Agenda

- About Con Edison

- Con Edison's Enterprise Data Analytics Platform

- Identifying Open Neutrals in Three Parts
  - Initial Ideation
  - Application v1
  - Machine Learning Approach

- Open Neutrals Identified

- Future Roadmap

About Con Edison

- Founded in 1823 as New York Gas Light Company

- Provide electric, gas, and steam to NYC and Westchester
  - 160k miles electric cable
  - 4.2k miles steam gas mains
  - 105 miles steam mains & pipes

- 4M+ accounts serving 10M residents
  - 3.5M electric accounts
 - 1.1M gas accounts
  - 2k steam accounts

Gas and Electric Service Areas

Electric Service Areas

Con Edison's AMI Program

- Full scale AMI implementation
    - 4.8 million meters
    - 3.6 million electric & 1.2 million gas meters
        - 15 and 5-minute readings for electric
        - Hourly reads for gas
    - Communications network
    - System upgrades and enhancements
- Cost: $1.285 billion
- Deployment started in 2017 and planned for capital project close out end of 2022
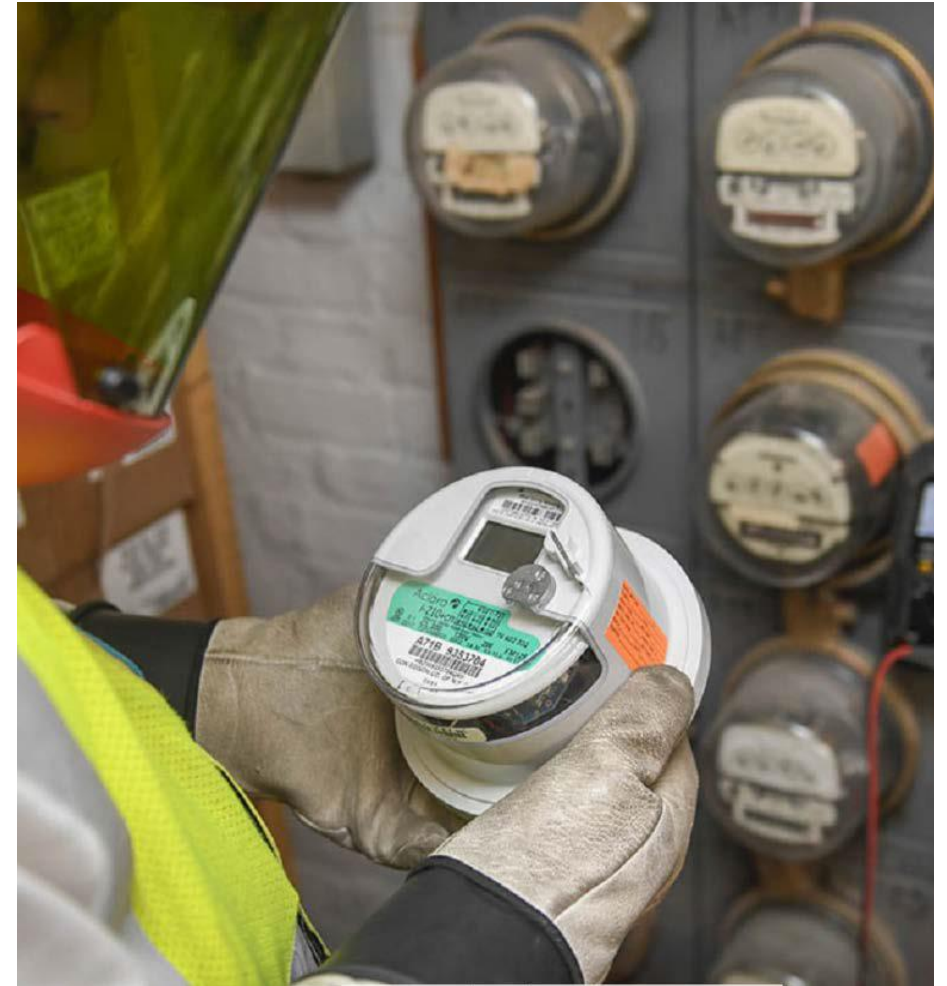
Enterprise Data Analytics Platform (EDAP) Overview

EDAP is comprised of a suite of technologies to deliver enterprise analytics capabilities.
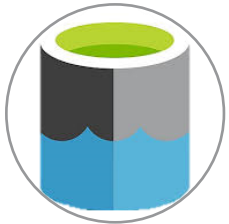The platform ingests data extracts from source systems, transforms it, and performs analytics
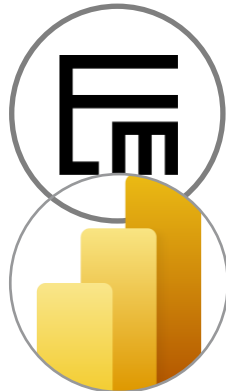
Big data processing & analytics

Modeling & forecasting at scale

Centralizing raw data / unified data model

Self-service analysis & reporting
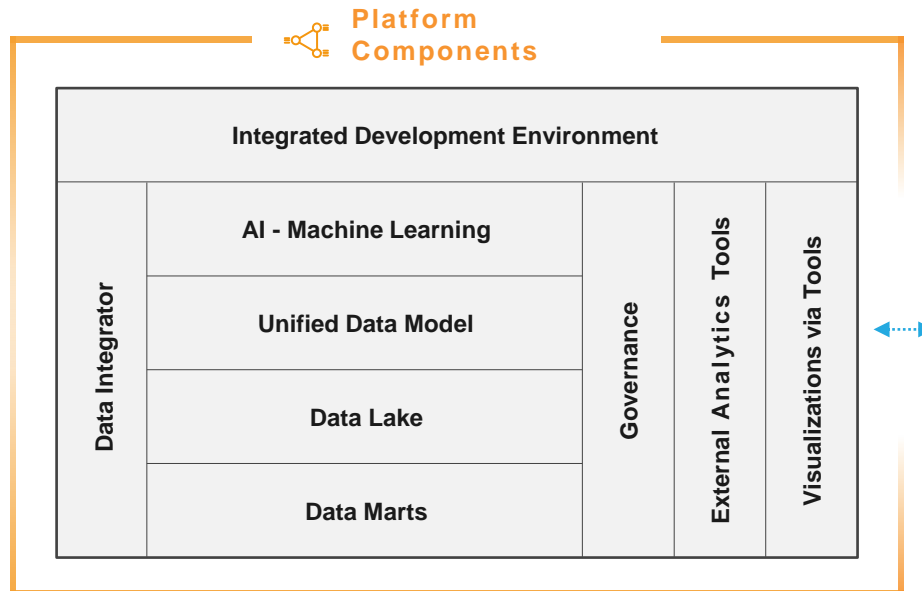
Build and share datasets

EDAP Conceptual Architecture & Data Domains

30 data sources integrated with EDAP empowering analytics across 10+ business units

| Customer Interval Timeseries |
| Customer Assets & Facilities |
| Customer Information |
| Customer Service Data |
| Customer Demographics |
| Meter Install Work Orders |
| Distribution System Connectivity Model |
| Distribution System Interval Timeseries |
| Grid Modernization |
| Public Data |

Platform Components

Integrated Development Environment

| Data Integrator | AI - Machine Learning | Governance | External Analytics Tools | Visualizations via Tools |
	Unified Data Model			
	Data Lake			
	Data Marts			

Technologies

- C3.ai
- Azure (Databricks, Data Lake, SQL)

Business Units

- Rate Engineering
- Energy Management
- Energy Efficiency
- Demand Management
- AMI Operations / Communications
- AMI Deployment
- Revenue Protection
- Customer Operations
- AMI Rate Pilots
- Usage Analytics
- Distribution Engineering
- Regional Engineering

Open Neutrals on the Secondary System

Normal Condition

Defective, Oxidized or Open Neutral
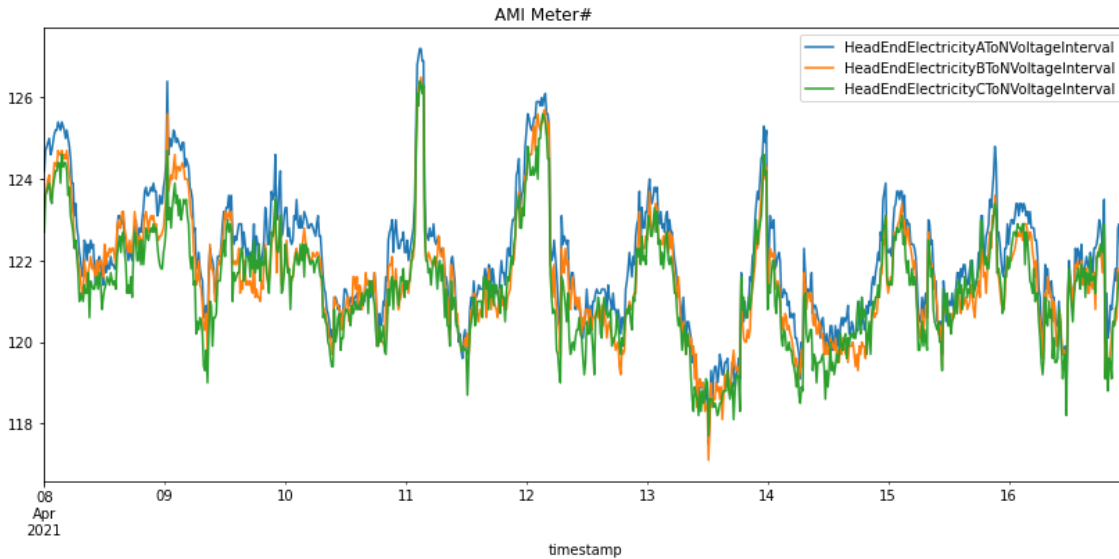
- Controls and regulates voltages
- Completes the circuit back to the source
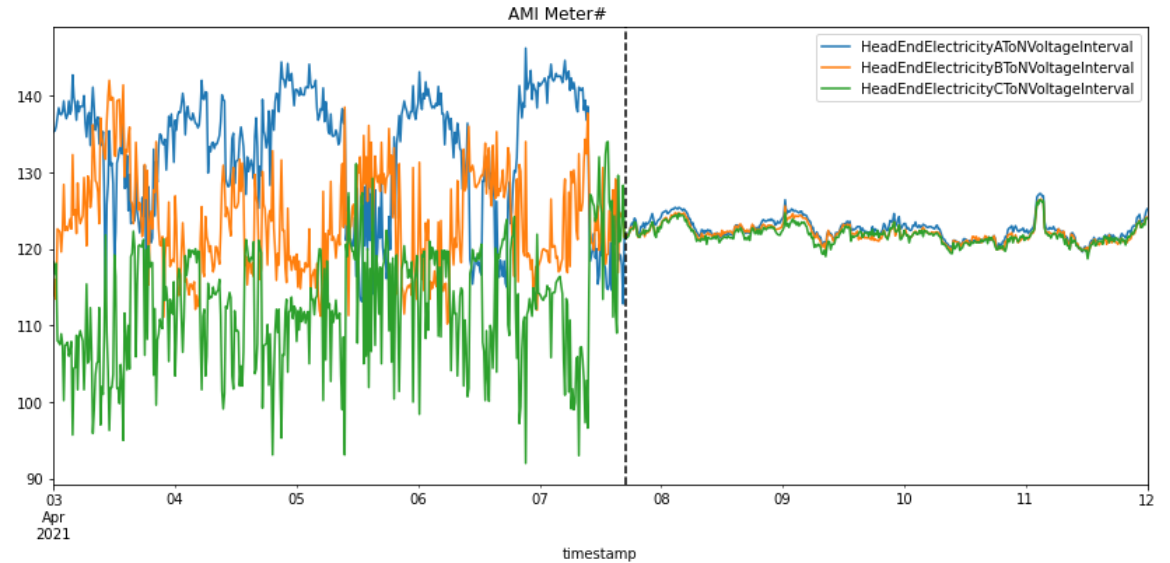
Customer Experience/Reliability
- Dim/No lights
- Home appliances damaged due to voltage issue

Safety
- Potential for electric shock

Previous process to identify open neutrals relied on field identification or customer complaint

Open Neutrals Identification High-Level Timeline

Run ad-hoc analysis over certain meter populations

Engineers begin triaging open leads; feedback given of high false-positive rates with no ability to prioritize

Dec 2020

Nov 2021

Mar 2022

Prototype → Productionize → Operate → Model Release 1 → Operate → Model Release 2 → Operate

2019

Jan-Oct 2021

Nov 2021-Feb 2022

Mar 2022-Sep 2022

Application deployed with rule-based algorithm

Con Ed team implements ML based predictions for 3-phase meters
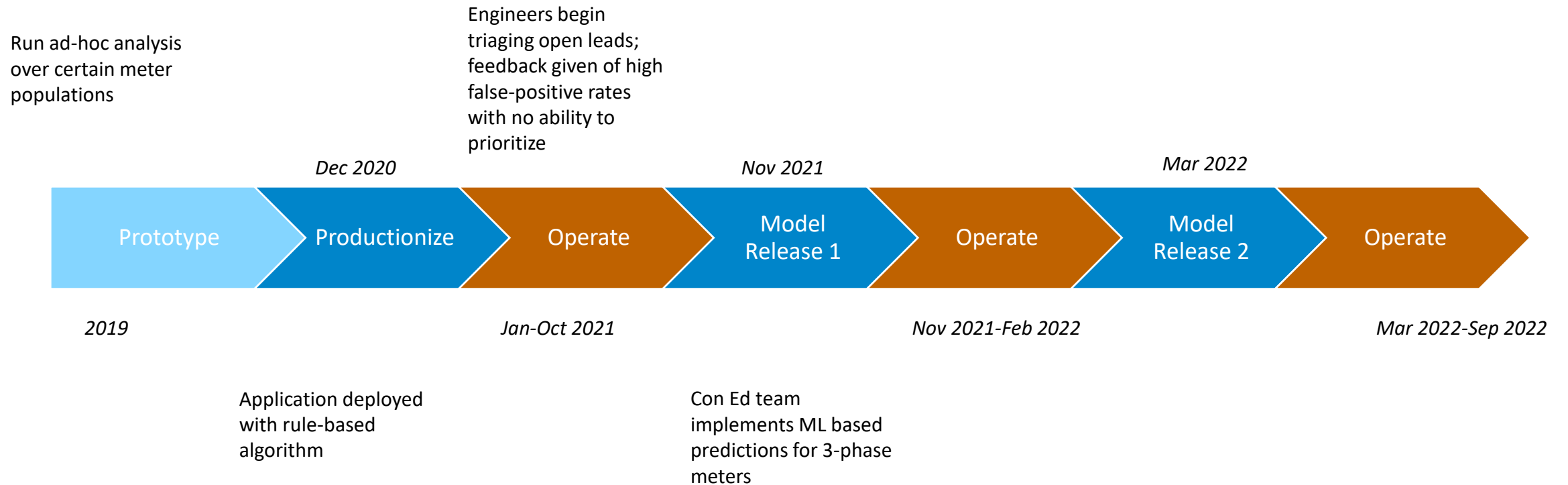
Open Neutrals Identification with AMI – Ideation

- Engineering team worked with the AMI Implementation group to understand available AMI data
  - Meter Diagnostic Events – Low/High Voltage, Phase-Shift, Last Gasp, Sags and Swells
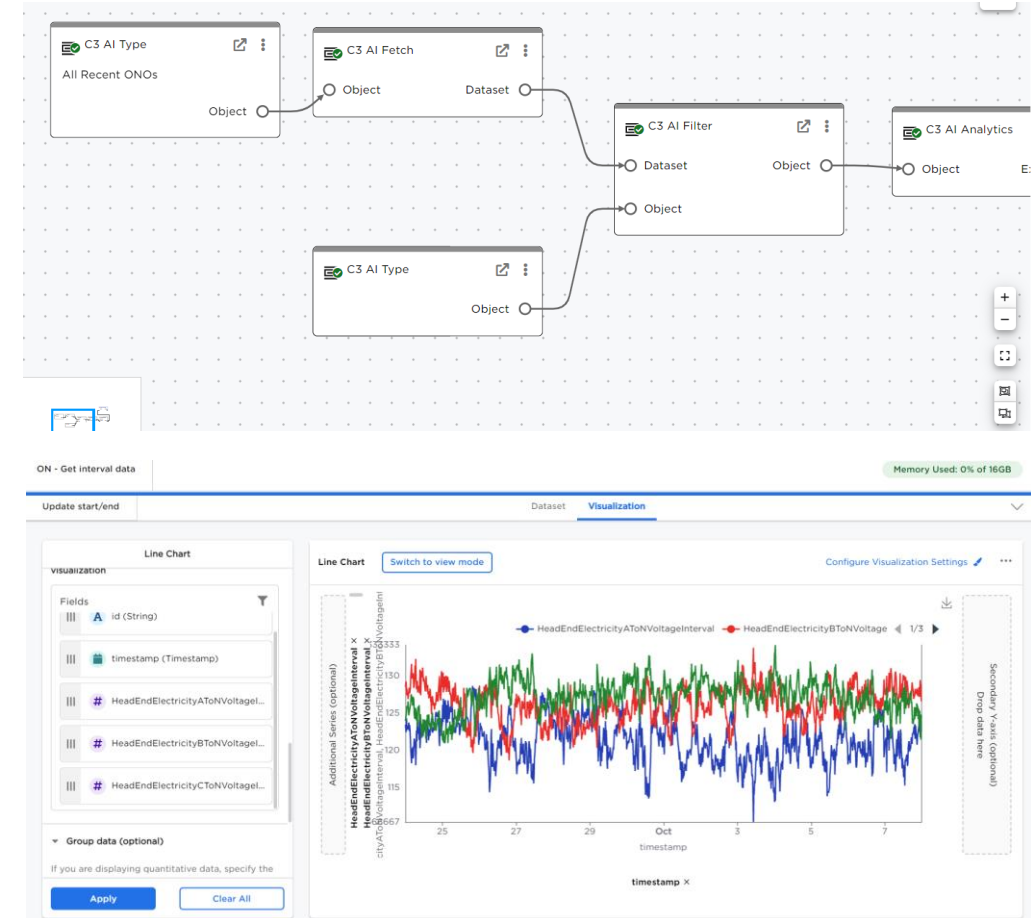 - Meter Information – Type, Form, Address, Service feed-point, etc.
  - **Times-series Data – Difficult to access and analyze**
- Enterprise Data & Analytics team assists ideation process by providing ad-hoc analysis
- Pro: **Allowed for initial exploration of the idea and to understand feasibility**
- Con: **Analysis not scalable**

Open Neutral Application v1

- Open Neutral application developed in EDAP as part of Grid Modernization Analytics project

- Application utilizes business logic to identify potential open neutral conditions on 3-phase meters for engineers to review
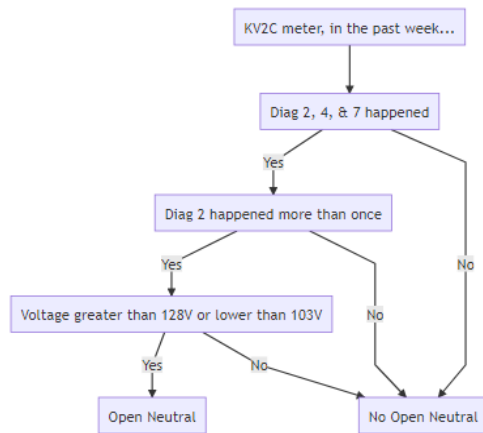
- Pro: **Surfaced large amounts of potential conditions for investigations. Easier UI/UX for investigation and tracking. Kicked-off field pilot.**

- Con: **high rates of false-positives and lack of prioritization**

Transition to Machine Learning Based Approach

Meter Data

Application Logic

Open Neutral Leads

(too many unprioritized leads)

Legacy ineffective rules-based approach

Meter 1	Meter 6
Meter 2	Meter 7
Meter 3	Meter 8
Meter 4	Meter 9
Meter 5	Meter 10

Rules-based Logic

1. **Specific meter diagnostic events occurred, and**
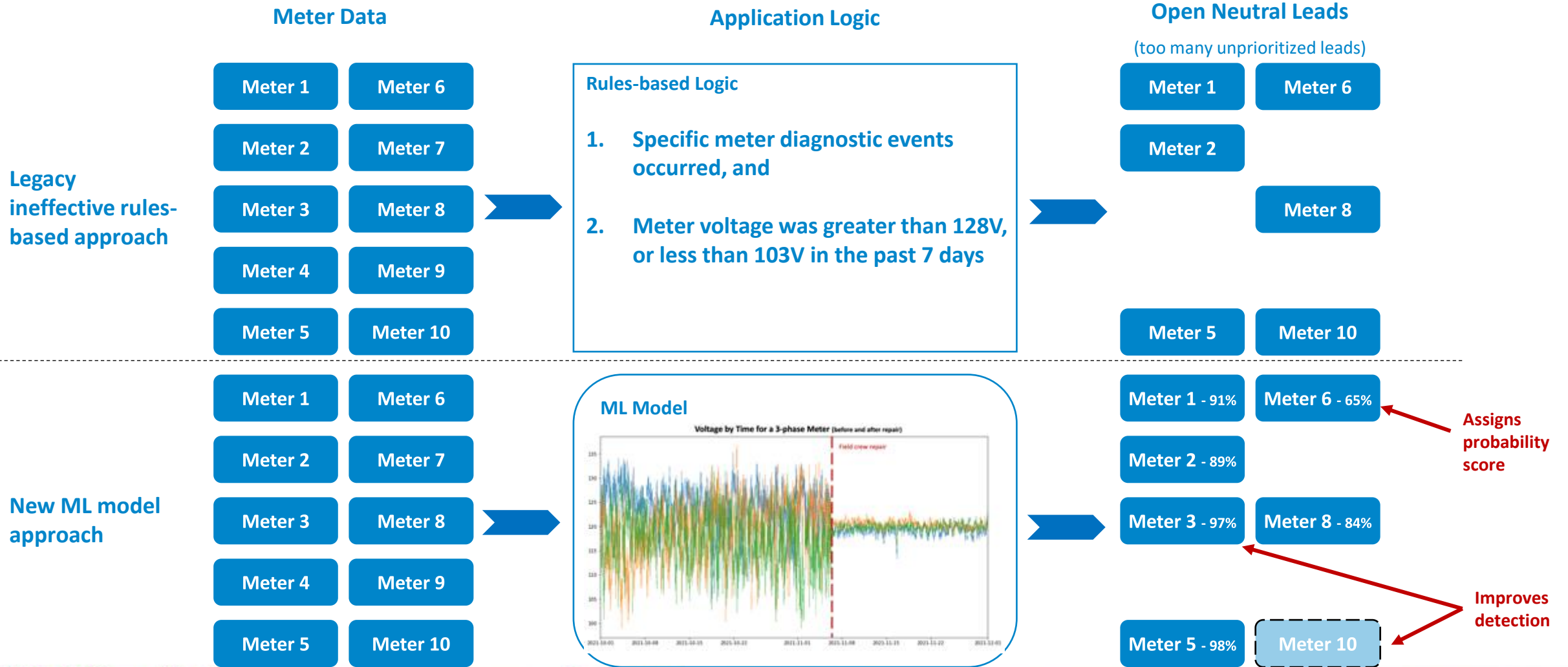2. **Meter voltage was greater than 128V, or less than 103V in the past 7 days**

Meter 1	Meter 6
Meter 2	
	Meter 8
Meter 5	Meter 10

New ML model approach

Meter 1	Meter 6
Meter 2	Meter 7
Meter 3	Meter 8
Meter 4	Meter 9
Meter 5	Meter 10

ML Model

Voltage by Time for a 3-phase Meter (before and after repair)

Meter 1 - 91%	Meter 6 - 65%
Meter 2 - 89%	
Meter 3 - 97%	Meter 8 - 84%
Meter 5 - 98%	Meter 10

Assigns probability score

Improves detection

Transition to Machine Learning Based Approach

- Moving from deterministic to non-deterministic and probabilistic can be jarring

- Business buy-in and stakeholder acceptance of approach is critical

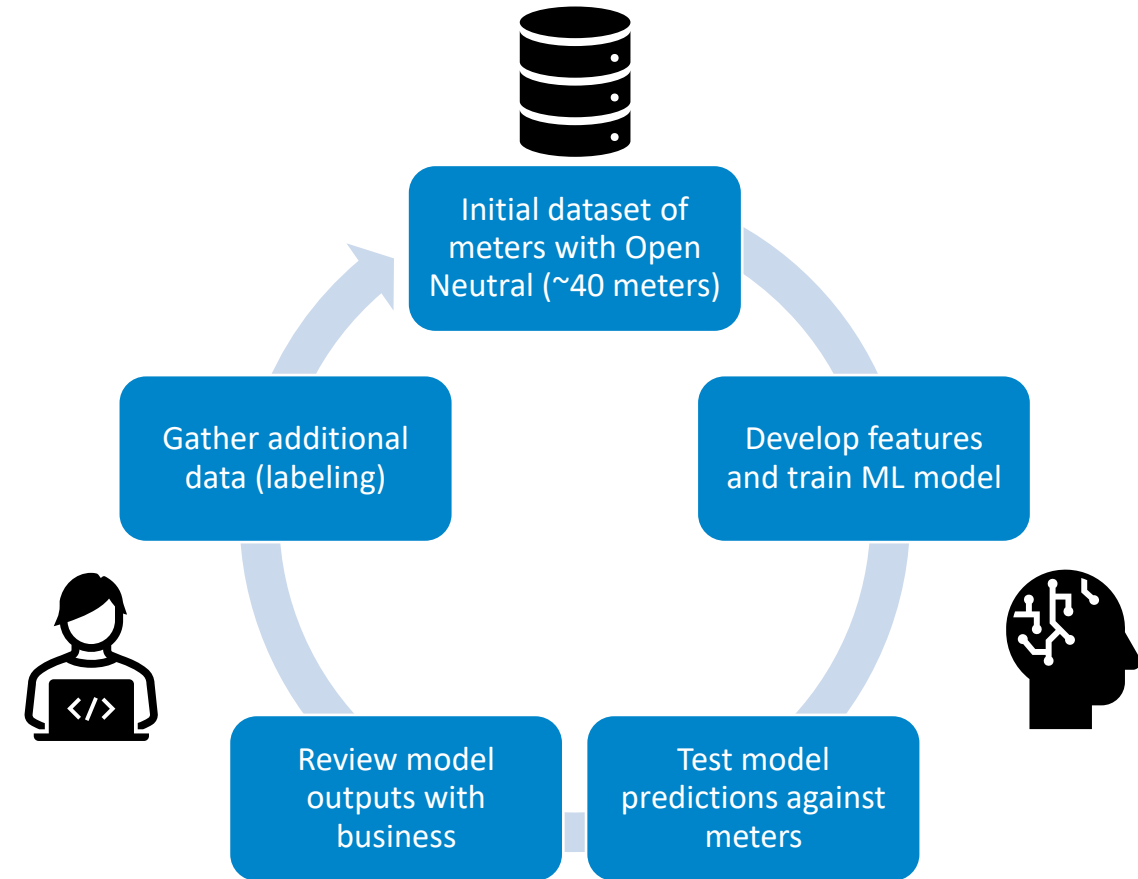 - Spending time for knowledge transfer and "unboxing" of the black-box

  - Setting realistic expectations given current "AI revolution"

  - Holding Machine Learning 101 sessions

- Taking the supervised learning approach requires a large enough **labeled** dataset that contains the target variable
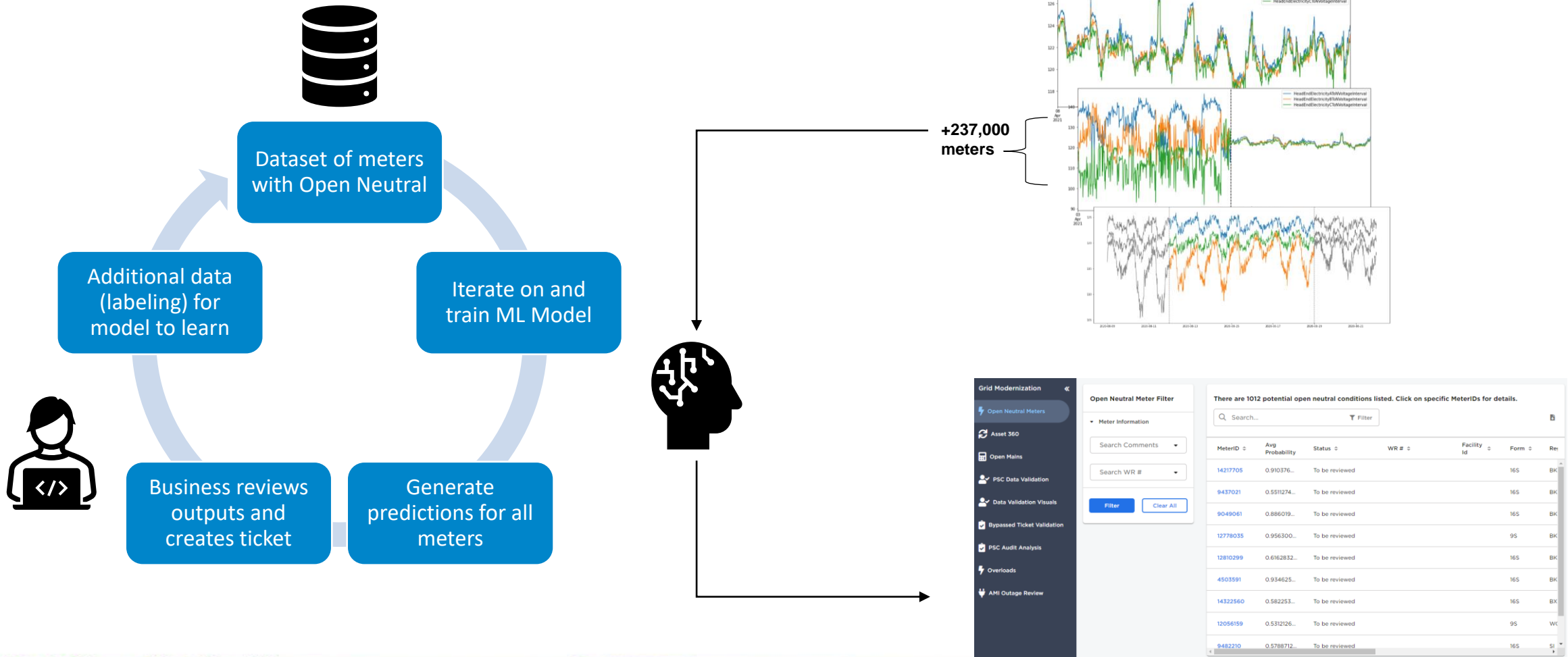
  - Business support is essential

The ML Challenge: Iterating on Few Examples

- The supervised learning approach requires a dataset that contains the target variable the model is trying to predict

- Very few examples to begin with:
  - We have billions of **unlabeled** datapoints
  - Handful of meters with labels
    - Generated from field pilot from application v1

- Took an iterative and collaborative approach to gather additional labels for the model to learn

Initial dataset of meters with Open Neutral (~40 meters)

Develop features and train ML model

Test model predictions against meters

Review model outputs with business

Gather additional data (labeling)

Open Neutrals Application v2 – ML at Scale

+237,000 meters

Open Neutrals Identified

Neutral not connected in the meter pan

Neutral broken/corroded in the service box

Open Neutrals Identification High-Level Timeline

Run ad-hoc analysis over certain meter populations

Engineers begin triaging open leads; feedback given of high false-positive rates with no ability to prioritize

Engineers select high probability leads for field investigation; ~60% results in true positive

Engineers select high probability leads for field investigation; ~95% results in true positive

Dec 2020

Nov 2021

Mar 2022

| Prototype | Productionize | Operate | Model Release 1 | Operate | Model Release 2 | Operate |

2019

Jan-Oct 2021

Nov 2021-Feb 2022

Mar 2022-Sep 2022

Application deployed with rule-based algorithm

Con Ed team implements ML based predictions for 3-phase meters

Con Ed team releases model improvements along with additional metrics for leads evaluation

Future Roadmap

Application Future Roadmap

- Model enhancements
  - Continued enhancements for three-phase meter model
  - Model for two-phase meters – expand population of predictions to another 770K+ meters
  - Model for single-phase meters – expand population to encompass all 120/208 meters
- Integration of application to the work management system
  - Fully complete the cycle for labeling
- Automated ticket generation for high-confidence conditions

Questions?