SCTE® CABLE-TEC EXPO®
FAST FORWARD 2021

ATLANTA, GA
OCTOBER 11-14

SCTE®
a subsidiary of CableLabs®

UNLEASH THE POWER OF LIMITLESS CONNECTIVITY

2021 Fall Technical Forum
SCTE® • NCTA • CABLELABS®

Operational Transformation

Machine Learning for RF Impairment Detection

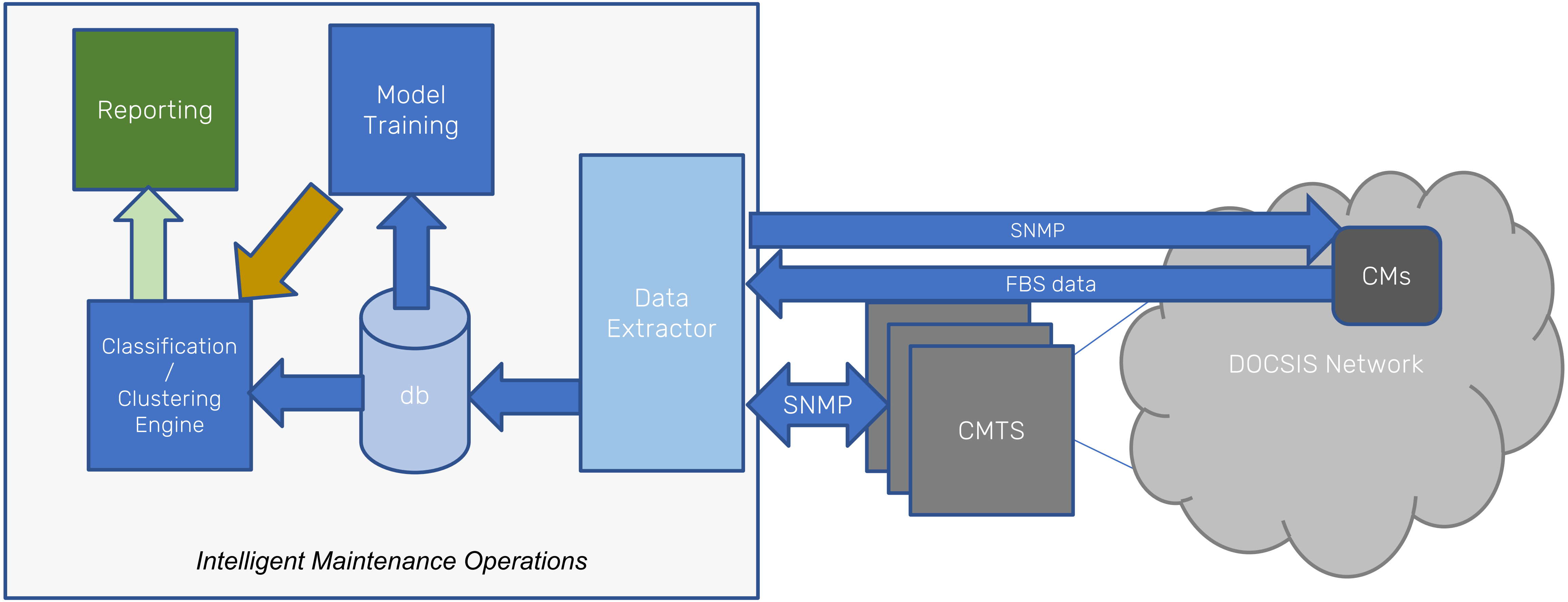
David Virag

Distinguished System Engineer
Commscope

Background

- Past 20 years explosion of data
  - Large scale web-services e.g. search, social media
  - New methods for organizing and processing massive data sets
- New tools and processing pipelines developed
  - Available to everyone
  - Can now leverage new techniques to different problems
- Simultaneously, new PNM methods becoming available for DOCSIS networks
  - Ability to gather large data sets to assess health of plant
  - Need to apply new techniques to large data sets for real-time plant assessment
- Automated methods to detect plant issues beneficial for industry
  - Increased customer satisfaction (fewer/less frequent problems)
  - Lower operational costs for MSO with improved automation techniques.
- Goal of research to evaluate the application of ML algorithms to available data.

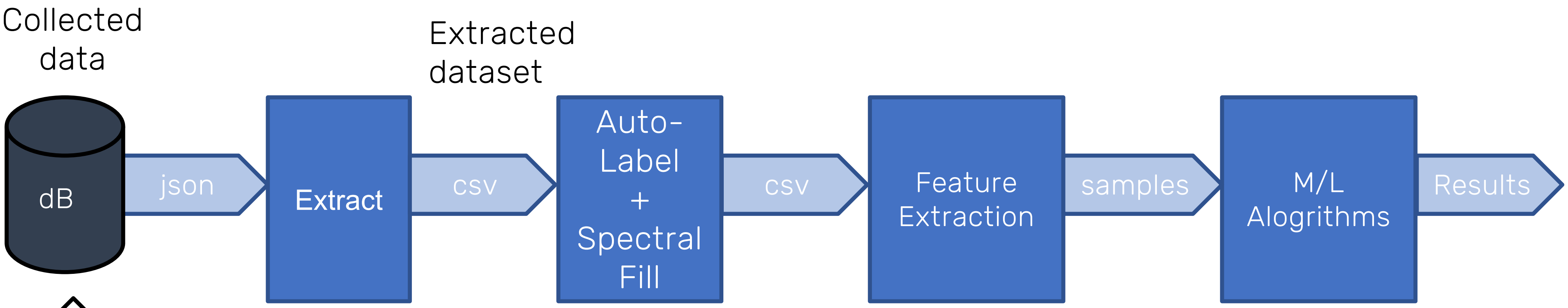
Implementing a Closed Loop System



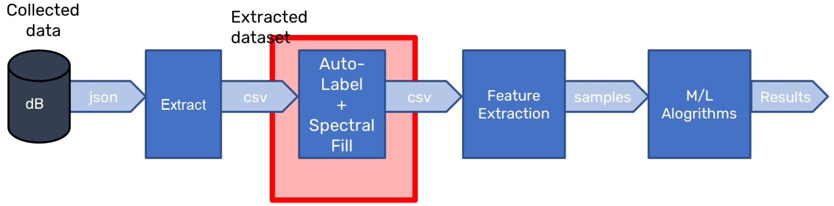
Challenges of using Machine Learning

- Need for Labeled Data
  - Supervisory based systems require lots of labeled data
  - Difficult to label data
    - manual process
    - Requires specific expertise
- Non-service spectrum
  - Full-band spectrum can contain gaps of unused spectrum
  - Gaps appear as noise, can make it difficult to detect wideband issues
- FBS captures include lots of data and require memory and CPU power for processing
  - Storage and compute costs today allow for mass processing of data at much lower cost than a decade ago.

A Machine Learning Pipeline



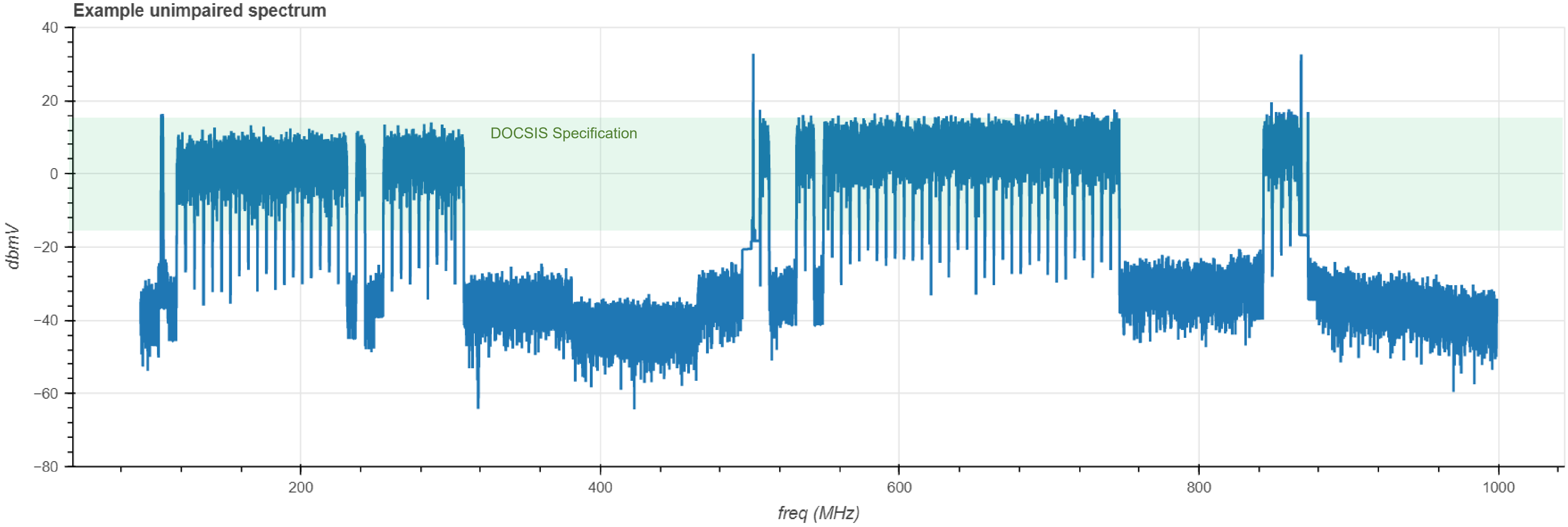
- Full band spectrum captures
- Collections from 2 CMTS Systems
- > 1 million samples in each
- 93-993 MHz (151 6-MHz channels
- 256 bins/channel



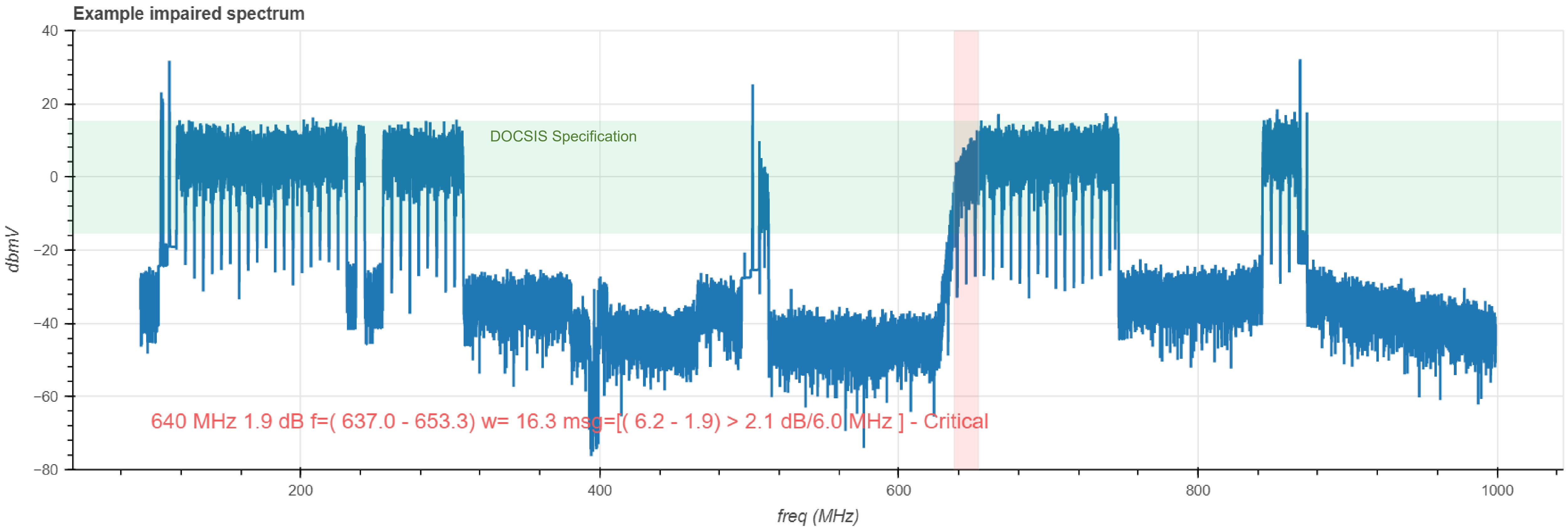
Auto-Label

- Supervised Learning – learning based on labeled data sets to map features to labeled output.
  - Labeling is a challenge for spectral based samples. How to obtain those labels?
  - Manual process requires specific domain expertise and is time consuming.
- Utilize auto-label based on traditional signal processing approaches.
  - Channel framing and detection based on received power levels within 6 MHz bands.
- Impairment Triggers
  - Power outside of DOCSIS specified range (-15 dBmv to +15 dBmv)
  - Adjacent power level mismatch
  - Delta power between successive 3 MHz ranges outside of threshold.
  - Delta power within a consecutive 6 MHz channel spectrum was outside threshold.
- Samples exceeding one or more impairment triggers labeled ‘bad’.
  - ‘impaired’ spectrum does not necessarily mean customer services are degraded.

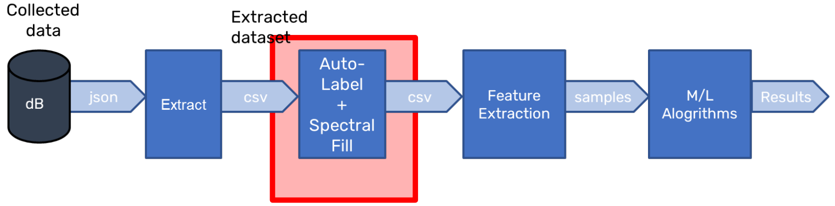
Full Band Spectrum Capture – 'Good' Example



Full Band Spectrum Capture – ‘Impaired’ Example



SCTE. CABLE-TEC EXPO
FAST FORWARD 2021
ATLANTA, GA > OCTOBER 11-14



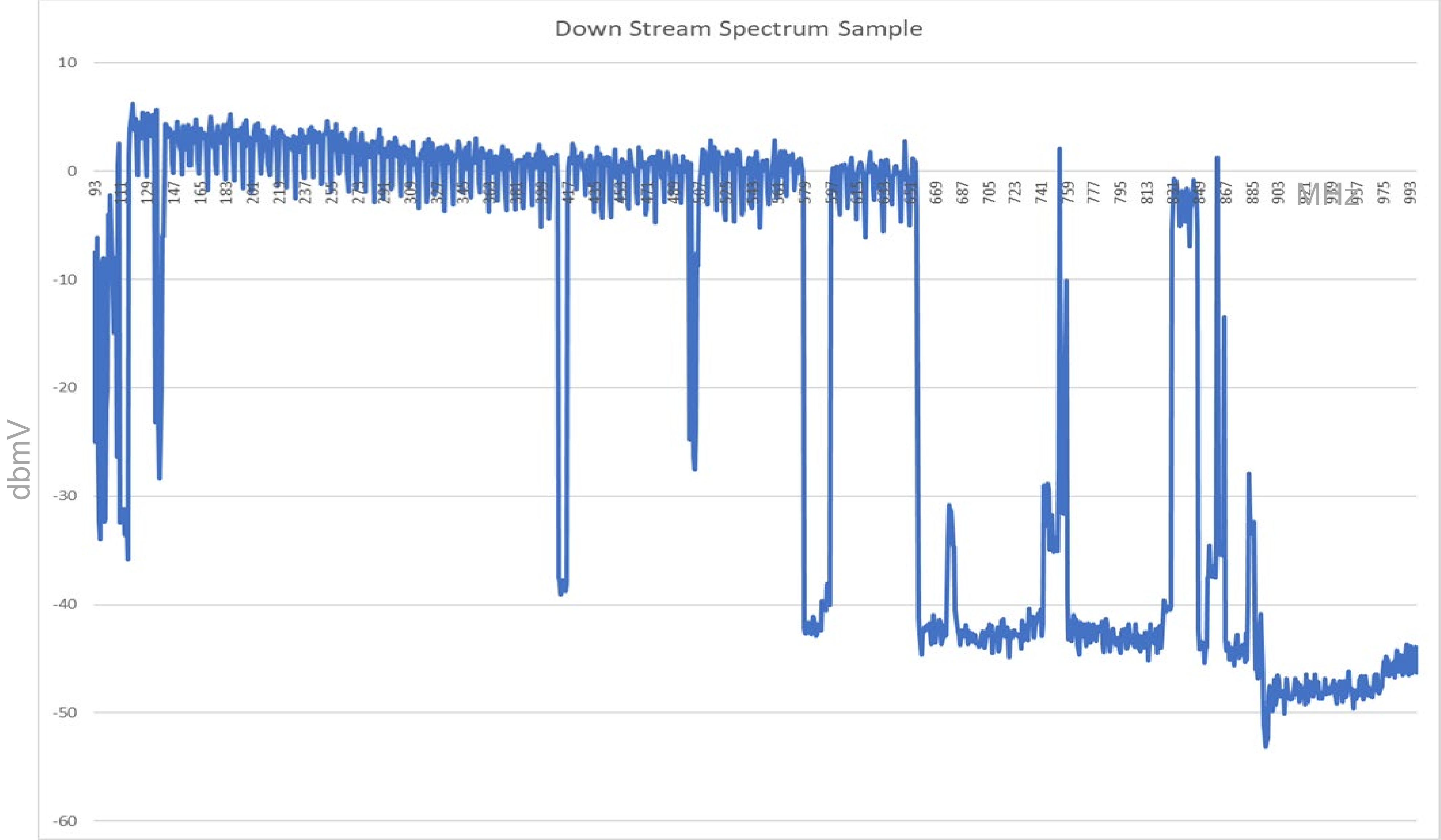
Spectral Fill

Full-band capture spectrum includes unused spectrum with no signal energy.

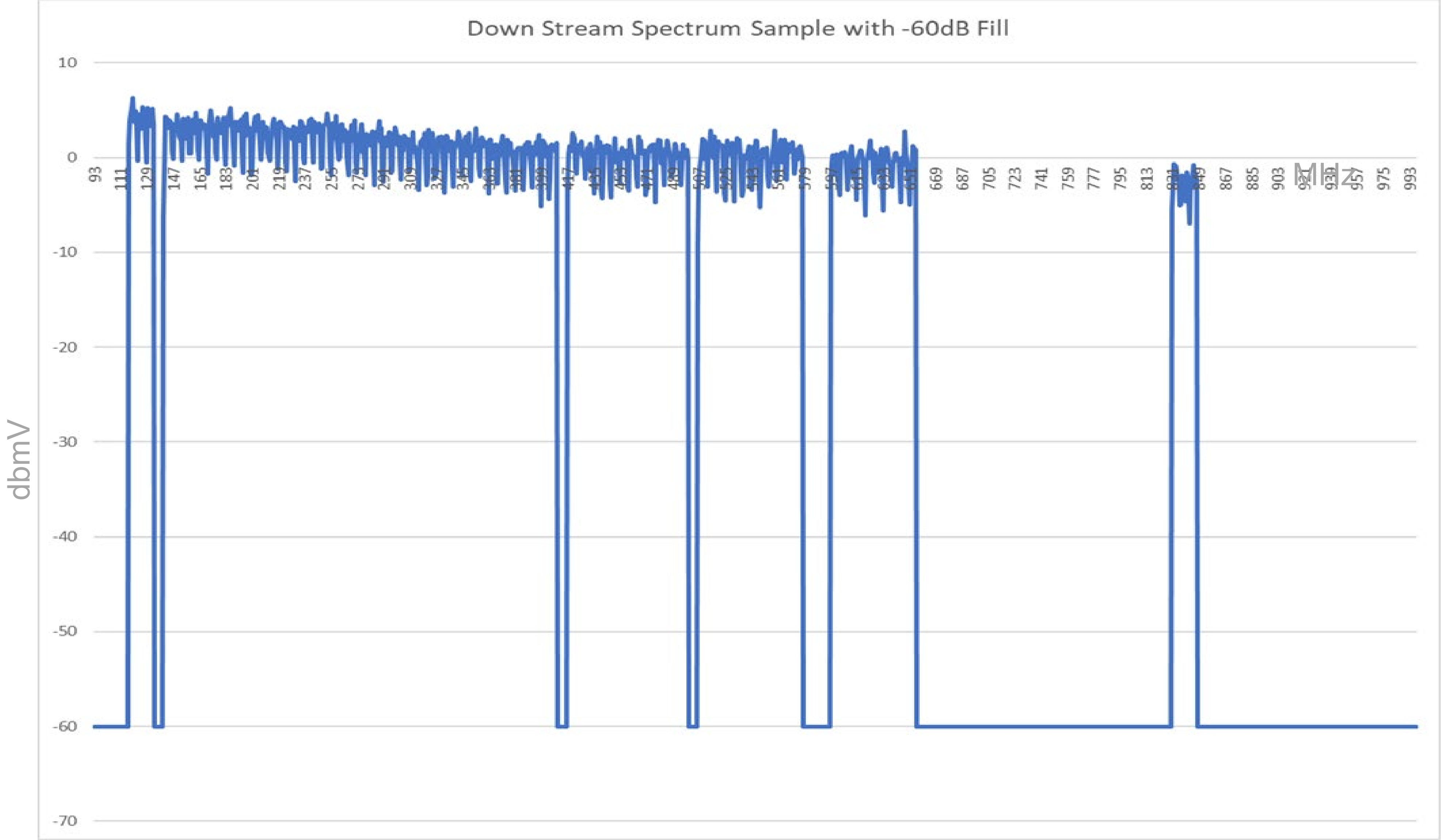
- These bands may not be relevant to discerning the overall health of the channelized spectrum.
- Would pre-processing the unused spectrum improve results of the M/L algorithms with greater separation of characteristics from the channelized spectrum?

Evaluated filling unused spectrum with a flat '-60db' value prior to running M/L algorithms to test this notion.

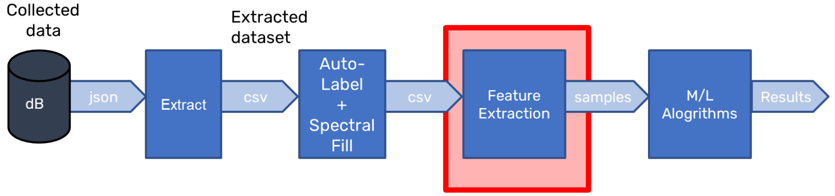
Example of Downstream FBS Capture



Example of Downstream FBS Capture



SCTE. CABLE-TEC EXPO
ATLANTA, GA > OCTOBER 11-14



Feature Extraction

Raw data sets have 151 channels x 256 bins/channel = 38k points data per-sample.

Feature extraction provides method to reduce the points per sample to smaller highly relevant set for M/L algorithms.

Utilize TSFresh library for feature extraction

- Generic feature extraction – no specific domain expertise incorporated
- Developed for time-series applications, applicable to uniformly sampled datasets
- Reduces each sample from 38k points to 788 ‘efficient’ features.

Example Features:

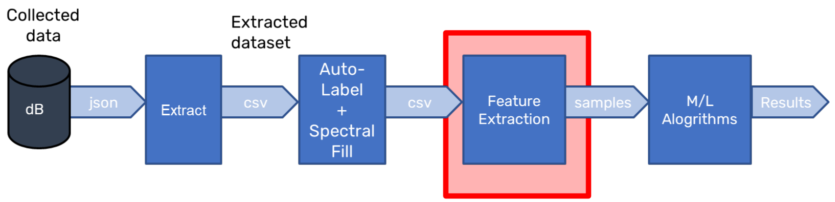
Absolute Energy => $\sum (x_i)^2$ for all X in sample

Max value => Maximum (x_i) value in sample

Autocorrelation => Sum of Autocorrelation of (x_i) with (x_{i-lag})

SCTE. CABLE-TEC EXPO
FAST FORWARD 2021
ATLANTA, GA > OCTOBER 11-14

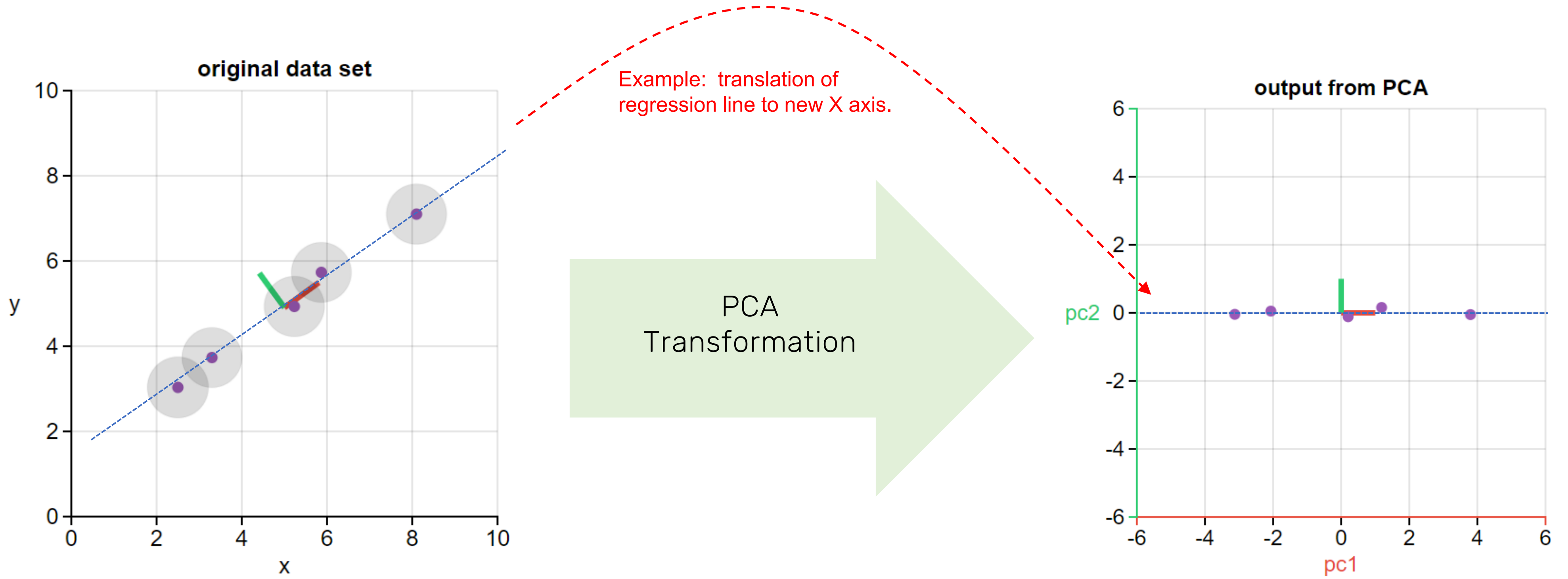
Feature Extraction



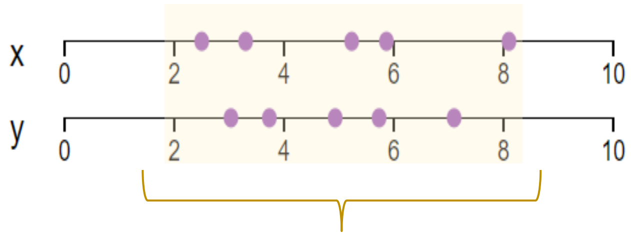
Principal Component Analysis (PCA) employed to further reduce feature set

- Technique based in linear algebra
- Computes linear combinations of features to maximize feature variance
- Calculates relevance factor for each 'transformed' feature
  - PCA = 99% used for this analysis

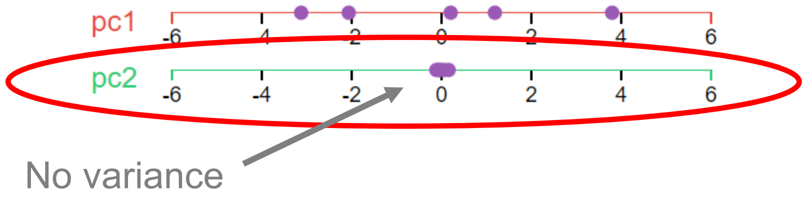
PCA Simple Example



Original data variance among each axis …

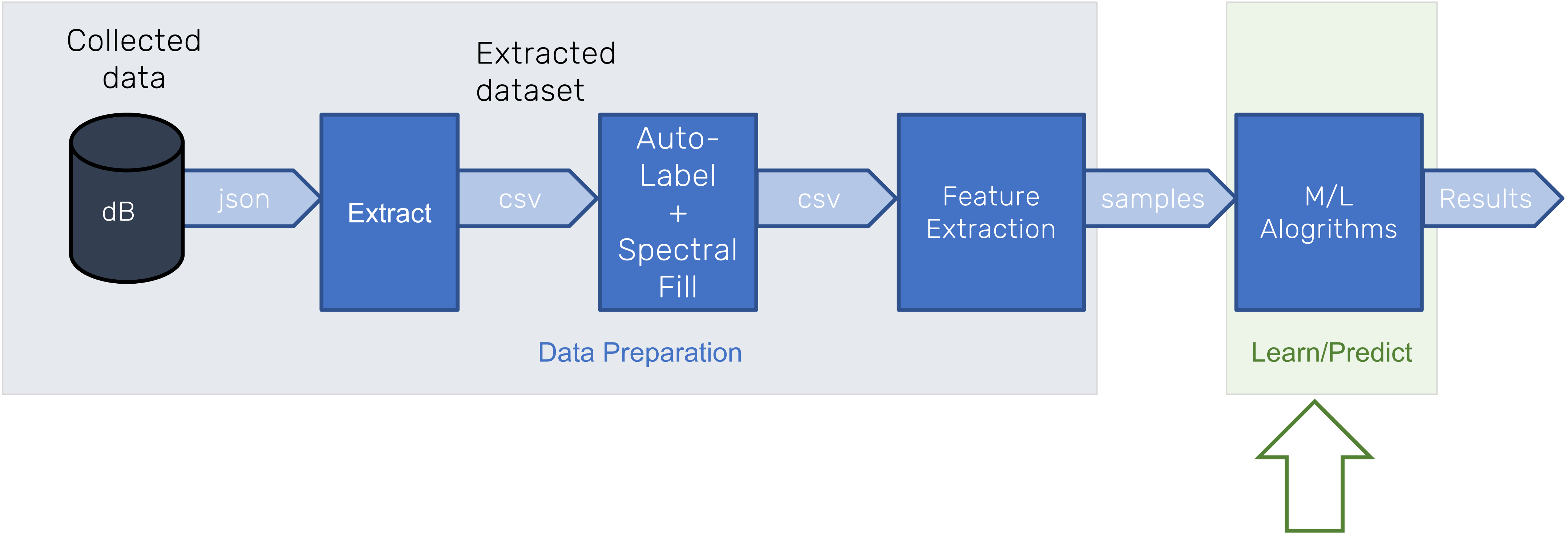


Little information on pc2, can drop this …



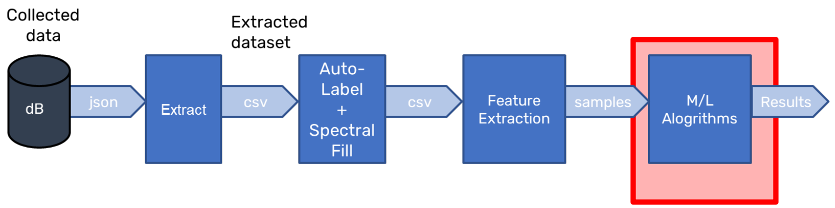
 Source: https://setosa.io/ev/principal-component-analysis/

A Machine Learning Pipeline



- Learning Models
- Tune/Train
- Test/Evaluate

SCTE. CABLE-TEC EXPO
ATLANTA, GA > OCTOBER 11-14



Adaboost

- A boosting algorithm that utilizes a Decision Tree as its base classifier. Final classifier based on linear weighted combination of base classifiers.

Logisitic Regression

- A classifier based on a linear regression model.

Multi-Layer Perceptron (MLP)

- A basic neural network classifier allowing for non-linear decision boundaries.

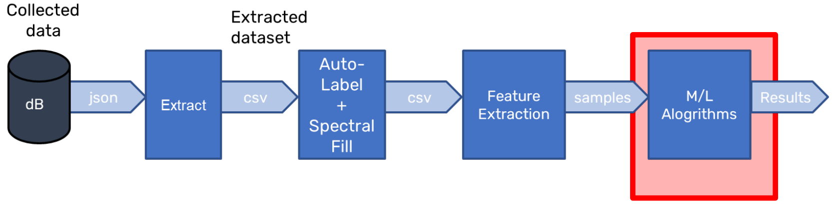
Convolutional Neural Net (ResNet)

- A convolutional neural network. Resnet utilizes the spectral samples directly and not the extracted feature set allowing for a simpler processing pipeline.

All M/L algorithms evaluated have associated hyper-parameters impacting performance.

- Some level hyper-parameter performance tuning was performed prior to running full dataset to optimize performance.

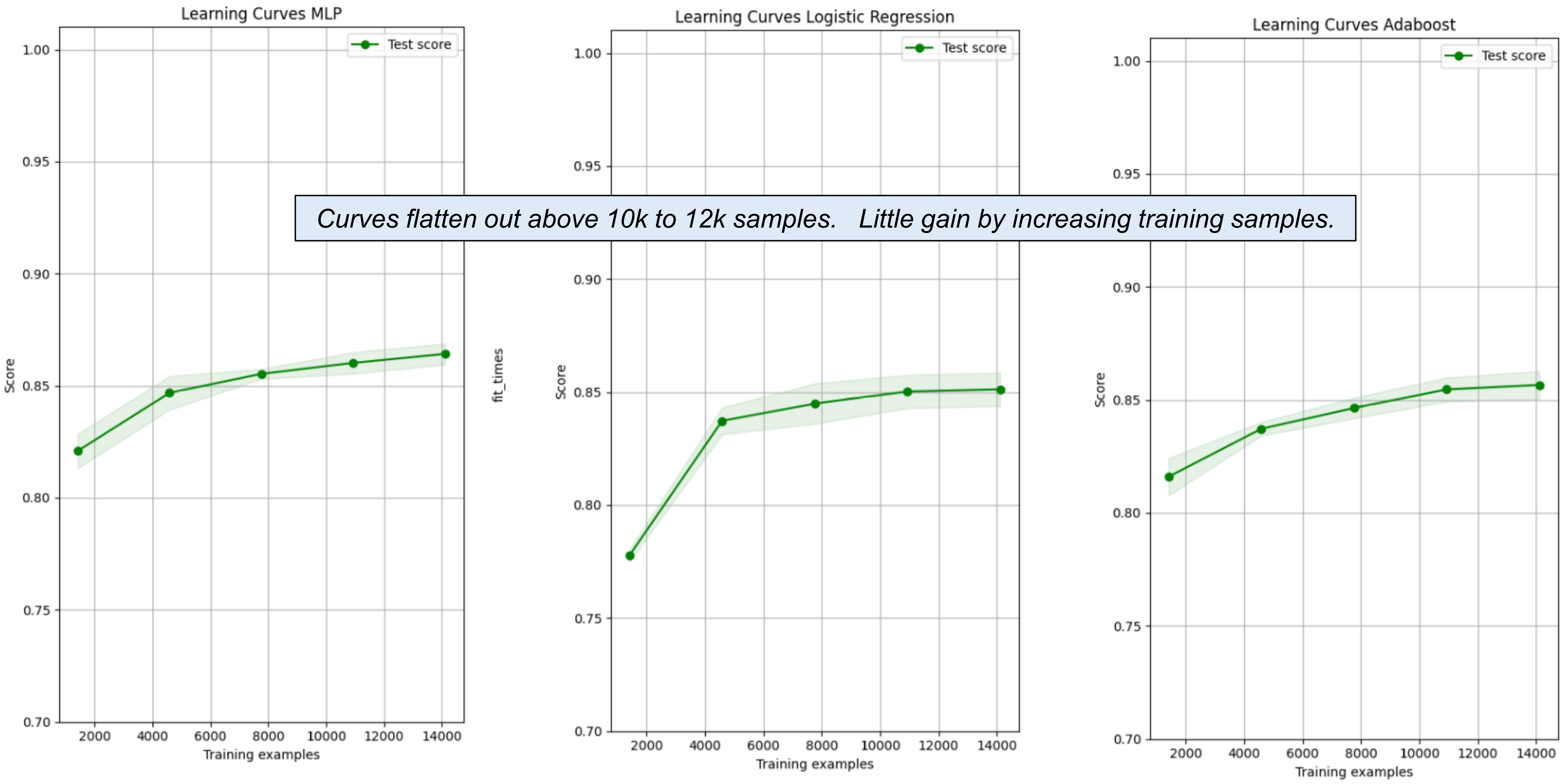
Learning Curves



- Evaluate algorithms across training set sizes
  - Indication on how much data to be effective
  - Appears diminishing returns at 10k samples or more
- Run with specific set of hyper parameters

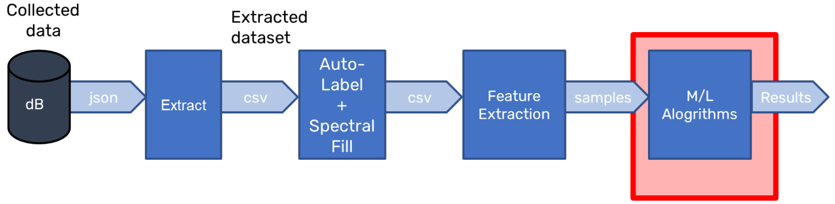
- Useful tool to provide guidance on the amount of training data needed to train algorithms.

Learning Curves

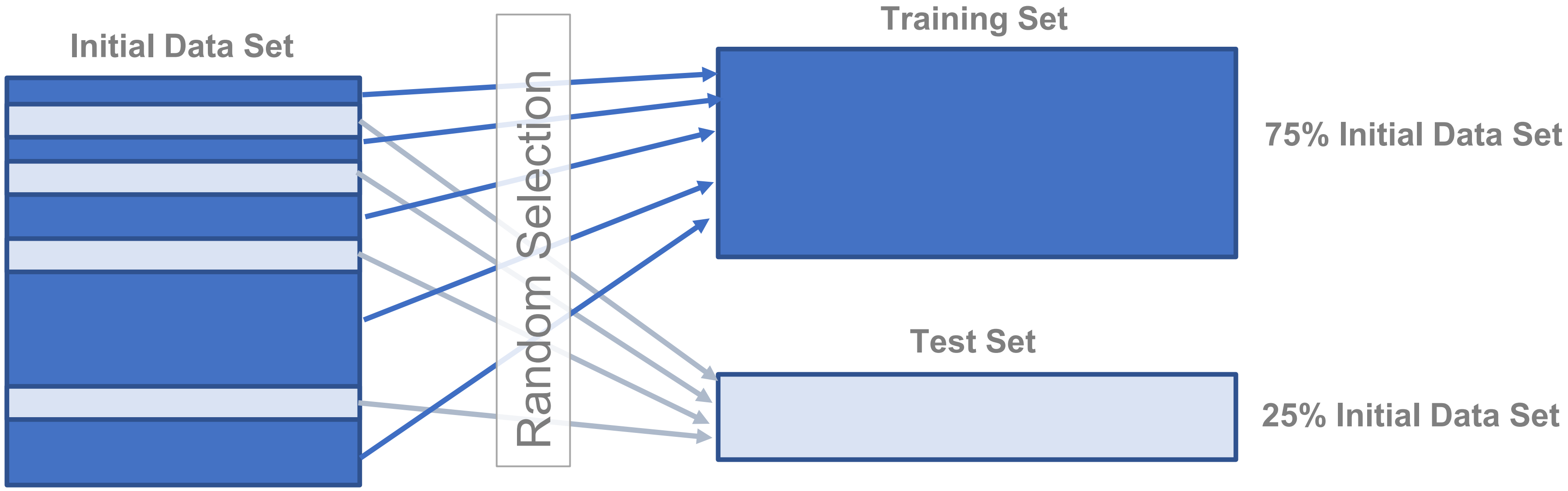


Curves flatten out above 10k to 12k samples. Little gain by increasing training samples.

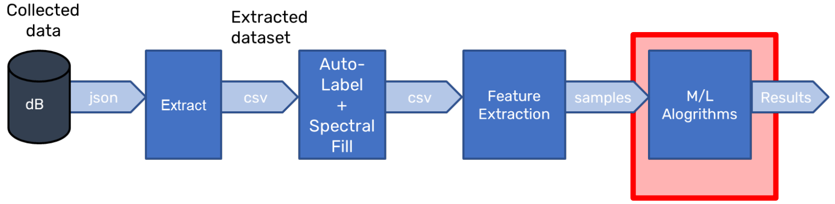
Data Splitting for Train and Test



- Data set is split between Train & Test
- Training set used to train algorithm
  - Normalized prior to training
- Test set used to evaluate algorithm after training
  - Accuracy and Confusion Matrix data captured



Data Sets



2 Data Sets evaluated – one from each CMTS network

- CMTS networks are unrelated

Raw data samples exhibit high bias toward 'good' samples

- Reduced data set to achieve close to equal bias or 'good' to 'bad' ratio close to 1

Data sets exhibited different levels of unused spectrum

Each data set includes two types:

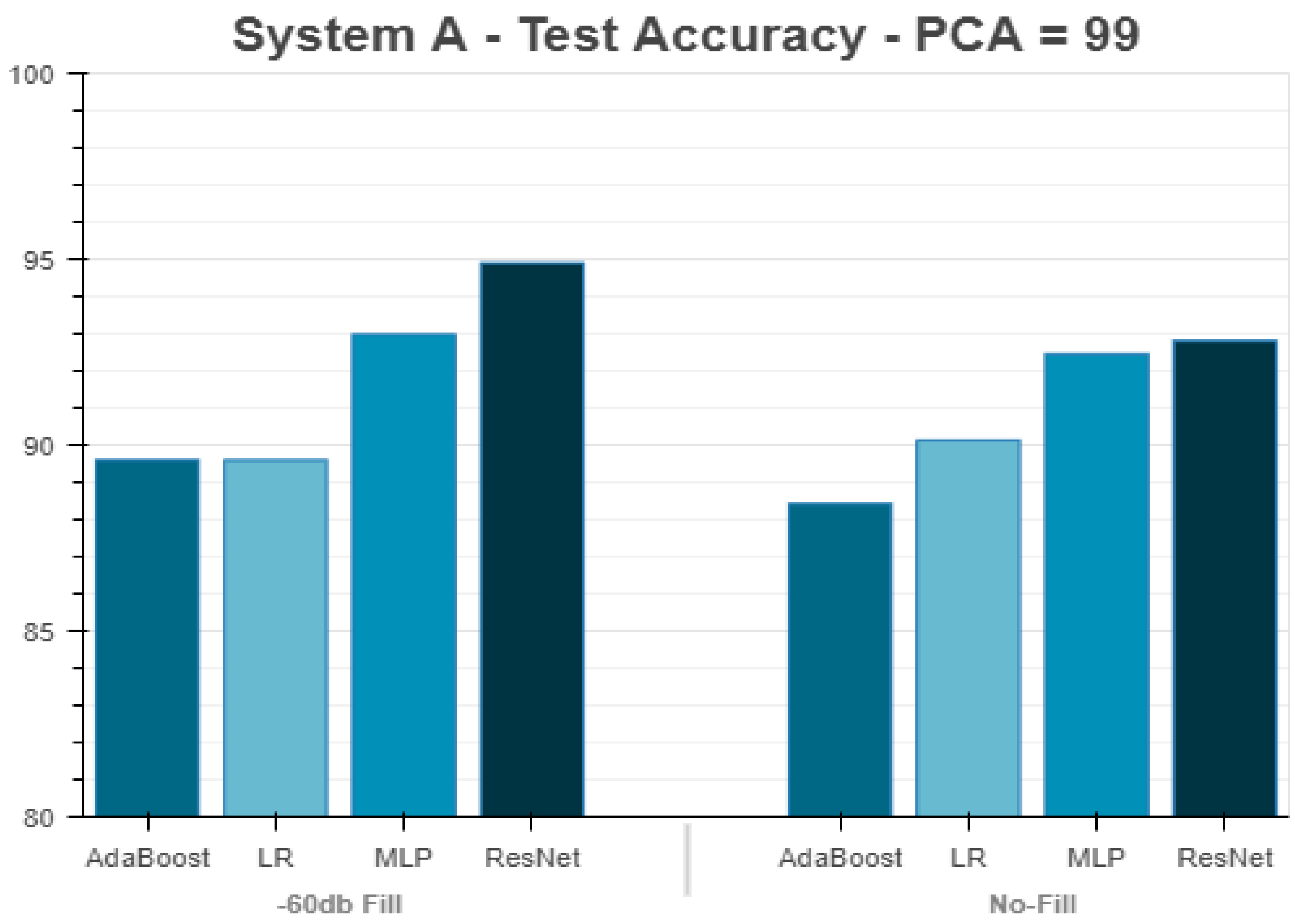
- Unused spectrum left as is in data sample - 'no-interpolation'
- Unused spectrum filled with constant -60dBmV value.

System	Samples	Unused Spectrum (%)	Bias (%)
A	11226	24.65	51.283
B	16296	32.01	51.283

Data set impaired samples and unimpaired samples nearly equal.

30% more unused spectrum in dataset B vs dataset A

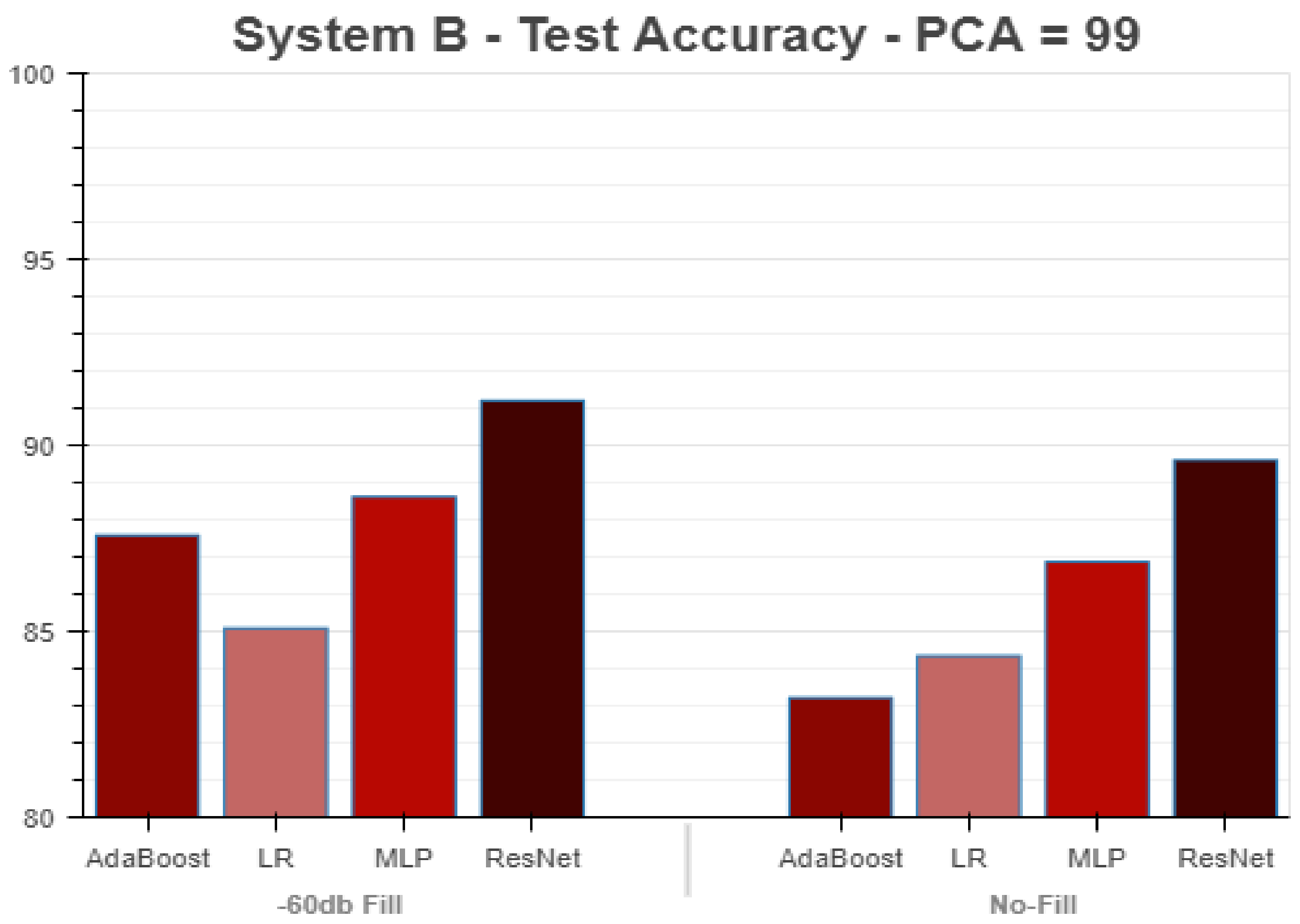
Results



System A	PCA = 99		
	NoFill	-60db Fill	Diff
AdaBoost	88.42	89.6	1.18
LR	90.13	89.6	-0.53
MLP	92.45	92.98	0.53
ResNet	92.8	94.9	2.1

- Strong Overall Results for all algorithms – Simply guessing would provide about 50% accuracy.
- Synthetic fill improved most results
  - Exception: Logistic Regression
- MLP & Resnet (Neural Networks) had best results.

Results



System B	PCA = 99		
	NoFill	-60db Fill	Diff
AdaBoost	83.21	87.58	4.37
LR	84.34	85.08	0.74
MLP	86.87	88.61	1.74
ResNet	89.6	91.2	1.6

- Strong Overall Results for all algorithms – Simply guessing would provide about 50% accuracy.
- More unused spectrum than system A.
  - Results not as good as System A
  - Synthetic fill greater improvements than system A.
- MLP & Resnet (Neural Networks) had best results.

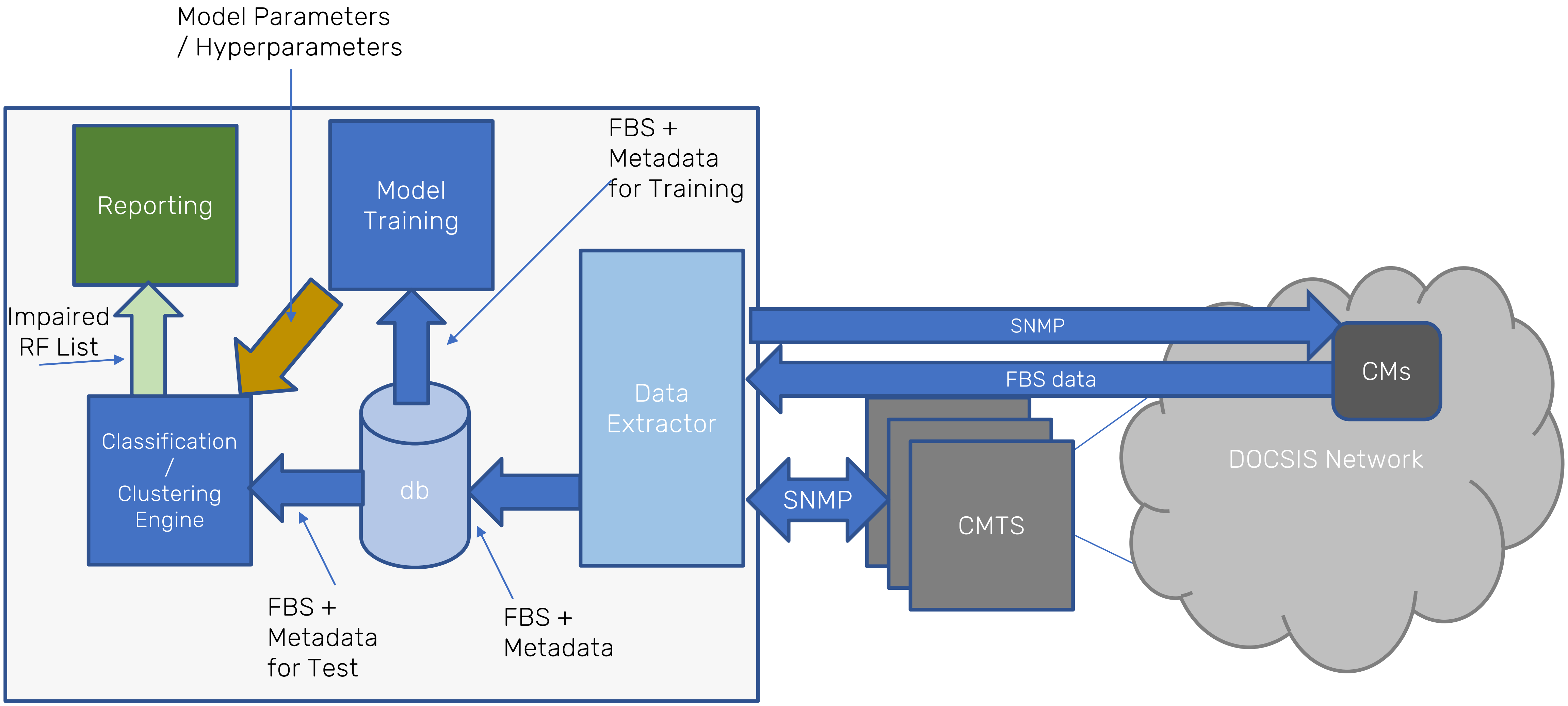
Results

- All ML Algorithms performed significantly better than a ‘dummy’ classifier
- System A and System B results were different
  - Possible due to difference in amount of un-used spectrum
  - System A had less un-used spectrum and better results
- Synthetic -60dB fill generally improved performance
  - LR on System A was exception
  - LR generates linear decision boundary, may limit performance
  - Larger improvements on System B using synthetic -60dB fill
    - System B had more unused spectrum
- ResNet and MLP algorithms had best performance
  - Capable of non-linear decision boundaries

Next Steps

- Use of Machine Learning to detect RF Impairments using FBS is very promising
  - Proactive maintenance focus, improved customer satisfaction
- Enhanced models - more and/or improved data points:
  - Multi-level categorization to identify specific types of impairments
  - Impairment grouping within topology
  - Feature development using specific RF domain knowledge
  - Customer experience/feedback, environmental data, operational data (e.g. plant modifications/changes)
- Model tuning to improve results
  - Hyperparameter adjustments
  - Node level aggregation

Implementing a Closed Loop System



ATLANTA, GA
OCTOBER 11-14

Thank You!

David Virag
Distinguished System Engineer
Commscope
david.virag@commscope.com