MACHINE LEARNING for DISASTER RISK MANAGEMENT

Using data to protect people across the globe

WHAT IS A MACHINE LEARNING ALGORITHM?

A machine learning (ML) algorithm is a type of computer program that learns to perform specific tasks based on various data inputs or rules provided by a human.

They can perform these tasks in two ways:

Supervised

A human user inputs a "labelled" training dataset that tells the computer which answer is correct and which is incorrect.

Think of it like flash cards. If you show someone a picture of a cat and one of a dog, and explain which is which, they can use that information to sort other pictures of cats and dogs into respective groups.

Unsupervised

Computers use pre-existing statistical methods to attempt to group data with similar characteristics together.

For example, consumer demographics and purchasing habits could be used as a dataset for an advertising program attempting to group the U.S. population into smaller, more specific markets.

ML IN THE REAL WORLD

Since disasters often affect poor and vulnerable areas most significantly, it's imperative to use the technology we have to protect those areas. For example:

Guatemala: Earthquake Vulnerability

In 1976, an earthquake decimated the Guatemalan town of Los Amates, causing:

23,000 deaths

Economic damage estimated at nearly 20% of the country's GDP

76,000 injuries

Hurricanes, volcanic activity, landslides, and other disasters have continued to hit Guatemala in the years since, raising questions about how the country can better prepare for hazards.

THE CHALLENGE

In areas of high seismic activity, identifying high-risk buildings can help prioritize retrofitting investments and, most importantly, save lives.

However, sending large teams of surveyors into the field is time-consuming and expensive.

THE SOLUTION

The World Bank used imagery from satellites, drones, and 360° street-view cameras to identify homes that were at high risk for collapse during an earthquake. Using imagery was about 70% cheaper than relying only on people for data collection and prevented human bias in data collection.

The algorithm identified high-risk buildings based on . . .

Slope of land

Large first-floor openings

Rooftop material

In the end, World Bank's method resulted in . . .

85% of buildings that field engineers deemed at-risk flagged.

Helping to identify modest homes that are prone to collapse and buildings that are good candidates for retrofitting.

Clearly, using data wisely can have tremendous results when it comes to disaster risk management. But—believe it or not—sometimes, numbers do lie.

PREVENTING BIAS

Since all sets of data are approximate representations of the real world, no model is free from bias. Whether due to a model from a data-poor area, selecting improper criteria, or another cause, bias can creep into projects with even the best intentions.

Steps can be taken to prevent bias:

1. Assess which preconceptions could present themselves in data
2. Ensure all data are directly relevant to the desired results
3. If necessary, collect additional data to build a larger, less biased sample

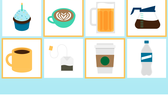
OPEN DATA AND THE FUTURE OF ML

Over the past decade, machine learning methods have seen enormous growth. This is largely due to the openness of information in the field—knowledge and information from teams around the world stack on top of one another to speed up innovation for the benefit of all. This is particularly essential in the DRM field.

The power of the crowd has also led to significant advancements. This strategy involves manual data mapping by networks of individuals, whether in a focused physical area or across the world. When humans provide much-needed training to ML algorithms, the results can be astounding.

If you've ever filled out a re-CAPTCHA, such as the one below, you've been a part of a crowdsourcing campaign—Google uses them to train image-recognition algorithms.

Select all squares that contain ***coffee***



Together, humans and computers have come together to make tremendous progress in the disaster risk management field, and as technology and datasets continue to grow, so will our results.

Read our full report on machine learning for disaster risk management and learn more at:

bit.ly/mlfordrm