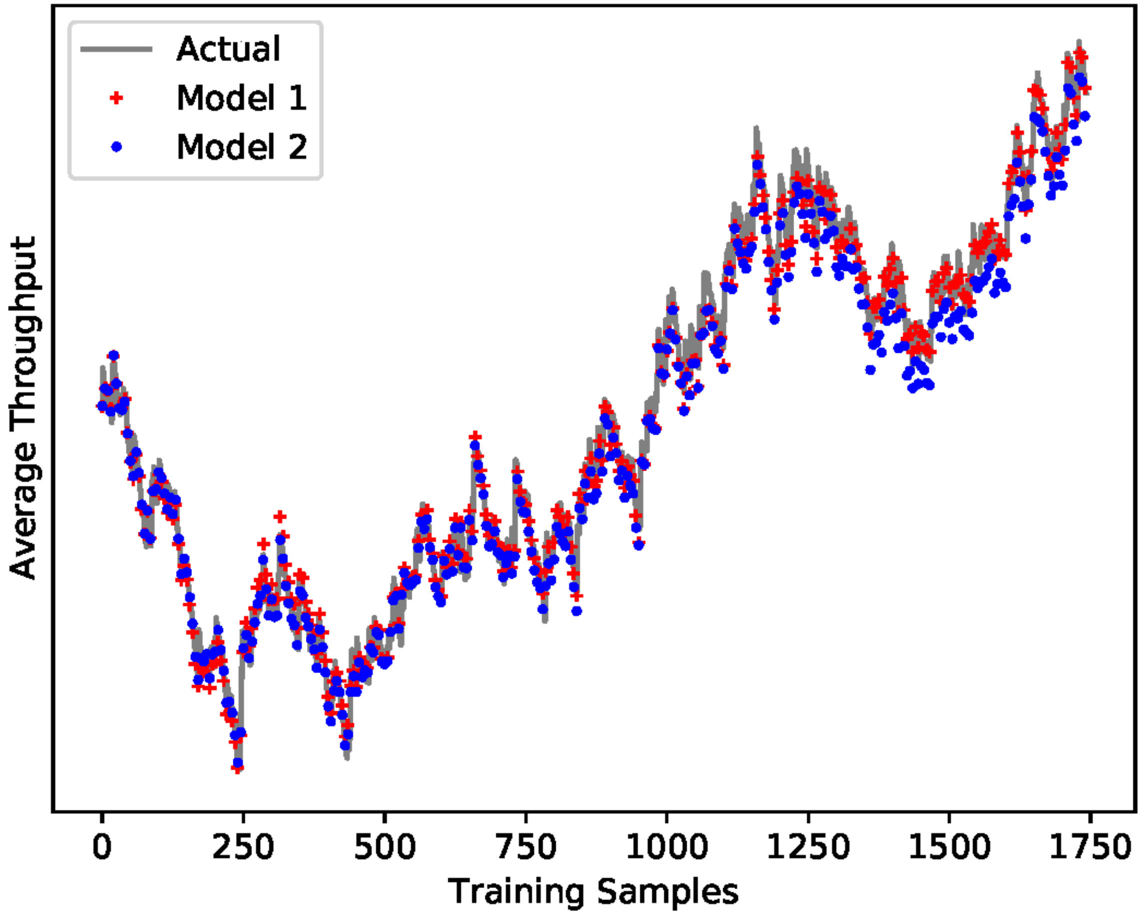
- **Raw traces** acquired through **simulations** and **network sniffing**
- **Feature extraction from packet attributes**
- Pre-processing of training data
- Fully-connected, feed-forward neural networks **Multi-layer Perceptron (MLPs)** are used.
- Exhaustive hyper-parameters search through random-search method to find best possible values of hyper-parameters
- Validate the model performance using metrics such as mean absolute error (MAE), mean squared error (MSE), R-squared (R^2).



Prediction (synthetic dataset)

Mean Squared Error:
Model 1 ~ 0.010
Model 2 ~ 0.012

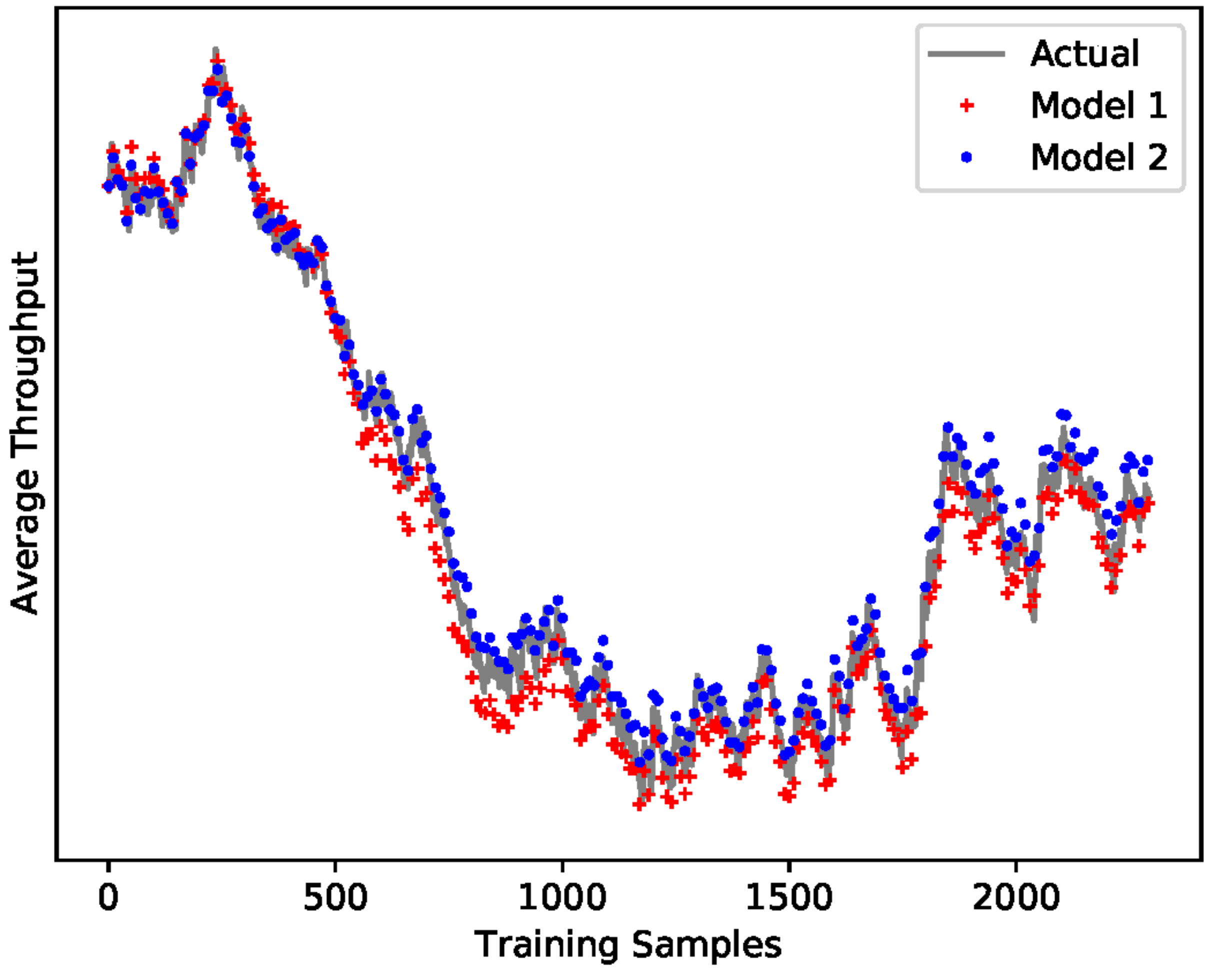
R-Squared:
Model 1 ~ 0.989
Model 2 ~ 0.987



Prediction (real dataset)

Mean Squared Error:
Model 1 ~ 0.010
Model 2 ~ 0.011

R-Squared:
Model 1 ~ 0.99
Model 2 ~ 0.98



Localization

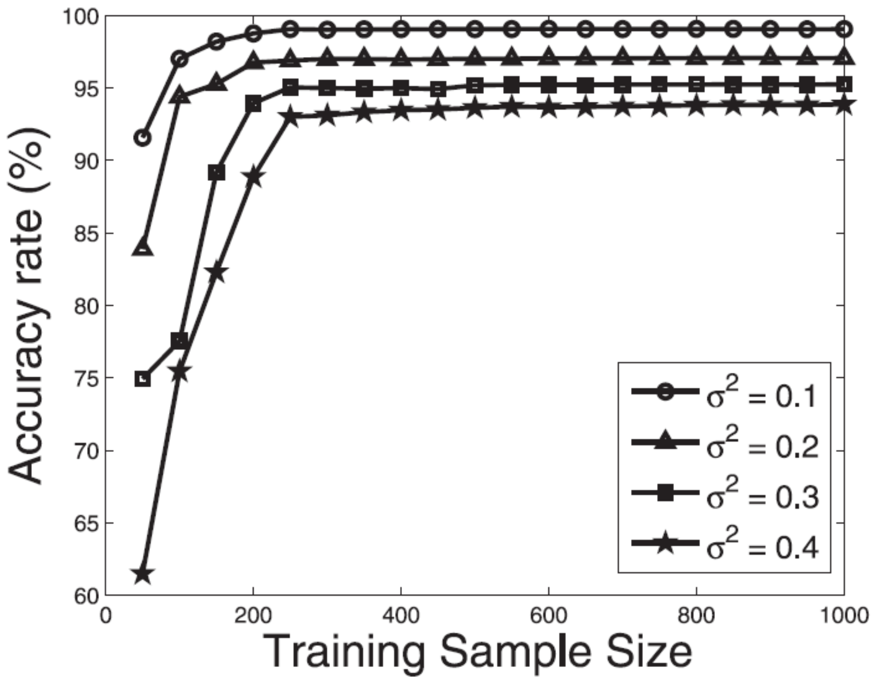
Accurate location information can be useful in several applications, including:

- Scheduling synchronization
- Topology control
- Power control
- Handover prediction
- Proactive caching
- Load balancing
- Network planning

Localization (cont.)

Estimating device's location using Hierarchical Support Vector Machine (H-SVM)

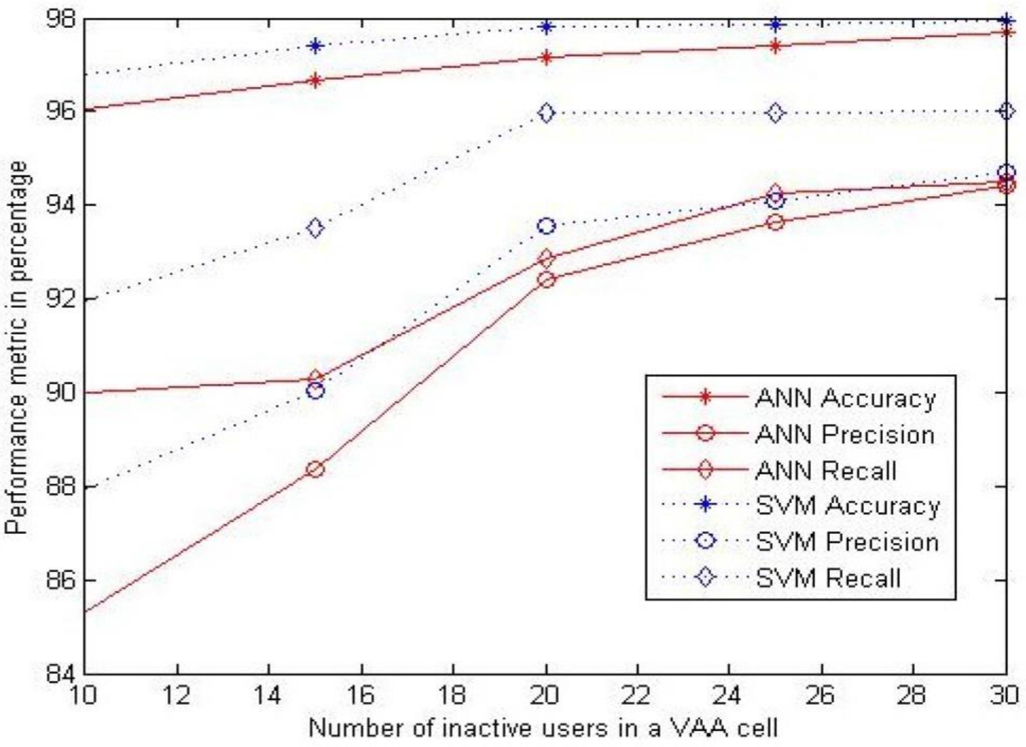
- Accuracy of model for different values of location variance σ^2.
- Accuracy increases with training samples.
- Accuracy decreases with increasing variance of user's location.



Relay Selection

Relay Selection using ANN and SVM[12]

- Reduced node discovery time by 29%.
- SVM shows high accuracy and less error.



SUMMARY

Supervised learning techniques	Unsupervised learning techniques	Reinforcement learning
• Massive MiMo channel estimation and data detection • Spectrum sensing • Inferring the mobile users locations and behaviors to improve the QoS • Hand-over decisions • Applications and services classification	• Cells clustering • Anomaly/ intrusion detection	• Inferring the mobile users decisions • Resource allocation under unknown resource quality

CHALLENGES AND RESEARCH DIRECTIONS

Unavailability of benchmark datasets !!

- Most of the published works uses synthetic data generated through network simulations.
- Simulation data typically pose high accuracy of ML models, which might not be true for real datasets.
- **Standard methods should be established to acquire real network data.**

CHALLENGES AND RESEARCH DIRECTIONS (2)

Data representations

- Common methods to represent **network attributes** in datasets.
- Is the representation equally significant for different ML algorithms?

Evaluation of ML methods over benchmark datasets

- Performance versus Complexity tradeoff !
- New context-specific metrics needs to be defined to evaluate ML algorithm's performance, rather than traditional metrics such as RMSE, precision etc.

CHALLENGES AND RESEARCH DIRECTIONS (3)

ML based Cognitive Radios (CR) and SON

- ML-based algorithms for network optimization needs to be implemented in SDN and CR architectures to enable self-organizing networks (SON).

References

1. P. Demestichas, A. Georgakopoulos, D. Karvounas, K. Tsagkaris, V. Stavroulaki, J. Lu, C. Xiong, and J. Yao, “5G on the horizon: Key challenges for the radio-access network,” IEEE Veh. Technol. Mag., vol. 8, no. 3, pp. 47–53, Sept. 2013.
2. E. Hossain and M. Hasan, "5G cellular: key enabling technologies and research challenges," in IEEE Instrumentation & Measurement Magazine, vol. 18, no. 3, pp. 11-21, June 2015.
3. X. Wang, E. Grinshpun, D. Faucher and S. Sharma, "On Medium and Long Term Channel Conditions Prediction for Mobile Devices," 2017 IEEE Wireless Communications and Networking Conference (WCNC), San Francisco, CA, 2017, pp. 1-6.
4. Z. Sayeed, E. Grinshpun, D. Faucher and S. Sharma, "Long-term application-level wireless link quality prediction," 2015 36th IEEE Sarnoff Symposium, Newark, NJ, 2015, pp. 40-45.
5. J. Liu, R. Deng, S. Zhou and Z. Niu, "Seeing the Unobservable: Channel Learning for Wireless Communication Networks," 2015 IEEE Global Communications Conference (GLOBECOM), San Diego, CA,
6. E. Nan, X. Chu, W. Guo and J. Zhang, "User data traffic analysis for 3G cellular networks," in 8th International ICST Conference on Communications and Networking in China (CHINACOM), Guilin, China, 2013.
7. X. Dong, W. Fan and J. Gu, "Predicting LTE Throughput in Traffic Time series," ZTE COMMUNICATIONS, vol. 13, no. 4, 2015.
8. Samba, Alassane and Busnel, Yann and Blanc, Alberto and Dooze, Philippe and Simon, Gwendal, “Throughput Prediction in Cellular Networks: Experiments and Preliminary Results,” in 1`eres Rencontres Francophones sur la Conception de Protocoles, lE´valuation de Performance et Experimentation des Reseaux de Communication (CoRes2016), 2016.
9. M.A. Khan, R. Hamila, M.S. Kiranyaz, “Accurate prediction of transmission throughput using machine learning” Submitted to IEEE Trans. of Mobile Computing.
10. V. s. Feng and S. Y. Chang, "Determination of Wireless Networks Parameters through Parallel Hierarchical Support Vector Machines," in IEEE Transactions on Parallel and Distributed Systems, vol. 23, no. 3, pp. 505-512, March 2012.
11. R. Daniels and R. W. Heath, "Online adaptive modulation and coding with support vector machines," 2010 European Wireless Conference (EW), Lucca, 2010, pp. 718-724.
12. Sankhe K, Pradhan C, Kumar S, Ramamurthy G. Machine Learning Based Cooperative Relay Selection in Virtual MIMO. arXiv preprint arXiv:1506.01910. 2015 Jun 5.
13. C. Jiang, H. Zhang, Y. Ren, Z. Han, K. C. Chen and L. Hanzo, "Machine Learning Paradigms for Next-Generation Wireless Networks," in IEEE Wireless Communications, vol. 24, no. 2, pp. 98-105, April 2017. doi: 10.1109/MWC.2016.1500356WC