NEW AMERICA

esri THE SCIENCE OF WHERE

FUTURE OF PROPERTY RIGHTS

Machine Learning and Property Rights

An Introduction to Machine Learning

Machine learning is an application of artificial intelligence (AI) that enables systems to programmatically "learn" and improve from past experience. Computers use algorithms and statistical models to "learn" patterns and insights from sample sets of data (often called "training data sets"), and apply those insights to make intelligent predictions and decisions about much larger sets of data.

Applications of machine learning have proliferated over the last several years. For example: GPS navigation apps like Google Maps and Waze use machine learning to better predict traffic patterns, and social media services like Facebook and LinkedIn use machine learning to tailor the content of a user's news feed.

In the realm of property rights, the most commonly used data sets include imagery from satellites, low flying aircraft, drones, and even cameras on cars. Here's how a typical process of developing an imagery-based machine learning model works:

1. First, a computer ingests a collection of sample images.
2. Then, each image is manually classified into a category (for example, vegetation, water, or soil) or broken into pixels, and then manually classified. The computer then attempts to predict the classes that are manually assigned to each image through an iterative process of generating a "prediction class" and then checking it against the manual "label." If the prediction and label match, the computer learns the pattern; if not, the computer tries to change its prediction to match the label. This iterative process continues until the computer reaches an acceptable accuracy in predicting the classes. This process is called "training."
3. Next, a new set of images with labels is used to validate the model accuracy. These images are similar to the ones used in the training phase, but the model has not been exposed to them before. This process enables the developer to assess the model's ability to extend beyond its training data. If the model fails on the new data, the training phase is repeated. The output of this phase is a trained model.
4. Finally, the trained model is used to predict classes or to segment the input to different classes using new imagery and scale the prediction to larger areas.

Why Machine Learning is Important for Property Rights

By automating multiple components of the property mapping, documentation, and transaction process, machine learning can vastly increase the scale and speed of property rights delivery, resource management, and land use planning. Machine learning replicates existing knowledge at scale, driving down the cost and time associated with labor intensive tasks like surveying, filing biographical information, and conducting background and financial checks.

For example: machine learning promises to lower the cost and time associated with mapping by predicting the boundaries of land parcels based on common property boundary features (for example, a lack of vegetation, the existence of a fence or a path, etc.) detected from a training set. Until recently, this sort of machine learning application was impossible; however, the recent proliferation of high-resolution satellite imagery puts it within reach. This capability has already been applied in slum mapping.

As another example: machine learning has been used to assist with property valuation, both in established markets and in thin real estate markets where comparable sales data is hard to come by. Not only can machine learning be used to predict a property's value, but it can predict market demand based on the type of property.

Machine learning can help automate registration processes by using natural language processing to scan documents for key terms or identify red flags. And within the land use and zoning use case, machine learning can use Google Street View and other street-level imagery sets to map gentrification and identify vulnerable housing.

The Strengths and Limitations of Machine Learning

Similar to other emerging technologies, machine learning possesses both strengths and limitations for property rights:

Strengths

Perhaps the biggest strength of machine learning is its potential for scale. By automating otherwise labor-intensive processes, machine learning can save time and money, and achieve new levels of efficiency throughout the property rights cycle.

Not only that—machine learning has the potential to take a certain degree of human subjectivity and error out of mapping and valuation efforts.

Limitations

Accuracy: While machine learning-based parcel mapping is a big step forward, its outputs are not yet 100 percent conclusive. As a result, the parcel maps must be thoroughly vetted by the parcel occupants to which they pertain.

Acceptance: Because it is abstract and remote, the process of machine learning-based mapping may be hard for both communities and governments to accept. Practitioners will likely need to do significant up-front work to socialize machine learning concepts prior to implementation.

Drawbacks of being remote: While it is possible to establish a parcel's location and attributes from the sky, it is more difficult to remotely link that parcel to the parcel holder, and to document the relationship that exists between the parcel and the occupant (e.g., ownership, rental, etc.). Therefore, machine learning is limited in the amount of data it is able to provide about a property right. For purposes of property registration, it must be paired with other methods.

Bias: Because machine learning models learn from training data, and their predictions are only as good as the quality of the training data, it is possible to introduce bias into machine learning-based predictions, if the original data set is not carefully selected.

Use Cases of Machine Learning for Property Rights

Machine learning is starting to be deployed for a variety of projects worldwide. Below are a few use cases:

- **Slum mapping:** A Duke University study released in 2018 used machine learning to map slums in Bangaluru, India. The study revealed the existence of 2,000 slums, almost four times the number officially recognized.
- **Property valuation:** Chilean researchers have used machine learning algorithms to predict housing values in Santiago, Chile. GIS company Esri has used forest-based classification and regression models to predict home values in California.
- **Agricultural parcel identification:** Researchers at Clark University are currently using a combination of machine learning algorithms and crowdsourced verification in an effort to map every agricultural parcel in Ghana in a seven-day period. The research builds on prior successful uses of machine learning to map agricultural parcels in South Africa, demonstrating the potential of machine learning for mapping heterogeneous land parcels.