Creating Infinite Possibilities.

Machine Learning and Telemetry Improves Outside Plant Power Resiliency for More Reliable Networks

Stephanie Ohnmacht and Matthew Stehman

Comcast

SCTE CABLE-TEC EXPO 22

SEPTEMBER 19-22 • PHILADELPHIA, PA

2022 Fall Technical Forum

SCTE • CABLELABS • NCTA

Overview

1. Introduction
2. Identifying Power Supplies
3. Asset Inventory Interface
4. Power Supply Telemetry
5. Modeling Power Supply Health
6. Conclusions

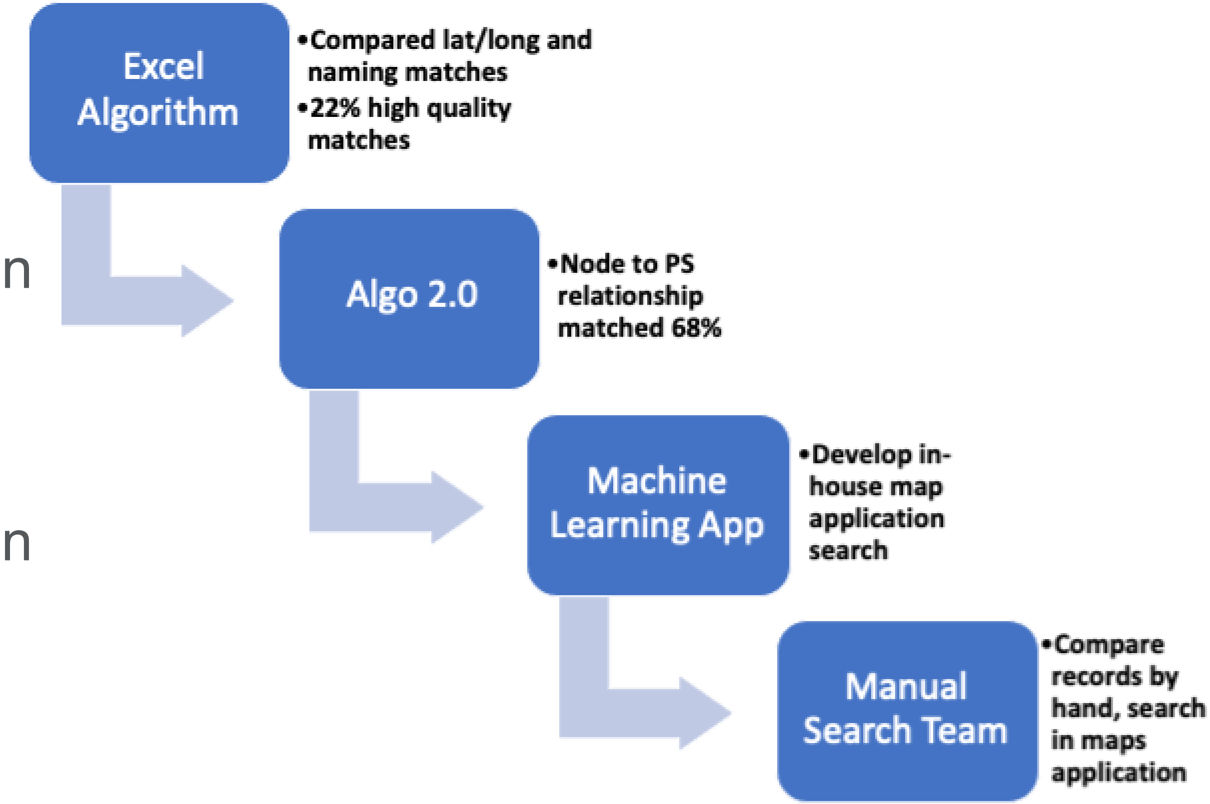
Problem Statement

- Comcast manages an extensive power supply network to power its HFC plant
- Resiliency of the power supply network is key to keeping customers connected
- Tracking and monitoring all aspects of the power supplies is paramount to ensure the power supply network is performing optimally
- Build predictive models to consume the rich data sets and provide actionable insights to the field

Build an architecture to monitor and expose all aspects of power supply information to support the next generation of predictive modeling

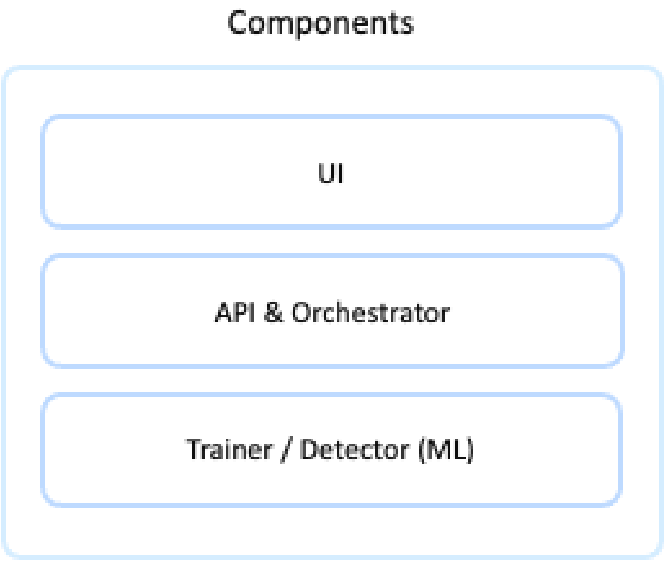
4 Step Process to Link Disparate Databases

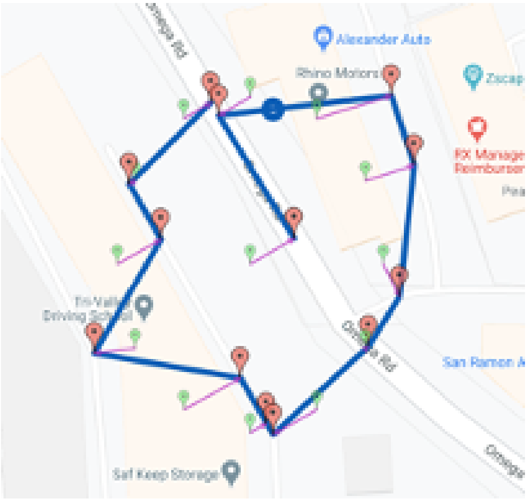
1. Compare linear feet between the two locations identified in databases of same power supply
2. Using node-to-power supply database relationship to solve for additional location matches
3. Identify PS in its natural environment (street or pole) using existing mapping applications to confirm actual real location
4. Manual team confirmed or denied proposed matches from the machine learning tool to update database

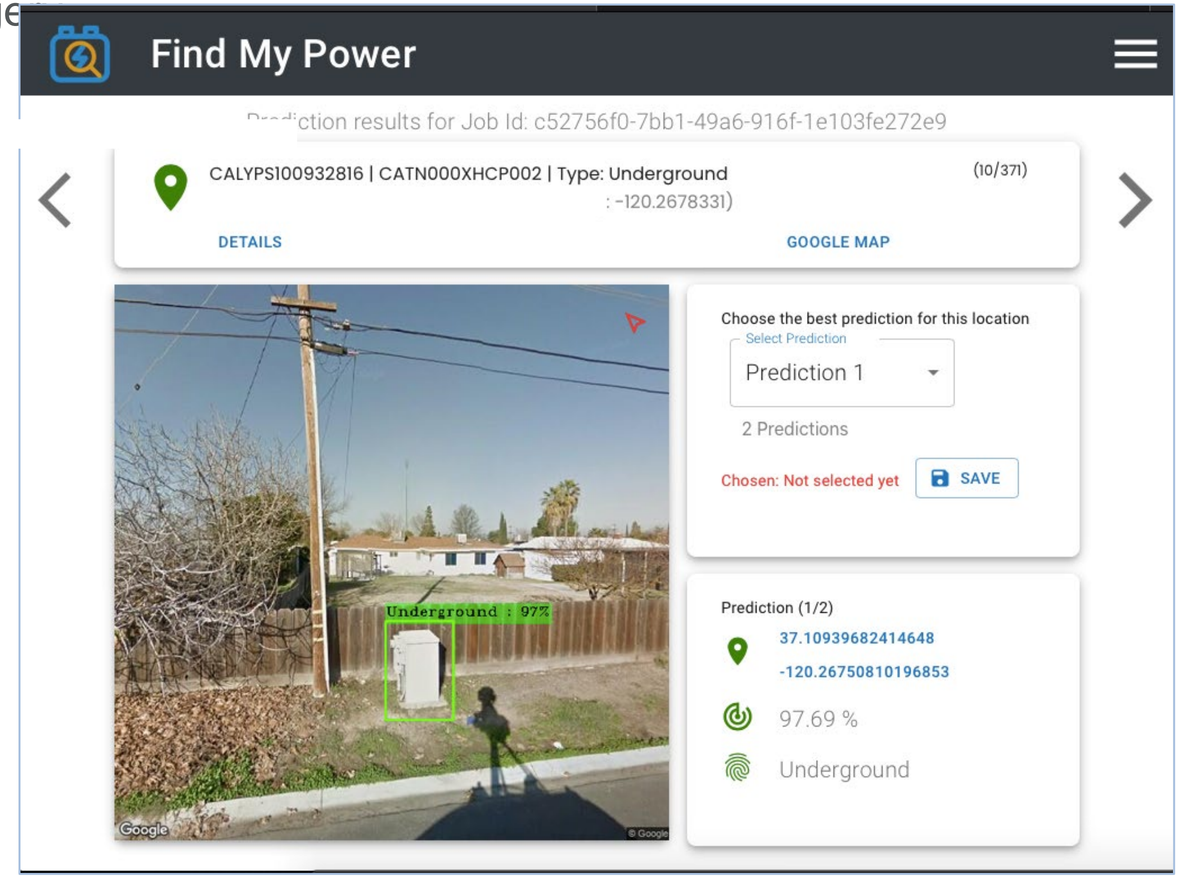


Machine Learning PS Identification Tool

- Created to identify PS in its natural environment and along using existing mapping software
- The tool determines a level of confidence match based on image likeness
- Found PS 650% of attempted locations
- Boosted manual team matching by 2X

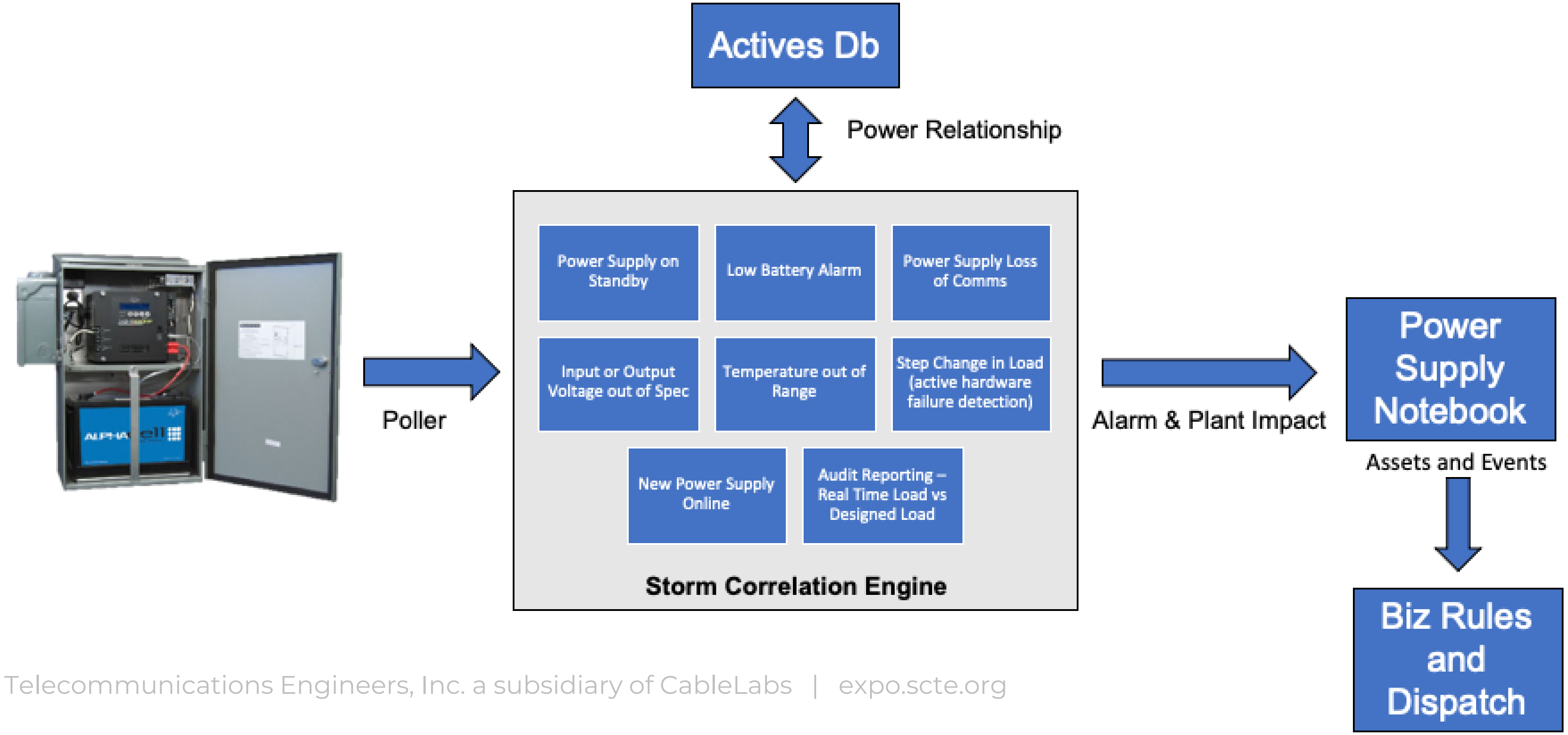






Tool Created to Manage Assets and Alarms

- Online asset management tool (Power Supply Notebook) to allow ongoing updates and changes by admins, technicians, business partners and help desk
- Power Supply telemetry flow through Storm, the correlation engine, and alarm events identified to be presented in Power Supply Notebook with associated actives



Disparate Power Grids

- Correlating power events to plant actives is giving more insight to all equipment downstream of the PS
- Data suggests that on average ~40% of customers stay online during a PS discharge scenario
- Teams are empowered by this additional layer of information to make more informed decisions about how and when to dispatch repair

PS Outage Impact Analysis Example

Alert Impacted Devices Analysis

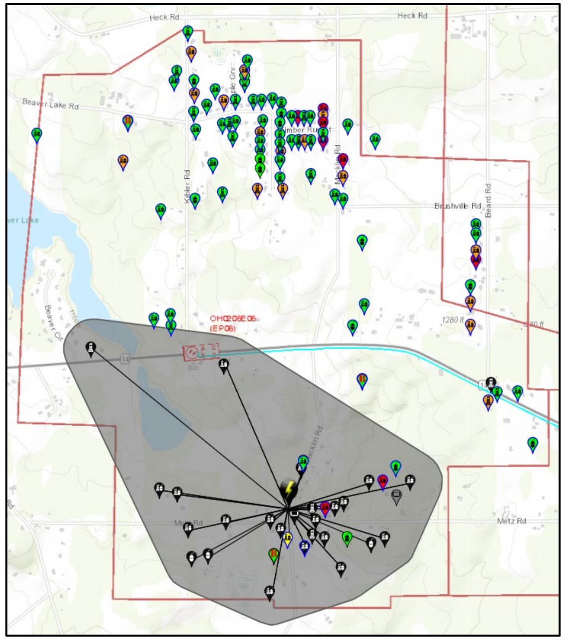
12:41:13 PM

Total Commercial Customer Offline:	0	Total CM Offline:	27	Total IVR Call Count:	0
Total Commercial Customer Online:	3	Total CM Online:	67	IVR Call Value:	NA
Total Residential Customer Offline:	51	Total MTA Offline:	24	Inverter Status:	2
Total Residential Customer Online:	137	Total MTA Online:	73	PS State:	on_battery
Customer Offline Value:	NA	Total Outage Event Found:	0	Task Name:	

Impacted Nodes

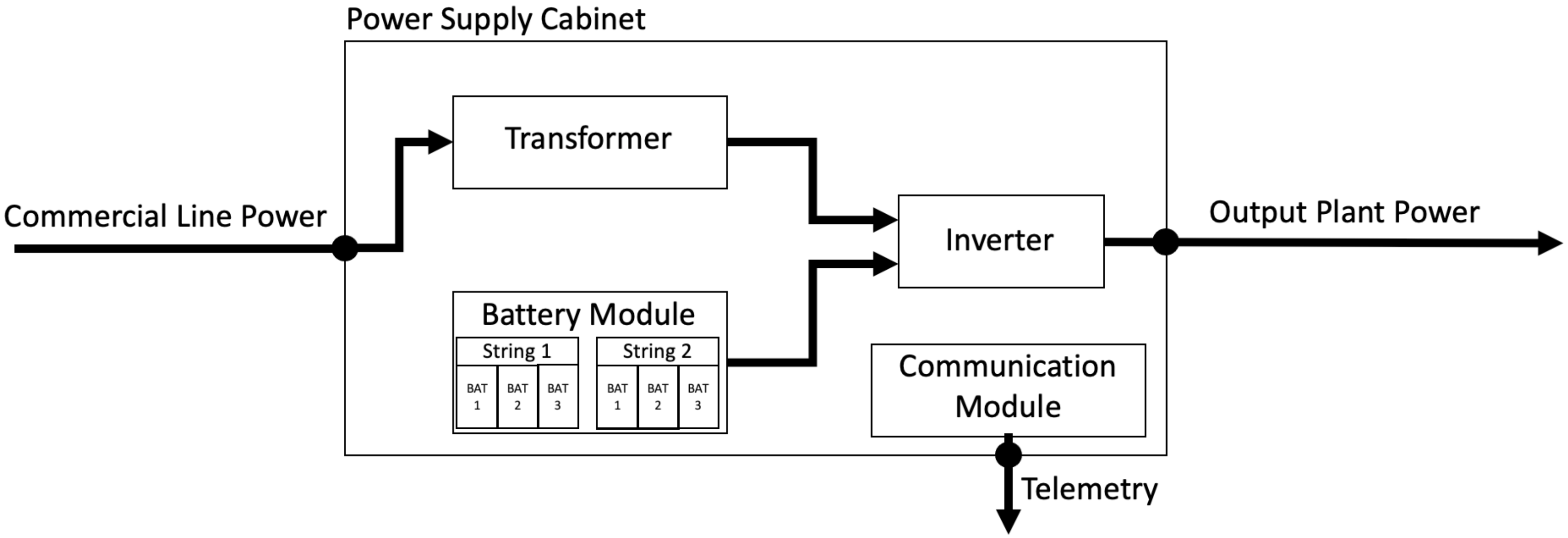
EP08A

Commercial Customer Offline:	0
Commercial Customer Online:	3
Residential Customer Offline:	51
Residential Customer Online:	137
Customer Offline Value:	NA
CM Offline:	27
CM Online:	67
MTA Offline:	24
MTA Online:	73
IVR Call Count:	0
IVR Call Value:	NA
Primary Node:	Yes
Storm Mode:	Disabled
Outage Event Found:	



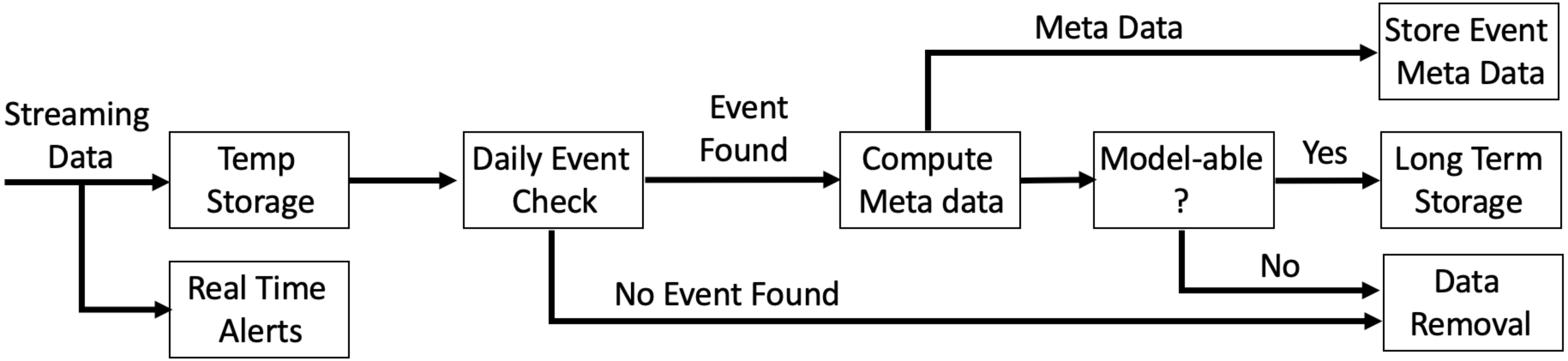
Overview

- Each PS has a rich set of telemetry related to the functional operation
  - Voltages, currents, temperature, battery levels, hardware, software, etc.
- Telemetry is polled from the transponder to form a live data stream
- Raw telemetry combined with asset inventory information has great potential for predictive modeling results



Data Pipelines

- Streaming
  - Real time alerts on raw telemetry
  - Scoring of PS's based on recent alerts
  - Historical accounting of alerts
  - Raw data is only saved for short time frame
    - Minimizes data storage costs
- Big Data
  - Process raw data and check for model-able events
  - Compute descriptive meta data on each event
  - Store raw data only for events of interest for modeling
  - Significant reduction in size of stored data
    - 0.25% of full raw data size



"We collect, store, and use all data in accordance with our privacy disclosures to users and applicable laws".

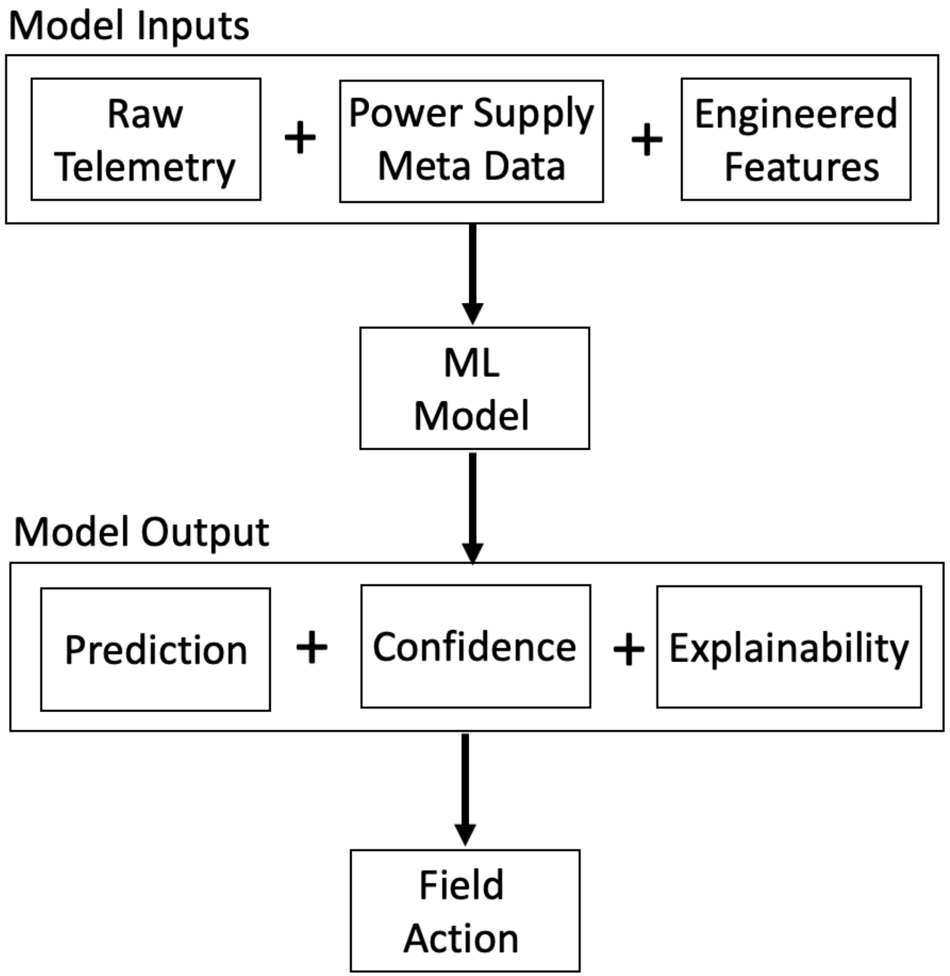
Objectives

- Use all available data to train predictive models
  - Inform field teams to efficiently prioritize resource allocation and maintenance activities
- Proactive Maintenance
  - Slow degradation leading to a potential customer impacting event
  - Example:
    - Batteries in PS are reaching end of life and may not be able to sufficiently power plant during a future commercial power outage
- Demand Maintenance
  - Active customer impacting event that needs resolution
  - Example:
    - Active commercial power outage requiring PS to operate on battery backup and field needs to decide whether action is needed to prevent a service outage

	Customer Impact	Resolution Window
Customer Impact	Now: Experiencing Impact Future: Impact Imminent	(Now, Now) Demand Maintenance
Resolution Window	(Future, Future) Proactive Maintenance	Now: ASAP Future: Months, Weeks

Run Time Prediction - Architecture

- Demand maintenance use case
  - Predict how long a PS can operate on battery back-up in event of commercial power loss
- Model Inputs
  - Raw telemetry
  - Relevant meta data
  - Engineered features of various forms
- Model Output
  - Predicted run time
  - Confidence level of prediction
  - Explanation of prediction
- Output needs to be understandable and useful to the field teams

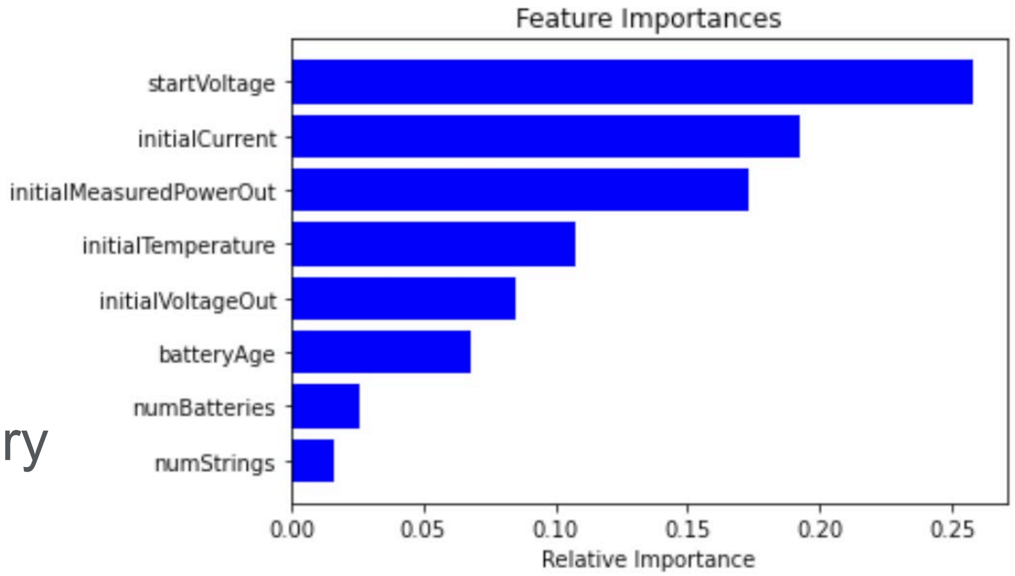


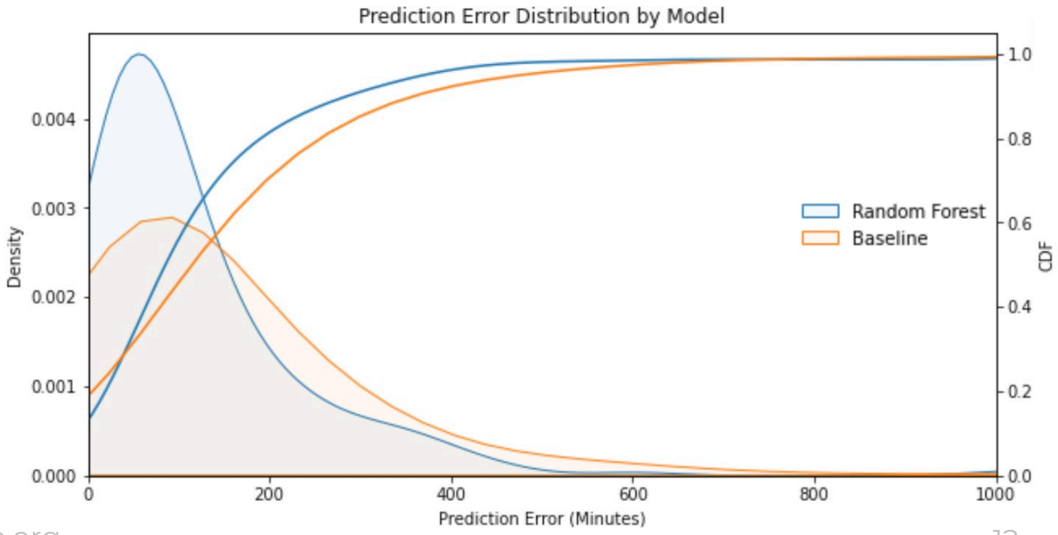
 | expo.scte.org

SCTE CABLE-TEC EXPO 22
SEPTEMBER 19-22 • PHILADELPHIA, PA

Run Time Prediction - Results

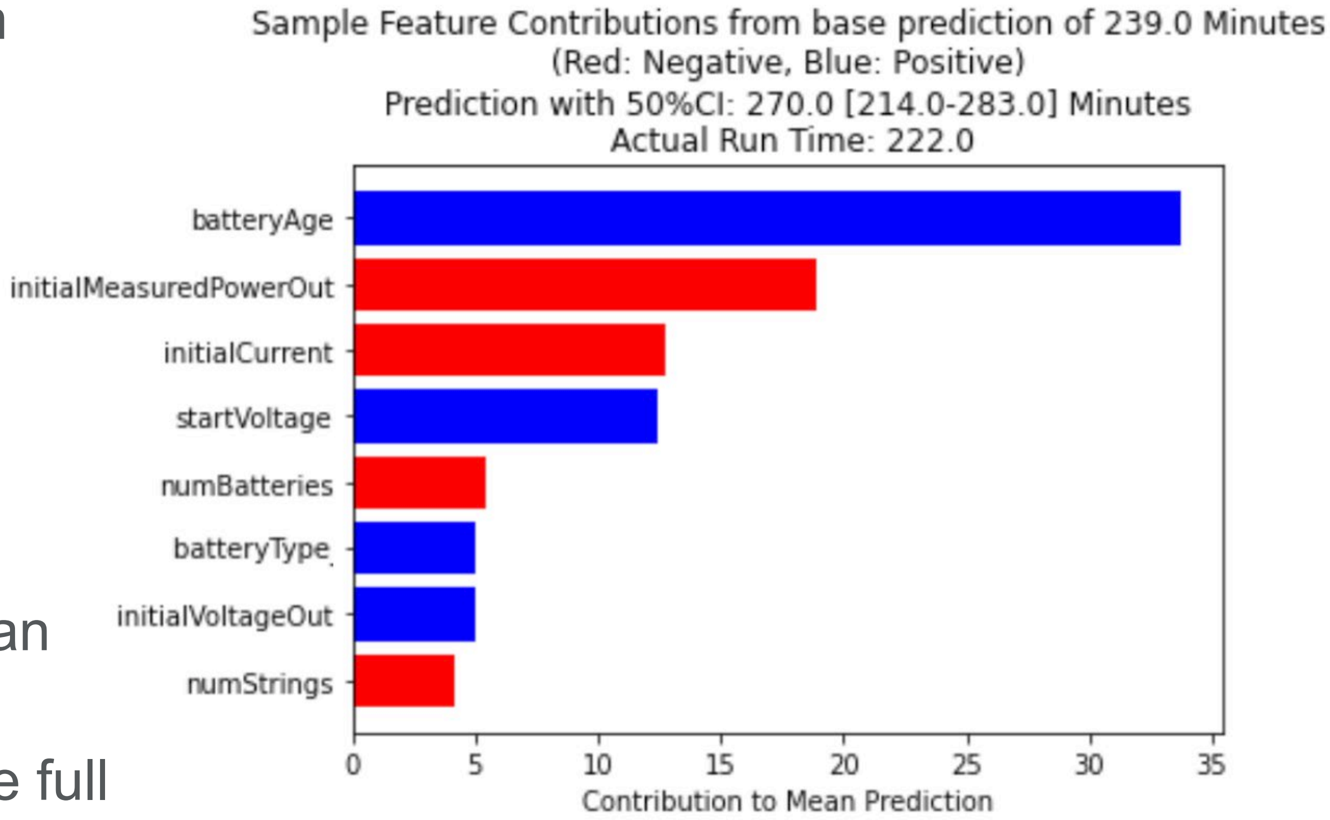
- Random forest model architecture
  - Very little constraints on inputs data
  - Inherently offers prediction confidence level and explainability
- Trained on ~2500 examples where PS ran on battery until it reached end of discharge
  - Feature importance indicates what inputs the model finds most useful in making predictions as a whole
- Test results compared to baseline model
  - Similar distribution of prediction errors
  - Slight improvement shown be distribution





Run Time Prediction - Explainability

- Individual model predictions include supplemental information to help field team prioritize
- Raw prediction (X minutes)
  - Same form as baseline model
- 50% Confidence interval
  - Width indicates model's confidence in prediction
- Individual feature contributions
  - What separates current prediction from an average baseline prediction
- Plan to keep expanding feature set to make full use of rich data sets



SCTE CABLE-TEC EXPO 22
SEPTEMBER 19-22 • PHILADELPHIA, PA

Findings

- A common PS identifier enables the Inventory Management System to gather network impact information from multiple sources
- Geography of the power grids from the utility company and broadband provider do not overlap
- Minute telemetry allows for additional granularity and improved response time in generating alarms
- Huge data storage savings by only saving data relevant to model-able events
- ML models perform at the same level as existing models with similar input data
  - Added benefit of supplementary information with each prediction

Future Work

- Automate dispatch for operational events
- Integrate predictive models into day-to-day proactive PS management and maintenance
- Enhance predictive models with more comprehensive feature engineering
- Develop proactive machine learning models

Contributors:

- Chris D’Andrea, Data Scientist, Comcast
- Ilana Weinstein, Data Scientist, Comcast
- Cory Thompson, Project Manager, Comcast
- Alex Falcon, Product Manager, Comcast

Creating Infinite Possibilities.

Thank You!

Stephanie Ohnmacht and Matthew Stehman

Comcast
Stehpanie_Ohnmacht@comcast.com
Matthew_Stehman@comcast.com