CABLE-TEC EXPO® 2017

SCTE · ISBE

THE NEXT BIG...

DEAL
CONNECTION
INNOVATION
TECHNOLOGY
LEADER
NETWORK

SCTE·ISBE CABLE-TEC EXPO 2017

DENVER, CO
OCTOBER 17-20

2017 Fall Technical Forum
SCTE·ISBE • NCTA • CABLELABS

SCTE · ISBE

Network Capacity and Machine Learning

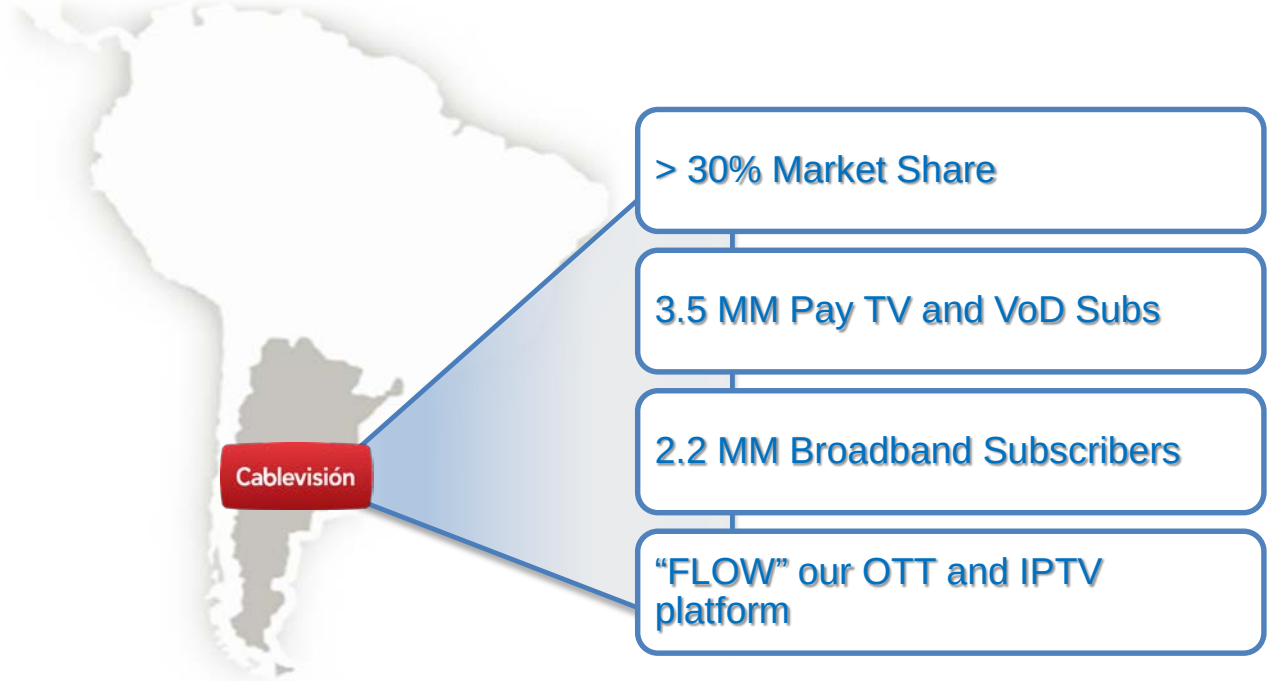
Claudio Righetti
Chief Scientist & Security
Cablevisión S.A.

SCTE·ISBE CABLE-TEC EXPO 2017
DENVER, CO
OCTOBER 17-20

Agenda

- Introduction to Machine Learning
- Exploring Traffic Utilization Data
- Principal Components Analysis
- Training our Neural Network

The largest MSO in Argentina and one of the biggest in Latin America



Agenda

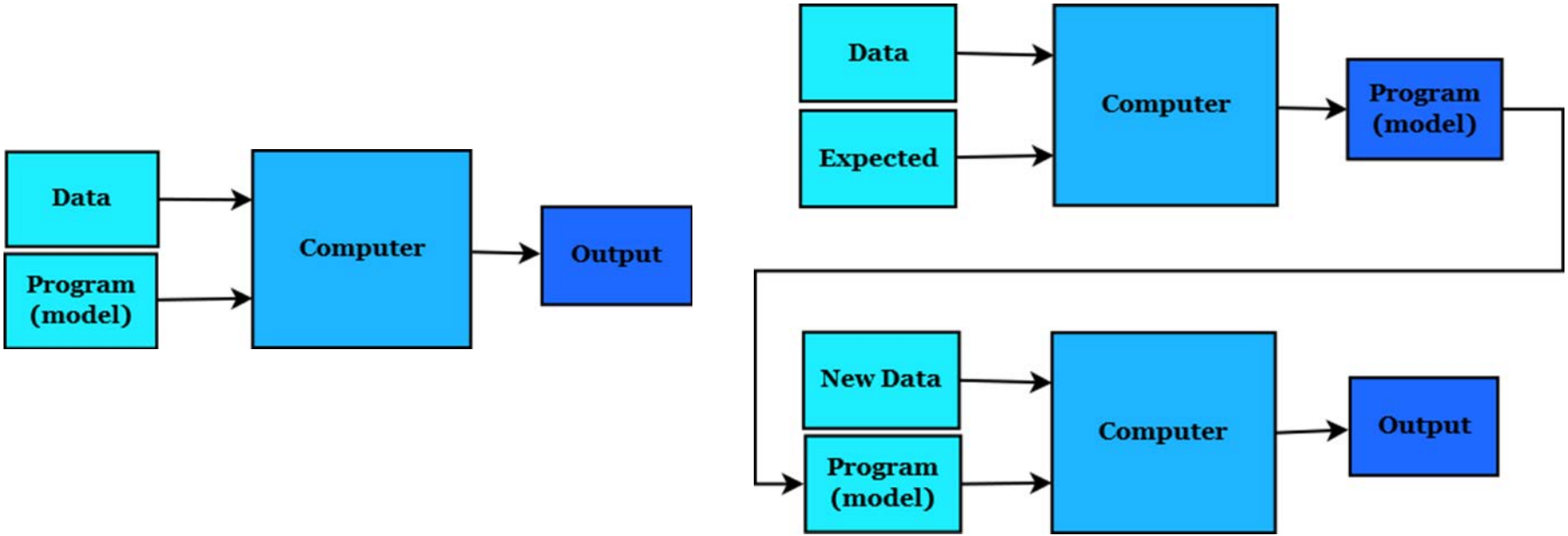
- Introduction to Machine Learning
- Exploring Traffic Utilization Data
- Principal Components Analysis
- Training our Neural Network

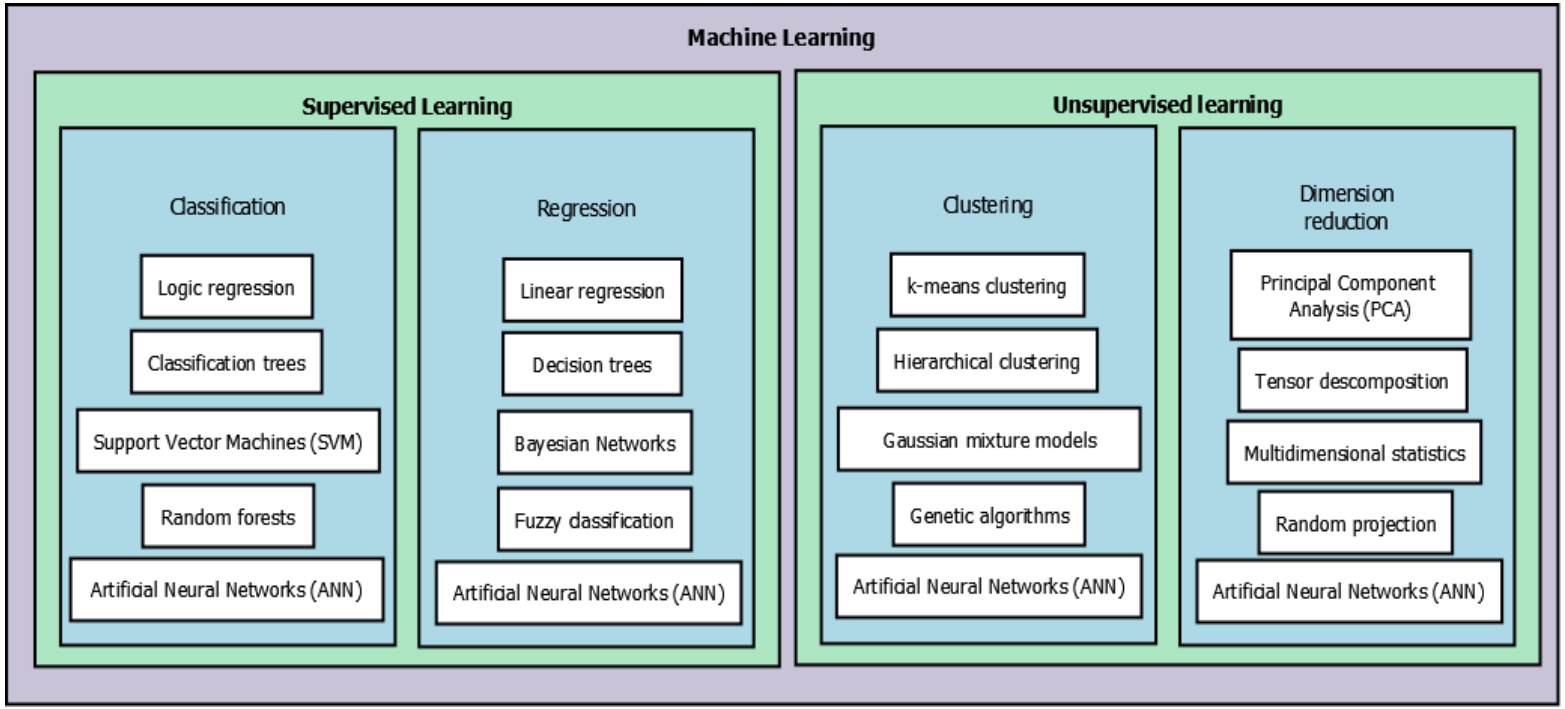
What is Machine Learning ?

“The field of study that gives computers the ability to learn without being explicitly programmed.“

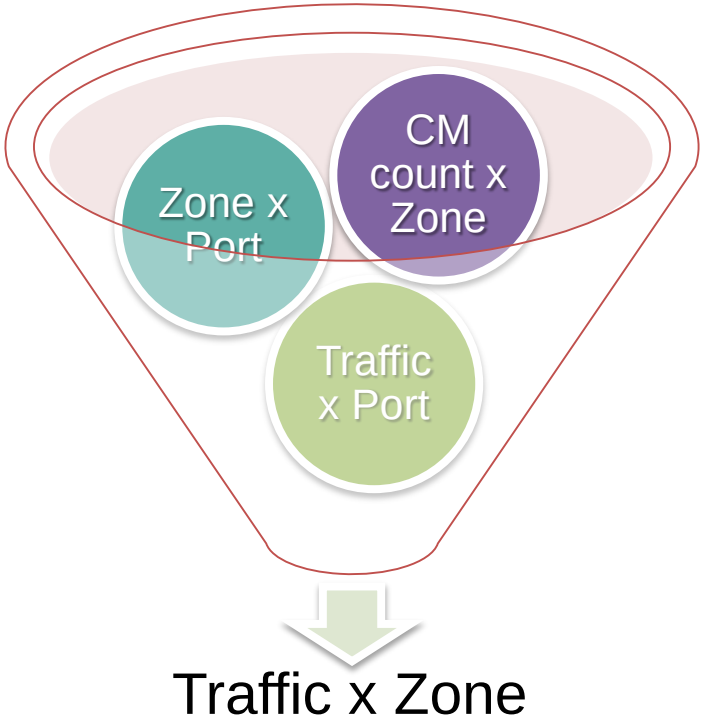
Arthur Lee Samuel, *Some studies in machine learning using the game of Checkers*," IBM Journal of Research and Development 3 (1959),

Traditional Programming vs Machine Learning



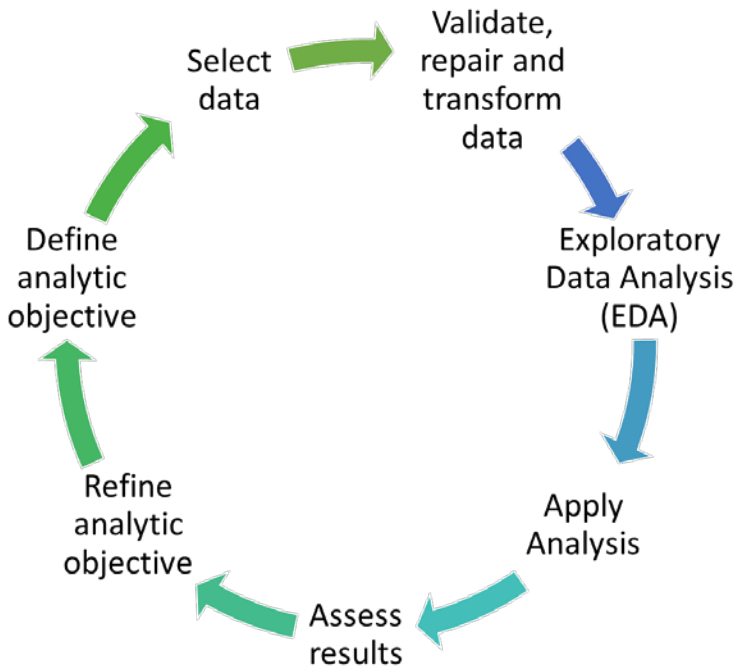


Datasets



Agenda

- Applied Machine Learning
- Exploring Traffic Utilization Data
- Principal Components Analysis
- Training our Neural Network



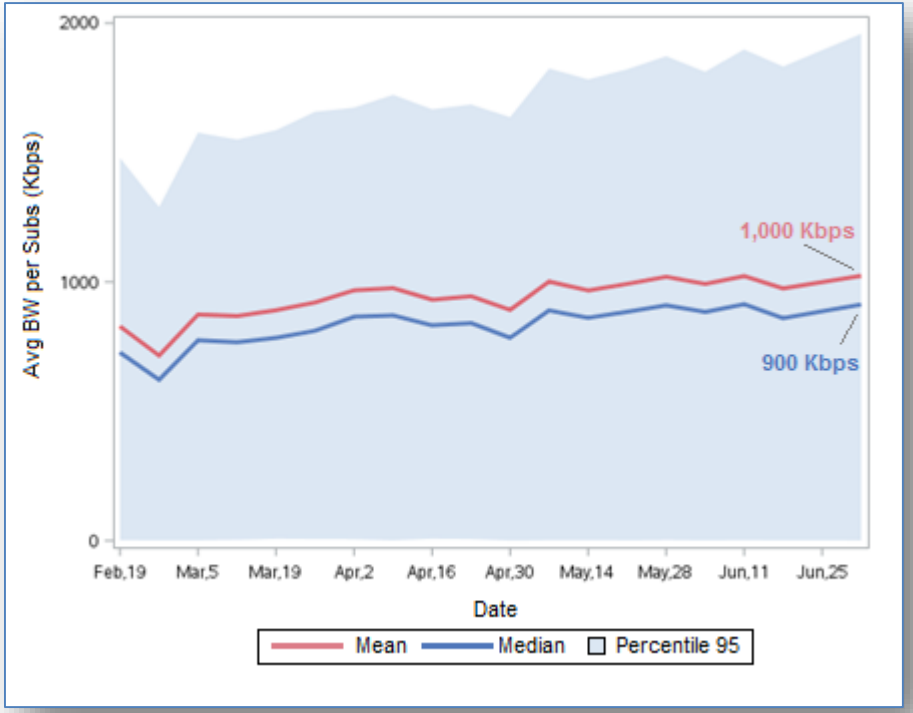
Exploratory Data Analysis

- Summarize main characteristics
- Suggest hypotheses
- Assess assumptions
- Select appropriate statistical techniques

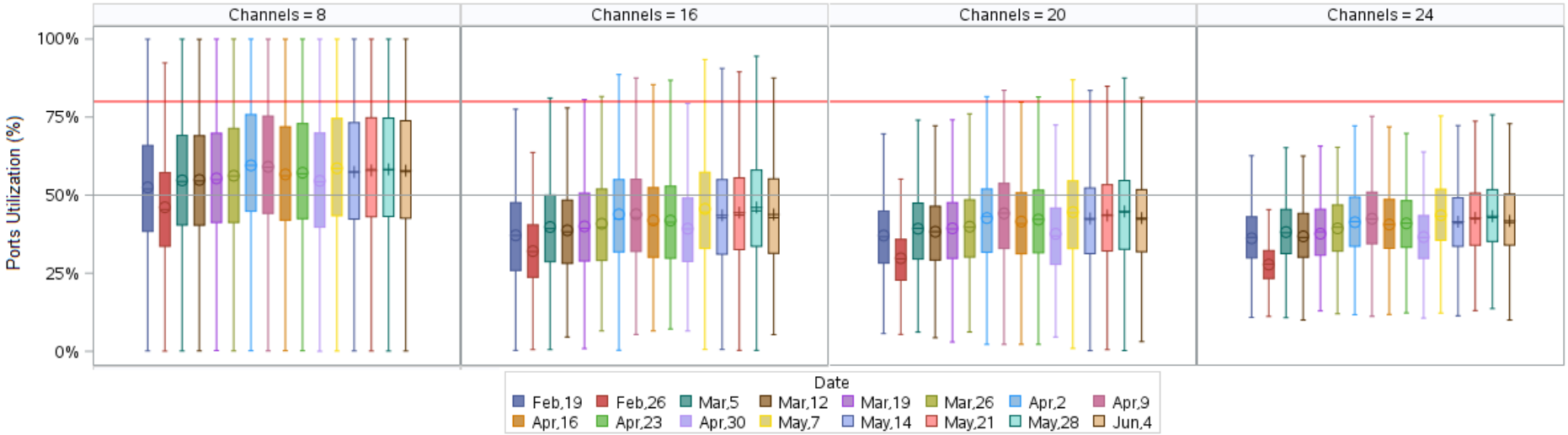
Average Bandwidth per Subscriber Evolution

New higher values of mean and median

$$Avg\ BW\ Subs\ (Kbps) = \frac{Port\ Traffic}{\#\ Subs}$$

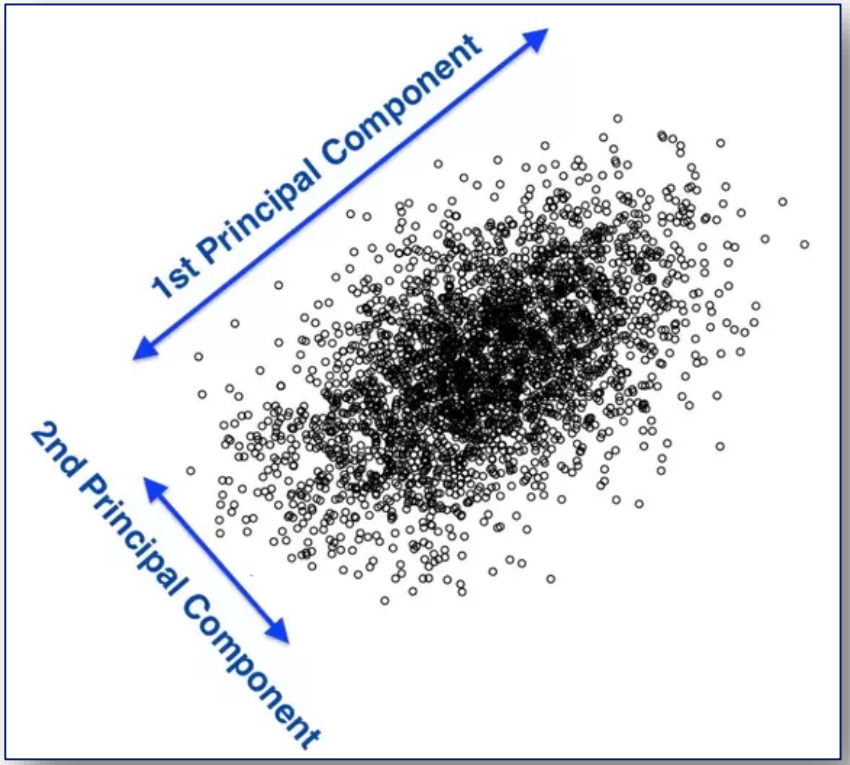


Utilization Evolution



Agenda

- Introduction to Machine Learning
- Exploring Traffic Utilization Data
- Principal Components Analysis
- Training our Neural Network



Principal Components (PCA)

- Statistical procedure to convert a set of correlated variables to linearly uncorrelated variables
- Used to reduce the dimensionality of a set of variables

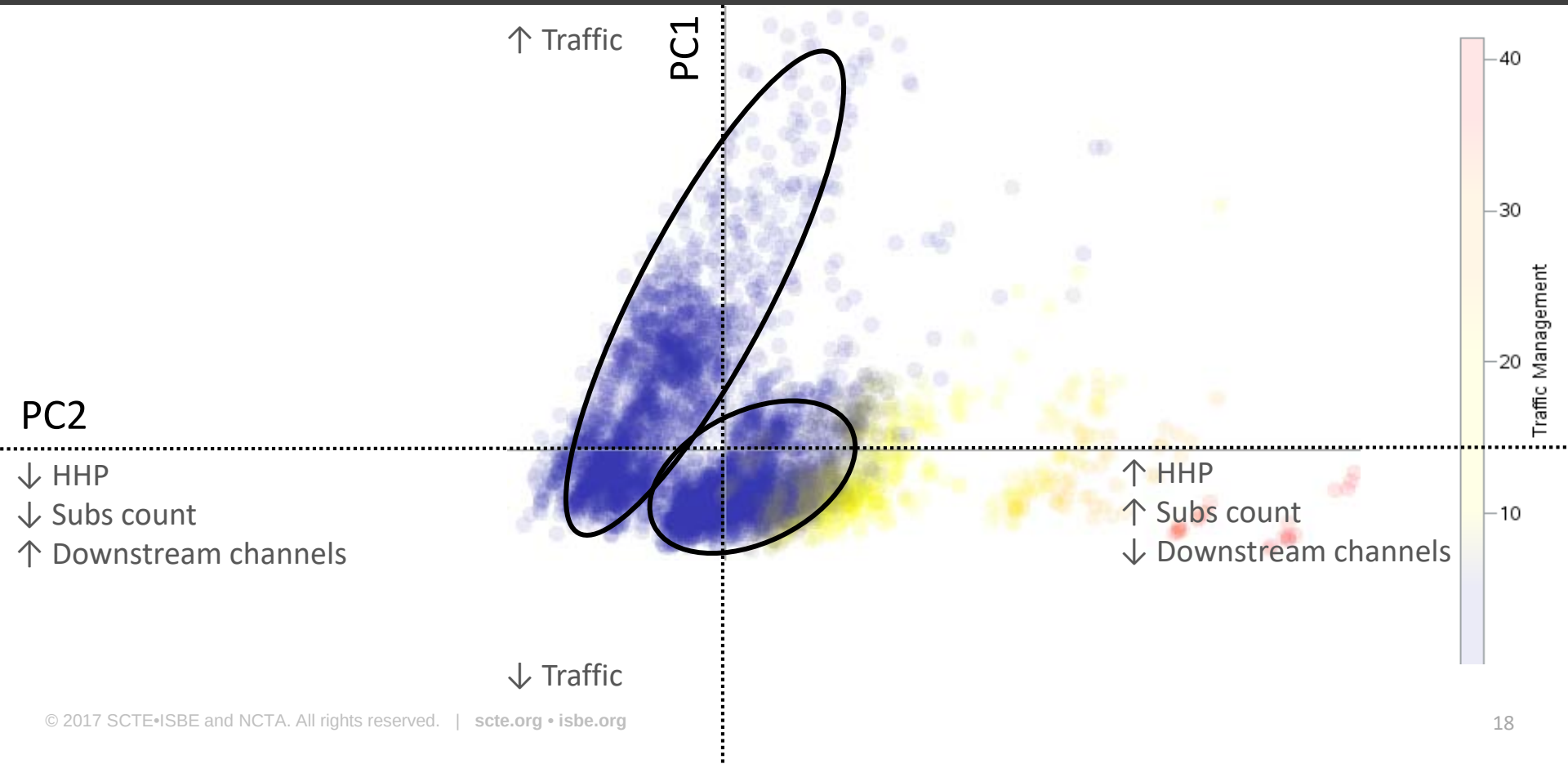
Variables

- Maximum traffic per port (evolution)
- DS channels per port
- Areas connected to each port
- Households passed
- Subscribers per port
- Traffic Management

Principal Components

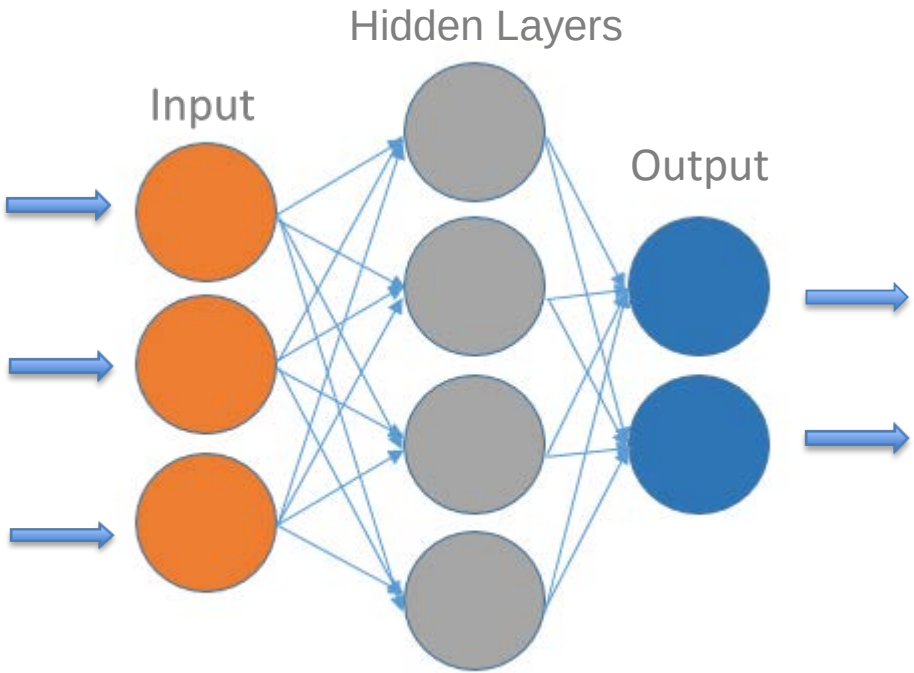
Principal Component	Proportion of Total Dispersion	Cumulative Percentage
1	77.2%	77.2%
2	5.7%	82.9%
3	4.3%	87.2%
4	3.7%	90.9%
…	…	…
21	0.1%	100%

Principal Component Analysis



Agenda

- Introduction to Machine Learning
- Exploring Traffic Utilization Data
- Principal Component Analysis
- Training our Neural Network



Artificial Neural Networks (ANN)

- Input Layer (Patterns)
- Hidden Layers (Processing)
- Output Layer (Computations)

Applying ANN: Network Access Strategies

Objective

Classification of nodes based on 4 Network Access Strategies:

1) Chassis Upgrade
2) Recombination
3) Node Segmentation
4) Node Split

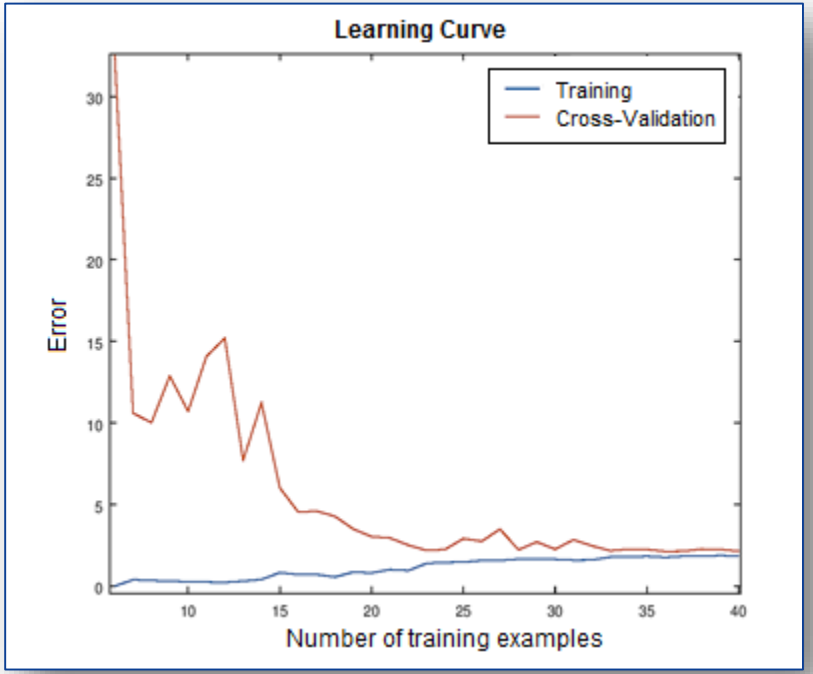
Data Sample

3 strata of nodes

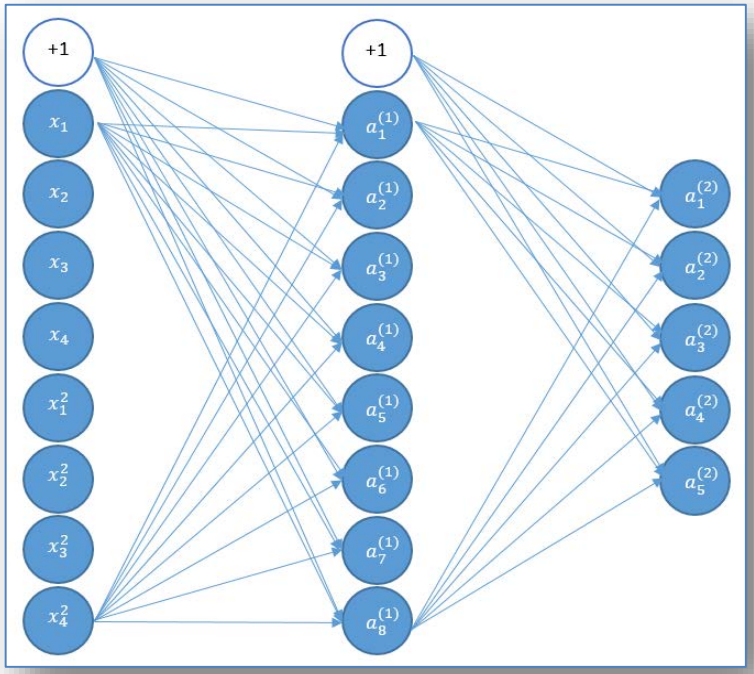
Utilization Range	Percentage of Nodes	HHP Mean
< 50%	51%	636
50% - 80%	40%	727
> 80%	9%	804

Dataset Treatment

- 3 subsets to prevent overfitting:
  - Training (60% of data)
  - Cross-Validation (20%)
  - Testing (20%)
- Stepwise technique to find the optimal combination of variables
- Necessity of including quadratic terms
- **90%** Accuracy



Our Neural Network Model



Input Layer: Eight Inputs

- Utilization
- DS channels
- Count of segments and nodes sharing a port
- Segmentation level
- Their quadratic terms

Output Layer: Five classes:

1) Chassis Upgrade
2) Recombination
3) Node Segmentation
4) Node Split
5) No action needed

Machine Learning and Network Capacity : STEM-ML

- Analyze the impact on a marketing campaign “Double your Speed”
- Include upstream ports
- Improve ANN accuracy
- Include Average Bandwidth per service Tier

Machine Learning and Network Capacity : STEM-ML

- The addition of machine learning algorithms in our STEM tool has proved to be powerful, accurate, and scalable in different scenarios, with many variables in our Access Network
- Machine Learning has taken a leap into our Industry
- Get raw data -> Clean data-> Build model -> Predict
- Parsimony Principle: the simplest model that fits the data is also the most plausible
- Applied machine learning = Art + Science

SCTE · ISBE

THANK YOU!

Claudio Righetti
crighetti@cablevision.com.ar
+5411 5530-4468

SCTE·ISBE CABLE-TEC EXPO 2017 | DENVER, CO OCTOBER 17-20

Cablevisión Fibertel