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

USING ENHANCED ANALYTICS TO ENHANCE CUSTOMER SATISFACTION

SCTE · ISBE

SIMPLIFYING FIELD OPERATIONS USING MACHINE LEARNING

Sanjay Dorairaj
Sr Director, Comcast Innovation Labs
Comcast

SCTE · ISBE CABLE-TEC EXPO 2017

DENVER, CO
OCTOBER 17-20

About the Speaker

- **Sanjay Dorairaj**
- Work at Comcast
  - 12th year at Comcast
  - Now @Comcast Innovation Labs
  - Past – RDK
- Education
  - MS Computer Science@Bharatidasan University
  - MBA Marketing @Temple University
  - MS Information & Datascience Student@UC Berkeley
- Adjunct @ San Jose City College

COMCAST

COMCAST LABS

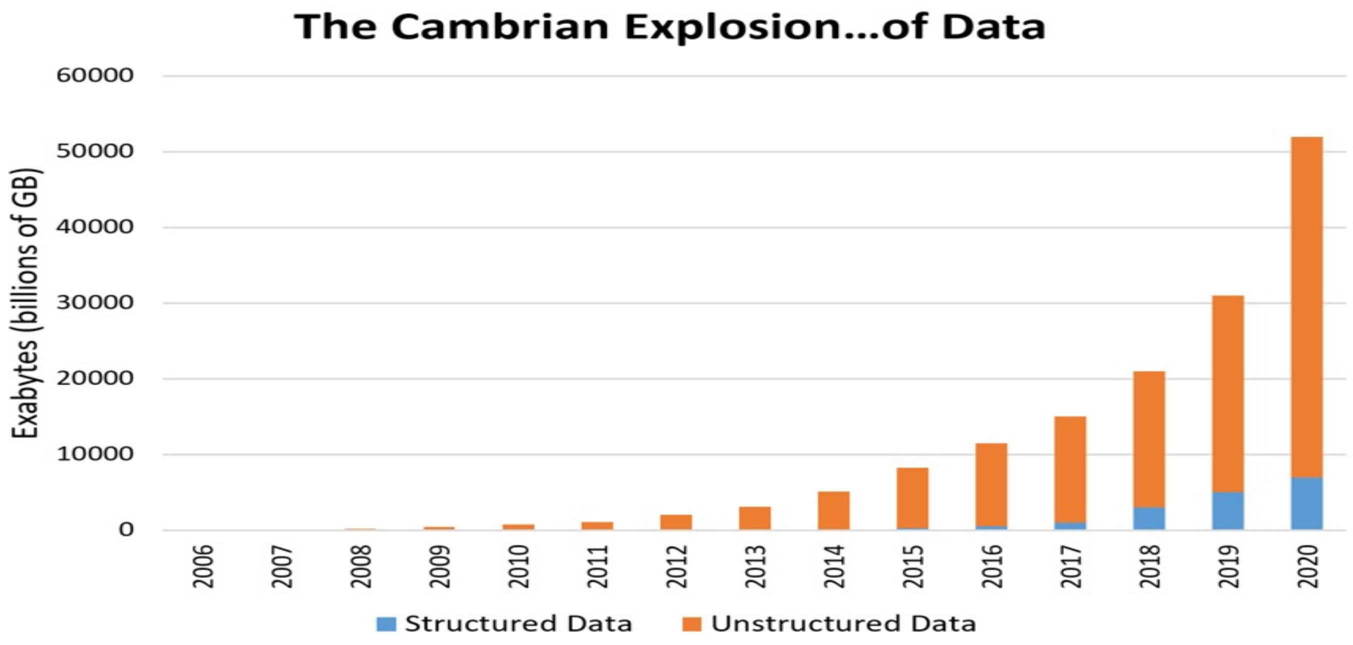
TEMPLE UNIVERSITY

SAN JOSE CITY COLLEGE

Machine Learning for MSOs

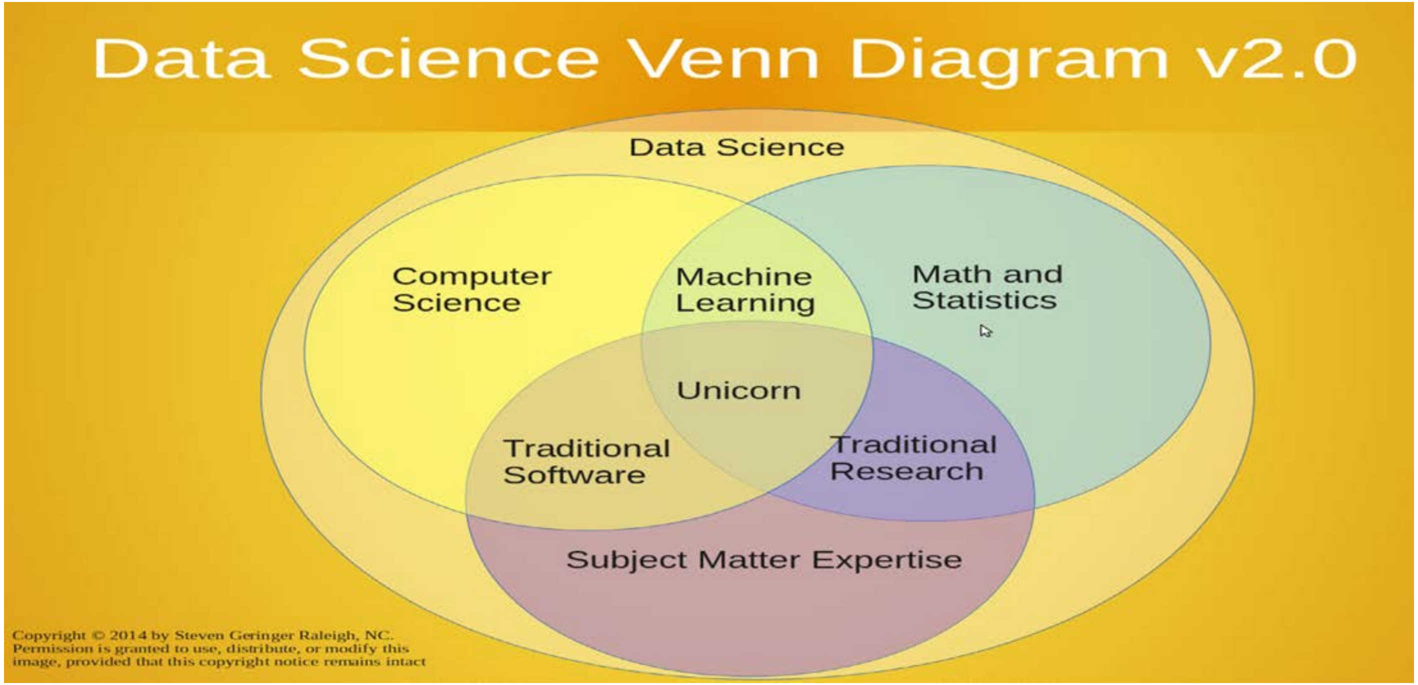
The relevance of machine learning for MSO operations

An Exponential Growth of Data



Source: Patrick Cheesman

The Advent of Data Science



Why does this matter for service providers?

- Predictive analytics can flag potential customer service issues early on
- ML tools can perform root cause analysis to identify underlying issues and recommend remediation actions
- ML insight in field tools drives down call volume and truck rolls, decreasing costs while increasing service quality
- And many more…

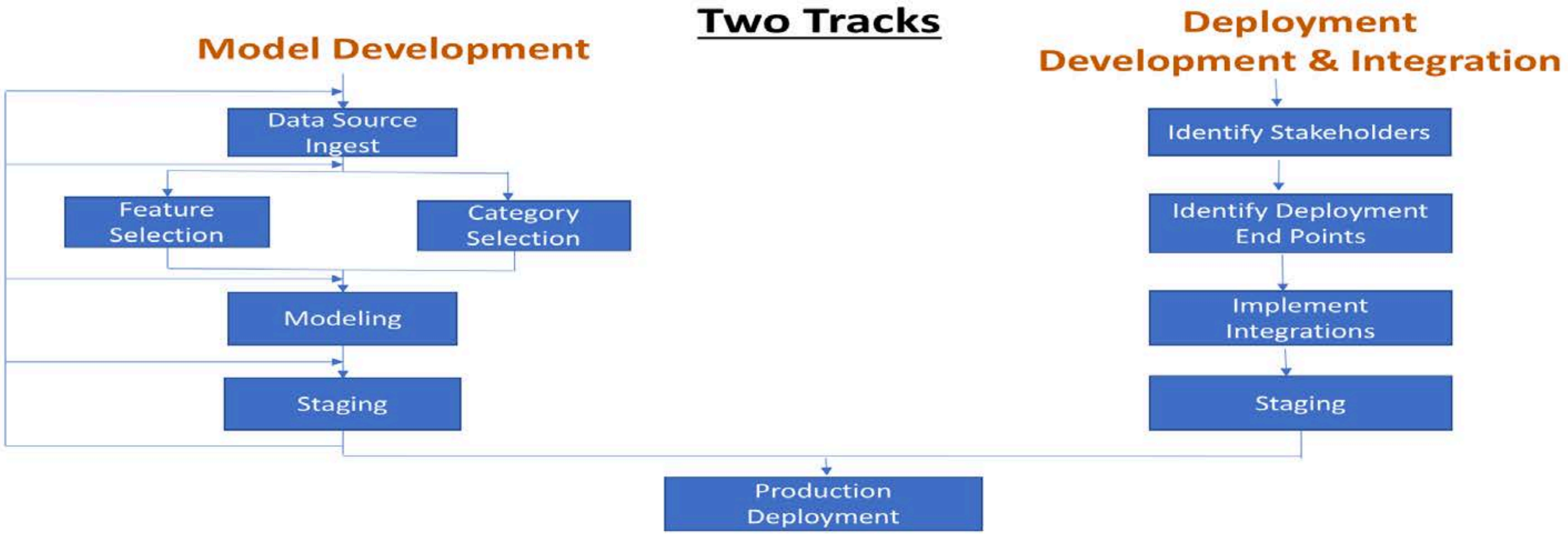
What is Machine Learning?

Class of Algorithms	Description	Technology Examples	Applications
Classifiers	Assigns data to categories based on similarity to other data.	Random Forest and Neural Nets	Sentiment Analysis, Image Classification
Clustering Algorithms	Groups similar data into clusters	K-Means, Hierarchical Clustering	User profiles and anomaly detection
Recommender Systems	Make recommendations based on historical data	Collaborative Filtering	Product recommendations
Anomaly Detection	Identify rare events	Joint Probabilistic modeling	Billing fraud detection
Linear Regression	Predict values for continuous variables	Linear Regression	Churn rate prediction

Concepts in Machine Learning

- Supervised versus unsupervised machine learning
- Training set, test set and Hyperparameters
- Features engineering
- Ensemble approaches
- Online versus offline algorithms

Development and Deployment Process



A Different Operating Paradigm

- Certainty involves a hundred percent guarantee that an event will occur
- Machine Learning **does not** deal with certainties
- What you get is an outcome and the probability of that outcome
- Example
    - There is a 95 percent chance that the spectral signature contains a wave impairment
    - There is a 90 percent chance that the customer will call today with a problem

Evaluating ML Outcomes

- Balancing cost reduction, customer satisfaction and model complexity
  - Large volume of repair calls implies that small improvements can yield sizeable cost savings
    - 1 million repair calls a month at a hypothetical $10 per call implies a monthly cost of $10m
    - A **1 percent reduction** results in a 100K monthly saving and an annual saving of approximately **1.2 million dollars**
  - False positives are tuned on a case by case basis to maximize cost reduction while minimizing disruption to the customer
    - A destructive self-healing action such as a reboot would require higher precision
    - A non-intrusive self-healing action such as a billing change would allow for lower precision

Spectral Impairments Detection

Case Study in Machine Learning

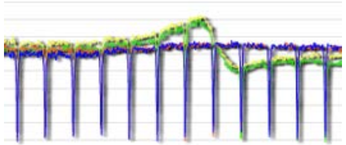
Spectral Impairment Detection

- Cable operations monitor the use of the spectrum for every device (e.g. cable modem). Such measurements give a state of the communication between the network infrastructure and the device.
- **Experts have identified 15 impairments** for which automatic detection would bring a competitive advantage. **Each of these impairments has an identified cause, and is linked to a repair action improving the performance of the RF spectrum.**

Goal of Spectral Impairment Detection

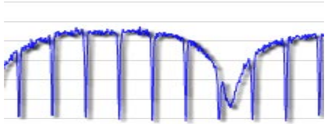
- Automatically characterize these spectra by labeling all their impairments.
- Key to assessing the performance of our RF spectrum, to standardize automation & detection or anomalies and remove subjectivity and manual interpretation by our technicians.
- Tasks
  - Detect 9 impairments instead of 5 currently
  - Reach higher accuracy in these detections

Spectra Classification - Examples



is a *resonance peaking* caused by an amplifier problem

adjacency/alignment caused by a head-end problem



is *a suck-out* caused by a home problem.

15 Impairments: Suck-outs, Notches, Tilt (and direction), Ripples / Waves, Off-Air Ingress, Foreign carriers, Wideband / Edison, Roll-off, Resonance / Peaking, Filters, Leveling, Adjacency / Alignment, Power Summary, Distortion / Intermod, Pilot to Channel ratio

Approaches

Method 1: Direct interpretation of the spectra data through traditional signal processing directly mapping features to impairments. Implemented, deployment in progress.

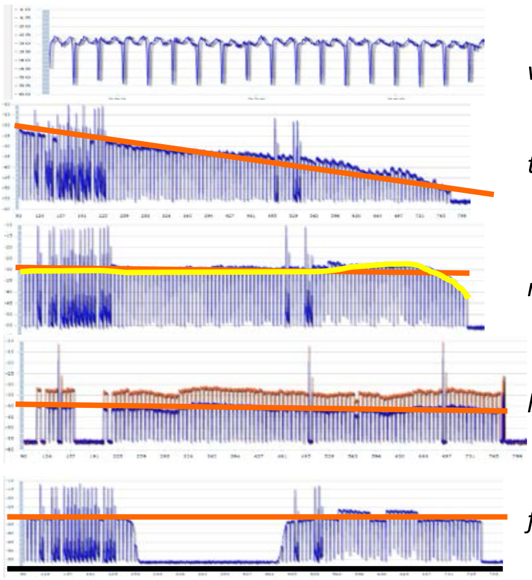
Method 2: Machine learning based approach that uses existing features and dynamically maps them to impairments, to reach optimality in predictions of new data.

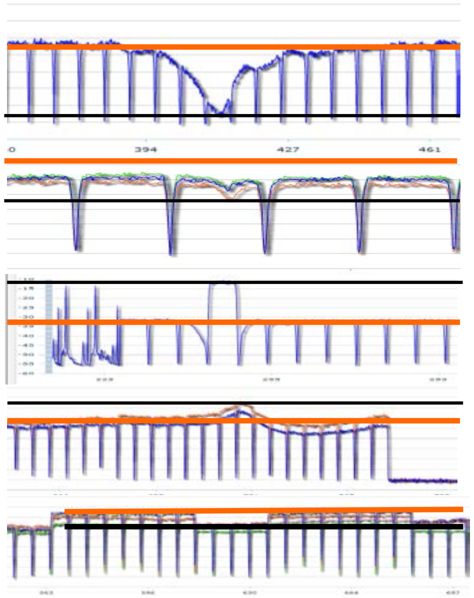
- **Comcast Labeling Engine:** Labeling spectra data is a key issue to address. Fortunately all field technicians are trained to label spectra data. A Turk was implemented to farm labeling task out. Implemented, deployed.

Method 1

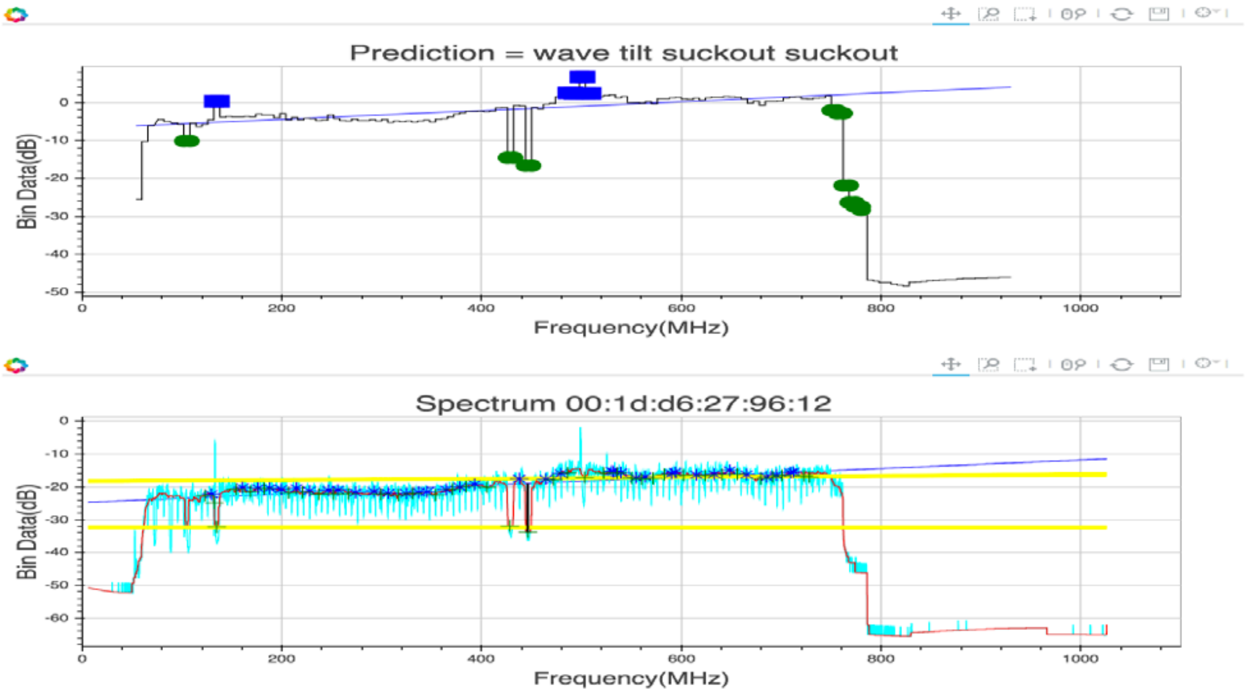
Mathematical Modeling of Spectral data

Spectrograph of Impairments





Basic Data Representation



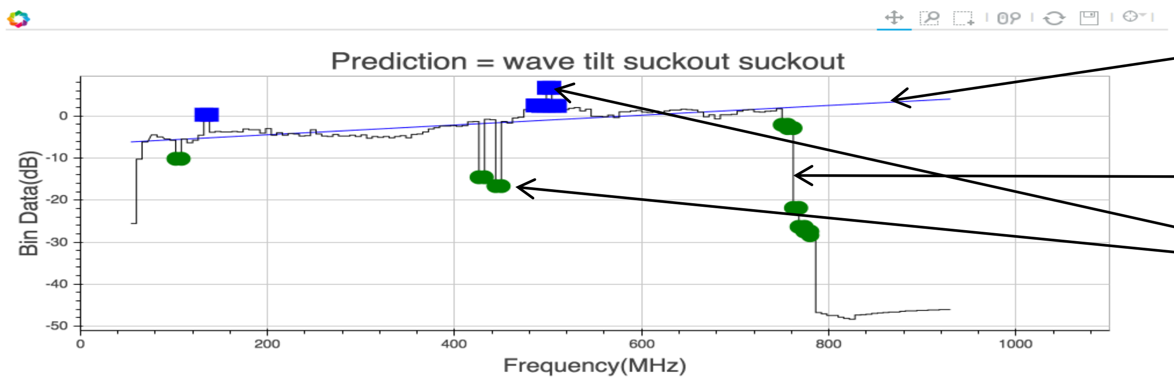
Channels

Represents each channel of 6MHz

Spectrum

Represents complete spectrum by samples of 117kHz

Feature Detection



plateau 1st order approximation
$y = ax + b$

channel signal

peaks (min and max)

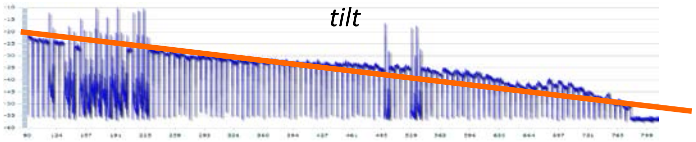
- Plateau 1st order approximation: $y = ax + b$ + residual
- Peaks
- Shape around peaks (single channel) or multiple channels
- Horizontal lines and steps

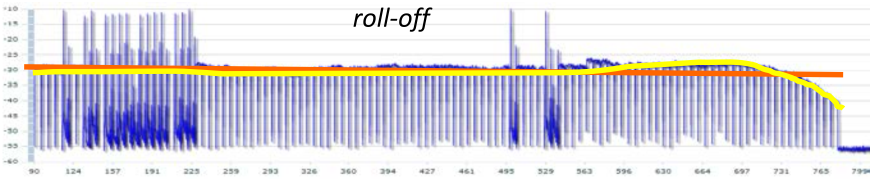
Detection of tilt versus roll-off

Spectrum: Plateau 1st order approximation: $y = ax + b$, res
Spectrum: Plateau 3rd order approximation: $y = ax^3 + bx^2 + cx + d$, res
Spectrum: residualRatio = 3rdorder residual/1st order residual

If residualRatio < Threshold then roll-off else tilt

The residualRatio tends towards 1 for a straight line





Detection of {suckout, notch} versus {foreign carrier and resonant peaking}

Plateau 1st order approximation: $y = ax + b$

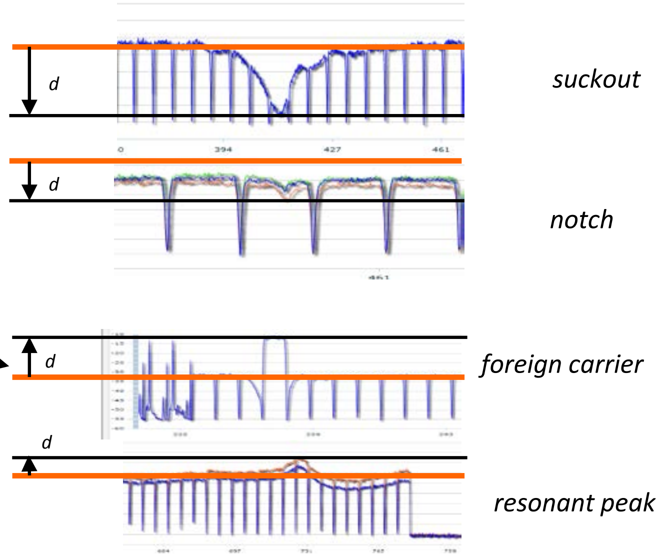
d = vertical distance peak depth/height to plateau

If d < 0

    then suckout or notch

else

    foreign carrier or resonant peaking



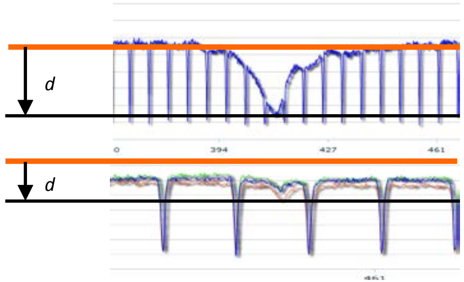
Detection of {suckout, notch}

Plateau 1st order approximation: y = ax + b

d = vertical distance peak depth/height to plateau

If d < 0

then suckout or notch



suckout

notch

Detection of {suckout, notch}

- **The suckout is a large dip spanning over several channels, whereas**
- **The notch is a tiny dip that cannot be seen in the channel view.**

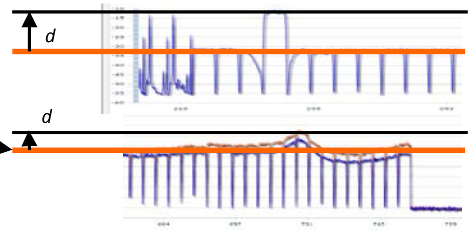
Detection of {foreign carrier and resonant peaking}

Plateau 1st order approximation: $y = ax + b$

d = vertical distance peak depth/height to plateau

If d > 0

then foreign carrier or resonant peaking



foreign carrier

resonant peak

Detection of {foreign carrier and resonant peaking}

- The foreign carrier is a sharp single channel peak in the signal, whereas
- The resonant peak is a shallow peak spanning over several channels.

Deployment Status

- The new impairment detection is not only more accurate, but can detect the existing 5 and 4 more: resonant peak, notch, foreign carriers, tilt and roll-off. With some work we can also detect leveling.
- There is still work to be done with this to fine tune the detection and reduce false positives. Technicians in the field **have no tolerance for false positives.** <u>**Meaning**</u>, failing to detect an issue is better than detecting false issues. These algorithms will need continued refinement before releasing them into beta.
- This work has been implemented with python and been adapted to a web service in Cloud Foundry (for now).

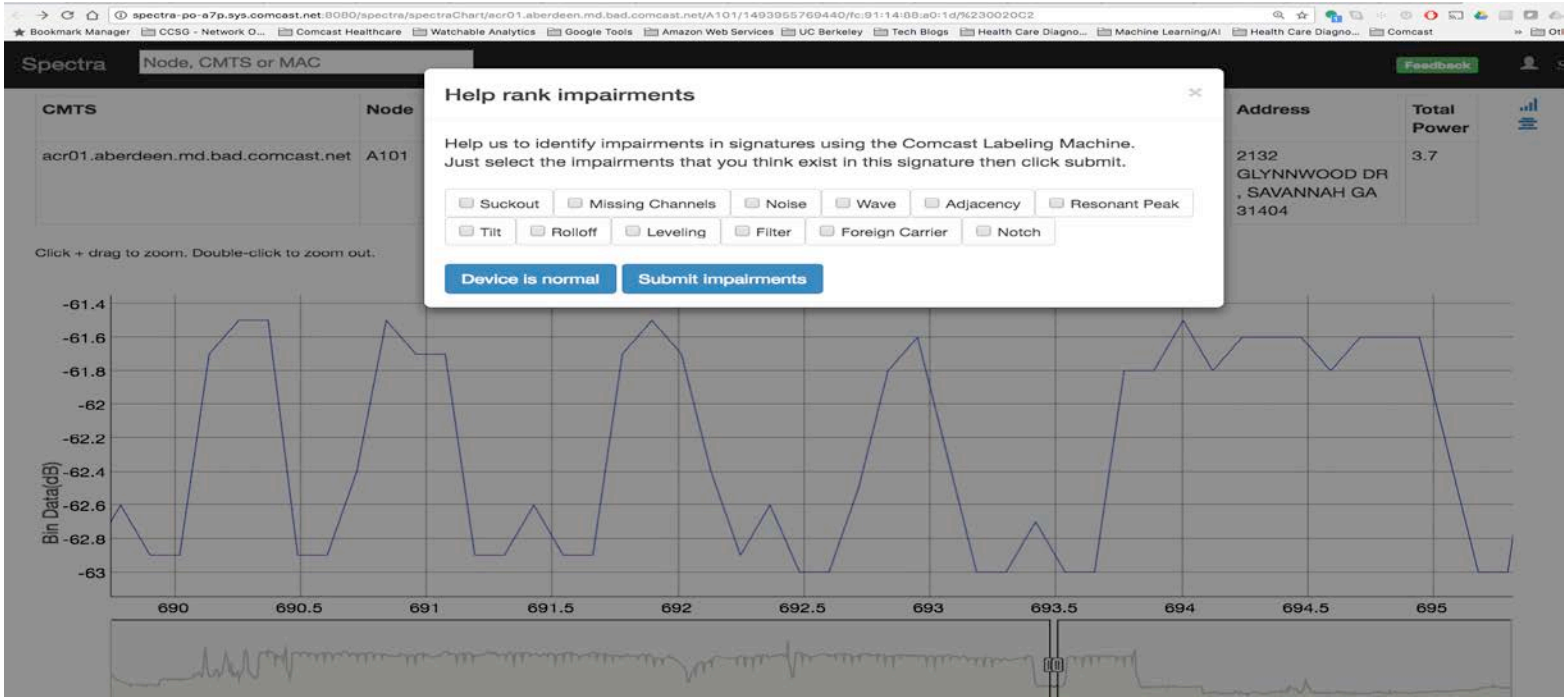
Method 2

Machine-learning based models: towards an optimal solution

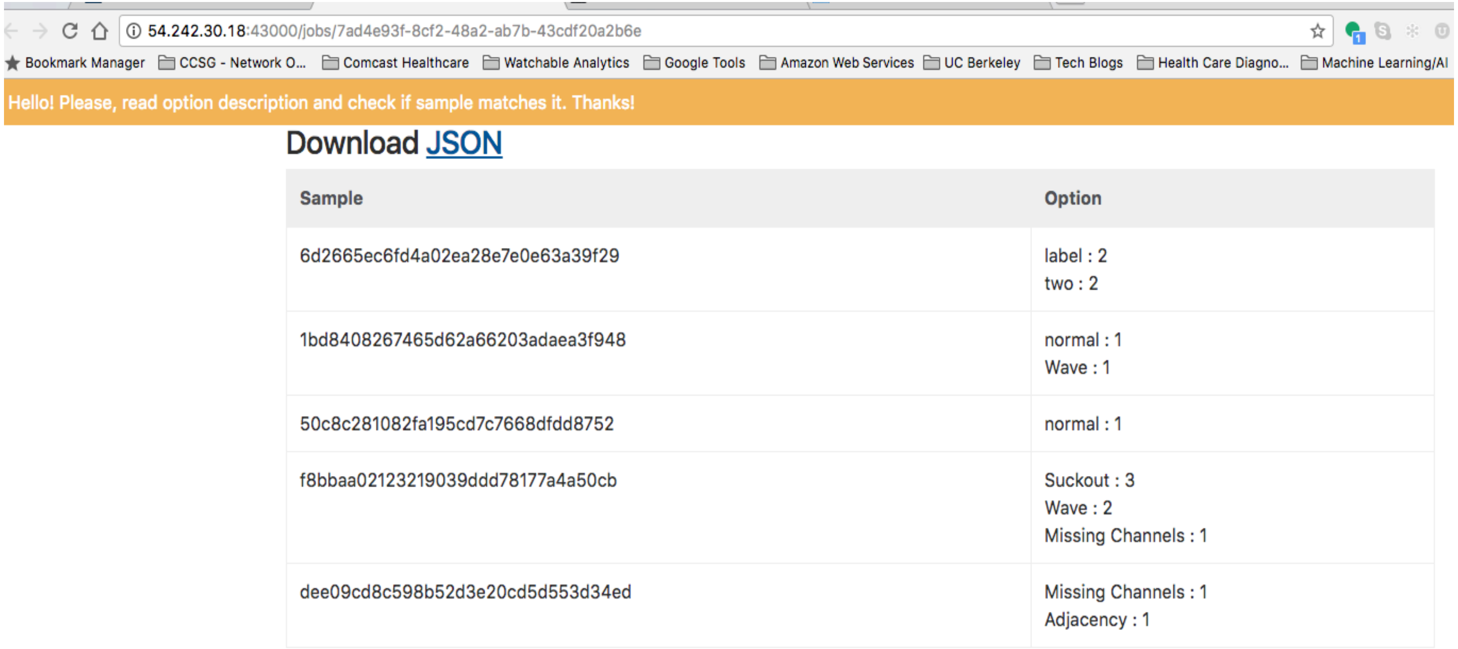
ML Model toward the Optimal Solution

- Supervised learning proves to be an efficient solution
- However, since impairments are non-exclusive, their combination leads to an exponential explosion ($15^{\text{something}}$)
  - ⇒ Requires labeling huge training sets, 10000’s of examples
- Solution is to farm labeling out to field technicians using the **Comcast Labeling Machine**

Spectra Tool - Impairment Identification



Comcast Labeling Machine



54.242.30.18:43000/jobs/7ad4e93f-8cf2-48a2-ab7b-43cdf20a2b6e

Bookmark Manager | CCSG - Network O... | Comcast Healthcare | Watchable Analytics | Google Tools | Amazon Web Services | UC Berkeley | Tech Blogs | Health Care Diagno... | Machine Learning/AI

Hello! Please, read option description and check if sample matches it. Thanks!

Download JSON

Sample	Option
6d2665ec6fd4a02ea28e7e0e63a39f29	label : 2 two : 2
1bd8408267465d62a66203adaea3f948	normal : 1 Wave : 1
50c8c281082fa195cd7c7668dfdd8752	normal : 1
f8bbaa02123219039ddd78177a4a50cb	Suckout : 3 Wave : 2 Missing Channels : 1
dee09cd8c598b52d3e20cd5d553d34ed	Missing Channels : 1 Adjacency : 1

Deployment Status

- Development of REST API backend and integration for spectral impairments is complete
- Currently in beta test in the Spectra tool pending validation and mass deployment
- Next Steps
  - Once labeled data is available proceed to modeling using CNN or similar approaches to predict spectral impairments in real-time

Conclusion

- Key Takeaways
  - Machine Learning requires a new operating paradigm
  - We must learn to work with uncertainties and use them to our advantage
- Case Study on Spectral Impairments
  - Beta testing of both features are currently underway
  - Next Steps
    - Training labeled data to build machine learning models to classify spectral signatures
    - Combine mathematical and machine learning methods to reinforce predictions
- For more information, please see the technical paper at the SCTE website.

SCTE · ISBE

THANK YOU!

Sanjay Dorairaj
Sanjay_Dorairaj@cable.comcast.com
408-900-8737