Machine learning can help better predict city gentrification

January 30 2023, by Nadia Razzhigaeva

Credit: Unsplash/CC0 Public Domain

A machine learning model developed and tested by researchers at UNSW City Futures Research Center may be able to better equip policymakers with knowledge and data about predicted neighborhood change—and with better predictive data, policymakers and government

can deliver more equitable city planning and outcomes.

The researchers recently published their initial findings about the model in journal *Cities*.

Gentrification is an urban phenomenon affecting many cities around the world—it's a type of neighborhood change that sees poor or historically working-class neighborhoods experience drastic changes to demographic, land-use and housing affordability.

Gentrification can displace and negatively impact residents such as low-skilled workers and vulnerable people, and governments and policymakers often struggle to address the associated harms.

"Gentrification is often identified when it's too late, and it can be costly to address the harms it has caused," says William Thackway, postgraduate researcher at UNSW City Futures Research Center. "The ability of policymakers to adequately tackle harms caused by gentrification rests on proactive strategies which prevent or mitigate displacement of vulnerable people before it becomes too expensive to do so."

Gentrification in Sydney

Mr. Thackway, Professor Christopher Petitt, Dr. Matthew Ng and Associate Professor Chyi Lin Lee developed the prototype machine learning model and tested various data from Sydney as a case study.

"A key finding from our work is that the gentrification frontier is predicted to expand outwards even further from the city center," says Mr. Thackway. "Previously, the rings of gentrification were in 5–10 kilometer rings around the Sydney CBD, but that is predicted to expand to 10–20 kilometers."

The study identified an effect the researchers call "spill over" as an across-the-board indicator of predicted gentrification in Sydney. "Spill over" is when displaced residents from gentrification hotspots move to neighboring suburbs where rents are slightly cheaper.

"In the early 2000s–2010s, the Inner West became a major hotspot for gentrification, but our analysis predicts deeper Western suburbs like Auburn and Bankstown could be new hotspots," says Mr. Thackway.

Eastwood beyond Ryde and Brookvale beyond Manly were other suburbs experiencing "spill over" effects and predicted to gentrify according to this machine learning analysis. The analysis also reaffirmed Botany and Maroubra beyond Randwick as gentrification hotspots.

Besides spill over effects, other potential drivers of gentrification and neighborhood change in Sydney included changes in crime levels, new transport developments and further changes in demographics.

Surprising indicators of gentrification

The advantage of this new machine learning model is that it can make links between variables that are otherwise overlooked in other methods of analysis involving just human expertise.

"Our study includes a wider range of predictor variables than previous machine learning studies, spanning socioeconomic, housing, business and Airbnb data," says Mr. Thackway.

The machine learning model was trained and tuned using over 80 predictor variables from a wide range of data inputs such as property reports, the census, business registry and Airbnb.
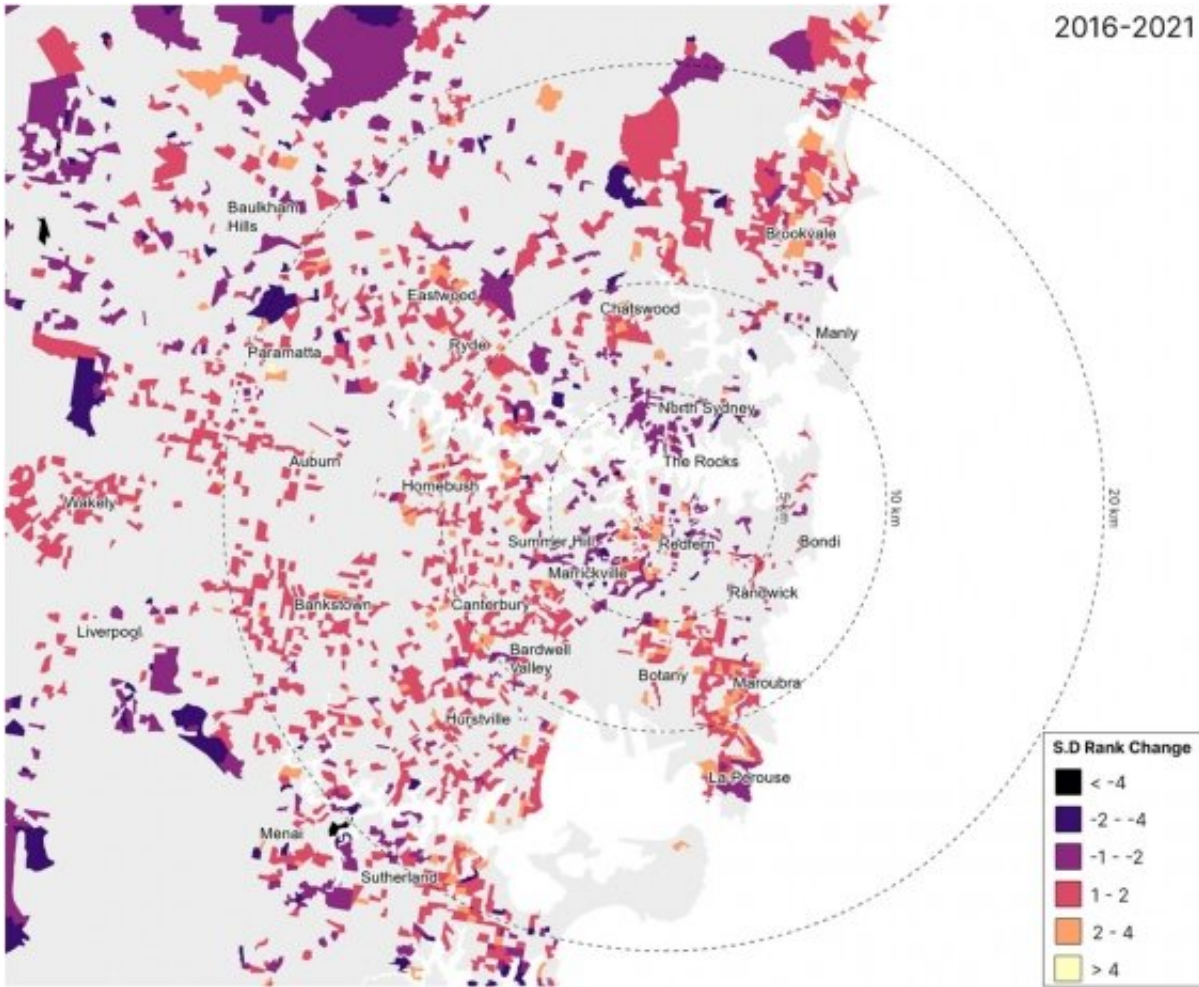
Map of predicted neighbourhood change in Sydney between 2016-2021. Areas in orange/red predicted to gentrify, areas in purple/blue predicted to decline in socioeconomic rank. Credit: Thackway et al. (2023)

To test its accuracy, the researchers retroactively applied the model to previously-ungentrified neighborhoods that ended up becoming gentrified.

Family compositions and relationship status were surprisingly important indicators of gentrification in some areas of Sydney, says Mr. Thackway.

"It was surprising to see that an increase in married couples in an area

lead to a higher prediction that the area will gentrify, while areas with more divorcees and one-parent families were less likely to gentrify according to our model."

In some cases, family and relationships were as important as house prices, education and employment in predicting gentrification for a suburb.

Toward better quantitative methods

Predictive modeling and machine learning tools in the urban policy spheres are still in their infancy.

"There is still skepticism among policymakers about the trustworthiness of such models," says Mr. Thackway. "Previous machine learning models have had a 'black box' element to them, meaning that we can't see how machine came to its conclusions. Because of this, the preference among policymakers is dominated by qualitative methods."

But this new machine learning model developed by UNSW researchers can predict gentrification with 87.3 percent accuracy and it eliminates the "black box" element by implementing a model explanation tool that interprets how the machine learning model came to its conclusions.

"Qualitative methods like the Neighborhood Change Warning System and Gentrification Index are easy for policymakers to understand," says Mr. Thackway. "But the downside is that they are quite simple and lack robustness.

"Our machine learning model incorporates tens, if not hundreds of indicators compared to qualitative methods. The advantage of using machine learning as opposed to basic indicators in qualitative methods is the model can identify interactions and relationships between variables

that one might not necessarily be able to do just from human expertise."

Overall, the UNSW team created a more holistic, robust, and explanatory machine learning model that improves best practice for predicting future gentrification hotspots.

Opportunity for future use

The tool is in its development phase and there is scope to test it to more extreme degrees to ensure its performance.

"Right now, the major implication of our work is that this model can produce meaningful and powerful results that will enable proactive policy decisions and interventions for urban planners," says Mr. Thackway.

"While we used Sydney as a case study to test the model, it can be applied to comparable cities by inputting similar data.

"In the past, most gentrification research has looked at what's already happened to analyze drivers of gentrification. This machine learning model enables predictive modeling of gentrification.

"With the 2021 census data release upcoming, forecasting vulnerable areas for 2026 will provide policymakers with an empirical tool to proactively intervene and design more equitable solutions for vulnerable communities."