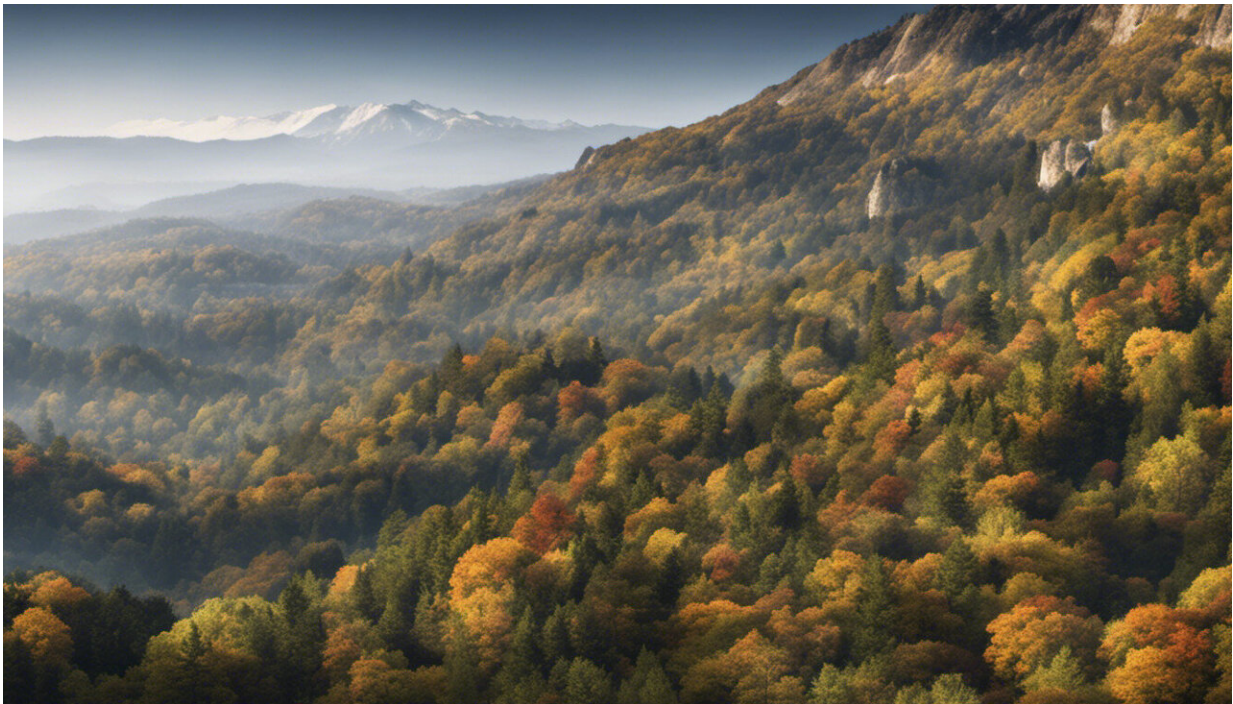# Researchers build a model that predicts business closures in cities with 80% accuracy

October 9 2018, by Krittika D'silva

Credit: AI-generated image ([disclaimer](#))

Over the past decade, changes in the way people shop have led more and more businesses to close their doors, from small music venues to book shops and even major department stores. This trend has been attributed to several factors, including a shift towards online shopping and changing spending preferences. But business closures are complex, and

often due to many intertwined factors.

To better understand and account for some of these factors, my colleagues at the University of Cambridge and Singapore Management University and I built a machine learning model, which predicted shop closures in ten cities around the world with 80% accuracy.

Our research modelled how people move through urban areas, to predict whether a given business will close down. This research could help city authorities and business owners to make better decisions, for example about licensing agreements and opening hours.

Pattern spotting

Machine learning is a powerful tool which can automatically identify patterns in data. A machine learning model uses those patterns to tests hypotheses and make predictions. Social media provides a rich source of data to examine the patterns of its users through their posts, interactions and movements. The detail in these datasets can help researchers to build robust models, with a complex understanding of user trends.

Using data about consumer demand and transport, along with ground-truth data on whether businesses actually closed, we devised metrics which our machine learning model used to identify patterns. We then analysed how well this model predicted whether a business would close, given only metrics about that business and the area it was in.

Our first dataset was from Foursquare, a location recommendation platform, which included check-in details of anonymous users and represented the demand for businesses over time. We also used data from taxis trajectories, which gave us the pickup and drop-off points of thousands of anonymous users; these represented dynamics of how people move between different areas of a city. We used historic data

from 2011 to 2013.

We looked at a few different metrics. The neighbourhood profile took into account the area surrounding a business, such as the different kinds of businesses also operating, as well as competition. Customer visit patterns represented how popular a business was at any given time of day, compared with its local competitors. And business attributes defined basic properties such as the price bracket and type of business.

These three metrics enabled us to model how closure predictions differ between new and established venues, how the predictions varied across cities and which metrics were the most significant predictors of closure. We were able to predict the closure of established businesses more accurately, which suggested that new businesses can face closure from a bigger variety of causes.

Making predictions

We found that different metrics were useful for predicting closures in different cities. But across the ten cities in our experiment – including Chicago, London, New York, Singapore, Helsinki, Jakarta, Los Angeles, Paris, San Fransciso and Tokyo – we saw that three factors were almost always significant predictors of a business's closure.

The first important factor was the range of time during which a business was popular. We found that businesses which cater to only specific customer segments – for example, a café popular with office workers at lunchtime – are more likely to close. It also mattered when a business was popular, compared with its competitors in the neighbourhood. Businesses that were popular outside of the typical hours of other businesses in the area tended to survive longer.

We also found that when the diversity of businesses declined, the

likelihood of closure increased. So businesses located in neighbourhoods with a more diverse mix of businesses tended to survive longer.

Of course, like any dataset, the information we used from Foursquare and taxis is biased in some ways, as the users may be skewed towards certain demographics or check in to some types of businesses more than others. But by using two datasets which target different kinds of users, we hoped to mitigate those biases. And the consistency of our analysis across multiple cities gave us confidence in our results.

We hope that this novel approach to predicting business closures with highly detailed datasets will help reveal new insights about how consumers move around cities, and inform the decisions of business owners, local authorities and urban planners right around the world.

This article is republished from The Conversation under a Creative Commons license. Read the original article.

Provided by The Conversation