FUTURE READY™

Case Study

HUNTER WATER MACHINE LEARNING

Project Overview

Hunter Water – a state owned corporation that provides drinking water, wastewater, recycled water, and some stormwater services to almost 600,000 people in homes and businesses across the Lower Hunter in NSW - engaged us to apply machine learning to improve decision making.

Machine Learning was used to review their historical data and see which pipes in their network were at risk of leaking. Traditionally, leaks are found by individuals or businesses in the community, who report them when water can be visibly seen on the service.

The Hunter Water's leak detection program aims to discover hidden or developing leaks through a survey schedule based on an analysis of leakage at a suburb level. As they have a vast storage of historical data, an opportunity arose to use machine learning to improve the efficiency of the programme in detecting leaks by surveying pipes that are higher risk of leaking more frequently. We improved this process by using a Machine Learning model to increase the probability of finding leaks earlier and reduce water losses and risk of pipe breakage.

What Future Trends Did We Consider?

Technology

Data, Transparency
& Cyber Security

How Did We Consider These Trends?

Technology

Data, Transparency & Cyber Security – There is an increasing use of digital tools and technologies to manage infrastructure more efficiency.

Using pre-existing accumulated data from Hunter Water allowed us to develop an accurate Machine Learning model to predict leaks. We also explored the effects of acute weather events and soil moisture on the maintenance and life cycle of pipes.

By combining 20 years of report leak events and network inventory data to define a leak case/control classification data set. After pre-processing the existing data, we generated a balanced case/control dataset using a sampling engine. This sampling engine creates cases by joining the events and network pipes data and creates controls by assigning a random date where a pipe has not yet had a reported leak. These datasets were then split into training and validation – this ensures that the inputs used in the validation phase was never seen by the model before.

How Was Our Approach Better?

Machine Learning is common in some industries including financial services but is less common in the management of infrastructure and in this case, utilities.

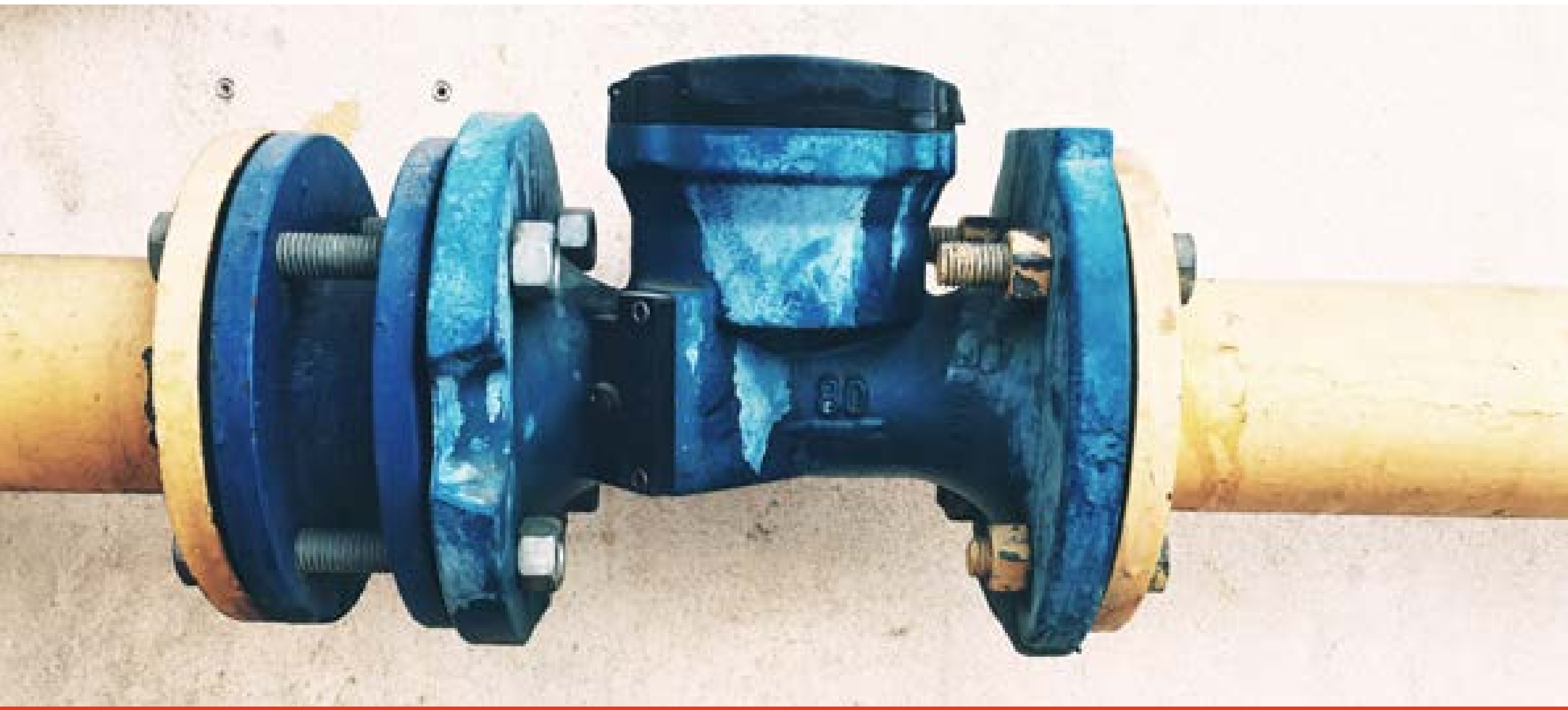
We developed two models: Logistic Regression and XGBoost - a decision-tree forest model. Both models were evaluated on a disjoint set of pipes and events that were not seen during the model training.

A prospective study was conducted with the XGBoost model. It predicted pipe leak risks for all pipes immediately prior to an exhaustive field survey of 41 suburbs and confirmed the model's ability to predict field-survey leaks with this new data.

The Outcomes

The field survey across the network uncovered a greater rate of leaks found per kilometre when using the Machine Learning model to guide the fieldwork which made the maintenance planning process more proactive. It also found the pressure and pipe age were the two most important variables when predicting leaks compared to pipe material and location.

The model allowed for additional data to be added to see the effect on the failure rate of pipes such as weather events and soil moisture. This has also been expanded into the exploration of how long a pipe survives until it's not viable to repair.



For More Information

Catherine Yu
Data Scientist / Rail Engineer

+61 2 9272 5180
Catherine.Yu@wsp.com

Find Us

WSP Australia Pty Limited
Level 27, 680 George Street
Sydney, NSW 2000