GENERATIVE DESIGN FOR PRECISION GEO-INTERVENTIONS

R. Wen [1], S. Li [1]

[1] Department of Civil Engineering, Faculty of Engineering and Architectural Science, Toronto Metropolitan University, 350 Victoria Street, Toronto, Canada - (rwen, snli)@ryerson.ca

KEY WORDS: Interventions, GIS, Decision Support, Generative Design, Machine Learning, AutoML, Geodesign, Hyperparameter Optimization

1. RESEARCH OBJECTIVE

The purpose of this research is to develop a Spatial Decision Support System (SDSS) that integrates Geographic Information Systems (GIS), Automated Machine Learning (AutoML), and Hyperparameter Optimization (HPO) to generate precision geo-interventions based on standardized geospatial data and user design constraints.

2. RELEVANCE

Geo-interventions, actions implemented in geographic space that alter specific outcomes (e.g., safe road design for reducing traffic collisions and hotspot policing for reducing crime), are an effective solution to reducing a large portion of injuries from traffic collisions and violent crimes, which typically occur in large urban settings. Common approaches (e.g., cluster mapping, cellular automata, and multiple criteria decision analysis) to modelling and analysing geo-interventions have focused on identifying target areas/risk factors, and simulating scenarios/impacts/theories to support decision-making. However, these approaches usually model/analyse a range of existing/pre-defined geo-interventions, as opposed to generating or exploring potentially new geo-interventions based on data and user design constraints (Mehaffy, 2008).

In addition, common approaches to modelling/analysing geo-interventions rely heavily on domain expertise (e.g., selecting/interpreting models/variables, data processing, model assumptions) and evaluate only a small number of alternative geo-interventions (typically ranging from perhaps 3 to 12 alternatives). With recent advancements in large-scale computing and data availability in urban settings, there is huge potential to explore hundreds to thousands of alternative urban geo-interventions (Li et al., 2016). This reduces the heavy reliance on domain expertise by exploring a larger space of alternatives with computing power and big urban data, which leads to substantially more comprehensive experiments and impact evaluations.

3. PROPOSED SOLUTION

AutoML has had great success in using large-scale computing and big data to automatically pre-process data, select important variables, and discover/compare accurate models across large search spaces (He et al., 2021). Hyperparameter Optimization (HPO) has also been effective at improving model performance in AutoML approaches through the optimization of model parameters given constraints such as time, parameter ranges, or desired performance criteria (Feurer and Hutter, 2019).

This research proposes a SDSS that integrates AutoML and HPO with GIS to leverage modern advancements in computing power and big data availability. Spatial binning, a GIS technique, is first used to standardize and aggregate the geospatial data into polygonal bins (e.g., cells and hex-grids). AutoML is then used to automatically pre-process and generate geo-intervention models based on existing geospatial data. HPO is finally used to optimize the most performant models from the AutoML process under user design constraints (e.g., applicable intervention areas, budget/resource constraints, desired impact). In the HPO process, model inputs represent the potential geo-interventions (e.g., road width, number of traffic speed cameras/schools/police stations), while the outputs represent the predicted impacts from the geointerventions (e.g., change in traffic collisions/stabbings/gun violence). By optimizing the inputs to the AutoML models, the HPO process explores and generates hundreds to thousands of possible geo-interventions based on the model inputs and outputs automatically. These potential geo-interventions are precise – locatable to each grid cell based on the spatial binning resolution (e.g., grid size), and quantitatively measured as a change in specific model inputs for each cell, which can be visualized as GIS map layers.

4. EXPERIMENTAL EVALUATION

An initial experiment was conducted to evaluate the proposed solution. Open geospatial data from the City of Toronto was spatially binned into a 40 by 40 cell grid to produce 1593 aggregated variables and 1302 cells with data values. The Python packages *auto-sklearn* and *tpot* were used to generate AutoML models for predicting the number of traffic collisions in Toronto based on open Toronto Police Service data from 2006 to 2021. These models had an average cross-validated R^2 between 0.8 and 0.85. Future work includes applying HPO techniques (Bayesian optimization, genetic algorithms, etc) to these models to generate alternative geo-interventions, and developing a web-based GIS visualization interface to automatically produce AutoML models for exploring the generated alternative geo-interventions.

REFERENCES

Feurer, M., Hutter, F., 2019: Hyperparameter optimization. *Automated machine learning*. Springer, Cham, 3-33.

He, X., Zhao, K., Chu, X, 2021: AutoML: A survey of the state-of-the-art. *Knowledge-Based Systems*, 212, 106622.

Li, S., Dragicevic, S., Castro, F. A., Sester, M., Winter, S., Coltekin, A., ..., Cheng, T, 2016: Geospatial big data handling theory and methods: A review and research challenges. *ISPRS journal of Photogrammetry and Remote Sensing*, 115, 119-133.

Mehaffy, M. W., 2008: Generative methods in urban design: a progress assessment. *Journal of Urbanism*, 1(1), 57-75.