EEA Conference & Exhibition 2018, 20 – 22 June, Auckland

Prioritising Predictive Maintenance Work Using Machine Learning

Author(s): Irvin Chew, Consultant to Transpower NZ LTD.
Mark Mitchell, Transpower NZ Ltd.
James Williams, Isogonal Ltd.
Alan Williams, Isogonal Ltd.

Presenter(s): Irvin Chew, Consultant to Transpower NZ LTD.
Mark Mitchell, Transpower NZ Ltd.

Abstract

Transpower's work order management system contains multiple descriptions entered manually by our service providers for equipment defects. These work order descriptions are rich in information, but they are unstructured and problematic for performing meaningful and reliable statistical analysis. With the large volume of open work orders (~60,000), manually processing and interpreting them is not practical. Recent advances in machine learning methods for extracting information from natural language have enabled us to automatically interpret work order descriptions, categorising them systematically and consistently.

In this paper, we will describe the process of manually entered textual data to estimate the likelihood of failure and the cost for undertaking each work order. This involves cleaning the free-text information, which may contain spelling mistakes, abbreviations, and a wide range of industry-specific vocabulary used to describe defects. We then constructed a set of hierarchical machine learning models to map work order descriptions to an asset and defect ontology created in parallel with the modelling process. The asset and defect ontologies describe the asset at risk (i.e. Circuit Breaker), the component at risk (i.e. Stand) and the risk keyword (i.e. Corroded). Once the ontologies were established, we assigned each asset and defect pair with a likelihood of failure for five different failure consequence dimensions – service performance, direct cost, public safety, worker safety, and environmental impact. We also used a machine learning process to predict the estimated cost for each work order. The likelihood of failure and failure consequence give a relative risk value for each work order, which along with the estimated cost was mapped to an overall work order priority. This enables risk-based prioritisation of work orders.

We also describe an innovative method for capturing organisation-wide experiential knowledge for use in training a machine learning system. This was used to aggregate subjective evaluations of the likelihood of failure for an asset/defect combination into a quantitative estimate.

**An ontology is a set of categories along with a description of their properties and the relationships between them, also referenced as "clusters".*

Introduction

The goal of the Predictive Maintenance (PDM) project described in this work is to determine a risk rating for each work order, which is then used for prioritisation of work orders. Work orders describe possible preventative maintenance actions and are typically generated by maintenance contractors.

Currently, our PDM work order priority is determined by service providers in collaboration with our Service Delivery Managers. This approach is subjective and lacks nation-wide consistency. Historically, there was also a lack of requisite knowledge/data related to the criticality (a.k.a. consequence of failure) caused by the defects on our network.

As per the Corporate Risk Matrix (refer to Figure 1), we introduce a two-dimensional risk-based approach to prioritising PDM work orders: by the likelihood of failure and by criticality.

Likelihood of Consequence	Insignificant	Negligible	Minor	Moderate	Major	Severe	Critical	Catastrophic
Almost Certain	HIGH	HIGH	VERY HIGH	VERY HIGH	EXTREME	EXTREME	EXTREME	EXTREME
Very Likely	MEDIUM	HIGH	HIGH	VERY HIGH	VERY HIGH	EXTREME	EXTREME	EXTREME
Likely	MEDIUM	MEDIUM	HIGH	HIGH	VERY HIGH	VERY HIGH	EXTREME	EXTREME
Possible	LOW	MEDIUM	MEDIUM	HIGH	HIGH	VERY HIGH	VERY HIGH	EXTREME
Unlikely	LOW	LOW	MEDIUM	MEDIUM	HIGH	HIGH	VERY HIGH	VERY HIGH
Very Unlikely	VERY LOW	LOW	LOW	MEDIUM	MEDIUM	HIGH	HIGH	VERY HIGH
Rare	VERY LOW	VERY LOW	LOW	LOW	MEDIUM	MEDIUM	HIGH	HIGH
Very Rare	VERY LOW	VERY LOW	VERY LOW	LOW	LOW	MEDIUM	MEDIUM	HIGH

Consequence Severity

Figure 1: Transpower's Corporate Risk Matrix

With the current relatively large volume of open work orders (~60,000), their manual processing and interpretation does not present a practical feasible option. While they contain structured data, most of the information is written into unstructured text fields, making them difficult to use for reliable statistical analysis. Recent advances in machine learning (ML) have enabled us to automate the interpretation of work order descriptions, performing this systematically and consistently.

Methodology

Overview

The risk rating for a work order requires two dimensions: Likelihood of Failure and Criticality. To generate quantitative values for likelihood and criticality for each work order, we first generated an asset ontology describing the asset at risk (i.e. Circuit Breaker) and the component at risk (i.e. Stand), and a defect (or risk) ontology describing the asset condition (i.e. Corroded).

Building this asset ontology was an iterative process, moving back and forth between text summarization algorithms and expert knowledge. We also developed a data-specific text normalization process, correcting spelling mistakes and standardizing a range of industry-specific abbreviations and vocabulary used to describe defects.

Each work order is related to an asset, and the asset's network location has a criticality associated with it. The criticality for an asset was developed outside of this workstream hence the calculations and details are not presented in this paper but available in [1]. In summary, criticality is the monetised cost following asset failure and the following dimensions were considered:

1. **Service Performance** – the lost load impact on the network when the asset fails
2. **Direct Cost** – the repair or replacement cost to reinstate the asset/network to its designed configuration following asset failure
3. **Public and Workplace Safety** – the impact on the safety of personnel or public that is near the asset following failure
4. **Environmental** – the impact on the environment when equipment fails causing a fire, oil leak or SF6 leak
5. **Compliance** – the financial implication following a regulation breach

We then generated a likelihood of failure for each work order, for five different failure consequence dimensions. The likelihood of failure for each dimension for each work order is derived by using a supervised machine learning system to categorise a work order into asset and defect ontologies, and to assign a likelihood of failure for each asset and risk combination. The risk rating is then calculated based on the likelihood of failure and its criticality.

The overarching process is described as per the following and in figure 2:

1. Training data is required as an input for supervised machine learning. Using the historical and open work orders, we defined an asset ontology and a defect ontology, and manually annotated the asset and risk categories for 10,000 historical work orders. This was an interactive process, using early models to target annotation to work orders not yet classified correctly.
2. Two hierarchical classification models were trained to determine the asset and risk categories of a worker order from the asset and risk ontologies. The models used 'cleaned' text from the work order descriptions and a handful of categorical features as input.
3. After training the models, all 60,000 open work orders and their descriptions were passed through the models to determine their asset and risk categories. In parallel, the models were also used on historical work order data. The historical work order data contains actual costs, and regression was used to determine the unit cost as a function of the asset category and location.
4. The asset and risk ontology combinations (e.g. Circuit Breaker Stand – Corroded) required engineering input to be converted into a Likelihood of Failure for each criticality dimension. First, the relative ranking of every pair of asset/risk categories was internally crowd-sourced, with engineers "voting" which asset ontology was higher, in terms of Likelihood of Failure, compared to other ontologies. Thousands of relative comparisons were then converted into an absolute value for each asset category/risk category combination. This was done for each criticality dimension as asset and risk combinations (i.e. Tower – Steel – Rusty) would have different likelihoods of failure depending on the criticality dimension (i.e. direct cost vs. service performance).
5. The risk rating is then calculated based on the likelihood of failure and its criticality.

Figure 2. Graphical depiction of process and methodology

Determining Asset and Defect Categories for a Work Order

Work Order Data

Work orders describe possible preventative maintenance actions and are typically generated by maintenance service providers. While they contain structured data, most of the information is written into free text fields. These free text fields are entered manually and are unstructured, making them difficult to use for reliable statistical analysis, but are the best source of information about the asset and defect described by the work order. This information is summarized in Table 1.

Field	Description
Work Order Identifier	Unique work order ID.
Summary Text Field	Data field for a summary of the work order. Entered by an engineer when the work order is created.
Description Text Field	Data field for a longer description of the work order. Can be changed and updated as the work order is completed.
System/Network Location	Location of the work order, which can be one of *Sub Device Position*, *Device Position*, or *Site*, depending on the level of detail.
Asset Category	The general category of the asset targeted by the work order, such as *Building*, *Tower*, or *Conductors & Accessories*.
Physical Location Fields	Used for cost estimation.

Table 1. Work order data. Only fields relevant to this work are shown.

Categorising Work Orders by Asset and Defect

Exploration of the text data using topic modelling showed clear groupings of assets and defects described in the work orders. Using these groups and expert knowledge, we created an asset ontology consisting of an *asset* and a *component*. We also created a *risk* or *defect* ontology describing the defect in the work order. In the rest of this work, we describe the asset and the component for a work order in full, and describe the risk or defect using a keyword identifier.

Work Order Description	Asset	Component	Risk Category	Risk Keyword
… STR … bolt … needs tightening	Tower	Bolt	Non-veg	Loose
… tower … bolts are corroded …	Tower	Bolt	Non-veg	Corroded
… corrosion on tower bolts …	Tower	Bolt	Non-veg	Corroded
CB insulator … mouldy	Circuit Breaker	Housing	Non-veg	Contaminated
Cracked insultor on circuit breaker	Circuit Breaker	Housing	Non-veg	Damaged
Conductor … spacer … corroded	Conductor	Spacer	Non-veg	Corroded
Spacer … missing	Conductor	Spacer	Non-veg	Missing
Conductor … trees … close to line	Conductor	Line	Veg	Fall Distance

Table 2. Example assignment of work orders to the asset and risk ontologies

We then built a machine learning system to determine the asset, component, and risk keyword categories for work orders, using the text descriptions as input. The machine learning system takes a set of training data (text descriptions for work orders and manually assigned categories) and uses this training data to build multiple models that can automatically generate categories for new work orders. The machine learning process is shown in Figure 3. The models are trained using training data and then deployed for inference on unseen work orders.

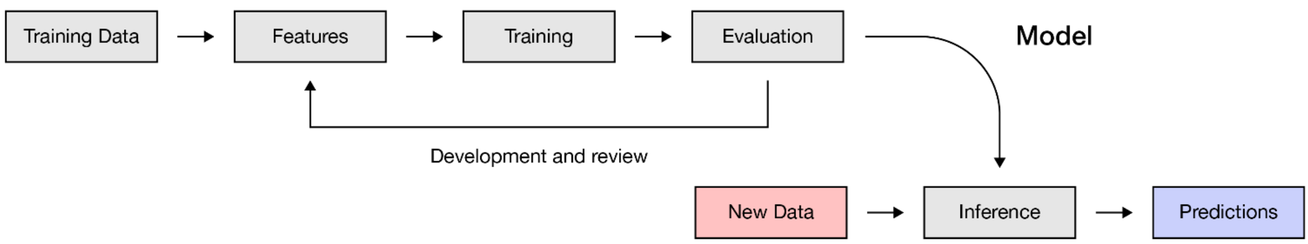

Figure 3. Machine Learning process

The free text fields contain many spelling mistakes and abbreviations, as well as inconsistent grammar. Therefore, the text was pre-processed prior to use in modelling. This pre-processing step included correcting spelling mistakes, replacing common acronyms, removing stop words and low frequency words, and lemmatization [2].

We then built a two hierarchies of classification models to predict the asset and component, and the defect for each work order. Each of the individual models used in this multi-stage approach embed the summary and description text fields in a vector space and use regularized logistic regression as the classifier.

The first defect model classified whether the work order was related to vegetation growth. If true, further models classified whether the work order was related to a risk keyword within the vegetation category, for example, sag distance, swing, or fall distance[1]. If false, further models classified whether the work was related to one of many other possible risk keywords. The second hierarchy of models classified which asset type and which component type was described by the work order.

Estimating Work Order Price

As well as the open work orders, the work order dataset also contains four years of completed work orders, containing additional information including completed date, total cost, material cost, and service cost.

We developed a regression model to estimate the costs for a work order based on a set of categorical features related to the asset and component category, the defect category, the service area, and region. Regional information is included due to the different rates different service providers charge. This model was trained on the completed work orders. The output from the regression model, predicted cost, is used alongside the risk rating to help guide decisions regarding budget allocation, as described below.

Estimating Likelihood of Failure

We then determined *likelihoods of failure* for each work order, should the work order be left untended. The likelihood is determined for five risk dimensions defined in the Transpower risk matrix: service performance, direct cost of replacement, public safety, worker safety, and environmental safety, and are based on expert judgement, as no training data exists to train a machine learning model.

The likelihoods of failure above are determined for each of the asset and defect combinations identified previously. As there are approximately 900 asset and defect combinations, asking individual engineers to assign each combination a likelihood, or rank all asset-defect combinations directly would likely lead to inconsistent outcomes as engineers are biased to their own area of responsibility.

To minimise this issue, we built a web application for collecting crowdsourced comparison tool to collect pairwise comparisons of randomly selected defects from many engineers, and used an algorithm based on the ELO rating system [0] to compute an overall ranking. The web application allowed the collection of multiple comparisons between each pair of asset and defect combinations, from dozens of engineers, as well as collecting metrics determining the time spent on each comparison.

The integer rank for each asset and defect combination was then mapped to a scalar interval for each risk dimension, representing likelihood of a failure in that risk dimension. The bounds of this scalar range were chosen manually to be within the corporate scale of 0.003 to 3 outcomes per/year.

[1] These terms relate to trees within the spacial limits of or the potential to be in physical contact with transmission lines as specified in vegetation regulations

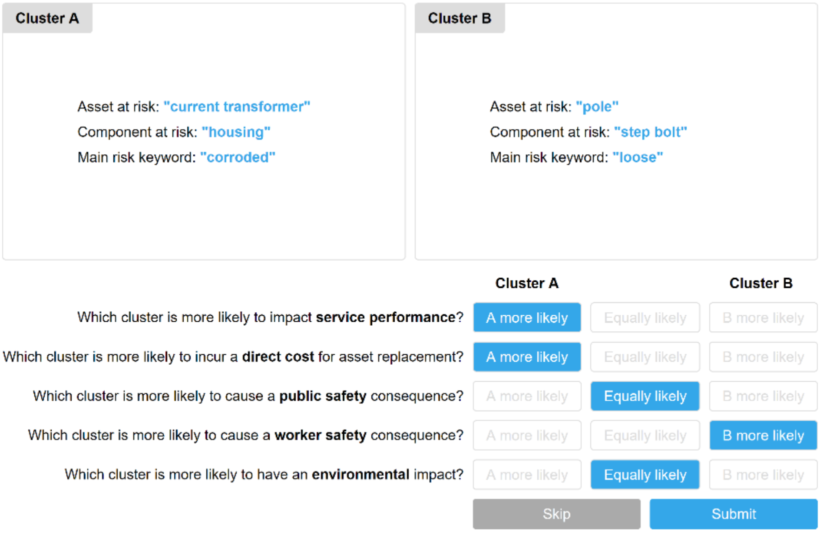

Figure 4. Cluster comparison tool web UI.

Assigning Failure Severity

Criticality Data

Asset criticality is the monetised cost of an asset failure across each of the five dimensions mentioned previously. Transpower has an existing criticality value mapped to a network location identifier. This criticality data was matched to the work order data through these network location identifiers. Work orders without any network location identifiers were assigned based on asset category or special cases such as bird related criticality or compliance criticality. Any other work orders with no matches are not assigned a criticality, we are working on reducing the number of these occurrences.

After all steps briefly described above, each currently open work order now has a criticality value, a likelihood value, and an estimated price.

Prioritisation

Work orders are prioritised according to a final priority score or risk rating, which is the “sum of likelihood x severity over risk dimensions”. We computed the final work order priority or risk rating based on the following formula for each risk dimension

Sum (all risk dimensions)
rescaled likelihood (on the scale of 0.003 to 3 rate of outcomes per/year) * log(1 + criticality) >
<typeset as formula>

The risk rating obtained was scaled to the range of 0 to 1000 to give a consistent and interpretable number that can be used to make decisions regarding work order scheduling based on risk rating / priority. This rescaling is relative to the maximum asset criticality, which is unlikely to change.

Results and Discussion

The capability to conduct rapid proof of concept and iteration on early prototype systems that accomplished the end to end process with limited accuracy and precision (days) was useful for achieving buy-in. Reaching the end of the prototyping and proof of concept phase required overall 3 months.

The asset and defect ontologies were developed as a way of grouping related work orders together to manually assign likelihood values. These ontologies were developed for this work, and the outcome was to categorise both the assets and the defects into two-level hierarchies. This size of the groups created affects the accuracy and precision of the supervised machine learning classification and the amount of manual work required to set up the system.

The defect ontologies were created in an iterative process, involving data exploration, ontology definition, modelling, examining model output, and repeating this process. Data exploration techniques included topic modelling such as latent Dirichlet allocation, clustering using distance metrics such as bag-of-words based BM-25 and cosine similarity using word embedding vectors. There was also some difficulty in assessing stations assets as each type has a wide variety of components compared to transmission line assets.

Using a machine learning system in practice involves creating a dependency on training data that is not present in a standard software system. Managing this data quality during the creation of a new system and as the system is used over time, potentially many years, is a crucial part of the overall project plan. Maintaining a training data audit trail was an important part of the model development process, both for understanding the biases and limitations in the model. During deployment of the model, a documented process was created allowing engineers to retire old training data, create new training data reflecting ongoing changes in the operational environment, and retrain the model using these new data sets.

Implementation

A software system based on the work described is currently being deployed within Transpower, with provisions for ongoing auditing of training data, retraining based on updated data as the underlying information shifts over time. Service providers are also being guided to provide accurate and more informative work order descriptions, giving the machine learning system more data to turn into more accurate category, cost, and risk predictions.

Future work will also include exploring expanding the defect ontology to include sub-components for substation assets and fine tuning it to infer different levels of corrosion.

Conclusions

We have demonstrated the utility of a machine learning approach for automatically assigning risk and cost to open work orders based on unstructured text, using a training dataset comprised of historical data and targeted data collection. As the information in new work orders changes over time new training data can be created, and old training data retired, and the model re-trained.

We have also described a method for collecting industry and model-specific training data using a crowd-sourcing web application. The data generated from this application was combined with a tournament ranking algorithm to generate an absolute likelihood value for each node in the asset ontology.

References

[1] Q. Boucher and S. Horn, “Asset Prioritisation Process - The Health & Criticality Journey Continues,” in *EEA Conference*, 2015.

[2] N. Indurkhya and F. J. Damerau, Handbook of Natural Language Processing, Second ed., New York: Chapman & Hall/CRC Press, 2010.

Biography & Contact Details

James Williams, PhD
CEO, Isogonal
+64 21 0837 4655
james.williams@isogonal.co.nz

James is a Director at Isogonal, a company building real world applications of AI technology. He also lectures at the University of Canterbury, and co-directs the Data Science programme at the university. He founded Isogonal in 2015 to pursue real world applications of AI technology, and is passionate about problem solving and machine learning. James holds a PhD in Applied Mathematics from Yale University, a BSc(Hons) in Mathematics from the University of Canterbury, and was a recipient of the Fulbright Science and Innovation Graduate Student Award.

Alan Williams, BEng (Hons), MSc
CTO, Isogonal
+64 21 672 526
alan.williams@isogonal.co.nz

Before embarking on a software engineering career, Alan worked on petroleum, geotechnical, and structural engineering projects around the world. In 2017, Alan co-founded Isogonal, producing machine learning systems for industrial asset operation and healthcare delivery. Alan has held CPEng registration and is a keen athlete and amateur Go player. He received a BE(Hons, first class) and an MSc(Distinction) from Canterbury University.

Mark Mitchell,
Transpower NZ LTD
+64 21 2427213
mark.mitchell@transpower.co.nz

Before leading this innovative project team, Mark has spent his career dedicated to the electrical industry spanning domestic and industrial installations to distribution and transmission electrical networks.