Pacific Gas and Electric Company

EPIC Final Report

Program	***Electric Program Investment Charge (EPIC)***
Project	***EPIC 3.20 Data Analytics for Predictive Maintenance***
Department	*Electric Operations - Asset Management*
Executive Sponsor Project Sponsor Project Technical Lead(s) Business Lead	Janisse Quinones Craig Kurtz Devon Yates, Maryam Variani Eric Schoenman
Contact Info	EPIC_Info@pge.com
Date	March 23, 2023
Version Type	Final

Table of Contents

Acknowledgements

Pacific Gas and Electric Company's EPIC program would like to recognize that this project would not have been a success without valuable contributions from various teams across the Utility. Members of Power Quality, Smart Meter Operations, Electric Operations, Standards and Asset Management, Asset Strategy, and the Asset Health and Performance Center provided important engineering and business insights. Members and alumni of the Data Analytics and Governance department created a culture of success and an environment that fostered collaborative development and mutual accountability. PG&E's Strategic Data Science team pushed the boundaries of what a utility can do and enabled others to do the same.

In addition to the above teams, there were many individuals critical to the creation, structure, and successful execution of this historic project. To all of you - Thank You.

Eric Schoenman

Principal Operations System Engineer

EPIC 3.20 Business Lead

Devon Yates

Principal Data Scientist

EPIC 3.20 Project Part 1 Technical Lead

John-Peter Dolphin

Director, Strategic Data Science Team

Maryam Variani

Expert Data Scientist

EPIC 3.20 Project Part 2 Technical Lead

Project Team and Key Contributors:
JP Dolphin, Devon Yates, Eric Schoenman, Maryam Hassani Variani, Sabrin Mohamed, Will McFaul, Aayushi Gupta, Rebecca Harcourt, Sandeep Kataria, Ana Maria Nungo, Aaron Lyons, Charles Scoma, Megan Kristovich, Barry Wong, Shane Buck, Michael Thibault, Jacinto Chen, Brinda Thomas, Greg Gillette, Fedor Petrenko, Dave O'Connor.

Abstract

Utility assets experience wear and tear, and eventually break down. Utilities have taken different strategies for asset management which directly affect service reliability, affordability, and the company's overall risk. The most passive strategy is to wait for the equipment failure to address the situation, or "run to failure". An already-implemented improvement on this strategy is scheduled maintenance using heuristics regarding expected useful life and level of utilization. EPIC 3.20 explored and implemented a predictive model to detect the signs of near-failure equipment.

The project's hypothesis for Part 1 was that an analytical model can be used in conjunction with existing PG&E data sets to predict distribution equipment failures. The results yielded that a machine learning algorithm can predict incipient failures on distribution transformers by using data sets (such as Advanced Metering Infrastructure [AMI] i.e., smart meter, asset location, and weather data) generally available to utilities through their ordinary course of operation. Furthermore, distribution assets have been proactively replaced based on the results of the model, reducing wildfire risk, and contributing to a more reliable and affordable service for customers.

For Part 2 of the project, the team hypothesized that precursor meter events could be used to predict sustained outages. An example of a precursor is vegetation contacting a line on windy days, animal contact, or a bad switch causing an event in the meter. In this minimum viable product (MVP) stage of Part 2, the model has developed early insights into precursor events. Continued Exploratory Data Analysis (EDA), model refinement, and Subject Matter Experts (SME) engineering reviews are planned to achieve a production-ready model.

Executive Summary

This final report summarizes the project objectives, technical process, initial results, and key learnings from *EPIC 3.20 – Predictive Maintenance*, as reported on in the EPIC Annual Report. It supports the EPIC program's goal of enabling the timely sharing of lessons learned across the entire energy industry and beyond the program administrator for a specific project.

Context

A utility's approach to asset management has an impact on reliability of service, system affordability, and the company's risk profile. "Emergency outage restoration" has been the primary approach employed by utility transmission and distribution operations for decades. This approach, where utilities wait for equipment failures and respond as quickly as possible once those failures occur, aims to minimize near-term costs passed on to customers by ensuring assets are utilized through their entire life.

The electric distribution sector of the utility industry has also adopted "scheduled maintenance" and "condition-based maintenance" approaches, both of which further reduce risk when compared to emergency outage restoration. This project aims to go a step beyond condition-based maintenance, implementing "predictive maintenance," which leverages sensor data and advanced analytics to identify signs of imminent failure. Predictive maintenance is most akin to a check engine light, which identifies when something in your vehicle is outside of normal operating conditions but has not yet led to your car being inoperable. This approach optimizes all three elements: reliability, affordability, and risk reduction. It identifies potential failures and outages before they occur, allowing for the re-routing of power or in some cases hot asset replacement (reliability), ensures assets are utilized through their entire useful life (affordability), and stops failures from occurring altogether - including those that could lead to wildfire ignitions or other safety incidents (risk).

Key Objectives

The core objective of EPIC 3.20 is to determine if machine learning models can be developed using existing utility data sets (i.e. data through SmartMeters™, asset location, and weather data) to predict electric distribution equipment failures and outages so that corrective action can be taken before either occurs.

While initially launched prior to the establishment of PG&E's Wildfire Mitigation Plans (WMP), the objectives of EPIC 3.20 directly align with the objectives of the WMP, specifically, "reduc[ing] the risk of catastrophic wildfires in Northern California." The accomplishments highlighted below directly support PG&E's ability to meet this objective in an affordable and data-driven manner.

Additionally, as discussed in more detail in PG&E's Wildfire Mitigation Plan, distribution transformers were prioritized as a key asset to target for predictive maintenance models because they are one of the top sources of PG&E equipment-caused wildfire ignitions.

Key Accomplishments

The following summarizes the key accomplishments of the project:

- A production-ready predictive maintenance machine learning model was developed, primarily using AMI voltage information, to predict failure of distribution transformers

- From April 2021 through February 2022, Engineering has reviewed over 270 model predictions, from which 64% were confirmed to be relevant transformer anomalies ("success") and were flagged for field investigation. An additional 26% were confirmed to be other issues in the distribution system.
- On multiple occasions near failing distribution transformers and meters have been proactively replaced based on the model's recommendations, in doing so reducing wildfire risk and improving reliability for customers.
- A MVP of a predictive maintenance machine learning model was developed, using meter event data, historical outage count, and weather data to predict outages caused by vegetation contact, equipment failure, animal contact and some unknown reasons.
- In April 2022, internal engineering teams established a process to review the top 20 ranked monthly outage predictions and establish a feedback loop back into the model to further optimize future predictions.

Key Learnings and Recommendations

The following are the lessons learned and recommendations from the project:

- The development of predictive maintenance algorithms is not only possible, but also has significant potential to transform asset management – improving reliability, aiding affordability, and reducing risk. However, the ability of a utility or any third party to develop such algorithms is dependent on the utility's underlying data management maturity and capabilities. This includes the entire data lifecycle including data acquisition, security, storage, organization, integration, governance, quality, and metadata.
- The project team highly encourages the continued utilization of open-source programming languages such as Python. These open-source languages provide increased functionality not capable in legacy tools like Excel or SAS without sacrificing transparency. The utilization of cloud computing resources can dramatically decrease compute time and allows for on-demand data insights when required.
- EPIC 3.20 has demonstrated the need to pair improvements in technology (data science included) with business processes, and people development. People, process, and technology can be thought of as a three-legged stool to improve business outcomes. Each leg is required for the benefits of the others to be fully realized. The project team encourages all those who work to develop data products to take a comprehensive approach to how said products are imagined, designed, developed, and implemented.
- Investments in advanced analytics, data science, and artificial intelligence will continue to be hamstrung if appropriate efforts are not made to treat data as an asset. There are several improvements that can be made to existing outage records. During Part 2 of the project, the team discovered numerous instances of inaccurate or missing data.

Conclusion

PG&E will continue to find ways to meet the challenges presented by operating the country's most connected smart grid in an affordable manner. Proving that the successful development of predictive maintenance algorithms is possible using AMI data is a key part of this journey. Not only do these findings represent an opportunity for a better PG&E, but given the propagation of AMI networks, represent an opportunity for safer, more reliable, and more affordable utility asset management practices across the country**.**

Part 1 of EPIC 3.20 has grown from an early-stage data science project to a technology demonstration that will ultimately result in an operational data product. The path to production will need to include continuous improvement and maintenance of the model and user interface, and the integration of a new business processes to ensure that predictions are acted on in a timely manner. The findings, learnings, and challenges from Part 2 of the project have also provided valuable insights for PG&E.

This project does not mark the conclusion of improvements to the Asset Management processes, or the end of utilizing industry leading data science to enable safer grid operations. PG&E looks forward to operationalizing this prototype and building additional capabilities in the upcoming years.

1 Introduction

This report documents the *EPIC 3.20 – Predictive Maintenance* project achievements, and key highlights and learnings. These findings provide industry-wide value and identify future opportunities for PG&E to leverage the results of this project. The CPUC passed two decisions that established the basis for this demonstration project. The CPUC initially issued D. 11-12-035, *Decision Establishing Interim Research, Development and Demonstrations and Renewables Program Funding Level*[1], which established the EPIC program on December 15, 2011. Subsequently, on May 24, 2012, the CPUC issued D. 12-05-037, *Phase 2 Decision Establishing Purposes and Governance for Electric Program Investment Charge and Establishing Funding Collections for 2013-2020*[2], which authorized funding in the areas of applied research and development (R&D), technology demonstration and deployment (TD&D), and market facilitation. In this later decision, the CPUC defined TD&D as "the installation and operation of pre-commercial technologies or strategies at a scale sufficiently large and in conditions sufficiently reflective of anticipated actual operating environments to enable appraisal of the operational and performance characteristics and the financial risks associated with a given technology."[3]

These decisions also required the EPIC Program Administrators[4] to submit Triennial Investment Plans to cover three-year funding cycles for 2012-2014, 2015-2017, and 2018-2020. On November 1, 2012, in A.12-11-003, PG&E filed its first triennial EPIC Application at the CPUC, requesting $49,328,000 including funding for 26 TD&D Projects. On November 14, 2013, in D.13-11-025, the CPUC approved PG&E's EPIC plan, including $49,328,000 for this program category. On May 1, 2014, PG&E filed its second triennial investment plan for the period of 2015-2017 in the EPIC 2 Application (Application (A.) 14-05-003). The CPUC approved this plan in D.15-04-020 on April 15, 2015, including $51,080,200 for 31 TD&D

1 CPUC Decision.11-12-035.

2 CPUC Decision.12-05-037.

3 CPUC Decision.12-05-037, p. 37.

4 PG&E, San Diego Gas & Electric Company (SDG&E), Southern California Edison Company (SCE), and the CEC

projects[5]. On April 28, 2017, PG&E filed its third triennial investment plan for the period of 2018 – 2020 in the EPIC 3 Application (Application (A. 17-04-028). The CPUC approved this plan in D.18-10-052 on October 25, 2018 including $55,600,000 for 43 TD&D projects[6].

Pursuant to PG&E's approved 2018-2020 triennial plan, PG&E initiated, planned, and implemented EPIC 3.20 - Data Analytics for Predictive Maintenance. Through the annual reporting process, PG&E kept CPUC staff and stakeholders informed on the progress of the project. This document is PG&E's final report on the EPIC 3.20 project.

2 Project Summary

2.1 Context

As outlined in PG&E's 2020 Wildfire Mitigation Plan, PG&E utilizes a variety of approaches to confirm the safety and operability of its equipment.[7] As part of PG&E's commitment to continuous improvement, EPIC 3.20 is one of several projects aiming to improve PG&E's approach to asset management – specifically the utility's approach to inspections and maintenance of distribution assets.

PG&E's current approach to inspections of distribution equipment is comprised primarily of on-site visual inspections. This includes highly trained inspectors who climb utility poles or otherwise inspect assets using aerial photography. These inspections are part of a proactive approach to asset management called "scheduled maintenance." Scheduled maintenance is implemented at many utilities with the goal of identifying asset degradation in advance of asset failures. In doing so, the risk associated with grid operations is reduced, when compared to traditional run to failure/emergency restoration approaches most common in the utility industry for decades. At PG&E, the cadence for this scheduled maintenance varies based on asset type. For example, the cadence for distribution transformers is approximately once every three to five years.

When compared to emergency restoration, scheduled maintenance does a better job of incorporating the risk of asset failure into maintenance and replacement schedules. However, it can be comparatively more costly to implement in the near term due to the cost of recurring inspections and the replacement of assets before the completion of their useful life. Continuing the car analogy, scheduled maintenance is most akin to inspecting and replacing your tires on a regular basis regardless of how many miles you drive or the condition of your vehicle.

An improvement over scheduled maintenance is "condition-based maintenance" in which heuristics regarding expected useful life and level of utilization are used to augment the duration of time between inspections. This is most akin to "1 year or 40,000 miles" when

---

[5] In the EPIC 2 Plan Application (A.14-05-003), PG&E originally proposed 30 projects. Per CPUC D.15-04-020 to include an assessment of the use and impact of EV energy flow capabilities, Project 2.03 was split into two projects, resulting in a total of 31 projects.

[6] [CPUC Decision.18-10-052, p. 115]

[7] [PG&E 2020 Wildfire Mitigation Plan]

recommending inspections for a car. The use of simple heuristics can be beneficial, but still leaves much to be desired across all three elements (reliability, affordability, risk). For instance, at PG&E, risk factors such as the location of an asset within a High Fire Threat District (HFTD) may be used as a heuristic. HFTDs have grown from accounting for 15% of PG&E's service territory to 50% of PG&E service territory.[7] As a result, what was previously considered high-risk, enhanced inspections are now the new normal, pushing Inspectors and PG&E to inspect more territory with increased frequency, while keeping rates affordable.

The development of predictive maintenance algorithms offers a potential release valve on these two competing forces. Like process improvements focused on the manufacturing industry, predictive maintenance could allow for "Just in Time" completion of maintenance. Such an approach would dramatically reduce the cost associated with constant inspections and could catch failure events that occur between scheduled inspections. The development of predictive maintenance using existing data sources would truly help maintain affordability, while reducing risk.

Following the CPUC's approval of the third EPIC triennial plan in 2018, EPIC 3.20 was initiated and the project team began focusing their efforts on reducing wildfire risk through predictive maintenance. As defined by the company's risk registry, wildfire is PG&E's largest risk. Due to climate change, California has experienced increased frequency and severity of wildfires, with the seven largest fires in the state's history all occurring in just the 3 years since the project was submitted for approval to the CPUC.[8]

In 2019, the year of EPIC 3.20 initiation, and therefore the year more detailed planning was completed for the project, there were more than 6,200 fires in California.[9] Of these, PG&E equipment was associated with 434, with asset failure accounting for 133. See Figure 1 for full details.

---

[8] CAL FIRE - Top 20 Largest California Wildfires

[9] CAL FIRE - Stats and Events

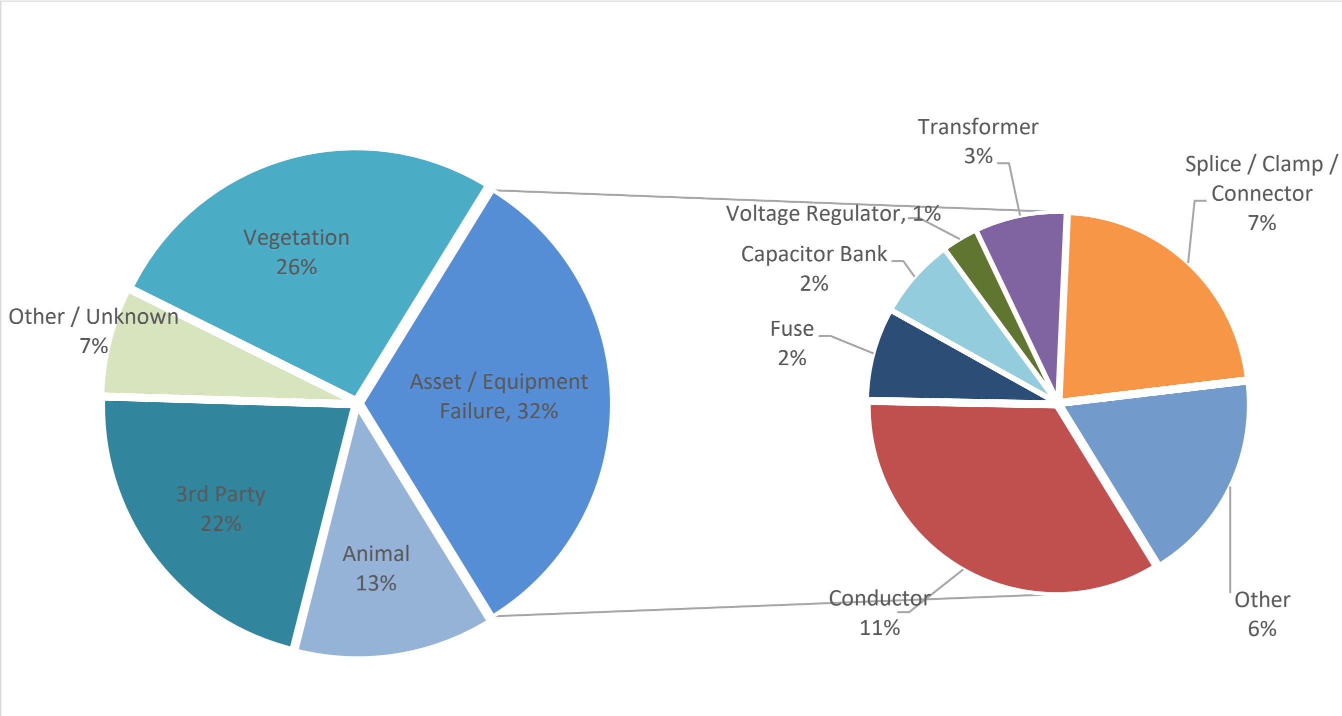

Figure 1: 2018 Suspected ignition initiating events and breakdown of asset-equipment failures

In Part 1 of the project, transformers, fuses, and capacitor banks were considered as the most addressable assets for EPIC 3.20's scope of existing data. This was expanded in Part 2 to include source side devices (circuit breakers, line reclosers, and fuses). The most recent figures can be found in the 2021 Wildfire mitigation plan[10]. Beyond the wildfire benefits, distribution transformers and fuses are also disproportionately associated with general reliability issues (a superset of all outages, regardless of if they are associated with ignitions), further emphasizing them as high priority targets for predictive maintenance.

Utility asset management is a rapidly evolving field, especially in California. PG&E hopes that by sharing this report, in a manner that is transparent, we will enable others to learn not only from our success, but also from our stumbles. By doing so we will improve reliability, enable affordability, and reduce risk at utilities across the industry.

2.2 Project Objectives

As described in PG&E's EPIC Triennial Plan (2018-2020)[11] , the project pursues the development of predictive maintenance algorithms for identifying conditions of impending asset failure in distribution equipment by using SmartMeter™ data and other existing utility data sources.

---

[10] PG&E 2021 Wildfire Mitigation Plan - Quarterly Data Report for Second Quarter 2021

[11] PG&E Electric Program Investment Charge Triennial Plan (2018-2020)

The goal is that algorithms detect and correlate data signatures associated with malfunctioning or failing system assets. These algorithms can be thought of as a "check engine light" for a vehicle identifying assets operating outside of normal parameters that have not yet failed. If successful, the performance of these algorithms would be evaluated against traditional condition-based maintenance systems. If appropriate, these data driven insights would then be incorporated into an asset management process, improving PG&E's reliability, affordability, and risk profile.

2.3 Scope

Using existing data available to PG&E, a model or algorithm(s) will be developed to detect incipient failures on the distribution system, in order to identify and mitigate the situation before it results in a failure, customer complaint, or ignition.

While multiple asset types were considered in the initial scope of this project, the primary focus was ultimately on distribution transformers, for Part 1, and source side devices (SSDs) for Part 2. Moreover, Part 2 of the project aims to use precursor momentary outages and events to identify an incipient sustained outage.

In the longer term, it is envisioned that outcomes from both parts of the project could result in an automatic dispatch of trouble man or field personnel through the work management system.

2.3.1 Time horizon of failure predictions

When developing a model for failure predictions, it is important to consider the time horizon. The relevant time ranges for distribution equipment can be thought of in the scope of years, months, days, or real time. Accordingly, the impact and probable actions could vary. For example, predicting failures that could take place in a year's time could help us in long term resource planning. The detailed impact of time scope on algorithmic approach is shown in Figure 2 below: Note that EPIC 3.20 focuses on real time and days to approximately a month before failure, and thus would be helpful in proactive/preventive maintenance.

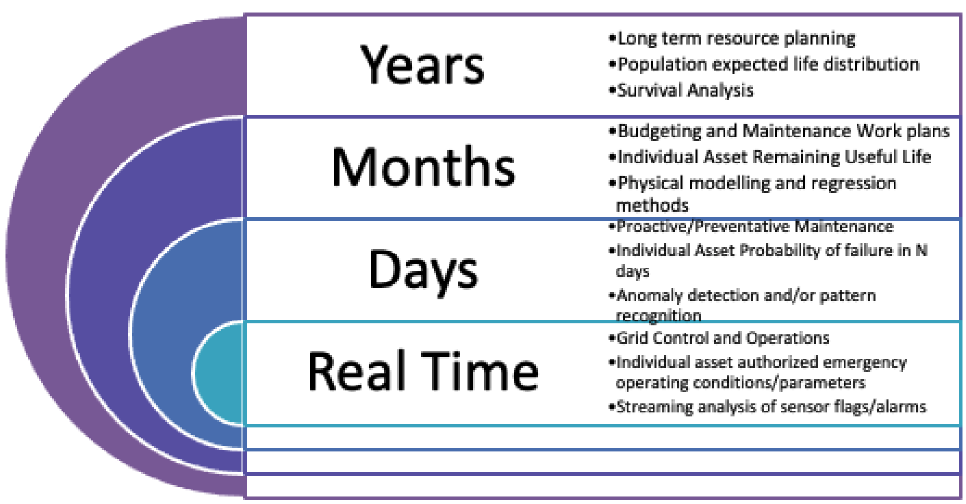

Figure 2: Impact of time scope on Algorithmic approaches

2.3.2 Out of Scope

Consistent with the Project Objectives and official application language, the following items were out of scope for EPIC 3.20:

- Acquisition of new data not already captured/available may occur but is not required.
- Model development for assets other than distribution transformers and source side devices may occur but is not required.
- Development of fully operational user interface, unless deemed necessary given algorithm results. A preliminary user interface may be developed for testing.
- Laboratory testing of distribution line equipment at ATS, unless the data discovery and failure mode exploration process determines that there would be significant value added from such testing.

2.4 Project Management

2.4.1 Agile-Scrum and Data Science Life Cycle

The project team implemented an Agile-Scrum project management approach, which allows the project team to adapt to changing business needs, avoid excessive sunk costs, and nimbly respond to project development challenges.

Paired with Agile-Scrum, the team used the Data Science Lifecycle as outlined below as a framework for their development. While these phases can be chronological, the project team regularly cycled back to a previous step based on new learnings or business priorities.

- Define the Problem
  - Problem Definition and Scoping
  - Literature Review
- Prepare Data
  - Data Discovery and Acquisition
  - Data Cleansing
  - Feature Engineering
  - Exploratory Data Analysis
- Build and Train Models
  - Model/Feature Development
  - Model/Feature Testing
  - Documentation and Code/Peer Review
- Deployment
  - Model Release
  - Model Deployment
  - Business Process Integration
  - Continuous Monitoring and Maintenance

3 Project Details: Part 1 – Distribution Transformers

3.1 Understanding the Problem

3.1.1 Distribution Transformer Descriptions

Transformers are electric equipment which are used to change the magnitude of electric voltage. They are composed of an iron core with two sets of copper windings wrapped around the two sides of the core. The current flowing in the windings on one side of the iron core induces a magnetic field in the core, which in turn induces a current and voltage in the secondary windings. The “turns ratio” of the number of turns in the two windings determines the ratio of the input and output voltage levels. A diagram of this process is shown in Figure 3:

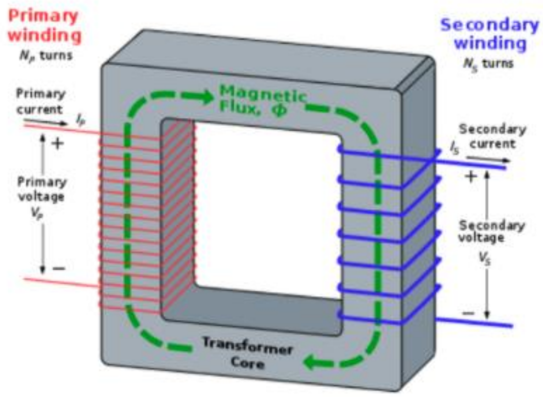

Figure 3: Transformer windings

Distribution transformers are the transformers which step down voltage from medium to low to supply service to customers, and for the purposes of this project excludes networked transformers and primary step-down transformers. Distribution transformers do not normally have any measurement equipment built in and it is difficult to observe their operating characteristics. In this project, AMI data has been used to estimate the voltage at the transformer low voltage side.

3.1.2 Distribution Transformer Failure Modes

As part of the fact-finding process, the common failure modes for distribution transformer have been tabulated and are described in Table 1. Note that these failure modes may in fact be intertwined. For example, a tank failure may result in low oil, which may cause overheating of the transformer, resulting in deterioration of the insulation, leading to winding shorts, which eventually makes the transformer more likely to fail due to voltage transients in a lightning storm.

Table 1: Distribution transformer failure modes

Failure Mode	Cause	Effect
Core failure	Direct Current (DC) magnetization or displacement of the core steel during transformer construction	Reduced efficiency which may manifest as a change in the impedance characteristics of the transformer.

Solid insulation failure Also known as Winding Failure	Transformer movement or forces generated during short circuits. Faults in insulating material may occur due to generation of gas or hot spots created due to low oil or transformer overloading	No insulation between windings, resulting in intermittent or complete short circuits between the windings. Short circuit may be preceded by a fuse failure.
Tank failure	Oil leakage, corrosion, internal arcing	Tank rupture
Cooling oil failure	Bad oil circulation or poor heat transfer to secondary cooling circuit; Oil contamination.	Overheating, short circuit
Bushings failure	Age, moisture, physical damage	Short circuit
Physical damage	Vandalism, structure failures, foreign objects, or animals	Various effects
Fuse trip	Lightning, short circuit	All or partial disconnection of phase.
Loss of neutral	Vibration, temperature, poor workmanship, poor balancing, network overloading	Conductor melts and ultimately breaks off. Safety risk from high voltage at secondary.

3.1.3 Definition of a Failure

The source records for failures in EPIC 3.20 are work management notification records which are composed of a start and end date, along with a notification type, equipment type, failure type and, in some cases, a cause. To create a useful classification label for supervised learning methods, these date ranges must be converted into labels which indicate a failure. The concept used to perform this transformation is shown in Figure 4:

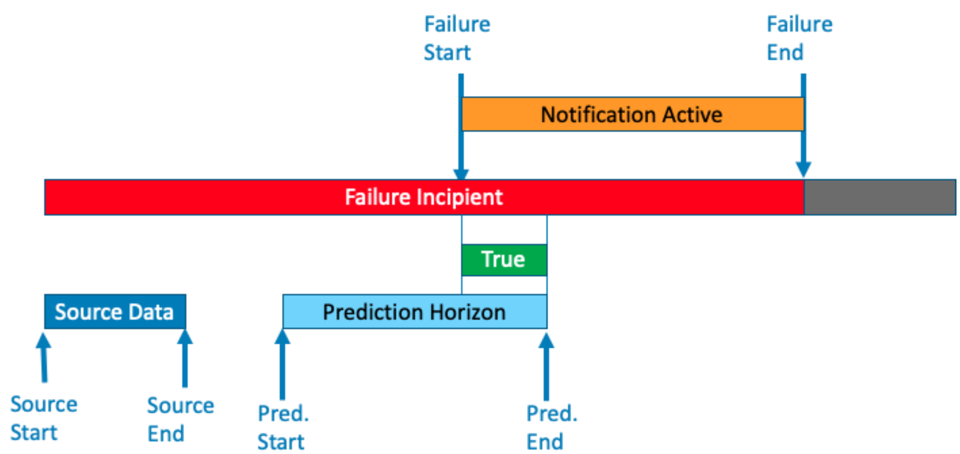

Figure 4: Failure definition

The source data bounded by the source start and source end represents the data used to develop the features used in the prediction. The prediction horizon is the length of time after the prediction start date, that a failure event would result in the prediction being labelled as a failure. The failure start date represents the date that the failure occurs or is first identified. The failure end date is the date at which the equipment is restored to full working order. In the case of unplanned transformer failures, the failure starts, and end dates are often the same, but in the case of voltage anomaly inspections or line regulator anomalies there may be a period between when the notification is initiated and completed. For voltage regulators, this can be a period where a trouble-man has identified an issue and bypassed the device as a temporary resolution until the equipment can be repaired and the notification completed.

The primary goal of this algorithm is to identify incipient failures using voltage data prior to the failure accelerating to a catastrophic or unplanned outage. Therefore, we want to observe failure characteristics in a time prior to the actual failure.

3.1.4 Characteristics of AMI Voltage Data

The AMI voltage data is a primary driver of the analytics in this project, so it is important to understand the characteristics and limitations of this data. A single line representation of a common configuration of a transformer is shown in Figure 5. For using voltage data, it is desirable to have measurements of the voltage on the high and low side of the transformer, $V_{T,H}$ and $V_{T,L}$. However, those measurements are not available in our system unless special diagnostic equipment is installed. Fortunately, PG&E does have the voltage readings at the Smart Meter™. In most cases, these will not match the voltage on the low side of the transformer, because some voltage drop is expected along the secondary and service conductor, as a function of current and impedance. In fact, if voltage, current, power factor and the impedance characteristics of the secondary and service drop conductors were known, then it would be possible to calculate the voltage at the low side of the transformer directly.

Though the voltage is captured at the Smart Meter™, in the existing AMI implementation, current and power factor are not available. Additionally, accurate secondary topology and therefore impedance are not known in most cases. At that point, there will be analytical methods which can be used to develop a reasonably accurate estimate of the voltage on the low side of the transformer, but until that point, simplifying assumptions must be made to estimate this value.

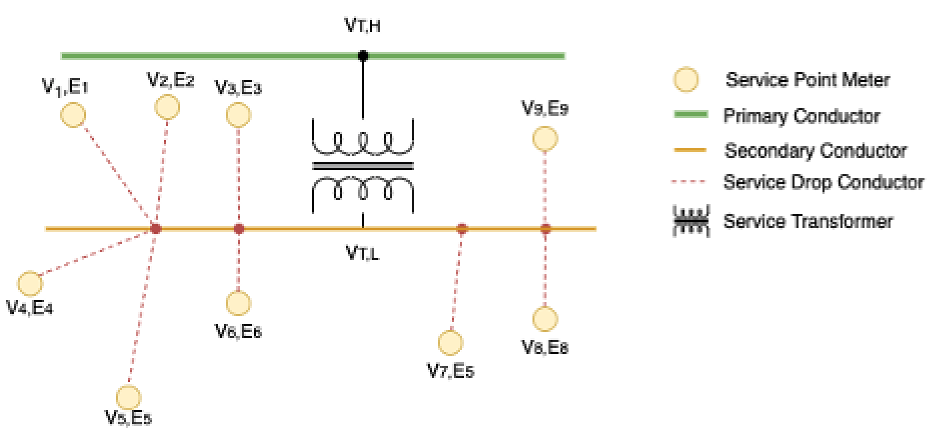

Figure 5: AMI data one-line diagram example configuration

A simplifying assumption which helped in comparing voltages from one meter to another is to normalize the voltage to a unit voltage to facilitate comparisons. This normalization is achieved by dividing the voltage read by the nominal voltage. Unfortunately, this nominal voltage is not provided by the measurement equipment or recorded on installation. Therefore, a process was developed to identify the nominal voltage of a given channel.

Finally, each AMI voltage measurement needs to be understood in the context of its circuit. When an anomaly is observed on an AMI voltage reading, there can be many causes. It could be a line regulator or capacitor operating or malfunctioning. It could be a result of extreme customer load behavior or weather conditions. It could also be a voltage event originating in the source voltage of the substation. When trying to isolate the root cause of the anomaly, a holistic view must be taken to isolate the signal of interest.

3.2 Analytics Platform

The primary data collection for EPIC 3.20 occurs on PG&E's customer data warehouse and GIS platforms. The data was integrated using a cloud platform customized for PG&E's requirements. This platform provided easy to use tools with detailed documentation and much better integration to better handle the data and create machine learning models using this data. This enhanced the efficiency with which the project progressed with the help of its spark distributed compute capabilities and other useful tools that made exploring, visualizing, and using the data a lot easier.

3.3 Benchmarking

The project team engaged in a fact-finding mission to identify learnings from prior work and benchmark against the state of the art. This search involved reaching out to other utilities, subject matter experts (SMEs) within and outside of PG&E and performing a literature review to identify relevant results and methods.

3.3.1 Industry Benchmarking

Interviews were held with ComEd, Southern California, and San Diego Gas and Electric to benefit from existing experience around predictive failure modeling, and to develop an ongoing dialog to foster collaboration.

ComEd has publicized their innovative approach in the article "Predicting Distribution Transformer Failures" [3]. Their work focused on several failure modes which could be identified through voltage anomalies. They identified that transformer windings failures could result in an abnormally high voltage as the turns ratio was modified by the windings failure. They also hypothesized that voltage could be used to observe a failure in the transformer core by observing a change in the voltage response of the transformer under load.

A key insight that the ComEd team discovered was that the transformer voltage cannot be looked at in isolation due to the complex dynamics of the distribution electrical network. Therefore, they isolated the anomaly by comparing transformers which were neighbors to each other, and on the same phase. For transformers A, B, and C, where B is between A and C, they created a feature which was the difference in voltage between transformer B and the average of A and B.

Given that each transformer may have multiple meters, a transformer voltage needed to be estimated using the meter voltages. They attempted to minimize the impact of the service drop by identifying the meters with the smallest difference between each other, with the assumption that these were the meters with the least voltage impact from load. One important input to this estimate was the phasing of the meter readings, as without this information phase imbalance can create a false signal.

The ComEd work reported impressive results. However, when testing the performance on data not seen in the training process, performance did not match that level. This is likely a result of the time series not being generalizable to different periods of time. The team at ComEd continue to develop promising models.

Southern California Edison and San Diego Gas and Electric also provided valuable insight into predictive failure models being developed. Details are not discussed here as results of those investigations have not been published.

3.3.2 Industry Vendors

The project team engaged with industry vendors in the space to identify if there were any existing available commercial solutions. Most vendors providing predictive maintenance for transformers were oriented towards substation power transformers. Power transformers tend to have more measurements than distribution transformers, as well as established preventative maintenance programs.

For distribution transformers, the commercial solutions for predictive maintenance were not well developed. This is likely due to the historically low appetite for this in the utility industry. Several vendors pointed out that the engineering standards, data structure, and data quality among different utilities made it difficult to make a generalizable product. Some vendors did have experience with distribution transformers and underground cables, but the project team recognized that using these products would result in an unsuccessful outcome without completing the hard work of developing good failure records for the assets. This effort was the fundamental problem that must be solved before the data could successfully be shared with outside consultants.

3.3.3 Internal SME Fact Finding

Fact-finding internal to PG&E focused around identifying the optimal project use case. Asset managers and operations staff were interviewed to understand each use case, business value, insights, and pitfalls. Internal SMEs were interviewed for each asset class under consideration. These potential use cases included secondary network devices, underground cables, transformers, and voltage regulators. This effort, as covered in the sections below, identified distribution transformers and voltage regulators as equipment that would benefit from predictive failure modeling due to the attractive blend of failure record availability, data availability, and business value. The distribution transformer use case had the added benefit of having an existing manually executed process to investigate voltage anomalies for potential failure.

3.3.3.1 Distribution Network Equipment

In urban areas, the PG&E distribution system is configured in a network configuration, as opposed to the radial configuration used in the bulk of the secondary systems. Secondary distribution network configurations are a redundant design where each component has multiple sources operating in parallel. Components on the network system tend to be much larger than those found on the radial circuits and tend to have much more measurement data available.

According to the SMEs, the best opportunity for predictive maintenance activities was on network protectors, devices used to automatically disconnect a distribution transformer in the case of reverse power flow or other anomalous behavior. In past failures, it had been possible to observe a change in the relationship of pressure and temperature to loading in these devices, and there was potential to automate this process. It was noted that the number of failures on these types of equipment were rare. Another component that could benefit from predictive maintenance are the power transformers, which tend to be large and well instrumented and annually inspected.

Underground cables in network systems also could benefit from predictive maintenance, as they can have a large impact, and historically failures occur on average about 10 times a year. However, cables have the challenge that there is very little measurement data on them. In network configured distribution systems, it is particularly challenging to estimate the loading for a particular cable due to the redundancy in the configuration.

3.3.3.2 Distribution Transformers

PG&E services distribution transformers according to a proactive maintenance and replacement strategy by, for example, identifying overloaded transformers. Data quality issues, however, have created problems for that process. Service points associated with the wrong transformer and inaccuracies in the transformer's recorded capacity cause the loading calculation to be inaccurate. Additionally, historical records of transformer loading are only stored for 3 years, leading to an inability to capture the full lifetime loading of the transformer.

Additional proactive mechanisms for identifying transformer failure exist within the power quality and voltage desk processes. These processes leverage information from customer complaints regarding power quality and monitoring of Rule 2 violations to trigger field reviews of distribution transformers.

The general process for replacement outside of the data-driven approaches above is a 1-3 year visual inspection cycle for transformers. A Standards engineer noted that many of the replacements were triggered due to external corrosion, which may or may not represent a functional deterioration of the transformer.

3.3.3.3 Voltage Regulators and Capacitors

Voltage regulators and capacitors are devices installed on the distribution system to mitigate voltage issues. They may have a schedule or control parameters associated with them, but these schedules are not available in a centralized database. In the past, these devices were maintained with a 10-year overhaul cycle, but now they are maintained with an annual inspection and are run to failure.

3.3.3.4 Underground Cable

The project team interviewed SMEs from the overhead and underground cable asset management groups and determined shortcomings in data recording that would hinder a machine learning approach. Cable segments are not tracked as assets in the equipment asset management database. While they are tracked in the geospatial database, there is no record of historical changes to the data making the task of associating asset and associated characteristics to a recorded failure difficult.

3.4 Literature Review

A literature review of predictive analytics uncovered three focus areas: predictive equipment maintenance, predictive failure detection, predictive stress case identification. Most of the predictive analytics models reviewed studied distribution transformers or substation power transformers, while one presented a generic approach with a use case in semiconductor manufacturing [2].

The literature review showed that projects developing machine learning (ML) methods typically evaluated multiple ML models and compared the results to historical data to identify which model performed best. ML models implemented included: Random Under Sampling with Boosting (RUSBoost) [1], random forest [1], Support Vector Machines (SVM) with Monte Carlo Cross Validation (MCCV) [2], feed-forward deep neural network (DNN) [3], DNN with digital signal processing [3], gradient boosting [3], and logistic regression [8]. Model outputs were compared against existing utility procedures and were found to perform with high precision [1,3]. While some models were executed periodically (e.g., weekly) [1], others were intended for real-time use [4].

Feature selection was a key part of model development and exposed the richness of data inputs to the ML models, as well as the development of tools to calculate features where recorded data did not exist. Features included equipment-related data (age, model, manufacturer, useful life, total operating cost), use-related data (loading, voltage, intensity of power spectrum of voltage), and environment-related data (region, subtype, weather (rain, humidity)) [1]. In [5] a tool was developed to predict hourly loading at any point between the customer and the substation. The ML models developed were trained over existing datasets, with the training sets showing a large variation in the size and time period. Proposed future work identified a need for abstraction from a time-period specific approach [3].

In addition to ML model research and development executed by electric utilities, asset analytics offerings by independent companies were reviewed [4]. These offerings leveraged real-time performance data and predictive algorithms to provide equipment maintenance insights, including the development of equipment maintenance dashboards.

Tying ML models back to physical equipment failures, an analysis of failed equipment [7] using an IEEE standard [9] served to identify distribution transformer failure modes and their frequency.

3.5 Data Collection

3.5.1 Data Sources

Developing a supervised model to predict asset failures requires finding sources for asset metadata, measurement data, and equipment failures. The following types of data were used:

- Smart Meter™ voltage data
- Distribution network topology information
- Asset characteristic data
- Equipment failure and outage data
- Historical Ambient Temperature data
- Transformer Historical Load data

3.5.2 Exploratory Data Analysis

Exploratory data analysis (EDA) conducted on the datasets identified in Section 3.5.1 is presented below.

3.5.2.1 Failure Data

Three types of failure data were used in developing the models. Each of these was filtered to identify failures tied to distribution transformers:

- Ignition failures: A list of ignition-related failures was obtained from an internal PG&E team that is focused on diving into each ignition and researching various data stream within PG&E to identify the true source and cause of the failure.
- Heatwave failures: These failures spanned the period of August 14th through 20th 2020 where the PG&E territory experienced a heatwave. PG&E's territory often experiences its highest temperatures in late June.
- Outages: PG&E records all outages in the system along with initiation and conclusion date, and cause. Unplanned outages which were associated with a transformer or transformer subcomponent and had a cause of Equipment failure or Environmental causes. Environmental causes were included, because frequently failures attributed to lightning have been identified to have incipient behavior prior to the failure, but the transients associated with lightning storms were the final cause of the failure.

A summary of the failure rates included in the curated failure list is shown in Figure 6. The month average rate of failure is 0.04%. There are frequently spikes in the summertime associated with heat waves, though those are not present in all years.

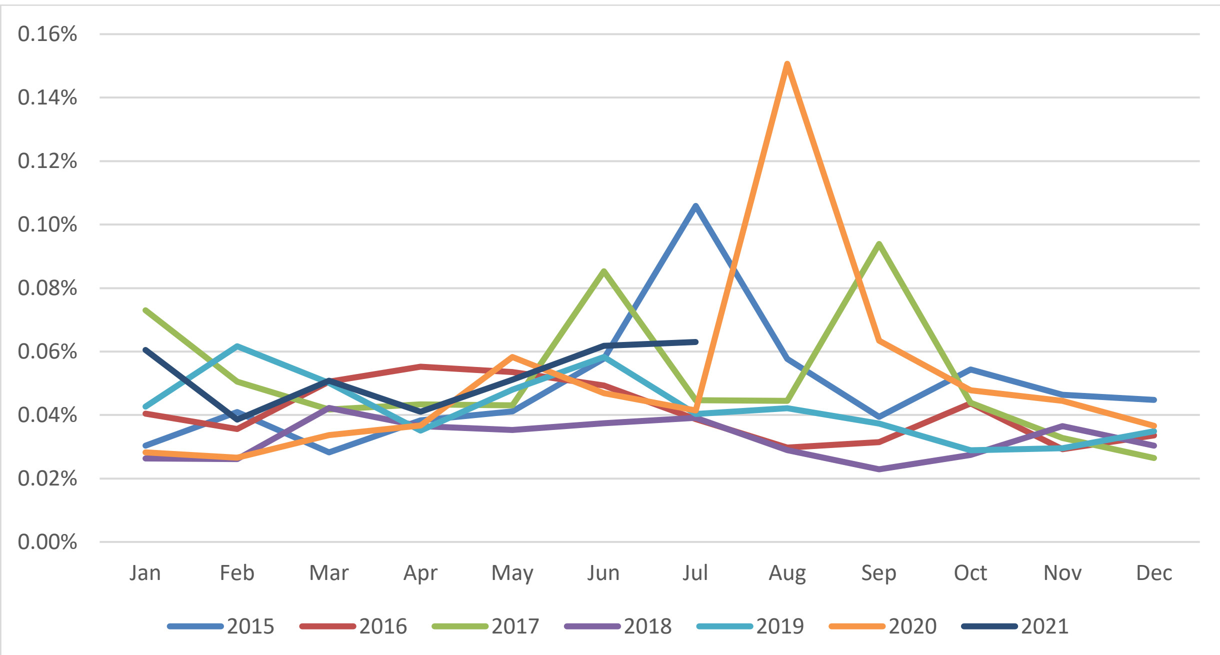

Figure 6: Failure Rate for selected failure events

3.5.2.2 Network Topology

The PG&E network includes electric power transmission and distribution systems. A substation at the head of a feeder steps down from a transmission-level voltage to a distribution-level voltage. Power travels along the length of a feeder, which resembles a tree with multiple branches reaching to individual service points. Upon arrival at the service point, distribution-level voltage is stepped down to the required service voltage by distribution transformers.

Electric meters installed at the service point monitor electric usage. The PG&E network includes legacy electric meters and SmartMeters™. The EPIC 3.20 project focuses on using SmartMeter™ data.

As shown in Figure 7, the PG&E network consists of approximately:

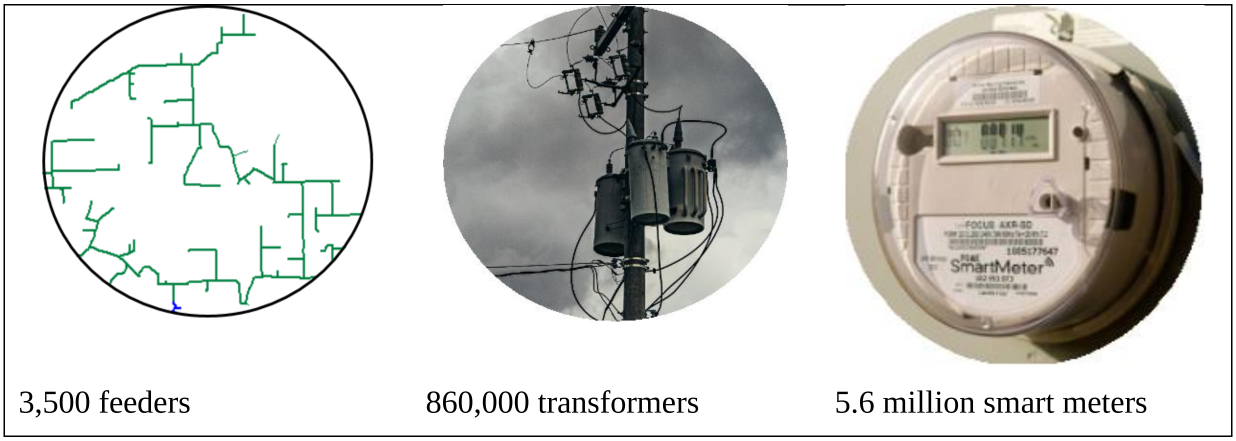

Figure 7: PG&E System characteristics

Transformers are “dotted” along the length of the feeder and its branches. All transformers except for the one closest to the substation will have an upstream transformer; all transformers aside from those at the end of a branch will have a downstream transformer. Of the 860,000 transformers:

- 98% had an identified upstream transformer
- 61% had an identified downstream transformer

Voltage regulators are installed at strategic locations along the feeder and serve to automatically modify downstream voltage in cases where it may begin to drop below desired levels. On average, there are approximately 100 transformers per regulator, with the minimum being 1 and the maximum being 1,500.

3.5.2.3 Asset Data

3.5.2.3.1 Transformer Installation

Distribution transformers may be installed overhead on poles, pad-mounted at street level, or underground. Figure 8 shows the distribution of transformer installation type across the PG&E network.

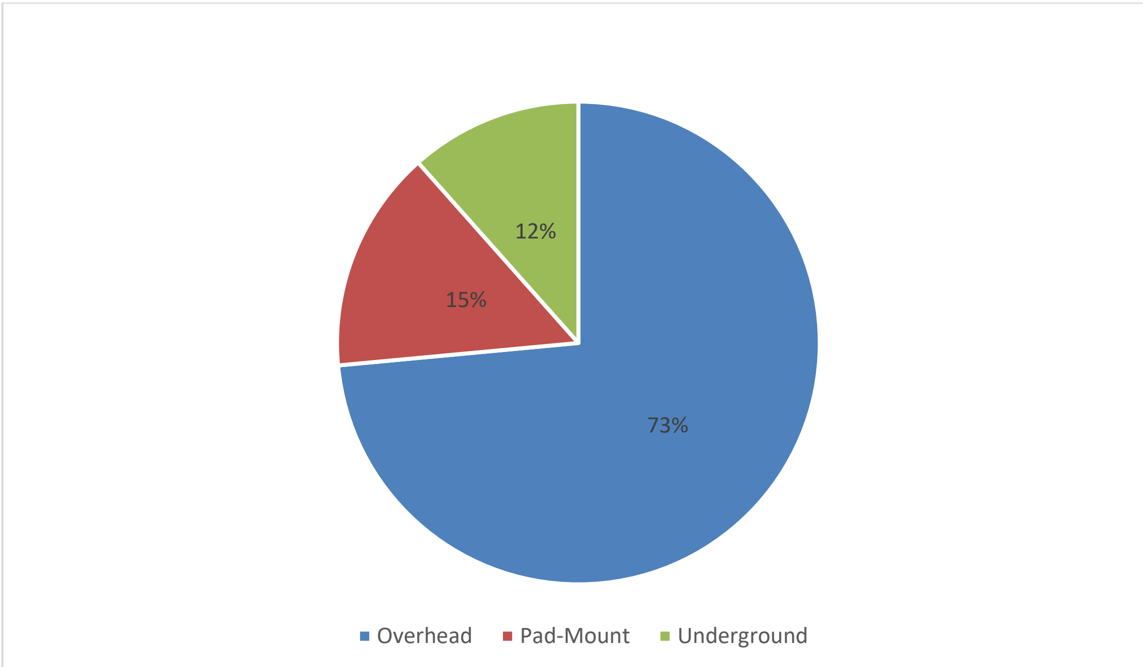

Figure 8: Transformer Installation Types

3.5.2.3.2 Transformer Rating

Of the 860,000 transformers, 96% had a rating of under 200kVA. Figure 9 shows the distribution of these transformers.

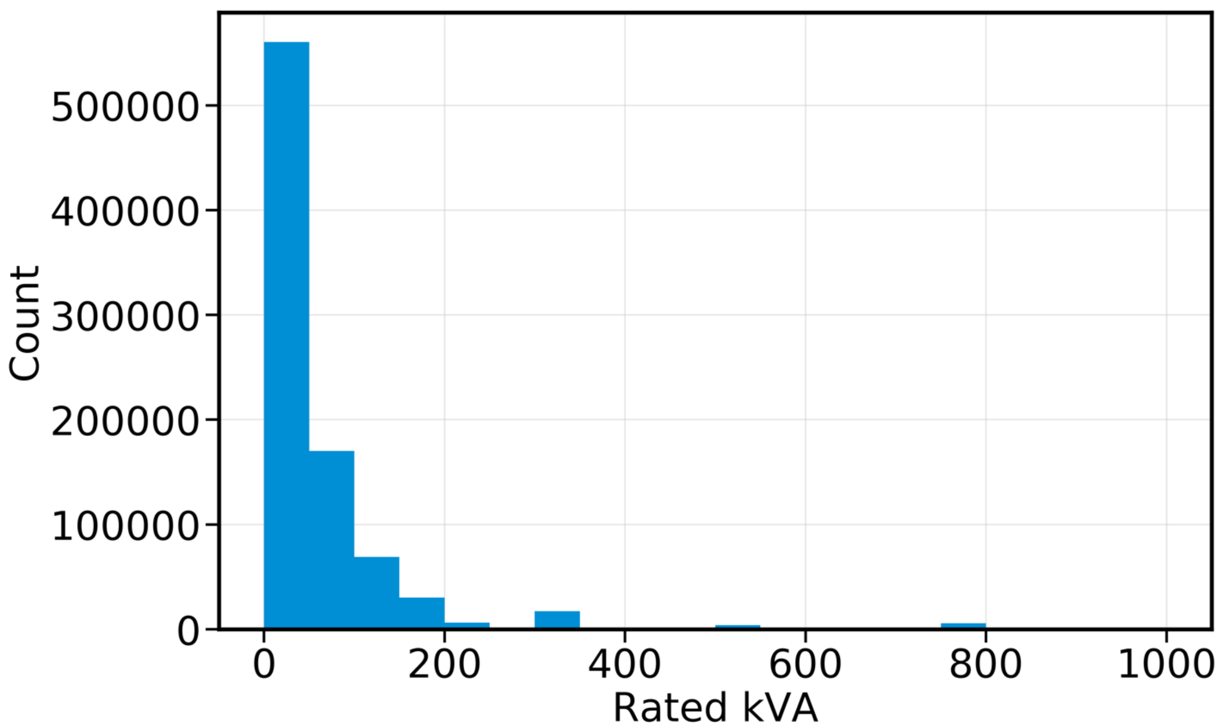

Figure 9: Histogram of transformer kVA rating

3.5.2.4 Voltage Data

3.5.2.4.1 Meter Form

The meter form determines what type of meter may be installed in a particular service. Figure 10 shows the distribution of meter forms across the PG&E network. Form 2s meters comprise about 80% of the PG&E network, with 12s- and 16s-meter forms having greater than 3% representation.

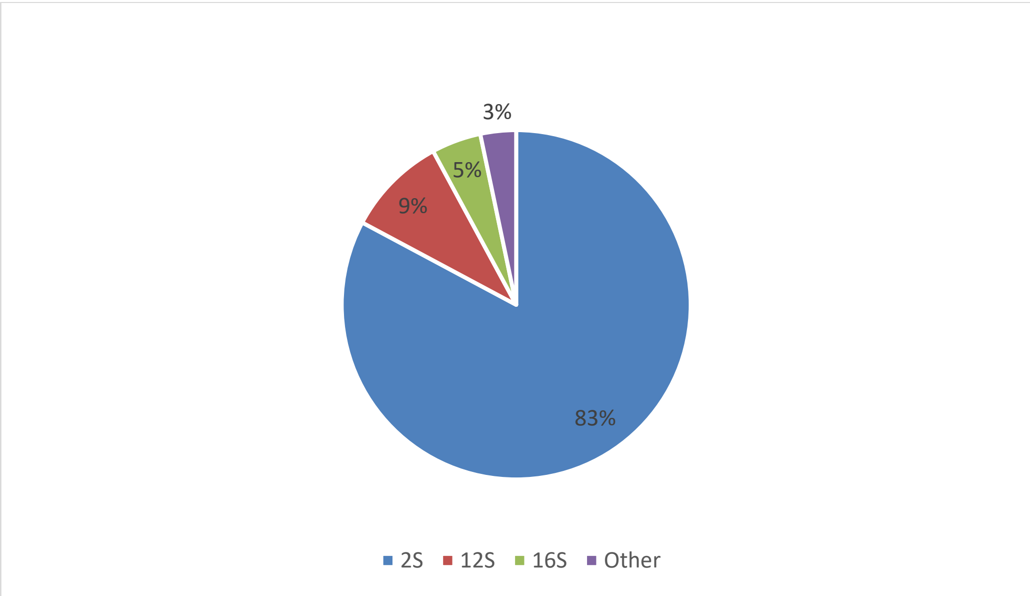

Figure 10: Meter forms across the PG&E network

3.5.2.4.2 Service Type

The service type involves the number of phases, number of wires, and the type of neutral present. The combination of these determines line-to-line, and line-to-neutral voltages and hence the service voltage. The ranges of service type include:

- Number of phases: single or three phase
- Number of wires: 2, 3, or 4
- Neutral present:
    - Wye systems have a neutral
    - Delta systems typically do not have a neutral

Figure 11 shows the distribution of service types across the PG&E network, as tabulated at the service point level. 92% of the services are 3-phase, 4-wire Wye. This type of service uses star connected phase windings with the fourth wire (or neutral wire) taken from the star point. This results in a line-to-neutral voltage of 120V, and a line-to-line voltage of 208V.

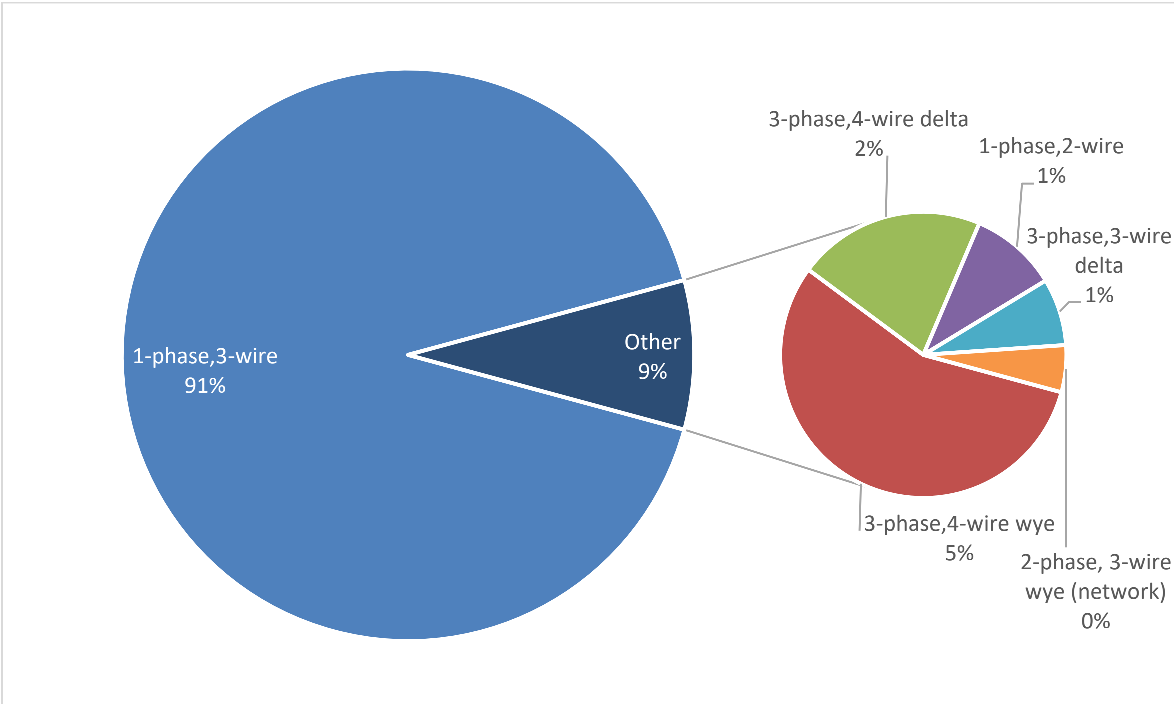

Figure 11: Service types across the PG&E network

3.5.2.5 Temperature Data

Ambient air temperature plays a significant role in the ability of a transformer to cool down and thus weather data was obtained and analyzed as a potential feature. The data consists of records collected every 15 minutes for years ranging from 2011 to 2020. We performed exploratory data analysis to find any potential quality issues and to understand the trends. As the data was plotted, non-physical outliers were identified and to overcome the issue an outlier filter was applied, as can be seen in Figure 12 and Figure 13. The meteorological tower data was further mapped to transformers to be made usable for the model. Due to the limitations of the dataset, ~0.5 % of the transformers could not be mapped to a near-by meteorological tower. As a part of feature engineering, a decision was made to use data for the years 2017 and 2018 to produce features.

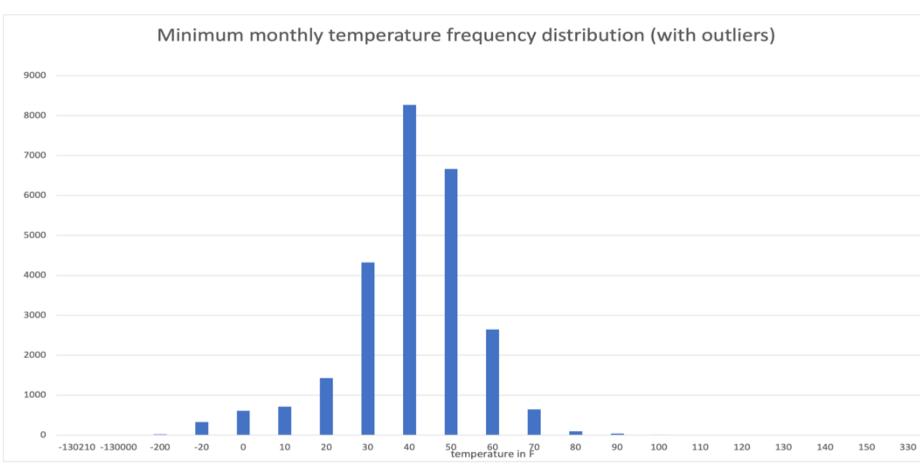

Figure 12: Minimum monthly temperature frequency distribution (with outliers)

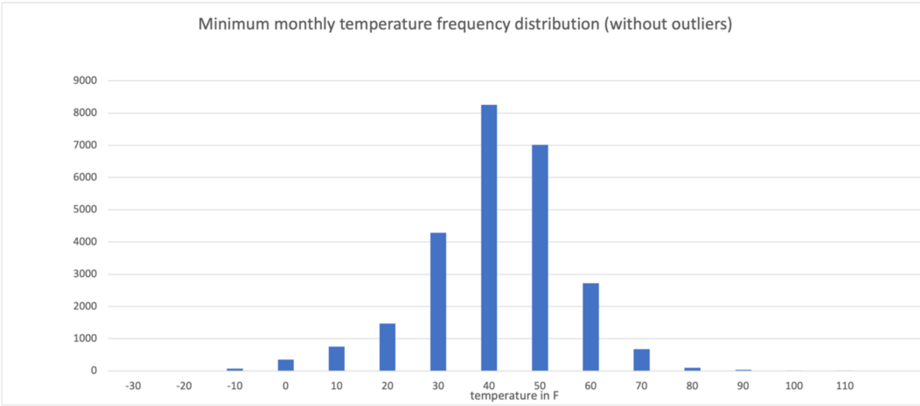

Figure 13: Minimum monthly temperature frequency distribution (without outliers)

3.6 Feature Development and Benchmarking Data

This section describes the data used in developing machine learning model features, along with the failure data used in training and testing the model.

3.6.1 Asset-level Voltage Anomaly Features

The nominal voltage data, consisting of voltage versus time tagged by service point and phase, was used to develop asset-level features by aggregating phase-level data up to the transformer level. Voltage data was aggregated over a day for sample dates selected for model evaluation. This aggregation resulted in anomaly features based on percent over- and under-voltages from nominal voltage, and was then summarized by service point phase (SPP), service point (SP), and transformer (TX).

3.6.2 Neighboring Transformer Features

The neighboring transformer features were formulated to determine if the voltage anomaly existed only on meters under the selected transformer or if neighboring transformers were also impacted. If the anomaly was seen on both the transformer and its neighbors, it was assumed to be caused by an issue on upstream network equipment, and not tied to an individual failure on a distribution transformer.

The neighboring transformer features were developed based on a representative voltage. For multi-meter transformers, this is the average voltage for two meters under a transformer identified as bellwether meters. For single-meter transformers, the transformer voltage was equated to the meter voltage.

3.6.3 Temperature Features

Temperature features were developed to provide the model with insight into how ambient temperature affected transformer function. Features developed included maximum and average temperature, and maximum number of consecutive days with minimum temperature greater than 80 degrees Fahrenheits to account for the cooling ability of a transformer depending on the ambient temperature.

3.6.4 Load Features

Load features were developed to provide the model with insight into transformer loading over the last month, last 3 months, and the last 12 months. Though it was desirable to include the lifetime loading of the transformer, given the data retention history and the train-test split constraints, only a maximum of 12 months of history was achievable. Average loading, maximum percent loading, and number of months with an overload incident were among the features provided to the model.

3.7 Model Development

The model delivered a list of ranked distribution transformer failure predictions using a planning horizon of 30-days, and an emergency horizon of 7-days and was developed in the open source Scikit-learn project library [11]. The model is presented below.

The model machine learning process was based off a “golden failure” dataset, which included transformers that experienced ignition, heatwave, or outage-related failures, combined with a subsampled dataset of non-failing transformers. Each of these transformers was assigned to either the train or test set as described in Section 3.7.1. A failure label column indicated the failure/non-failure state of the transformer. Transformers in the training dataset were split into separate groups with one of these groups making up the test dataset, one group making the training set, and one group making the calibration set.

Tree-based ensemble classification models were the primary class of algorithm used for this project as they were appropriate for tabular data, and because multiple failure modes were anticipated, so an ensemble method was more able to make simple models for multiple failure modes. Also, tree-based algorithms accomplish automatic feature selection. The project team analyzed the resulting feature importance to interpret the features which were being heavily weighted and to determine if the model was picking up features that were expected to be important given an understanding of the problem. Key features that had high importance in the models were the magnitude and direction of the voltage difference between a transformer and its neighbors, the percent of low Rule 2 violations, the range of voltage anomalies for all service points on the transformer, the loading characteristics over the prior 12 months, the voltage level and capacity rating of the transformer, and the 75% quantile of the average historical temperature.

3.7.1 Train/Test Split Strategy

Separate train and out-of-sample test data sets were developed by querying failure records for the period of analysis, grouping by feeder, sorting by number of failures, and then selecting alternating feeders from this sorted list to achieve an approximately even failure distribution. The feeder-level separation between train and test was implemented to avoid any data leakage related to feeder-level dynamics, such as malfunctioning regulation devices. To account for weather patterns resulting in many failures, for example heat storms, it was important to use at least a full year of data in both the train and test data sets. The training period was collected from January 2019 through April 2020, the test period spanned from May 2020 through May 2021. Feeders affected by major non-weather events resulting in a high number of failures were excluded to ensure that the failure metrics were not skewed by system-related failure events.

3.7.2 Pre-calibration Model

The pre-calibrated model was trained on data across two cross validation groups. For the planning and emergency scope, the days within the scope horizon prior to the set of failures were included in the positive sample, and a random sample of non-failure samples were included in the negative data set. This model used the XG-Boost machine learning algorithm. The hyperparameters were tuned using Bayesian optimization. To identify if any of the parameters was leading to overfitting, a learning curve was plotted for a sensitivity of each hyperparameter, and parameters which resulted in an excessive separation of performance metrics for cross validation splits were manually constrained. The developed model was evaluated against the test set. Results are discussed in Section 3.9.

3.7.3 Calibration

Some machine learning algorithms such as logistic regression produce true or 'calibrated' probability scores but most of them produce 'probability like' numbers or a class label for classification. Also, some of the data sampling techniques such as under-sampling might affect the model output by disturbing the sample distribution. For some use cases, class label is important, and probability is important for others, example ‘a model predicting that you don’t have cancer’ vs ‘a model predicting that you’re 49% likely to have cancer’.

For algorithms that do not produce a true probability and/or in case of under/over sampling, we need to calibrate model results to obtain true probability. Thus, calibration is the process of producing probability that reflects the true likelihood of events by rescaling their values, so they better match the distribution observed in the training data.

In EPIC 3.20, we are using a Gradient boosting ML algorithm, which does not produce true probability. Also, the input data is highly imbalanced and so we under sampled the majority class while keeping all the failure labels (minority class) to present more failure examples to the mode. Thus, we needed to calibrate our model results. The calibration model was trained across the cross-validation groups excluded in developing the uncalibrated model. The calibration set consisted of all failures in the train set, and a weighted subsample of the remaining data, limiting the total number of rows used in building the model to 1,000,000. The scikit-learn CalibratedClassifierCV was leveraged to implement calibration to our model results.

3.7.4 Primary Metrics

The primary metrics used for this model were Area Under the Precision Recall Curve (PR AUC) for the continuous model evaluation and hyper parameter tuning. For the discrete metrics, a probability threshold was selected, and the precision, recall, and F1 score were evaluated.

3.7.5 Benchmark Results

It is important to create a target for performance when developing a model. In this project, two benchmarks were evaluated.

First a simple model which used the rate of failure to randomly select and predict a total number of failures per prediction period was used. This Naïve benchmark provides an anticipated floor for performance.

Second, to capture the status quo, records of interventions associated with current business processes were captured and used as baseline performance metric.

When evaluating the impact of interventions, it was important to not only track outage events, but also to track successful interventions. Records of transformer outage and interventions were captured on an ongoing basis, and the performance of the model was compared against these records, in addition to outage events.

3.7.5.1 Naïve Benchmark

The naïve benchmark represents the floor of our anticipated performance. This is a simple model which randomly predicts a number of failures equal to the recorded failures per prediction period. For a number of transformer N, and K failures per prediction period, a randomly selected set of transformers, n, are predicted to fail. The number of true positives is expected to be TP = n*K/N. We can expect the precision (TP/n) of such a model to be K/N, and the recall (TP/K) to be n/N. For all models, the relative improvement over this naïve benchmark is calculated, to provide a relative improvement over this simple model.

3.7.5.2 Status Quo Benchmark

To develop the status quo data set, both historical and ongoing events needed to be assessed. In particular, as interventions were made in the field, those needed to be captured as successes even though they did not result in an outage. To capture the performance of the status quo business processes, work management tickets for dispatch of ad-hoc inspections associated with voltage complaints or power quality inspections were utilized as a proxy to predictions.

3.7.5.3 Power Quality Heuristics

In the first phase of the project, the benchmark function was developed from a set of heuristics developed by the power quality team. These enabled the power quality team to discriminate between issues caused by primary service issues, transformers, and individual meters. When responding to rule 2 violations or customer complaints, this mechanism has been very successful in identifying and mitigating issues.

This heuristic and that process have also been extremely useful in developing the features used in the final model. However, a major learning of the project was that most of the corrections made as a result of the power quality process were the correction of partial voltage conditions as a result of fuse failures. Fuse failures can sometimes be associated with equipment failures, but more frequently they are a result of transient events, where the fuse successfully protects the equipment.

To evaluate the performance of the heuristics used by the power quality group, it is valid to include the resolution of a partial fuse failure as a success. However, when evaluating historical EC tags or outages, a replaced fuse is not captured by these records, and therefore is treated as a failure. For the purpose of predicting and preventing equipment failures, identifying a partial fuse failure from a transient event should not be treated as a successful prediction. Ideally, fuse failures could be identified and excluded from the train and evaluation set. Efforts were made to do this, however, there were many challenges resulting from complexities in preparing the data which made it challenging to correctly label meter nominal voltages, fuse failures and real equipment failures.

The outcome of the Part 1 effort was a model that was reasonable at identifying fuse failures, but not good at identifying unplanned transformer failures. In addition, the transformer failure models developed in this effort were not much better than the heuristic models, though the probabilistic aspect of the model showed potential for prioritizing the transformers for inspection.

Major improvements were made to the processes which determined the nominal voltages on the meters and the ability to exclude fuse failure behavior from the training. Rather than benchmarking and training against the power quality process, only transformer equipment caused unplanned outages were used for training and evaluating the model.

For line regulator failures, which were only evaluated in the first phase of the project, the benchmark used was the heuristic model used to identify whether voltage anomalies might be attributed to a line regulator.

3.7.6 Model Inspection

Models were first evaluated against historical failures to identify data issues and improve model performance. Once the project team was satisfied with the results, the models were used to predict future failures. The predictions are generated for a single day each month and the predicted failures are delivered to an external team for review. The predicted failures were also investigated by the project team using a clustering algorithm as described in Section 3.7.6.1.

3.7.6.1 Failure Clustering

As discussed previously, the EPIC 3.20 model works to identify transformers that might fail within the next 30 (or 7 in case of emergency failure model) days. This would help in planning resources and in proactively checking or replacing the transformers that are highly probable to fail. The next step is to divide the predictions by failure modes so that root cause of failure can be understood, and different departments can take care of the issues falling under their domain and expertise proactively.

To serve this purpose, we performed clustering on high probability failure predictions and further built a decision tree to identify the characteristics of each of these clusters. Traditional clustering techniques such as k-means suffer from assumptions such as largely spherical clusters and the requirement to define number of clusters in advance. To overcome these shortcomings, we used feature reduction and density-based clustering techniques (specifically UMAP [17] and HDBScan [16]. The clusters were trained and labeled on the training data from the failure models as shown in Figure 14. A decision tree was then trained using the feature data to predict the cluster label and the resulting tree was used to develop meaningful labels as shown in Figure 15. A unique capability of the HDBScan library is the ability to apply a trained model to unseen data. This makes it possible to use this analysis to automatically label future predictions to support more automated failure mode labelling, and dispatch of investigations.

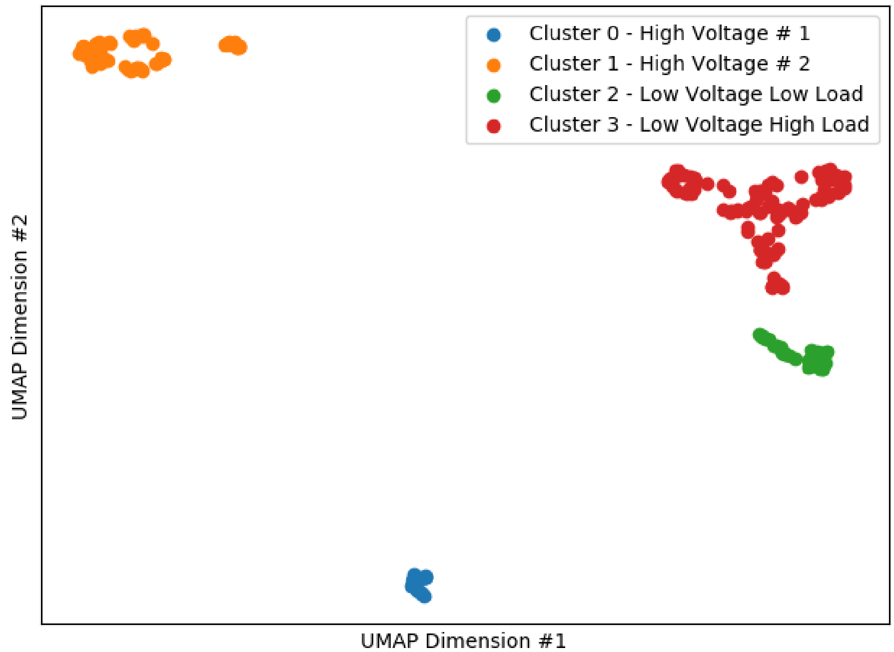

Figure 14: Failure clustering

There are two clusters, High Voltage #1 & #2, which have high voltage characteristics, in investigating these in more detail, it was observed that they were either ~5% or ~10% deviation from nominal, which is likely due to different layers of the windings failing. Though this might not impact the operational process associated with these types of failures, it provides useful information on the failure modes for the transformer standards group.

Cluster 3 and 4 are both failure modes which have low voltage behavior, but the distinction within them is primarily the magnitude of the deviation from nominal. Because cluster 2 is low voltage without low reported load, it may be a scenario where evaluating whether there is a metering problem such as energy theft or a parallel secondary failure mode where transformers are configured in parallel such that a fuse failure on one transformer will not result in an outage, but will overload the remaining transformers on the line.

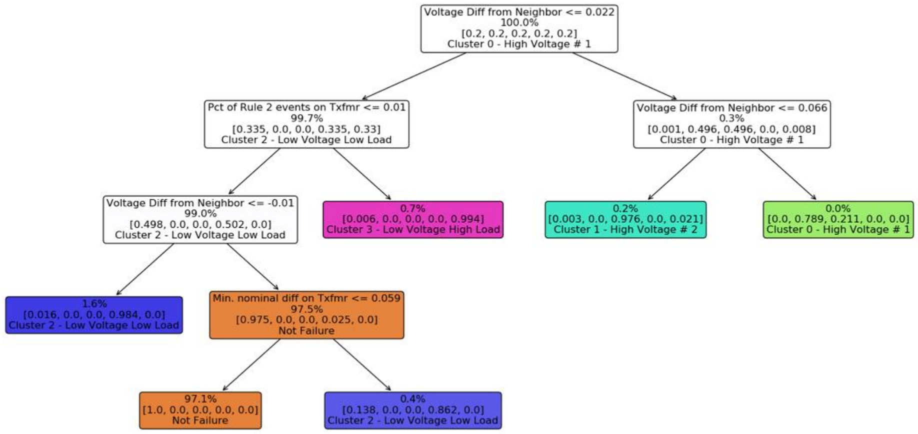

Figure 15: Decision tree for cluster labels

Note that the above clustering is just a demo of how we applied the clustering to our predictions and there could be more clusters to represent different failure modes depending on the parameters that we provide to our clustering algorithm.

3.7.6.2 Incipient Model

An incipient failure is an issue in the condition of the hardware such that the device can be expected to fail in the future unless corrective action is taken. A major challenge in this project has been the ability to know if incipient failure was present. If a failure occurs, it may or may not have been observable before the failure occurred, at the time that the prediction is being evaluated. On the flip side, incipient behavior can be present for a very long time, so it might be present before the sample is categorized as a failure, given the prediction horizon. The incipient model attempts to get the best sample of incipient vs non-incipient behavior by using the assumption that if incipient behavior is present for a failure event, it will be observable before the failure event, and not observable after.

The model was developed by taking sample data from two days before and two days after a failure and using it to train a XG-Boost classification model which predicted whether the data was ‘pre-failure’ or ‘post-failure’. The model was then used to predict for the failure events as far back as 120 days prior to the failure. While output from this model was not included in the final packaged model, it served as a useful benchmark of the upper bound on the features that were possible to predict and influenced model design.

Figure 16 shows the confusion matrix from predicting whether a transformer data point was pre or post failure for samples of +/- 2 days. The percent of True Positive predictions in the lower right quadrant indicates that in short horizons, about 28% of failures show some predictable signal prior to the failure event. The 16% false positives in the upper right quadrant indicate that there are either some failures which continue to have incipient behavior after failure, or there is some fuzziness in the incipient behavior that the model is unable to consistently distinguish.

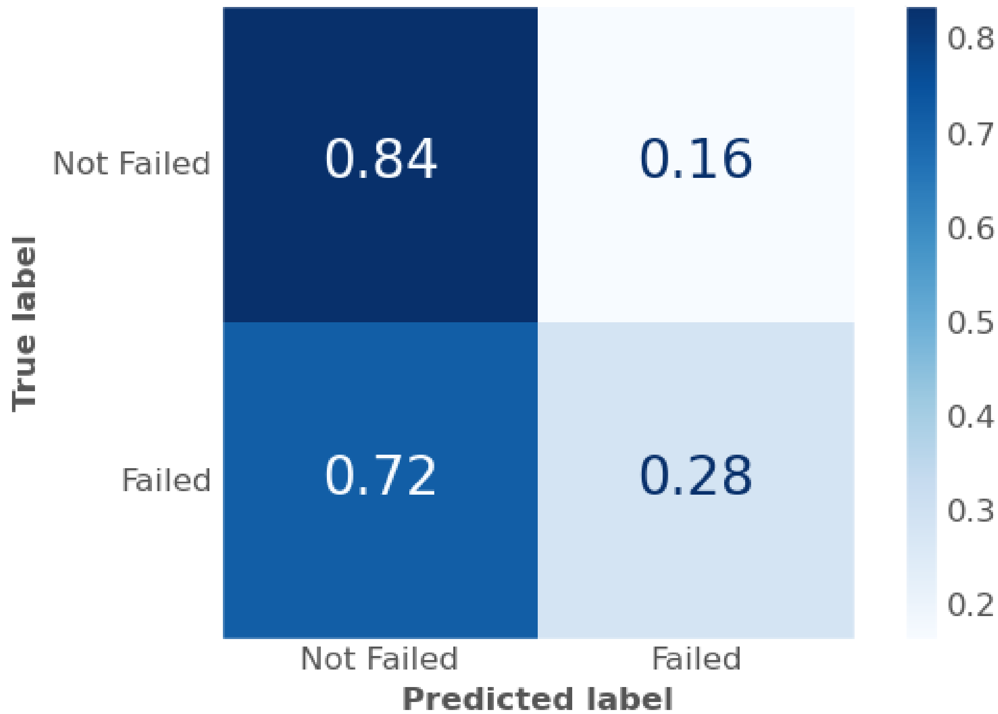

Figure 16: Incipient failure model confusion matrix for two days pre and post failure

One insight that can be gained using the incipient model is by applying the model to sample for time leading up to known failures. By observing the predictions of incipient behavior in the time leading up to known failures, one can estimate how long incipient behaviors are sustained prior to failures. The results of doing this are shown in Figure 17, from 120 to 1 day prior to the failure. From this chart, it can be observed that samples marked as incipient are present at a rate of almost 18% up to 1 year prior to the failure. In the 30 days prior to the failure, it is present in about 24% of the samples. In the days before the failure it increased, about 32%. In the days before the failure, it increased to about 28%. This shows that there should be an increase in the ability to identify failures in the 'emergency' time frame. However, it also indicates that there are a substantial number of incipient behaviors that are long lived.

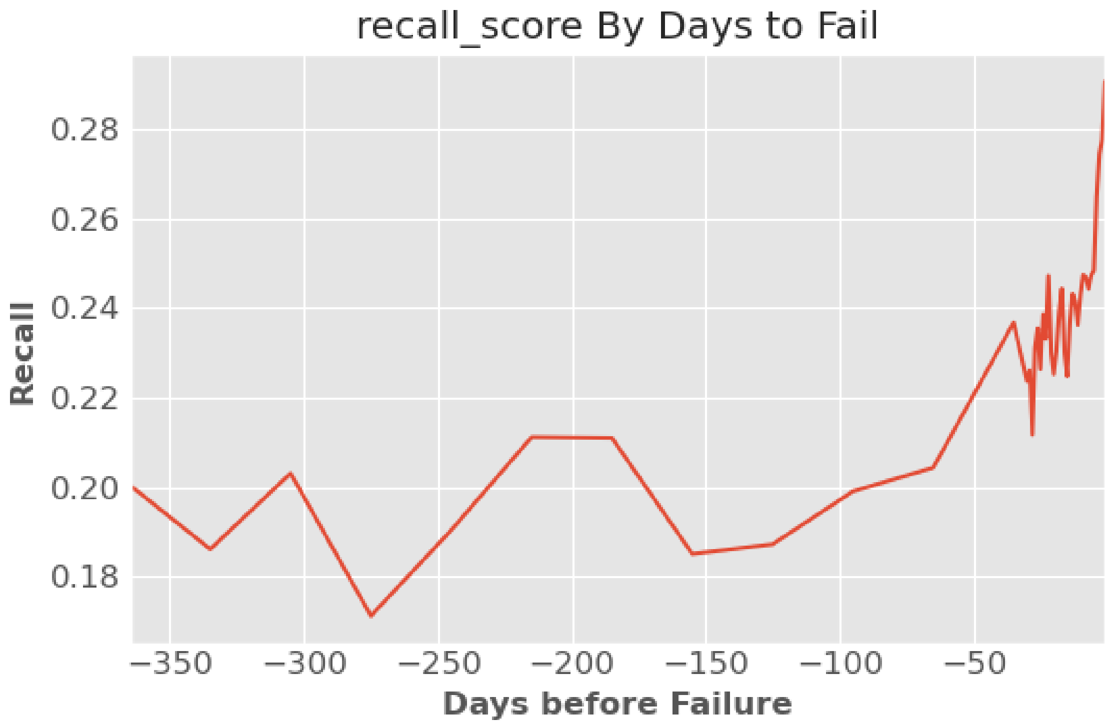

Figure 17: Incipient Model recall by days prior to failure

3.8 Demonstration Deployment / Field Validation

An invaluable component of this project was to leverage the insights of the PG&E engineering and field staff. After predictions were developed, a user interface was developed to support visibility into the underlying voltage and load data for the transformers, and to track the engineering review. The predictions above a selected threshold were delivered on a periodic basis to an engineer who performed a desktop review of the prediction to determine if it should be escalated to a field inspection or to a proactive replacement program. A flow chart of this process is shown in Figure 18. This project leveraged the insights of the PG&E engineering and field staff, and predictions identified as false positives were used to improve the model.

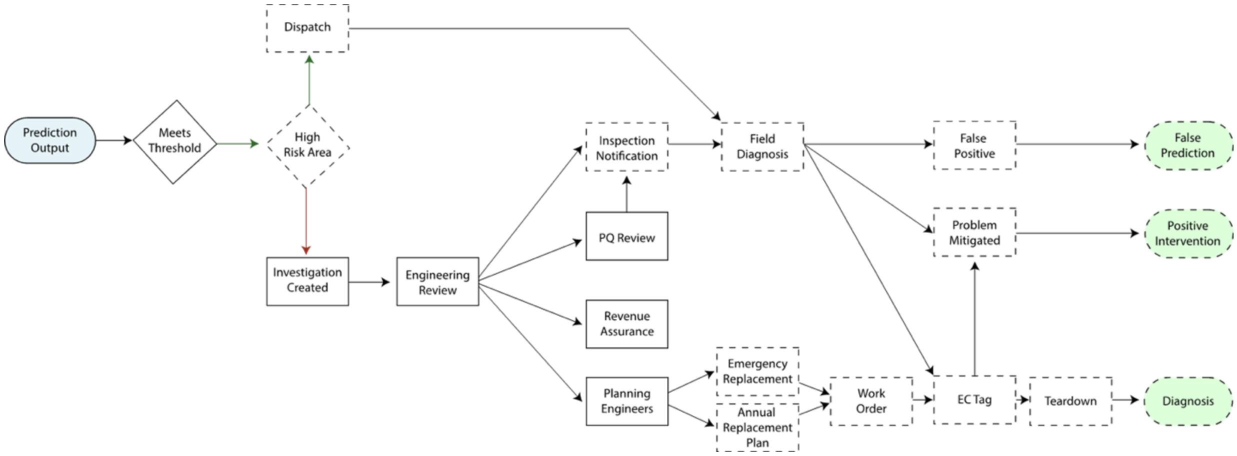

Figure 18: Desktop review process

3.9 Results

3.9.1 Transformer Failures

3.9.1.1 Transformer Planning Scope

The planning scope model for transformers was developed with a horizon of 30 days. A threshold of 1% probability of failure was used.

Table 2: Planning Scope and Benchmark calibrated test set results

	TN	FP	FN	TP	Precision	Recall	F1_Score
Planning Model	12,012,420	3,715	6,564	152	0.039	0.023	0.029
Ad hoc Benchmark	11,920,527	225	8,449	29	0.114	0.003	0.007

The result above represents the performance of the model for predicting a failure in a specific 30 day time frame. If the metrics are reformulated to evaluate the performance of the predictions for identifying a transformer that failed at any point during the evaluation period, then the results are significantly improved. These metrics are shown in Table 3.

Table 3: Planning Scope performance for predicting any transformer failure in the period.

	TN	FP	FN	TP	Precision	Recall	F1_Score
Planning Model	170,250	778	1,081	89	0.103	0.076	0.087

This finding is aligned with the fact that many incipient behaviors can be sustained for a very long time, so predicting exactly when they fail can be difficult. With this understanding, it should be recognized that the relatively un-impressive results in the model evaluation stage don’t represent the much better results observed in the desktop and field evaluation.

3.9.1.2 Transformer Emergency Scope

The emergency scope model was developed with a horizon of 7 days. The results are shown in Table 4. It is notable that the recall is higher than the planning scope predictions, though the precision is a bit lower.

Table 4: Emergency Scope calibrated test set results

	TN	FP	FN	TP	Precision	Recall	F1_Score
Emergency Model	11,458,143	443	1,435	9	0.02	0.006	0.009
Ad hoc Benchmark	11,920,527	225	8,449	29	0.114	0.003	0.007

Table 5: Emergency Scope performance for predicting any transformer failure in the period.

	TN	FP	FN	TP	Precision	Recall	F1_Score
Planning Model	171,038	100	1,039	21	0.174	0.02	0.036

3.9.2 Engineering and Field Investigations Results

The engineering field results were evaluated by first having an engineer perform a desktop review of the predictions. In the first phase of the project, this review was performed by the Power Quality team, using their existing process. The predictions were made in June and December of 2021. The outcome of these reviews identified several problems with the model. Many of the predictions were determined to be false positives resulting from deficiencies of the process used to prepare and filter the smart meters, or mis-mapped meters. Additionally, several were identified to be non-critical issues, such as blown fuses, or smart meter failures. There were 17 predictions that did identify problems in the field, but those were primarily fuse failures.

The results from this engineering review are shown in Table 6.

Table 6:Phase 1 Engineering Review Outcomes

Values	Cannot Determine	False Positive	Non-critical issue	Primary issue	Dispatch to Field	Grand Total
Outcome Count	11	20	13	10	100	154
Percent	7.14%	12.99%	8.44%	6.49%	64.94%	100.00%

The team took the learnings from Part 1, phase 1 and used them to take a second attempt at the problem. Predictions were made starting in April of 2021 and refreshed monthly through July of 2021. The outcomes of the planning reviews in Part 1, phase 2 are summarized in Table 7 and Figure 19.

The results of the engineering review proved to be very successful. In many cases, even if the transformers were not evaluated to be at risk of imminent failure, the engineer agreed that there was a clear problem in the operation of the equipment. Though the goal is to prioritize imminent failures, and the model is built to optimize that capability, in practice a transformer which only has a 1% probability of failing in the next 30 days may have a 75% probability of being flagged as malfunctioning in an engineering review.

The engineering review also helped to identify several false positives, which resulted in further improvements to the model. For example, some false positives were occurring when a neighboring transformer had anomalous voltages, resulting in a false positive on the wrong transformer.

Table 7: Detail of engineering review outcomes for phase 2 planning scope predictions

Outcome	**Success Criteria**	**Count**
Cannot Determine	Cannot Determine	3
False Positive	False Positive	18
Suspect Blown Fuse	Other issue	3
Suspect Overload - Bypass	Dispatch to Field	10
Suspect Overload - No Bypass	Dispatch to Field	47
Suspect Parallel Secondary Blown Fuse	Dispatch to Field	5
Suspect Primary Issue	Other issue	3
Suspect Windings Failure	Dispatch to Field	4
Suspect Wiring Issue	Dispatch to Field	2
Suspected Cycling AG load	Dispatch to Field	15
Suspected Meter Issue	Other issue	5

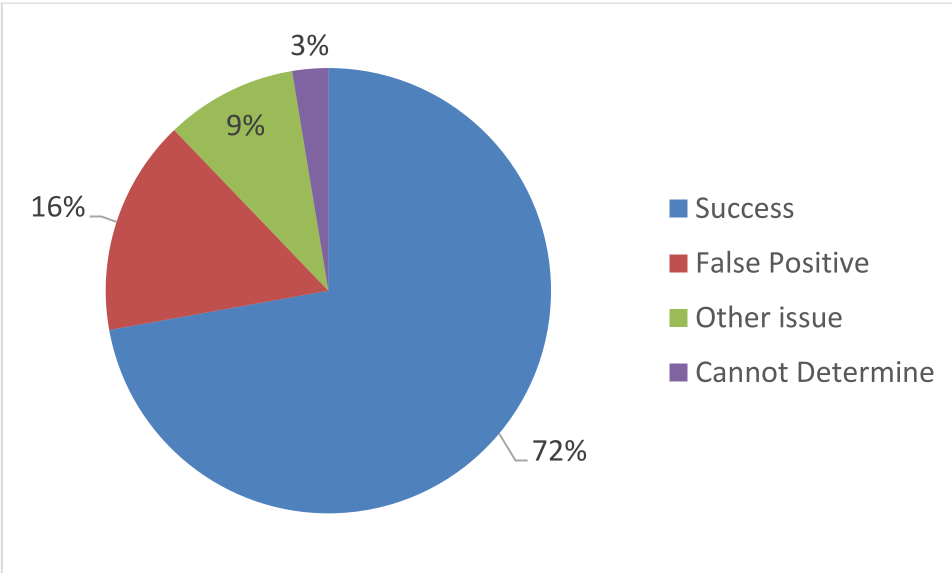

Figure 19: Summary of engineering review outcomes for phase 2 planning scope predictions

3.9.3 Field Investigations

If desktop reviews identified a problem, these investigations were escalated to a field investigation. As of the time of this report, over a dozen transformers had been investigated in the field and had identified 11 transformers which presented signs of incipient or imminent failure.

Most of the identified issues were windings failures, however, the process also identified other types of failures. In one case, a transformer which was overloaded with a customer bypassing the meter was found to have a secondary wire which was smoking due to the unmetered load. In other cases, transformers which were in a parallel secondary configuration were tripped off-line, leading to the other transformers in the parallel configuration to be overloaded. This scenario has previously resulted in catastrophic transformer failures, so it was valuable to identify.

In addition to that, several of the issues which were identified as not transformer problems, were other operational issues, such as customer wiring or meter problems. Though they may not represent the targeted equipment risk, the value of mitigating them is a good consolation in dispatching the resources to investigate them.

Table 8: Outcome of completed field dispatch inspections

Category	Investigation Outcomes	Percent
Fuse Failure	1	6.25%
Meter Issue	3	18.75%
Other	1	6.25%
Overload - Bypass	1	6.25%

Parallel Bank Fuse Failure	1	6.25%
Windings Failure	6	37.50%
Wiring issue	3	18.75%
Grand Total	16	100.00%

Table 9: Outcomes Categories for Field Interventions

Field Intervention Category	False Positive Prediction	Non-Transformer Problem	Successful Intervention	Grand Total
Outcomes	1	4	11	16
Percent	6.25%	25.00%	68.75%	100.00 %

3.9.3.1 Forensic Review

Of the transformers removed from the field, 3 were identified for a forensic teardown in PG&E's Emeryville repair facility. The results of the first of these teardowns are shown below, at the time of this report, the others had not yet been completed.

3.9.3.1.1 Teardown #1 – Windings Failure

In May of 2021, a transformer was identified with a 10% probability of failure. The characteristics of the failure were that the voltage on multiple meters had suddenly jumped to about 110% of nominal voltage, with no change in load. This behavior was not observed on neighboring transformers. The transformer was replaced on May 16th, and taken to Emeryville for a teardown. The voltage behavior of the transformer service points before and after the failure and replacement is shown in Figure 20.

During the teardown, the windings ratio was tested and confirmed to be out of specification. Upon opening the case, a distinct smell of burning oil was observed and the oil was noted to be a dark color.

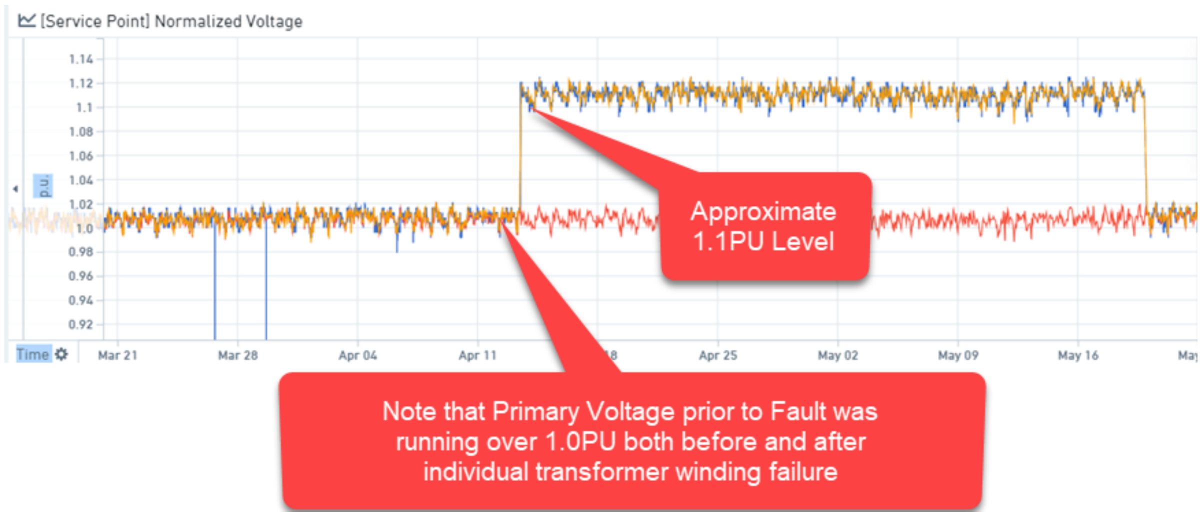

Figure 20: Voltage behavior of windings failure

Figure 21: Photos from windings failure teardown

3.9.4 Summary of Results

The final summary of results for the two transformer models are shown in Table 10.

Table 10: Results summary for all models

	Precision	Recall	F1 Score	Improvement Over Status Quo	Improvement over Naïve Benchmark
Transformer Planning	0.023	0.023	0.029	434%	5142%
Transformer Emergency	0.046	0.004	0.007	105%	5093%

4 Project Details: Part 2 – SSDs and Outage Detection

4.1 Understanding the Problem

4.1.1 Definition of an Outage

Specific types of outages were considered for this part of the project. These selected outages are ILIS outages filtered by unplanned, sustained outages with the following causes: Equipment Failure, Vegetation, Animal, or Unknown.

In order to map selected outages to meter event clusters, a Source Side Device (SSD) for each outage is located. For outages where the equipment is a protective device (fuse, recloser, circuit breaker, or transformer), the device ontology data is used for this mapping. For the outages where the equipment is a switch, the upstream SSD is found.

The prediction horizon is the length of time after the prediction start date, that an outage event would result in the prediction being labelled as an outage. The outage start date represents the date that the outage occurs or is first identified. The outage end date is the date at which the equipment is restored to full working order.

The primary goal of the Part 2 model is to predict outages using event data prior to a catastrophic or unplanned outage. Therefore, the objective is to observe meter events which precede an outage in a time prior to the actual outage. These events are called precursor events which are meter events that have occurred in the past 90 days before an outage. The precursor events are used as features to predict outages that may occur 30 days into the future from a given date. For example, if January 1, 2022, is the prediction start date, then the precursor events would be all of the events between October 3, 2021 and Jan 1, 2022. The future predictions would be for outages that may happen within the next 30 days (until February 1).

4.1.2 Definition of a Source Side Device (SSD)

Source side devices are protective devices that operate when an outage happens. In Part 2 of the project, the model built is used to predict outages by finding the SSD that would operate during the outage. This model considers four types of source side devices which are circuit breakers, line reclosers, fuses, and transformers. Even though transformers may not technically be considered "SSDs", they are included in our SSD list because transformers usually come with cutout fuses as protective devices between them and the primary. These cutout fuses do not appear in EDGIS, so the transformers serve as a proxy for them for the purpose of listing all protective devices that may operate when an outage happens. A main list of SSDs is produced and each SSD is represented by a "global_id", which is used as the primary key for the datasets.

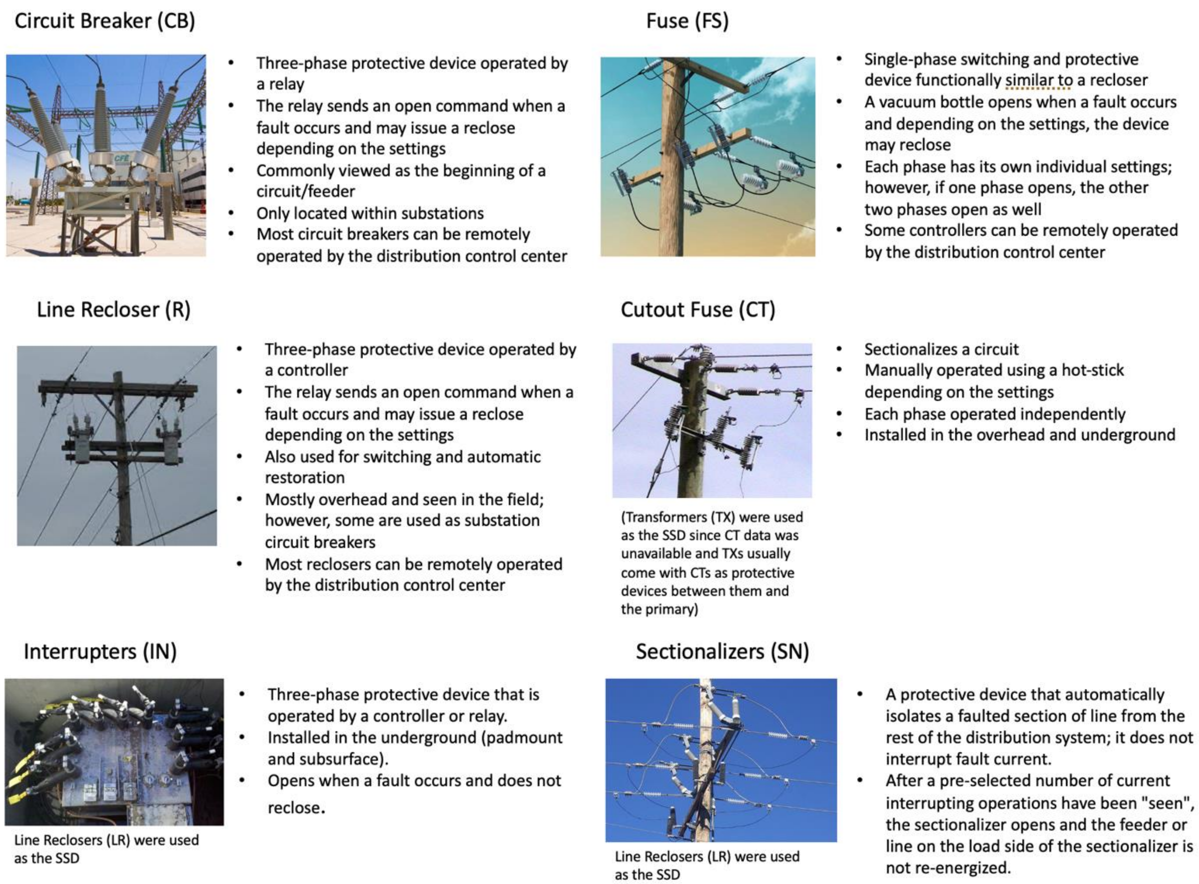

Figure 22: Diagram of Circuit Breakers, Fuses, Line Reclosers, Cutoff Fuses (Transformers), Interrupters, and Sectionalizers

4.2 Literature Review

A Literature review was performed for fault/outage prediction using smart meter event data [18]. The literature review found that there was early outage detection using real-time data from SmartMeter^{tm} data by filtering and clustering meter events to detect outages, but this was used for early detection rather than prediction [18]. This framework is still relevant to the project, but because it was not used to create predictions, the framework cannot be applied. The next literature review looked into a Predictive Maintenance model (with Natural Language Processing (NLP) techniques) in production with real-time Internet of Things (IoT) data [23]. This review is relevant to the project scope but more details about the implementation are needed since they are not included in the paper.

There were two frameworks that were considered but ultimately found to be irrelevant due to failure to include real-time predictions. They were: an anomaly (outage) detection framework that assigned a Poisson distribution to events and found parameters through an optimization [19], and a sliding-window classification approach, using a bag-of-words event representation that also made use of random forest models [20]. Fault analysis of smart meters using a fault tree was another framework examined, although fault localization is not in the scope for that project and phasing was not considered [22], so it is also irrelevant. Another framework considered detection of abnormal events in the smart meters [21] but the preventive maintenance application portion was proposed for future work and not included in the paper.

4.3 Data Collection

4.3.1 Data Sources

Developing a supervised model to predict SSD equipment outages requires finding sources for meter event data, weather data, fault current data, and outage data. The following types of data were used:

- Meter Event Dataset
- Feeder Network Trace Dataset
- Outage Dataset
- PI Dataset
- Weather Dataset

4.3.2 Exploratory Data Analysis (EDA)

4.3.2.1 Meter Events Data

A "meter event" is a type of message generated by a SmartMeter^{tm} which signifies certain specific information about the meter itself or about the voltage and current sensed by the meter. For example, a SmartMeter^{tm} might communicate a meter event when its voltage sensors detect an anomalous voltage or when the SmartMeter^{tm} updates its firmware. There are many different types of meter events. Part 2 of the project focused on the subset of meter event types which are expected to have predictive value (e.g. including anomalous-voltage events, but not including firmware-updating events). Each individual "meter event" corresponds to a particular type of thing happening to a particular meter in time. The time granularity of meter event timestamps varies with event type.

4.3.2.2 Outage Dataset

The Outage Dataset includes all the outages that have occurred and which source side device the outage was at. The dataset also includes other columns with information such as cause details, location, start and end times, whether it was planned, and whether it was sustained or momentary. For this project, this dataset was filtered to include only some data that was needed to build the model.

The first filter is on the types of outages. Outages can be sustained or momentary, and have various causes such as company initiated, equipment failure, or wildfire mitigation. For this project, the data was filtered to only include sustained outages that were caused by equipment failures, vegetation, animals, or unknown causes.

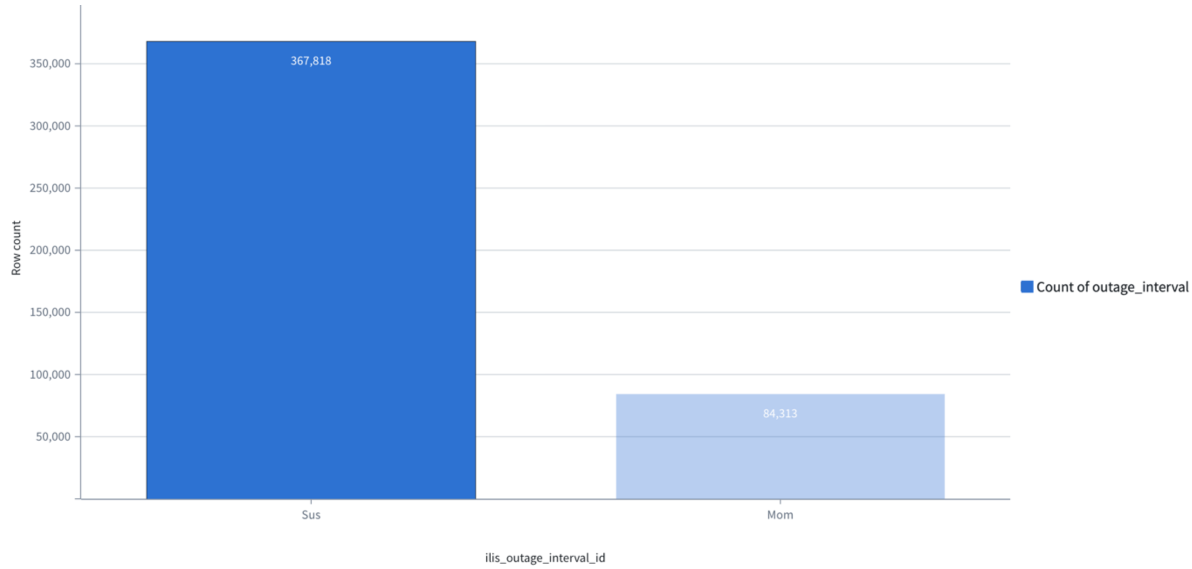

Figure 23: Type of Outage

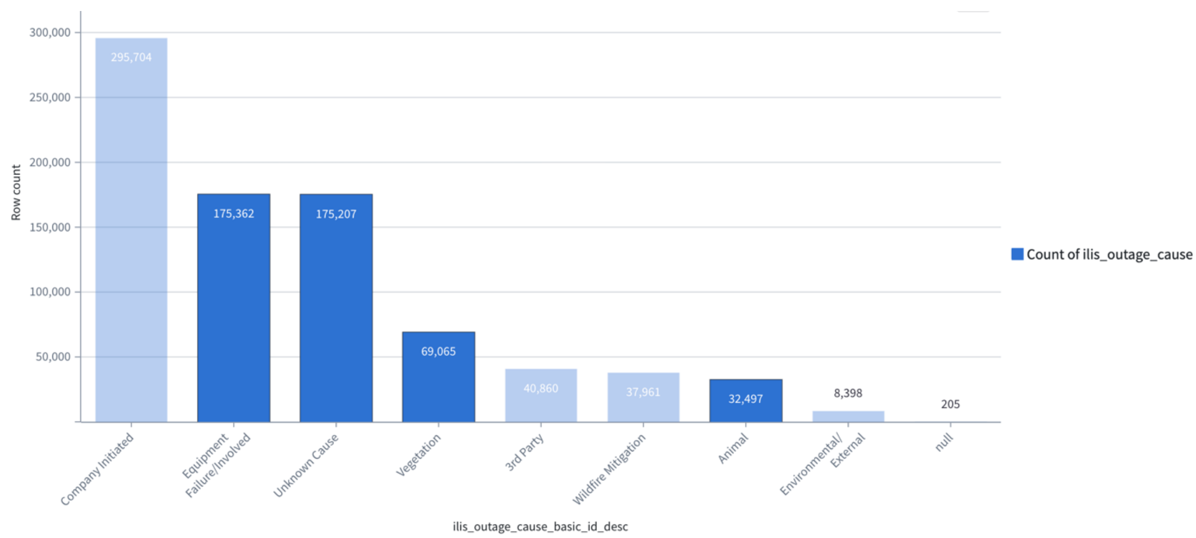

Figure 24: Outages Filtered by Cause ID

4.3.2.3 Weather Datasets

Part 2 of the project used three different weather datasets (High Resolution Climatology, PG&E Operational Mesoscale Modeling System (POMMS), and Weather Station (Synoptic).

The High Resolution Climatology dataset includes 2 x 2 km, hourly resolution data from 1989 to August 2021. The POMMS dataset is a mesoscale meteorological model that runs four times per day and produces hourly 2 x 2 km resolution weather forecasts out to 129 hours in the future for PG&E's service territory. The data is used to calculate wind components in the x- and y-axis using wind speed (mph) and wind direction (degrees) as input to calculate wind speed in meters per second (m/s), wind direction in radians, wind component along the x-axis in m/s, and wind component along the y-axis in m/s. The nearest POMMS point to each SSD point is calculated using H3, a hexagonal hierarchical geospatial indexing system developed by Uber Technologies[24]. This effectively sets the hexagonal area around a POMMS point over which the weather calculated at that point is considered to be in effect. The weather at the SSD point is then considered to be the POMMS-based hexagon that intersects with the SSD-based hexagon.

Weather Station or Synoptic Dataset is point data by weather station. Station coverage is coarser farther back in time and measurements between networks are not consistent.

4.3.2.4 Fault Current Dataset

The Fault current dataset that was used in the EPIC 3.20 project to create features for the Line Recloser SSD type. When a recloser opens due to a downstream fault, a fault current value is recorded. The data is available as a timeseries.

4.4 Feature Development

4.4.1 Event Features

There is a significant amount of data engineering involved in the meter events and their connection to outages. Certain instances of meter events were identified as phenomena of a qualitatively different type than other instances of meter events which had the same putative event type. These differences were identified and leveraged into filters to only include meter events which are expected to be of predictive value (i.e. increasing the signal-to-noise ratio).

An algorithm was developed to aggregate the meter-specific meter events, which are associated with a specific meter to the SSD level. That is, meter events are algorithmically associated with the SSD which we expect will operate if the meter event is predictive of a future outage.

Finally, in order to create event features, certain aggregations of the meter events over 90-day windows were taken so that the resulting aggregations reflect events on a particular SSD over a particular 90-day window. Some of the 90-day aggregations included a time-dependent weighting, in which the general principle used was that more recent events should be weighted more heavily than older events.

4.4.2 Historical Outage Count Feature

It was hypothesized that the count of historical outages on each SSD could be used as a feature because it may indicate SSD behavior based on its location, etc. For all other features the training data is taken from March 2019 to December 2021 but historical outage count is the historical value of our data label and may give the model information about the label beforehand. Thus, to prevent data leakage, we used data from the year 2018 to create the historical outage count feature. When included in model training, this feature turns out to be the most important feature. This feature remains the most important feature for the model with a feature importance value one order of magnitude higher than the rest of the features.

4.4.3 SSD type Features

Primarily four different kinds of SSDs are being considered in this project: fuse, circuit breaker, line reclosure and transformers. The rate of failures is different for each of these, specifically the rate of failure for transformers is much less than other SSDs. To consider these specific differences amongst different types of SSDs, SSD type was considered as a feature for the model. These features are in the top 5 important features, with the 5th ranked SSD type being transformers – followed by line reclosures, fuses, and circuit breakers in descending order.

4.4.4 Season Features

Based on season, weather conditions would change and factors such as wind gust or wind direction in a particular season would play a role in causing outages. The Season Feature was created to capture this hypothesis. The feature consists of four values: summer, winter, fall and spring corresponding to each example used in the training data. These four values were one-hot encoded and used in the model. These features do not currently appear at the top of the model's feature importance list.

4.4.5 Weather Features

One of the types of failures being considered in this project is those caused by vegetation. A scenario would be when due to high wind gust, a tree collides with a line and causes an outage. Another scenario is when wires flap together due to a high wind gust in a particular direction. To consider such scenarios, the decision was made to include various weather features. These weather features capture the same weather aspects for the model as the season features described above, but they can bring some extra intricacies such as the numerical values of wind gust, etc. Hence, weather features such as wind gust, wind direction, and ambient temperature were formed into features and used to train the model. Climatology weather data, as described in section 4.3.2.3 contains weather data recordings as granular as 2 x 2 kms from an SSD and also gives a forecast of weather into the future. These features appear in the top 20 in terms of feature importance, with 90 days min of wind speed at the 9th position indicating that a low wind gust would lead to less outages.

4.5 Model Development

The machine learning model delivers a list of ranked source side device (SSD) outage predictions using a horizon of 30-days and was developed using the open source Scikit-learn library.

The data label is based off an outage dataset, which includes SSDs that experienced outages. This is combined with a subsampled dataset of non-outage SSDs to form a balanced set for model training. Balancing is done at the substation level, meaning outage SSDs under one substation are matched with an equal (or as per the imbalance factor) number of non-outages under the same substation. This ensures a level playing field for both types of records in case there is an issue with the substation. Each of these SSDs is then assigned to either the train or test set as described in Section 3.5.1. SSDs in the training dataset are further split into separate groups with one of these groups making up the test dataset, one group making the training set, and one group making the calibration set. Further, various datasets are explored to form the model features, meter events being the focus of our data exploration. Some additional features are used as described in the previous section.

Tree-based ensemble classification models were the primary class of algorithms used for this project as they were appropriate for tabular data and because of their efficiency with this kind of data. Also, tree-based algorithms provide automatic feature importance. Specifically, Random Forest, Gradient Boosting and XGBoost classifiers were explored as model types. The model parameters were determined through hyperparameter tuning using Bayesian optimization. Various feature combination scenarios were simulated to determine if the model was picking up features that were expected to be important given an understanding of the problem. Key features that had high importance in the models were historical count of outages, Voltage anomaly events, meter power cycling events, type of SSD and wind speed. To identify if any of the parameters was leading to overfitting, a learning curve was plotted for a sensitivity of each hyperparameter, and parameters which resulted in an excessive separation of performance metrics for cross validation splits were manually adjusted.

4.5.1 Outage Labeling

In order to train the outage failure model, outages on SSDs are used to label the training data. The first step was to find SSDs that operated for the selected outages mentioned in 4.3.2.2. The SSDs included are transformer cutout fuse, fuse, Dynamic Protective Device (DPD), and circuit breakers. If the outage was forced out using a switch or a jumper, the upstream SSD is found as the SSD for the outage. Outages were to be predicted for 30 days from the date of running the model. Therefore, outage label 1 was given to those SSD/date select combinations where the date selected was between the outage date and 30 days prior to it (or the horizon date).

Based on some deep dives into the initial results, it was found that some of the outages that did not see a precursor event because the outage actually happened on a different SSD than the log reports. Label refinement was performed with the help of meter number/transformer numbers in the fault location. To help the model learn as much as possible, an outage label was provided to the SSD from an outage record as well as the one derived from fault location.

4.5.2 Train/Test Split Strategy

Separate train and out-of-sample test data sets were developed by using a weighted count of outage types in the original outage dataset. Vegetation, Animal, Equipment failure and Unknown were the types of outage causes used for the purpose of this project as shown in Figure 25. For each substation, for the period of analysis, outages were sampled according to their ratio in the original outage dataset to ensure uniformity over outage cause. An 80-20 split was made to create train and test sets out of this dataset using stratified sampling. The Train set was further split into five cross validation sets using the same weight as above. Train and test sets were further divided based on time to include 2019 and 2020 data for training the model and 2021 data for model testing.

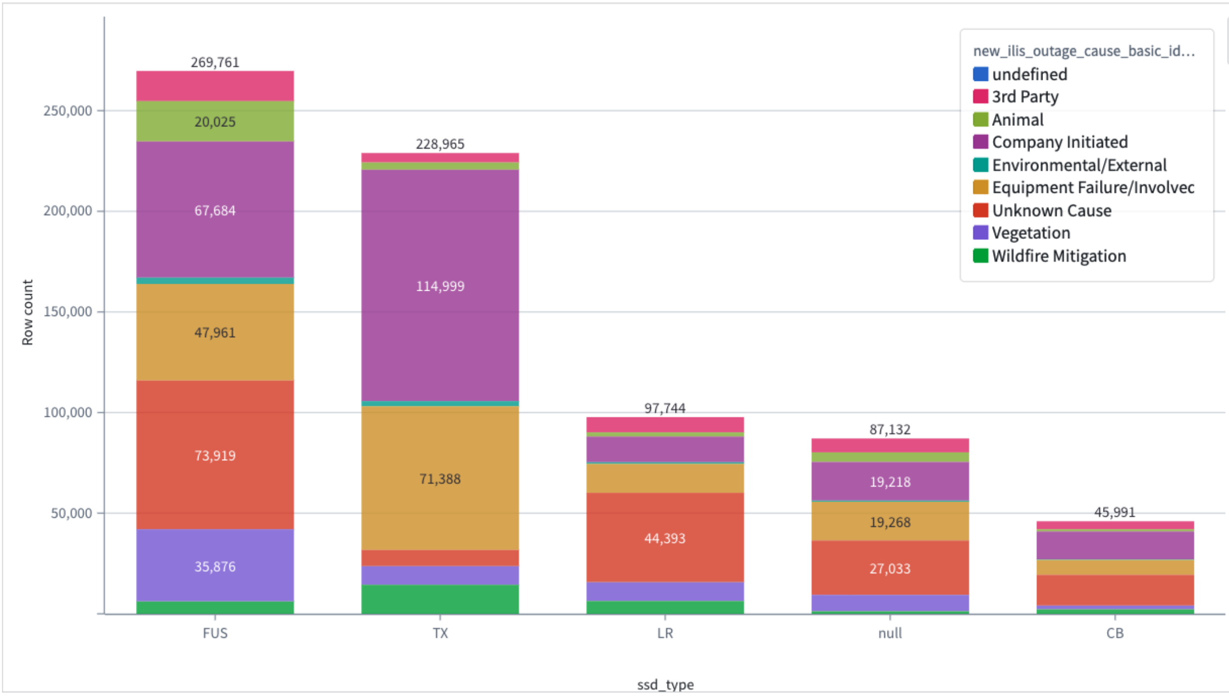

Figure 25: Train set by outage cause

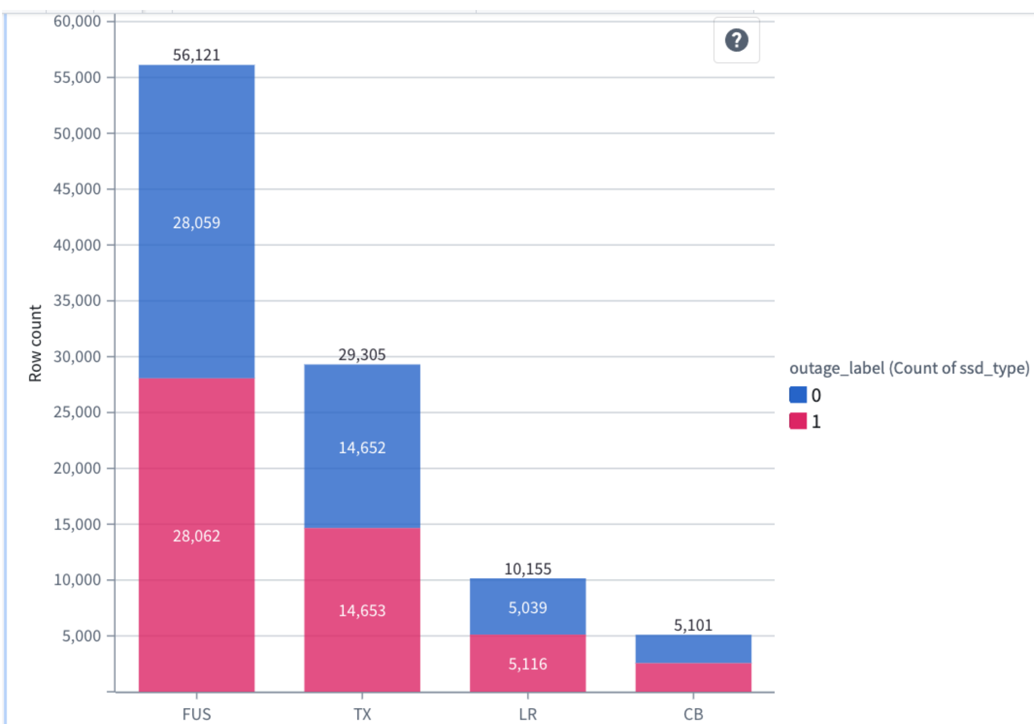

Figure 26: Outage cause by SSD type

4.6 Validation Process

After initial model development, there was some post processing done to validate the model's performance on new data. First, a date was selected to create future predictions, such as for

November 9th, 2021. Then, features were created for this date, the model was run, and predictions were created. The predictions were made up to 30 days from November 9th, for every SSD and had corresponding probabilities of likeliness for positive predictions. Finally, to validate these predictions, a check was made on the top 300 or so rows with the highest probabilities of an outage to see how many of these SSDs that had predicted an outage were in fact true positives. These true positives can be thought of as the number of outages the model was able to predict in the future that actually happened.

An invaluable component of this part of the project was to leverage the insights of the PG&E engineering staff. After predictions were developed, a user interface was developed to support visibility into the underlying voltage and meter event data for the SSDs under investigation. The predictions above a selected threshold are delivered on a periodic basis to an engineer who performs a desktop review of the prediction to validate the model results. This component moved the project into the MVP stage via leveraging the insights of PG&E engineering staff to identify false positives and continue to improve the model.

4.7 Primary Metrics

The primary metrics used for this model were Area Under the Precision Recall Curve (PR AUC) for continuous model evaluation and hyper parameter tuning. For the discrete metrics, a probability threshold was selected for each SSD type, and the precision, recall, and F1 score were evaluated.

4.8 Heuristic Approach

A heuristic approach was developed to investigate a simple way of predicting an outage. For this approach, the focus was on sag and swell events that were believed to be good precursor events to an outage. A year of outage data was used to create these meter event features by aggregating the events by 90 days and 1 year for each row based on the SSD global_id. Then, the sag and swell counts were compared for 90 days and 1 year before the prediction date, and if there was an increase in the counts, an outage was predicted.

Overall, this approach had a lower number of true predicted values than the machine learning models (7 true positive predictions made on November 17, 2021), which was expected. In conclusion, these types of meter events have high predictive value as features for the model.

4.9 Model Calibration

Our base model (the tree-based models) produces ‘probability like’ numbers as outputs. Knowing the true likelihood of occurrence of an outage is important to understand the gravity of the prediction, so we turned to model calibration. The Scikit-learn CalibratedClassifierCV was leveraged to implement calibration. The calibration data consisted of all the samples that were used to train the base model. Additionally, it included the majority class data, which was dropped for the pre-calibration stage model to create a balanced training set. Using the complete data, the model was given a sense of how the data distribution is in the real world and it could therefore predict the true probabilities of the outages.

4.10 Results

The outage prediction model for the protective devices and transformers was developed with a horizon of 30 days, which means that at any day the model is run, it makes predictions for the next 30 days. A threshold of 50% probability of outages was used.

	True Negative	False Positive	False Negative	True Positive	Precision	Recall	F1_Score
Gradient boosting model – pre-calibration test results	28164	7326	20180	15522	0.679	0.435	0.53

After training the model, predictions for a particular month showed ~5% of the top 300 model predictions were true positives for a given month. Top predictions are looked at and analyzed for further model improvements.

5 Challenges

The team encountered numerous challenges in developing the predictive models throughout both parts of the project.

Data Challenges

Unsurprisingly, significant data challenges were to be navigated as historic data recording methods and procedures were not designed for such data applications as existed in this project. While standards existed for data entry, they were not always followed or may have changed over time, and several opportunities existed for data entry issues to occur.

In many cases, records were available only as free text, which made leveraging them difficult and inaccurate. Data was not accurately labelled and associated. For example, the data did not capture the case where a failure was identified and resolved prior to causing an outage. In addition, failure records did not necessarily identify the specific device that failed and were instead associated with an asset that was removed from the failure, e.g., a failure record may be associated with a pole rather than the failed transformer.

Misconfigured meter cases existed where an incorrect meter form was installed at a service point, resulting in an incorrectly normalized voltage in the data. Phasing information, which would have allowed a more accurate voltage comparison to be made across multiple meters under a transformer, was not accurately recorded. Cases existed where the physical wiring of the meters was flipped during field service, causing a new nominal voltage level to run across a particular service. Finally, GIS mapping of meters to transformers was found to be error prone.

The project team identified challenges related to underground cable and regulator data records while exploring which use case to focus on.

Underground cable locations and events were not consistently tracked. In addition, individual underground cable segments were not tracked as assets, and the data did not identify the segments that had failed during an outage. Asset characteristics such as cable type were lost for equipment that was replaced and could not be used to train failure models. Regulator settings, which have a direct effect of the voltage pattern on the circuit and that would have provided the project team with useful information in identifying the historical cause and resolution of voltage problems on the primary, have not been consistently tracked.

In Part 1, weather features such as maximum temperature, wind speed, etc. were created using the historical weather data that came from weather stations. In Part 2, the aim was to use POMMS data to create weather related features as this has better granularity (2 x 2 kms) as compared to weather stations. However, the POMMS data was unavailable for the last 6 months as it is predicted quarterly for the past and then undergoes third party validation, and that takes time. Also, the future predictions from POMMS data, that were considered to be used for making predictions had mismatches with the historical data and thus can be used only after additional processing. All this led to the project being unable to fully leverage the POMMS data.

Modelling Challenges

The project team encountered several model-related challenges. The project team faced the challenge of an imbalanced data set, for example in Part 1 only 0.05% of transformers were failing on a monthly basis. Capturing the few that did fail was highly important, but any model that predicted zero outages might be correct 99.95% of the time. To account for this, the project team evaluated the model using imbalance-friendly metrics.

Definitively determining whether anomalies observed in voltage readings were due to incipient failure rather than meter failure, customer misassignment, or secondary wiring issues proved challenging until the field investigations were initiated. Appropriately labelling failures was challenging; in some cases, fuse failures were identified as equipment failures, and sometimes they were not, incipient behaviors were found to be sustained for long periods of time leading to challenges in identifying the specific time frame that a failure might occur. For many failures, no incipient behavior could be observed prior to failure, setting an upper bound on the performance of the model in a way that was difficult to correct for.

Process Challenges

The project team encountered challenges in internally promoting and driving changes to improve the quality and structure of the data sets used. These challenges were somewhat mitigated as the project reached maturity and the team was able to demonstrate effective results. Certain failure types, e.g. fuse failures, identified by the project team did not have a clear owner that would be driven to execute a field review and respond to a failure prediction.

Inaccurate outage start time

In ILIS, outage time is the time when an outage is manually recorded in the database and this may happen substantially after the outage actually occurred. Thus, to estimate the real time of the outages, these ILIS records were matched with meter events. The Meter power up/down and partial voltage and GMI low loss (events related with unusually low voltage) were used as indicators of actual outages. For this purpose, the meter outage dataset was combined with ILIS outage dataset to match the corresponding events. These estimates were used for Transformer and Fuse outages because transformers and fuses are generally not SCADA-enabled, so the recorded outage start times may not reflect the actual time when the outage physically started.

6 Accomplishments

Developed a machine learning model to predict distribution transformer failure and demonstrated success

The project team successfully executed the machine-learning model to provide a ranked-list of distribution transformer predicted failures to other PG&E teams. Of the 270 model predictions that were reviewed by engineering experts from April 2021 through February 2022, 64% were confirmed to be relevant transformer anomalies requiring further investigation. An additional 26% were confirmed to be other issues in the distribution system.

These investigations resulted in an average of one successful intervention with a near-failing asset per week (avoiding catastrophic failures) in the first ten months of running the model. In the process of developing the final model, the project team also demonstrated the potential benefits of automating PG&E's Power Quality Rule 2 compliance process. Failure category clusters were developed to categorize groupings of failure predictions for more efficient investigation.

Fused, cleaned, and strengthened ties between disparate PG&E data sources, while developing algorithms to leverage the data

The project team focused on leveraging existing PG&E data sources, including meter events, outage SSD, and weather aggregate data. The mechanisms developed to operationalize this data will continue to serve the PG&E analytics community. These newly developed processes enabled the project team to collect historical asset failures from records, infer nominal voltage for smart meters, determine historical temperature records for individual distribution transformers, apply fuse failure heuristics to identify these in existing data, distinguish between voltage anomalies due to potential distribution transformer failure as opposed to metering configurations such as incorrect wiring or changes in the winding order, identify neighboring transformers, and to develop a bellwether voltage for distribution transformers which incorporated a mechanism to correct for solar generation.

Deployed a user interface for users to view predicted distribution transformer failures

The project team deployed a user interface that allows external PG&E teams to view and comment on the predicted distribution transformer failures along with a deep well of data associated with the predicted assets. Multiple views exist including viewing the voltage time series on the transformer as well as on its neighbors, map-related views, transformer metadata including manufacturer, model, and year of manufacture. The user interface is currently being used by other PG&E teams to validate and leverage monthly model predictions.

Deployed a preliminary user interface for testing to investigate features of predicted SSD outages

The project team created a user interface for engineers to view the predicted SSD outages along with a deep well of data associated with the meter events and weather measurements. For the MVP user interface to be ready for external users, multiple future tasks are planned to include additional functionality such as a process for checking off predictions for validity and an option to comment on predictions that have been reviewed. A map-related view to show the outage and associated SSDs is also planned to be added. Multiple views have already been added such as a view of the meter events over time for the SSD global ID that is predicted to be an outage and weather measurements such as daily maximum wind speeds by SSD global ID.

Predicting that line equipment will require maintenance means that maintenance can be scheduled within normal operating workflow and avoid expensive unscheduled maintenance

Permitting targeted inspections as opposed to periodic inspections will better optimize field resources. The reliability benefits of 1.2 million customer-minutes per year results in an estimated economic benefit to customers of $3.1 million per year.

Leveraged modern project management tools to drive project resiliency and relevance

The project team employed agile project management and demonstrated resiliency by navigating staff turnover while continuing to drive forward progress. The project team prioritized stakeholder engagement via outreach meetings to a broader audience, and biweekly meetings with the project’s business sponsor.

7 Learnings and Recommendations

Improving failure records and root cause tracking

The key data element for any failure analysis is the record of the failures and, ideally, a root cause analysis of the failure. The project team recommends standardizing and improving failure recording and root cause data collection.

Tracking fuse failures and fuse resets

When looking at meter voltage data, fuse failures are very distinct. In some cases, a fuse failure will be temporary and will be resolved after reset, and in others it will be a symptom of a transformer failure. Records of cut-out and primary fuse resets would be valuable information in distinguishing between these cases, but these are not reliably tracked. The project team recommends that actions to reset fuses be tracked in a systematic way, to enable better capture of the impact and resolution of these events in the system.

Maintaining a history of grid configuration

Reliably identifying the transformer and grid topology relationships for historical failures proved challenging. While the project team found mechanisms to mitigate this, it would be preferable to develop centralized mechanisms to track these changes systematically and enable point in time reconstruction of the grid for analytics.

Improving smart meter data

Smart meters do not provide meter nominal voltage, which must be inferred from the data. This is a relatively complex process, and events such as wires being reordered can create data quality issues. Smart meter vendors should be encouraged to develop mechanisms to capture and report the nominal voltage.

Leveraging analytical techniques for mitigating label errors and data quality issues

Label errors were a particular challenge for this project but are not unexpected in any real-world application. The use of confident learning for modeling and label inspections, and other methods such as clustering of failure predictions to identify potential data quality issues are good mechanisms to mitigate this issue.

Investing in analytical infrastructure

Though many improvements have been made to the data and IT infrastructure to support analytics, challenges were still encountered in enabling end to end analytics solutions. The project team recommends that PG&E continue the development of its analytical infrastructure.

Leveraging internal talent

Utilities should actively invest in internal talent when considering new analytical projects and software development. Internal teams can be more cost-effective, nimbler, and can reduce business continuity risks to improve data quality within business systems when compared against external vendors.

Using agile project management

Agile development continues to show success in managing data science for PG&E EPIC projects and provides teams with the resiliency and nimbleness to manage change while driving growth.

Improving outage records (outage start time, equipment id, fault location)

There are a number of improvements that can be made to existing outage records. The project team discovered inaccurate outage start times, missing equipment IDs, and incorrect fault location records. It is recommended that these records are cleaned at their source so that they are accurate.

Creating and maintaining documentation or a data dictionary on meter event data

Meter event data was used for the first time during this project, but there is not currently any form of documentation on the dataset. The project team reached out to vendors for additional information but did not receive a response. In lieu of this, some information was provided from stakeholders. The project team recommends creating documentation or a data dictionary to better understand the dataset. This would help with feature development and assist future work. It would also be beneficial to create more direct relationships with vendors who provide data to reduce ambiguity.

8 Path to Production

The project's products were developed with repeatability and maintainability in mind. With the implementation of the new analytics platform, a production-ready model for Part1 and a MVP for Part 2 have been developed, which supports updates of the source data sets, retraining and evaluation of models, and deployment of updated models. In addition, user interfaces were developed during Part 1 and 2 which can support evaluation and track outcomes of the predictions.

Because the products developed in this project have continued to display a clear ability to identify assets nearing failure, internal organizations that manage and monitor assets have committed resources to support processing and acting on these predictions going forward, after the conclusion of the EPIC project. In addition, some of these organizations have expressed intent to provide product ownership roles.

To support the path to production, some additional work will be needed to transition to a production level product. Additional features are desired to support ease of use and to incorporate the process into a risk-based decision-making process by incorporating consequence model outputs into the prioritization process. On an ongoing basis, resources need to be maintained to support continuing operation and improvement of the products.

9 Technology Transfer Plan

9.1 IOU's Technology Transfer Plans

A primary benefit of the EPIC program is the technology and knowledge sharing that occurs both internally within PG&E, and externally with other IOUs, the California Energy Commission (CEC) and the industry. In order to facilitate this knowledge sharing, PG&E will share the results of this project in industry workshops and through public reports published on the PG&E website. Specifically, below are information sharing forums where the results and lessons learned from this EPIC project were presented or plan to be presented:

9.1.1 Information Sharing Forums Held

- The Utility of the Future - Transforming Utilities Through Innovation
  Organized by EUCI
  Phoenix, AZ – December 2019
- AI: A Reverse Pitch Virtual Event.
  Organized by EPRI
  Virtual – May 2021
- PG&E and SCE - Session on Predictive Maintenance
  Organized by PG&E
  Virtual – October 2021
- 2021 EPIC Symposium
  Organized by EPIC Program
  Virtual – December 2021

9.1.2 Adaptability to other Utilities and Industry

The following findings of this project are relevant and adaptable to other utilities and the industry:

- Smart meter voltage and loading data can be used to identify incipient transformer failures.
- Meter event data, weather data, and historical outage counts can be used to predict outages with precursor events.

9.2 Data Access

Upon request, PG&E will provide access to data collected that is consistent with the CPUC's data access requirements for EPIC data and results.

10 Conclusion

EPIC 3.20 has successfully demonstrated that a machine learning model can be used to identify and mitigate equipment failures utilizing smart meter voltage data and other information. Both an emergency scope and planning scope have been evaluated. The planning scope model has demonstrated a 4 fold improvement over existing processes for mitigating incipient problems on transformers. As a result, some of the predictions have identified incipient failures in distribution assets, mainly transformers, and these have proactively been replaced, leading to improvements in reliability, affordability and risk.

Despite being at different stages in the development lifecycle, with the distribution transformer model being used operationally and closer to production and the source side device outage model only recently reaching MVP staging – both models hold the potential of growing into full-fledged production models. The success of this project opens the door to continue to use data in order to improve existing utilities' asset management practices, which translates into a safer, more reliable, and more affordable service for customers across the country.

11 References

11.1 Bibliography

1. Kabir, B. Foggo, N. Yu in *Data Driven Predictive Maintenance of Distribution Transformers*, in Proceedings of the 2018 China International Conference on Electricity Distribution, Tianjin, Sept 17-19, 2018.
2. Susto, A. Schirru, S. Pampuri, S. McLoone, & A. Beghi. (2015). Machine Learning for Predictive Maintenance: A Multiple Classifiers Approach. *IEEE Transactions on Industrial Informatics*, vol. 11, issue 3, June 2015, pp. 812-820.
3. M. Kimmet. "Predicting Distribution Transformer Failures." *Predicting Distribution Transformer Failures | T&D World*, T&D World, 19 June 2018, https://www.tdworld.com/grid-innovations/asset-management-service/article/20971387/predicting-distribution-transformer-failures.
4. DNV-GL. "Asset analytics and electric utility asset performance management software with equipment maintenance dashboards providing actionable intelligence from Cascade" *Asset Analytics | equipment maintenance dashboards | Cascade Insight – DNV GL*, DNV-GL, https://www.dnvgl.com/services/asset-analytics-and-equipment-maintenance-dashboards-for-utilities-cascade-insight-101896
5. R. C. Sonderegger, D. Henderson, S. Bubb and J. Steury, "Distributed asset insight," in *IEEE Power and Energy Magazine*, vol. 2, no. 3, pp. 32-39, May-June 2004.
6. R. Moghaddass and J. Wang, "A Hierarchical Framework for Smart Grid Anomaly Detection Using Large-Scale Smart Meter™ Data," in *IEEE Transactions on Smart Grid*, vol. 9, no. 6, pp. 5820-5830, Nov. 2018.
7. J. Singh, S. Singh. "Transformer Failure Analysis: Reasons and Methods." in *International Journal of Engineering Research & Technology*, Special Issue – 2016, vol. 4, no. 15, pp. 1-5.
8. M. Li and Q. Zhou, "Distribution transformer mid-term heavy load and overload pre-warning based on logistic regression," *2015 IEEE Eindhoven PowerTech*, Eindhoven, 2015, pp. 1-5.
9. IEEE Guide for Failure Investigation, Documentation, and Analysis for Power Transformer and Shunt Reactors, IEEE Standard C57.125

10. Curtis G. Northcutt and Lu Jiang and Isaac L. Chuang, Confident Learning: Estimating Uncertainty in Dataset Labels, 2019
11. Scikit-learn: Machine Learning in Python, Pedregosa et al., JMLR 12, pp. 2825-2830, 2011.
12. Lemaître, Guillaume & Nogueira, Fernando & Aridas, Christos. (2016). Imbalanced-learn: A Python Toolbox to Tackle the Curse of Imbalanced Datasets in Machine Learning.
13. Breiman, L. Random Forests. *Machine Learning* **45,** 5–32 (2001). https://doi.org/10.1023/A:1010933404324
14. Electric Rule 2. PG&E. Advice Decision 1303-E. 1990. https://www.pge.com/tariffs/assets/pdf/tariffbook/ELEC_RULES_2.pdf
15. Takuya Akiba, Shotaro Sano, Toshihiko Yanase, Takeru Ohta,and Masanori Koyama. 2019. "Optuna: A Next-generation Hyperparameter Optimization Framework" in KDD
16. L. McInnes, J. Healy, S. Astels, hdbscan: Hierarchical density based clustering In: Journal of Open Source Software, The Open Journal, volume 2, number 11. 2017
17. McInnes et al., (2018). UMAP: Uniform Manifold Approximation and Projection. Journal of Open Source Software, 3(29), 861, https://doi.org/10.21105/joss.00861
18. S. Gupta and A. Tiwari, How to Generate Greater Value from Smart Meter Data. Cognizant, April 2021
19. R. Moghaddass and J. Wang, "A Hierarchical Framework for Smart Grid Anomaly Detection Using Large-Scale Smart Meter Data," in *IEEE Transactions on Smart Grid*, vol. 9, no. 6, pp. 5820-5830, Nov. 2018
20. Viegas, Joaquim & Vieira, Susana & Melicio, Rui & Matos, Hugo & Sousa, João. (2016). Prediction of events in the smart grid: Interruptions in distribution transformers.
21. Vafeiadis, Thanasis & Alexiadis, Anastasios & Dimaridou, Valia & Krinidis, Stelios & Kitsikoudis, Kostantinos & Makris, Lampros & Davidovic, Daniel & Ioannidis, Dimosthenis & Tzovaras, Dimitrios. (2019). Anomaly detection in smart meters.
22. Jiang, Jianfeng & Zhang, Yin & Chen, Jintao. (2019). Application of Fault Tree Analysis on Smart Meter Fault Diagnosis. Journal of Physics: Conference Series.
23. Hsieh, Eric (2020, Nov. 23). Climate Change, Machine Learning, and the Power Grid [Conference presentation]. https://www.youtube.com/watch?v=HfAZ0G96xY4
24. Uber Engineering H3: Hexagonal Hierarchical Spatial Index, Brodsky, et al. (2018)

12 Appendix

12.1 List of Figures and Tables

Figures

Tables

12.2 Definition of Terms and Abbreviations

AMI	Advanced Metering Infrastructure. This represents the fleet of Smart Meters and the mesh network that enables them to report back to a central database.
ATS	Applied Technology Services, a department at PG&E for laboratory for engineering analysis.
Distribution	The distribution portion of the electrical grid that facilitates the movement of lower voltage electricity. Majority of overhead conductors are part of the distribution grid.
Distribution Transformer	The transformer that steps voltage down in the final phase of the electric distribution from medium voltage to low voltage which is the voltage level provided to the customer.
DPD	Dynamic Protective Device, a device that opens when a fault is detected and recloses to attempt to re-establish service. Reclosing is designed to reduce or eliminate the effects of temporary faults
EC	Electric Corrective
EDGIS	Economic Development Geographic Information System
Feeder	Electrical line segments of the distribution grid.
GMI	Generic Meter Interface provides meter reads, data logs, and event logs of the meter data stored on Itron
GIS	Geospatial Information System
Grid	The electric grid
HFTD	High Fire Threat District. As defined by the CPUC in Decision 17-12-024[12], HFTDs are

[12] CPUC Decision 17-12-024

	regions where there is an elevated or extreme risk of destructive wildfire.
ILIS	Integrated Logging Information System
IOU	Investor Owned Utility
Line Regulator	A system used to maintain a constant, steady voltage level.
Low Voltage	0 - 600 Volts
LVR	Line voltage regulator
Medium Voltage	5,000 – 35,000 Volts
Meter	Device that collects electric and natural gas usage data from homes or businesses.
MW	Megawatt. 1,000,000 watts; MW is the standard unit of measure for describing feeder capacity.
MWh	Megawatt hour
NLP	Natural Language Processing
PG&E	Pacific Gas and Electric Corporation.
Power Transformer	Device which permits changing the high voltage needed for distribution down to the lower voltage for customer use. Overhead transformers are usually bolted to a wood pole and connected by overhead high voltage cables to individual customer service meters. Underground, transformers are usually green steel cabinets mounted on concrete pads and connected by underground high voltage cables to individual customer service meters.
Primary	High voltage or input side of a transformer. Includes the circuit that feeds into the transformer.
POMMS	PG&E Operational Mesoscale Modeling System
R&D	Research and Development
RUL	Remaining Useful Life
Sag	Decrease to between 0.1 and 0.9 pu in rms voltage or current at the power frequency for durations of 0.5 cycles to 1 minute.

SAP	System Applications and Products in Data Processing
SCADA	Supervisory Control and Data Acquisition. A software and communications system which enables capture and archiving of measurement data.
Secondary	The output side of a transformer and the circuit connected with it. Voltage delivered between 0 and 750 volts. Also referred to as service delivery voltage.
Service Point	Specific location at a premise where PG&E supplies service. (e.g., electric meter, gas meter)
SIQ	Sensor IQ, a project to update the firmware of a subset of our Smart Meters™ to enable additional and higher resolution measurements, along with customer event trap configurations.
SME	Subject Matter Expert
SSD	Source Side Device Source are protective devices that operate when an outage happens. Examples of these protective devices are circuit breakers, line reclosers, and fuses.
Swell	Increase to between 1.1 pu and 1.8 pu in rms voltage or current at the power frequency durations from 0.5 to 1 minute.
TD&D	Technology Demonstration and Deployment
Volt	Unit used to measure electrical potential of pressure.
WM	Work Management