PREDICTIVE ANALYTICS FOR IoT NETWORK CAPACITY PLANNING

Constant WETTE
Ericsson

AGENDA

- Project Goals
- The Problem of Network Capacity Planning
- Machine Learning Process
- Data Collection and Processing
- Traffic Modeling and Capacity Prediction
- Analytics Framework
- Summary

PROJECT GOALS

- Machine Learning in IoT Network Capacity Planning
  - IoT Traffic Modelling
  - Accurately Predict future IoT Traffic load and Network Capacity requirements
  - Network Resources Optimization in IoT context

- .

MACHINE LEARNING PROCESS

Business Understanding

- The Problem

Data Collection

- Offline Traffic data
- Live Traffic data
- Traffic forecast
- OSS & BSS data

Data Preprocessing

- Cleaning, Parsing
- Integration
- Simulation

Traffic Modelling

- Features Engineering
- Model Training
- Predictive modelling
- Model Evaluation

MACHINE LEARNING PROCESS

Business Understanding
- The Problem

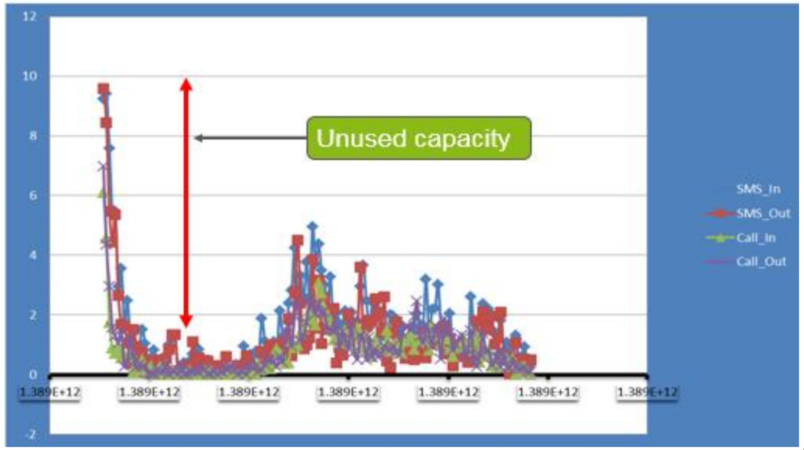
Data Collection
- Offline Traffic data
- Live Traffic data
- Traffic forecast
- OSS & BSS data

Data Preprocessing
- Cleaning, Parsing
- Integration
- Simulation

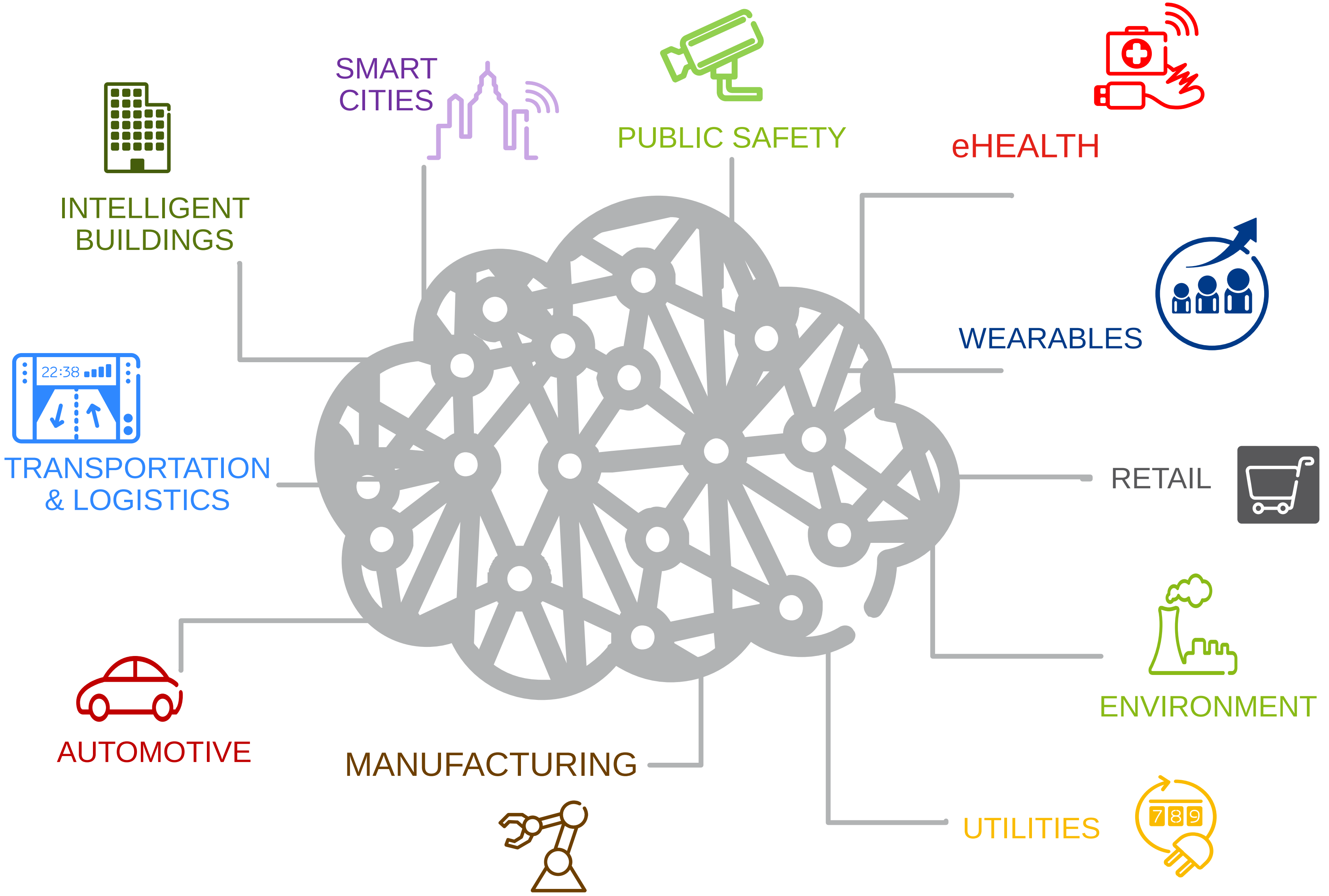
Traffic Modelling
- Data Exploration
- Features Engineering
- Model Training
- Predictive modelling
- Model Evaluation

NETWORK CAPACITY PLANNING

- A methodology used to
  - Model current network state and traffic behavior
  - Project future traffic load and sizing the network resources
  - Predict future capacity issues - congestion, resource shortages
  - Suggest when, where and which resources to add or reconfigure
- Today Network Capacity planning
  - in-house & proprietary tools, spreadsheets
    - not regularly updated
  - based on KPI over the peak period
    - resources are overprovisioned

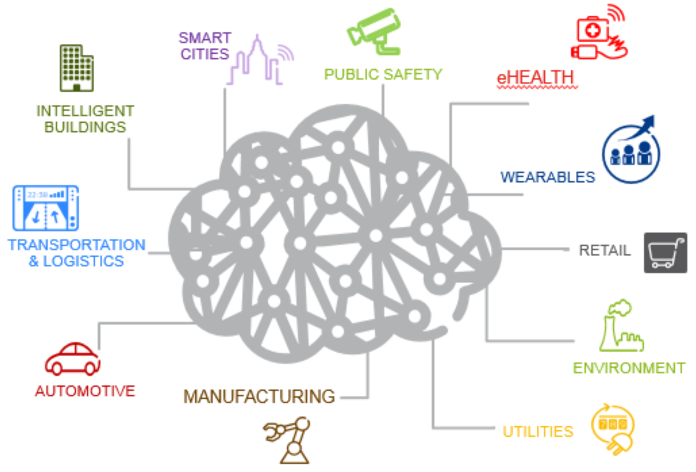


THE INTERNET OF THINGS NETWORK



THE INTERNET OF THINGS NETWORK

- Heterogeneous networks
  - Integration of multiple network standards and protocols
  - Wide range of verticals with specific network requirements
  - Wide variety of devices of different capabilities
- Various traffic models not explored yet
- Traffic volume: IoT devices (25 billions) generate a lot of data



THE PROBLEM

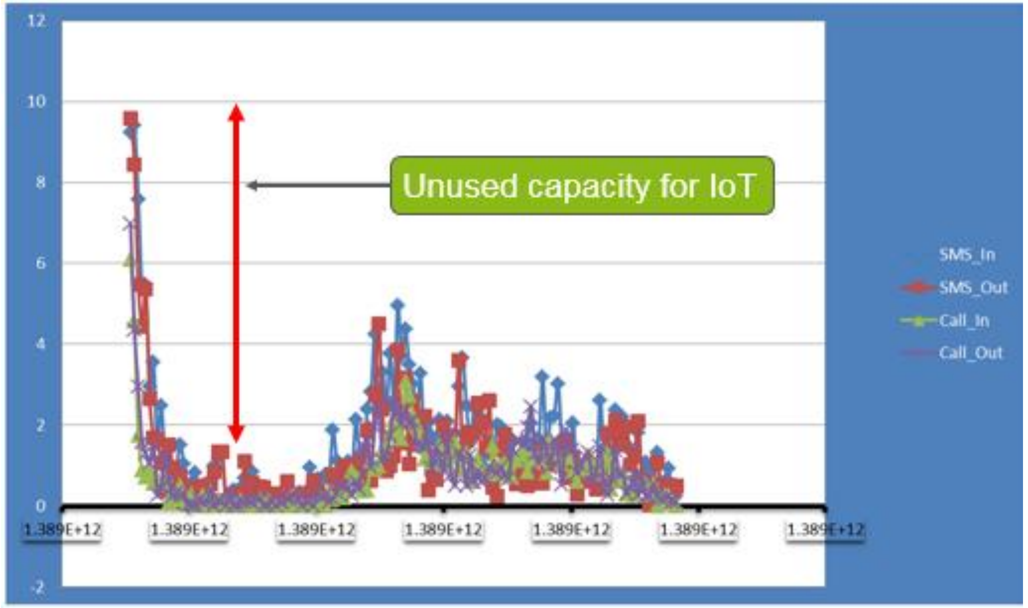
- IoT impacts on the network
  - new traffic patterns that change more frequently
  - QoS requirements and dedicated networks
  - Low Average Revenue Per Device (ARPD)
- Carrying IoT traffic on existing cellular network requires more advanced analytical in capacity planning
- This is possible by leveraging
  - Big Data and Analytics - accurate predict of the capacity for each network resource
  - Cloud Computing - faster and frequent Resources Provisioning

BUSINESS VALUE

› Accurate prediction of network resources requirements - reduce CAPEX and overprovisioning {type, time, location, capacity}

› Manage existing cellular network resources to carry multiple traffic types instead of a dedicated network for IoT

› Optimal network resources utilization - increases ROI



MACHINE LEARNING PROCESS

Business Understanding

- The Problem

Data Collection

- Offline Traffic data
- Live Traffic data
- Traffic forecast
- OSS & BSS data

Data Preprocessing

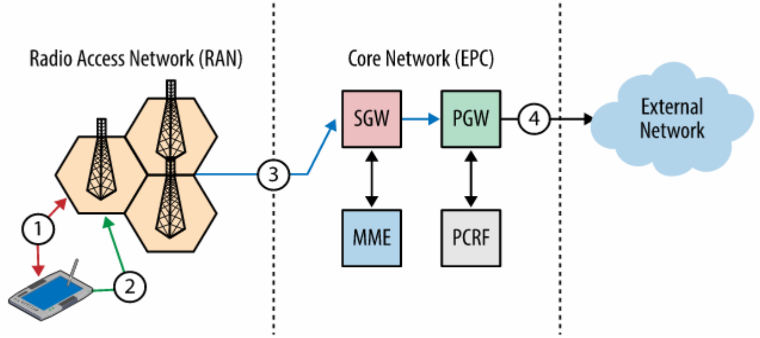
- Cleaning, Parsing
- Integration
- Simulation

Traffic Modelling

- Data Exploration
- Features Engineering
- Model Training
- Predictive modelling
- Model Evaluation

DATASETS

RAW DATA - 2, 6, 12 months



› RAN data	2G, 3G and 3G logs (Terabytes)
› Core Network data	anonymized CDR, Voice, SMS, Internet, IoT
› OSS, BSS data	Cells position, CM, PM, FM
› Environmental data :	Weather Station, Precipitation, Air Quality and Social Pulse Data
› Smart Cities Data	Road Traffic, Parking, Pollution, Weather, Cultural Event, Social Event
› Connected cars data	
› Smart Buildings data -	Ericsson office; City of New York
› IoT Traffic forecast	Number of users, verticals, network technologies, regions (Machina)
› Simulated data	Traffic generator

DATASETS

TARGET FEATURES

- Number of connected devices
- Transmission frequency/duration (daily)
- Packet size

CAPACITY PREDICTION

- Computing – CPU, memory
- Storage
- Connectivity – cells, bandwidth, etc.
- QoS requirements
- Availability – time, duration

Traffic volume predictions for next month, next year, per vertical, location or RAT (2G, 3G, 4G)

Dynamic configuration of network resources with accuracy at the right time

MACHINE LEARNING PROCESS

Business Understanding

- The Problem

Data Collection

- Offline Traffic data
- Live Traffic data
- Traffic forecast
- OSS & BSS data

Data Preprocessing

- Cleaning, Parsing
- Integration
- Simulation

Traffic Modelling

- Data Exploration
- Features Engineering
- Model Training
- Predictive modelling
- Model Evaluation

PREPROCESSING

- Identify key features in the datasets
- Select subsets of Data to train the model
- Clean redundant data
- Replace missing attributes values with zero or the column’s median
- Remove outliers
- Data transformation
- Data Fusion
- Normalization

MACHINE LEARNING PROCESS

Business Understanding

- The Problem

Data Collection

- Offline Traffic data
- Live Traffic data
- Traffic forecast
- OSS & BSS data

Data Preprocessing

- Cleaning, Parsing
- Integration
- Simulation

Traffic Modelling

- Data Exploration
- Features Engineering
- Model Training
- Predictive modelling
- Model Evaluation

IoT TRAFFIC CHARACTERISTICS

	WATER METERS	ELECTRICITY METERS	GAS METERS	VENDING MACHINES	BIKE FLEET MANAGEMENT	PAY-AS-YOU-DRIVE
Typical message size	100 bytes	100 bytes	100 bytes	150 bytes	150 bytes	150 bytes
Message interval	12 hours	24 hours	30 minutes	24 hours	30 minutes	10 minutes
Device density	10,000/km²	10,000/km²	10,000/km²	150/km²	200/km²	2,250/km²

https://www.ericsson.com/assets/local/mobility-report/documents/2016/emr-november-2016-massive-iot-in-the-city.pdf

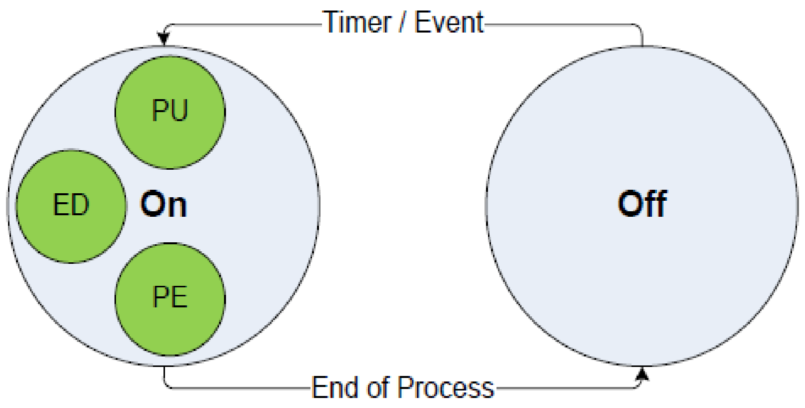
IoT TRAFFIC CHARACTERISTICS

- Transmission at anytime of the day
- From locations not accessible to humans
- All machines running the same application behave the same
- Transmission may be coordinated, synchronized
- Periodic or Event-driven
- Short and small packets
- Real time and non-real time
- Sleep time
- Uplink-dominant traffic

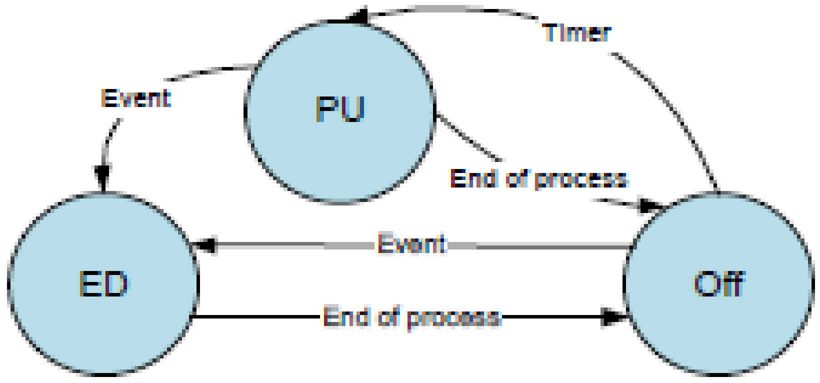
IoT TRAFFIC MODELLING

- IoT has 3 elementary traffic patterns
  - Periodic Update (PU)
  - Even-Driven (ED)
    - Includes Query Driven
  - Payload Exchange (PE)

  - On – ED, PU and PE states
  - Off - artificial traffic type



M2M Traffic Modeling Framework



Example: Sensor based Alarm

http://www.eurecom.fr/en/publication/4080/download/cm-publi-4080.pdf

IoT TRAFFIC MODELLING

- A device ***n*** *is represented by* ***a Markov chain***

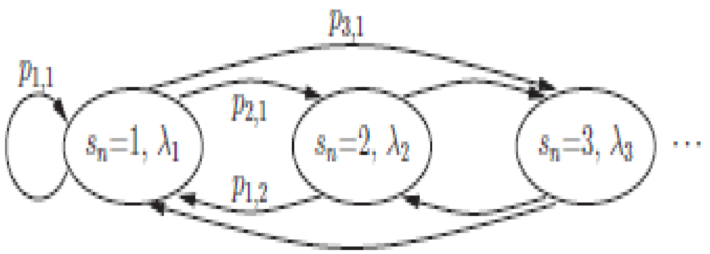
- Traffic is a Poisson process modulated by the rate *λi[t], determined by the state of a Markov chain Sn[t]*

- ***Pi,j*** state transition probabilities
- ***Πi*** state probabilities
- **P** state transition matrix

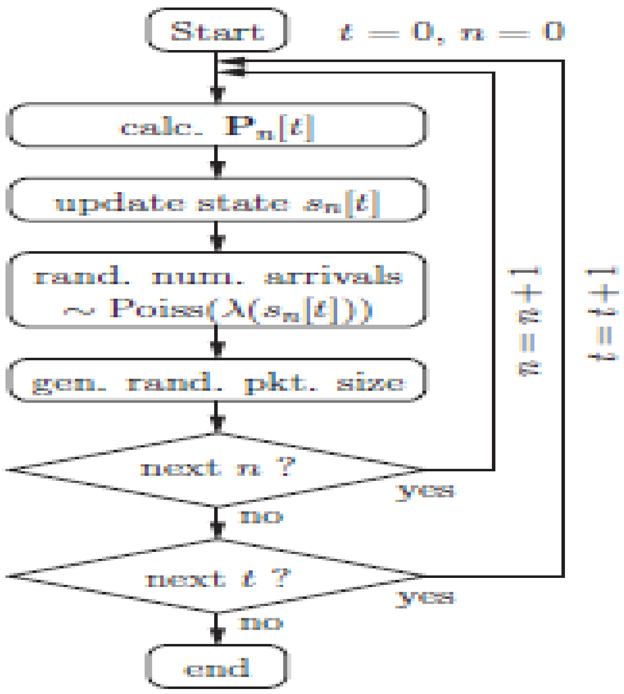
$$\mathbf{P} = \begin{pmatrix} p_{i,1} & p_{1,2} & \cdots \\ p_{2,1} & p_{2,2} & \\ & \vdots & \end{pmatrix} \qquad \boldsymbol{\pi} = \begin{pmatrix} \pi_1 \\ \pi_2 \\ \vdots \end{pmatrix}; \qquad \boldsymbol{\pi} = \boldsymbol{\pi}\mathbf{P}$$

- The global rate

$$\lambda_g = \sum_{i=1}^{I} \lambda_i \pi_i \text{ , } I \text{ is the total number of states.}$$



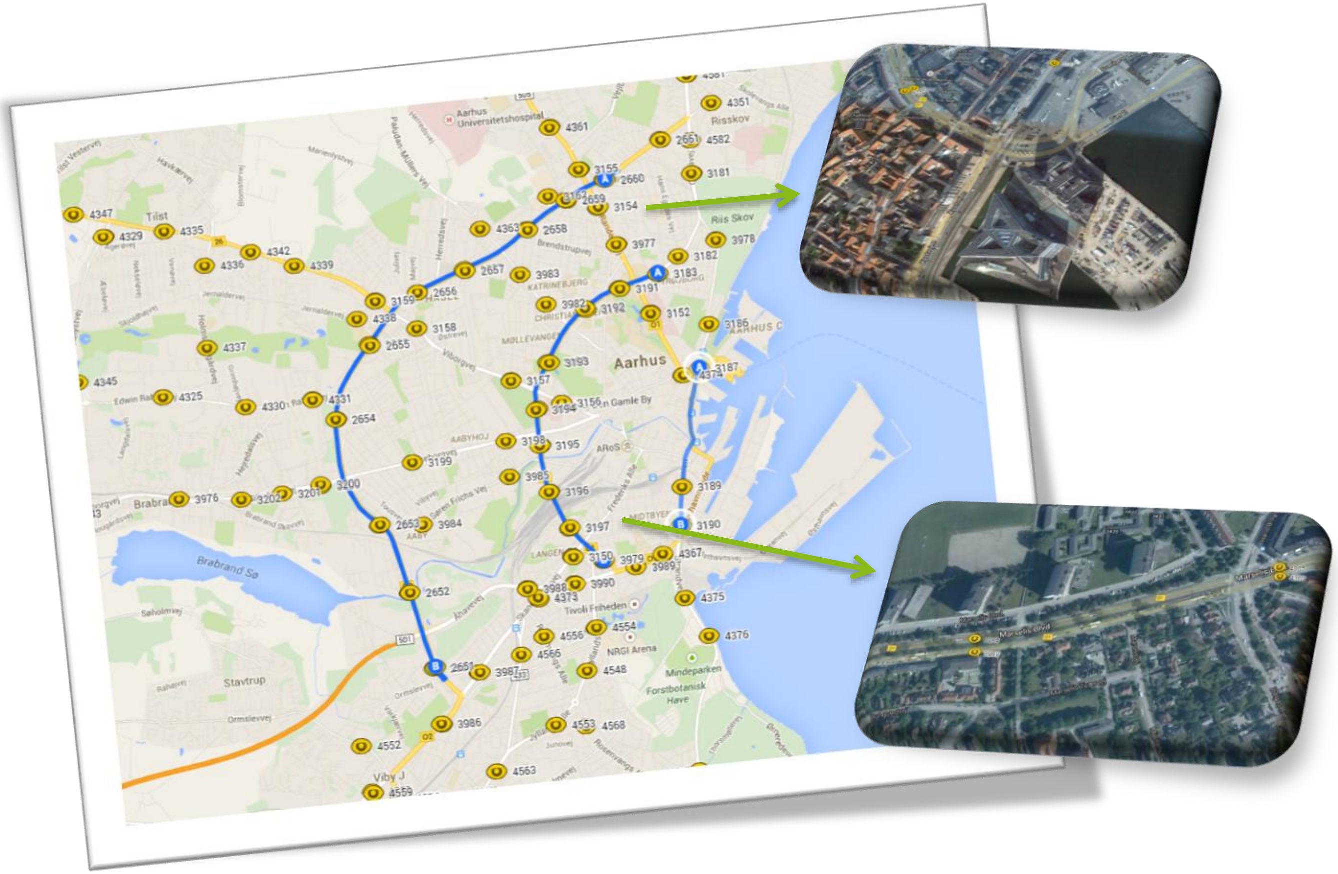
MTC device (*n*)

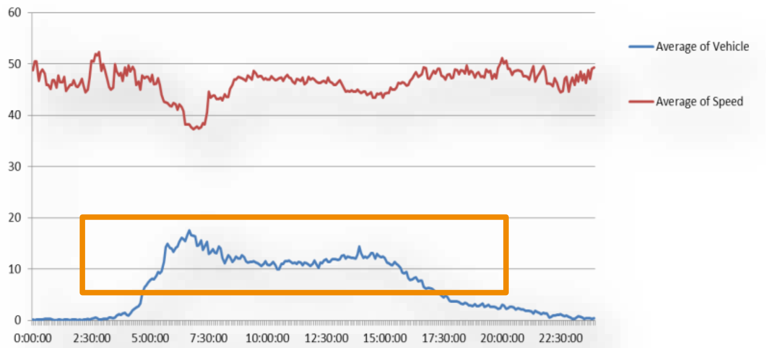


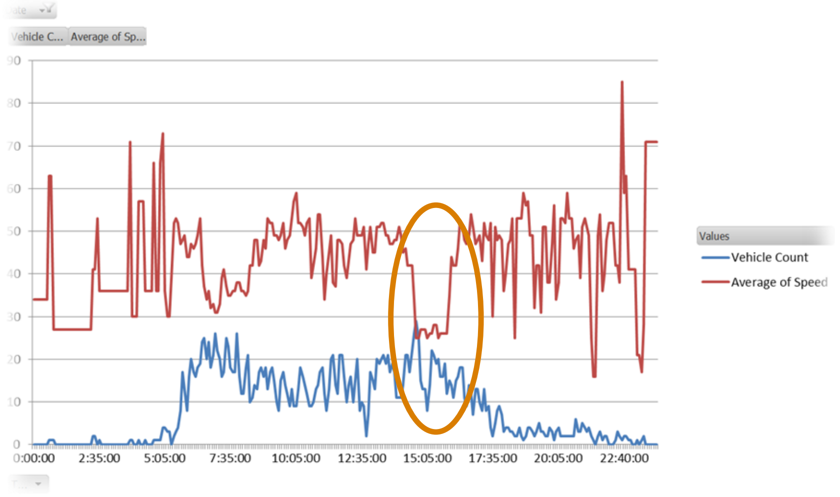
MMPP Traffic Model

http://www.eurecom.fr/en/publication/4265/download/cm-publi-4265.nikaein.02.08.14._final.pdf

AARHUS VEHICLES TRAFFIC DATA ANALYSIS



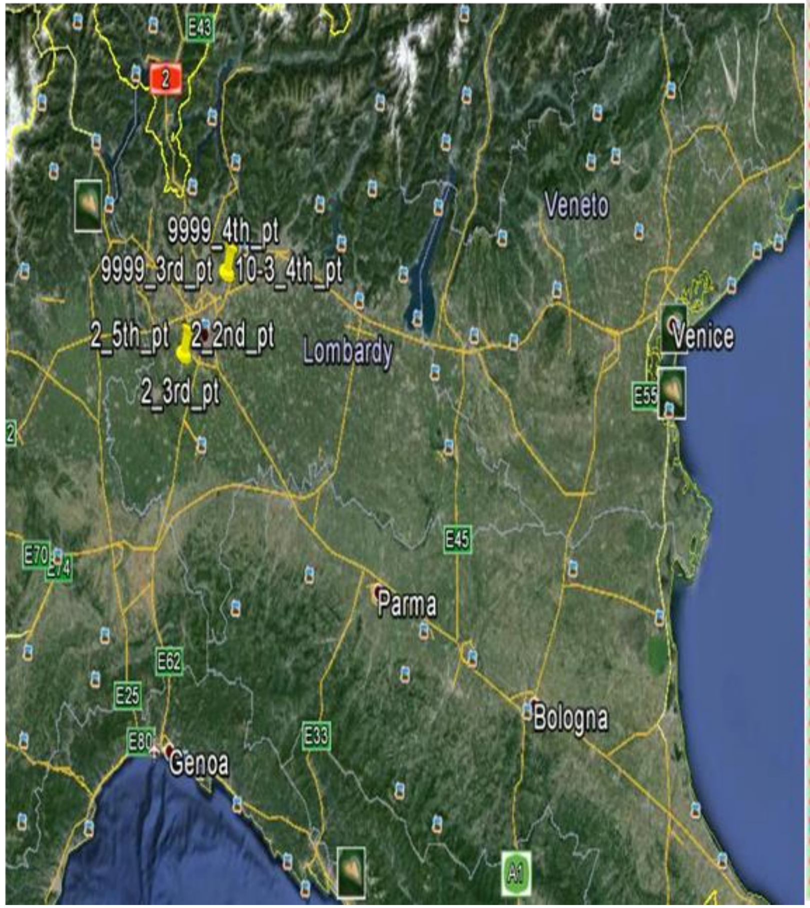




http://iot.ee.surrey.ac.uk:8080/datasets.html

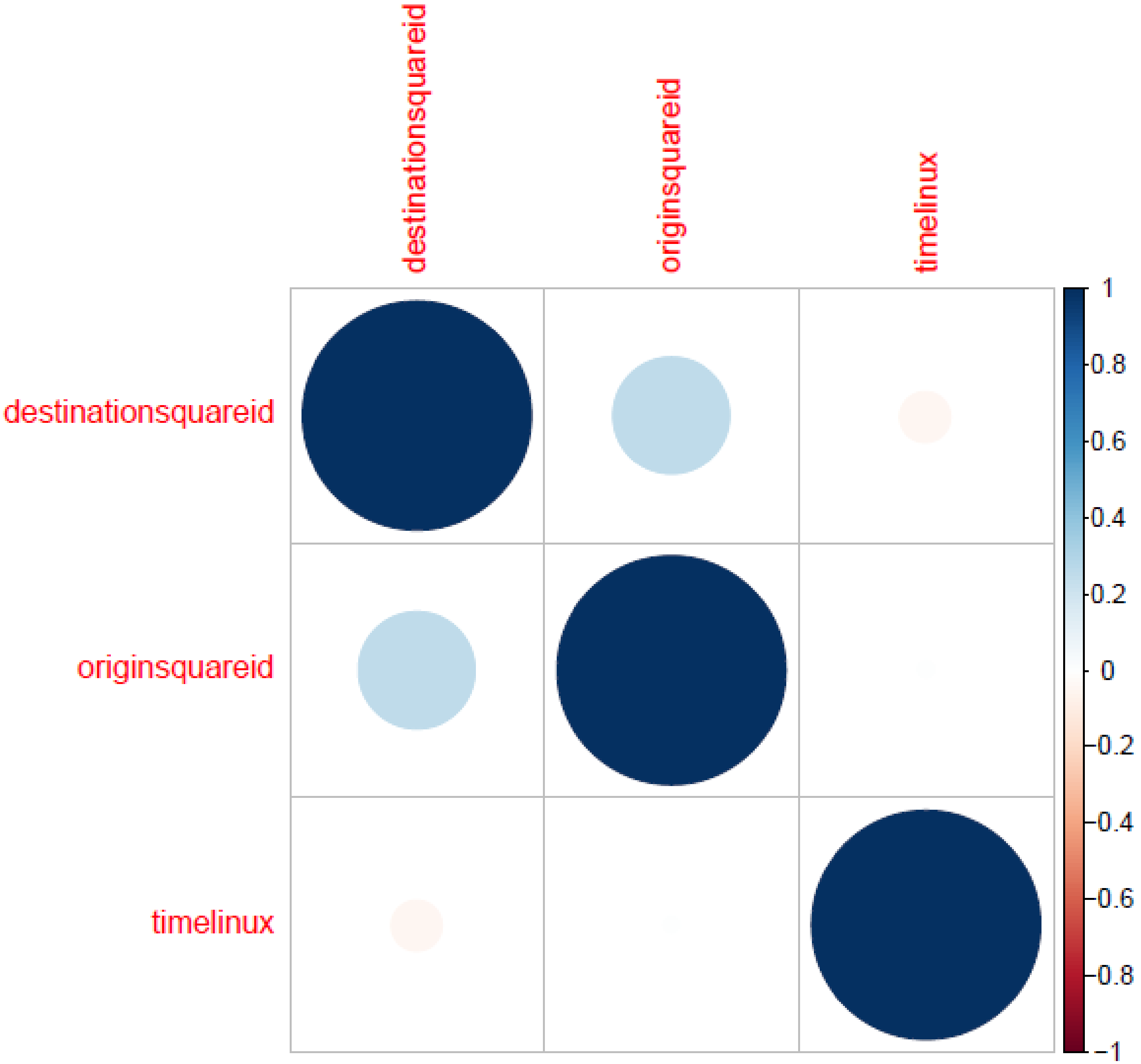
IoT TRAFFIC – LOCATION DATA ANALYSIS

Day of Experiment	Cell-or-square_Id	Country_Code	Total observations (Input)
2014-01-01	**9999**	**39**	145



https://dandelion.eu/datamine/open-big-data

DATA EXPLORATION - VARIABLE CORRELATION WITH PEARSON



ANALYTICS FRAMEWORK ARCHITECTURE

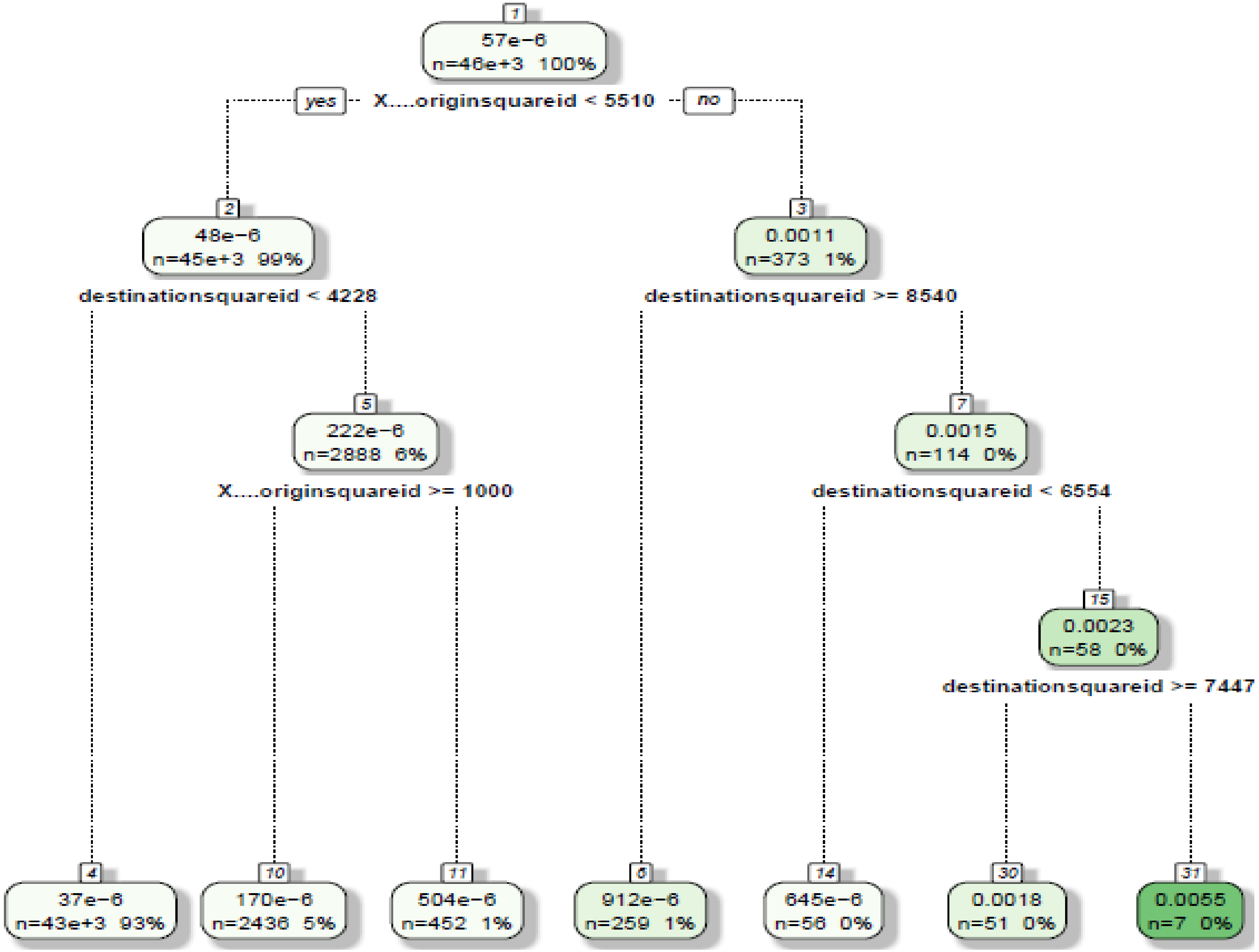
Visualization, APIs

Batch | Real time | Predictive | Simulation

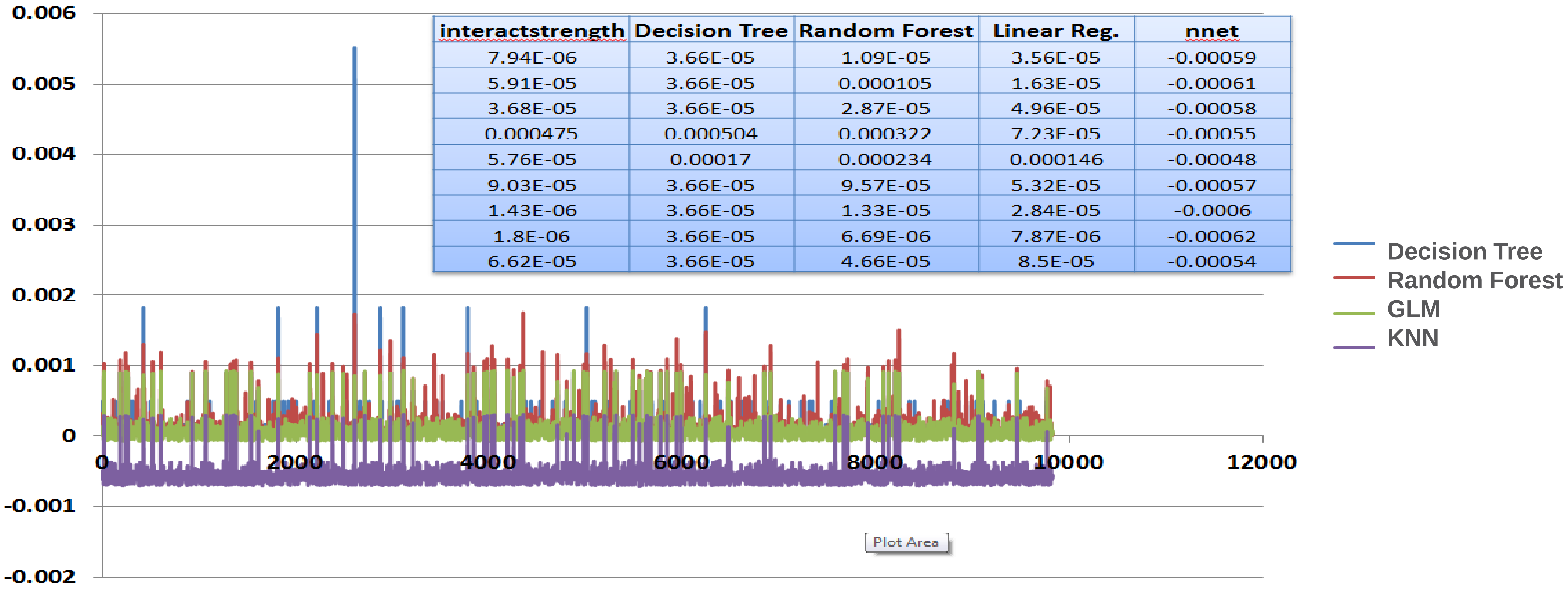
ETL, Integration

Data Repositories – Historical data and Data stream

MODELLING - DECISION TREE



MODELLING AND SCORING



interactstrength	Decision Tree	Random Forest	Linear Reg.	nnet
7.94E-06	3.66E-05	1.09E-05	3.56E-05	-0.00059
5.91E-05	3.66E-05	0.000105	1.63E-05	-0.00061
3.68E-05	3.66E-05	2.87E-05	4.96E-05	-0.00058
0.000475	0.000504	0.000322	7.23E-05	-0.00055
5.76E-05	0.00017	0.000234	0.000146	-0.00048
9.03E-05	3.66E-05	9.57E-05	5.32E-05	-0.00057
1.43E-06	3.66E-05	1.33E-05	2.84E-05	-0.0006
1.8E-06	3.66E-05	6.69E-06	7.87E-06	-0.00062
6.62E-05	3.66E-05	4.66E-05	8.5E-05	-0.00054

SUMMARY

› IoT traffic modelling and prediction

› Accurate prediction of resources requirements {type, time, location, capacity} - reduce overprovisioning and CAPEX

› Optimal network resources utilization - increases ROI

Thank You

ERICSSON