



A proposed model using AI to reduce randomness in the distribution of service and commercial facilities in residential neighborhoods

Prepared By

Name: Majed Humud Alhumuda

ID No.:253002679

Course Title: Research Methodology in Computer Science

Course Code: AI800

Assignment 1:

Weighting: 20%

Lecturer: Dr. Mohammed Anbar

Hand-out Date: 20 Jun 2026

Hand-in Date: 04 July 2026

Hand-in Method: Online Submission via E-Learning.

Assignment 2:

Weighting: 30%

Lecturer: Dr. Mohammed Anbar

Hand-out Date: 27 Jun 2026

Hand-in Date: 11 July 2026

Hand-in Method: Online Submission via E-Learning.

Table of Contents

Problem Statement:	4
Research Objectives:	5
Summary research paper's:	6
Synthesis Matrix for Literature Review:	8
Critical Analysis:	9
Connections:	9
References:	10

Problem Statement:

With a growing economy, continuous development, and population expansion, the presence of service and commercial facilities within neighborhoods remains of paramount importance. The availability of such facilities and services enhances the quality of life and fulfills essential human needs. Consequently, the establishment of these commercial services and facilities must be driven by precise data that ensures their accurate distribution. However, it has been observed that a significant gap in services and facilities exists among various neighborhoods within the same global cities, posing a substantial challenge to the quality of life for individuals, who are the cornerstone of global renaissance and development.

This challenge is highly critical; hence, the licensing of facilities and services should only be permitted based on the current situation as well as future forecasting driven by accurate data. This includes relying on spatial and temporal data, as demonstrated by (ISPRS Int. J. Geo-Inf., 2025), and utilizing logistic regression and Cellular Automata-based modeling, as illustrated by (T. Munshi, M. Zuidgeest, M. Brussel, and M. van Maarseveen, 2014).

Furthermore, another challenge lies in neglecting the actual, real-time, and instantaneous behavior of the population, alongside relying on linear distances rather than actual distances, as pointed out by (Xinxin Zhou, Yuan Ding, Changbin Wu, Jing Huang, and Chendi Hu, 2019), as well as the lack of understanding of satellite imagery, as highlighted by (Yaxin Li, 2024). Previous research has focused on certain algorithms that do not cover all the requirements necessary to make a correct and accurate decision. Although researchers have employed various models—such as GraphSAGE, Word2Vec, several other algorithms, and Geographic Information Systems (GIS)—these tools remain insufficient for effective decision-making. Leaving this unaddressed will inevitably lead to numerous gaps in individuals' quality of life and cause spatial inflation of services.

Today, the world relies on various electronic services that store vast amounts of data, which can significantly improve decisions and forecasting. Examples include commercial operating licenses, primary education data, healthcare center records,

and restaurant delivery traffic, among others, in addition to leveraging Artificial Intelligence (AI) tools for prediction

Research Objectives:

The solution proposed in this research is a framework designed to optimize and advance current models by enhancing available data quality, refining artificial intelligence models, and expanding the scope of collected data. This approach aims to facilitate decision-making that effectively and efficiently addresses the problem, based on the following:

1. To develop an Artificial Intelligence model that utilizes the latest state-of-the-art algorithms in forecasting, thereby enhancing the quality and accuracy of the outputs.
2. To improve input accuracy and quality by leveraging the latest available technologies in mapping and spatial data.
3. To expand the scope of data fed into the AI model to contribute to the precision of the final outcomes.

Summary research paper's:

Analysis and Optimization of the Spatial Patterns of Commercial Service Facilities Based on Multisource Spatiotemporal Data and Graph Neural Networks: A Case Study of Beijing, China

This paper is based on the analysis and optimization of the spatial patterns of commercial service facilities. It employs an analytical methodology that utilizes spatial and temporal data from multiple sources, employing the GraphSAGE model and exploring spatial and temporal relationships. The study was conducted in Beijing, the capital of China, due to its high population density. The researcher analyzed a wide range of data, employing several criteria, including the average distance between commercial establishments and the total number of features, based on a proposed area within the study (ISPRS Int. J. Geo-Inf. 2025).

Logistic regression and cellular automata-based modelling of retail, commercial and residential development in the city of Ahmedabad, India

This study developed a hybrid simulation model combining logistic regression and cellular automation-based modeling to create a model applicable to Ahmedabad, India. The model serves as a predictive tool for development based on the current division of activities into zones at the district level, relying on a chronological sequence (T. Munshi, M. Zuidgeest, M. Brussel, and M. van Maarseveen, 2014).

Sensing spatial distribution of urban land use by integrating points-of-interest and Google Word2Vec model

This study focuses on the mechanism of urban land use, which is the most important element in planning processes. With the emergence of Geographic Information Systems (GIS) and large geospatial datasets, many tools have appeared that help in using land with high accuracy. The study presented an innovative framework for monitoring urban land use distributions at the district level and used Google Deep Learning models, as well as several other models such as Word2Vec and K-Means, and several algorithms (Yao Yao, Xia Li, Xiaoping Liu, et al. 2017).

Measuring the Spatial Allocation Rationality of Service Facilities of Residential Areas Based on Internet Map and Location-Based Service Data

This study focused on measuring the rationality of the distribution of facilities and services. Previous studies relied on population size and some additional information, neglecting the actual, real-time behavior of residents and relying on linear rather than actual distances. Therefore, the researchers used map data and location-based service data. The study was conducted in Wuxi City, China(Xinxin Zhou, Yuan Ding, Changbin Wu, Jing Huang, and Chendi Hu.2019).

Understanding of the predictability and uncertainty in population distributions empowered by visual analytics

This study focuses on the mechanism for understanding satellite imagery, as in-depth analysis of this data contributes to a correct understanding of the location, thus increasing classification accuracy. It addresses the challenge of similar-looking models by studying the surrounding spatial context, which helps in monitoring unplanned urban sprawl(Yaxin Li .2024).

Synthesis Matrix for Literature Review:

Study	Authors	Year	Research Design	Key Findings/Results	Methodology	Limitations
Analysis and Optimization of the Spatial Patterns of Commercial Service Facilities Based on Multisource Spatiotemporal Data and Graph Neural Networks: A Case Study of Beijing, China	ISPRS Int. J. Geo-Inf	2025	Experimental	- Spatial Distribution Characteristics - Spatial Autocorrelation Analysis - Description and Classification of Commercial Service Patterns Based on GraphSAGE	- POI - GraphSAGE - GIS	This model is not without its limitations. Being a black box in unsupervised graph neural network modeling, the model's behavior lacks interpretability, necessitating a semantic interpretation of the output results.
Logistic regression and cellular automata-based modelling of retail, commercial and residential development in the city of Ahmedabad, India	T. Munshi, M. Zuidgeest, M. Brussel, and M. van Maarseveen	2014	Experimental	The results of population growth and activity projections for Ahmedabad, which were incorporated into the logistic regression model of the Urban Development and Growth Simulation Model, are based on the first step and extrapolation of past trends. These projections indicate that Ahmedabad's population will exceed 9 million by 2030, approximately one and a half times its current size.	logistic regression model	The study focused only on aspects related to momentary factors and the actual movement of the population.
Sensing spatial distribution of urban land use by integrating points-of-interest and Google Word2Vec model	Yao Yao, Xia Li, Xiaoping Liu, et al	2017	Experimental	a K-Means-based clustering model to analyze correlations between the POI/TAZ vectors and deploy TAZ vectors to identify urban land use types using a random forest algorithm (RFA) model. Compared with some state-of-the-art probabilistic topic models (PTMs), the proposed method can efficiently obtain the highest accuracy (OA = 0.8728, kappa = 0.8399). Moreover, the results can be used to help urban planners to monitor dynamic urban land use and evaluate the impact of urban planning schemes.	- TAZs - Word2Vec - POIs	The data used does not represent all segments of society.
Measuring the Spatial Allocation Rationality of Service Facilities of Residential Areas Based on Internet Map and Location-Based Service Data	Xinxin Zhou, Yuan Ding, Changbin Wu, Jing Huang, and Chendi Hu	2019	Experimental	The ASFC-RA Results The PDC-RA Results The SARSF-RA Results	- Scrappy - LBS - GIS	the incompleteness of the data acquisition still limits it, and there are some shortcomings. The following points are worth further improvement and development: (1) The experimental area could be expanded to the whole city, and the scope of the experimental area can be enlarged in the later stage to obtain more stable law. (2) The number and scale of the service facilities are

						two important indicators for measuring the supply of the service facilities. Due to the lack of scale data, this study has certain limitations. It would also be worthwhile to study the method for supplementing the scale data to POI through other Internet datasets. (3) The spatial subdivision method of Tencent LBS data can improve to template-based GIS computation
Understanding of the predictability and uncertainty in population distributions empowered by visual analytics	Peng Luo	2024	Experimental	- Model accuracy assessment of population estimation - Explaining population distribution - Explaining the uncertainty of population prediction	- GeoVisX - Visual Analytics - Predictability - Uncertainty	- Visualization as a high-dimensional approach to understanding geographical variables - The spatial process of population distribution

Critical Analysis:

All the research addressed the use of maps and other functions, but it did not take advantage of the governments' readiness to provide other helpful data such as operating licenses for shops and services, primary education data, health center data, and restaurant order movements, etc., to add additional parameters. It also did not include delivery requests for those neighborhoods through the use of Uber, Careem, and other applications.

Connections:

1. Using GeoVisX
2. Refining data captured from satellites
3. Using government resources for data

References:

ISPRS Int. J. Geo-Inf (2025) . Analysis and Optimization of the Spatial Patterns of Commercial Service Facilities Based on Multisource Spatiotemporal Data and Graph Neural Networks: A Case Study of Beijing, China .<https://doi.org/10.3390/ijgi14010023>

T. Munshi, M. Zuidgeest, M. Brussel, and M. van Maarseveen . (2014) Logistic regression and cellular automata-based modelling of retail, commercial and residential development in the city of Ahmedabad, India . <https://www.sciencedirect.com/science/article/abs/pii/S0264275114000298>

Yao Yao, Xia Li, Xiaoping Liu, et al (2017) . Sensing spatial distribution of urban land use by integrating points-of-interest and Google Word2Vec model .
<https://doi.org/10.1080/13658816.2016.1244608>

Xinxin Zhou, Yuan Ding, Changbin Wu, Jing Huang, and Chendi Hu.(2019) Measuring the Spatial Allocation Rationality of Service Facilities of Residential Areas Based on Internet Map and Location-Based Service Data . <https://doi.org/10.3390/su11051337>

Peng Luo .(2024). Understanding of the predictability and uncertainty in population distributions empowered by visual analytics .
<https://doi.org/10.1080/13658816.2024.2427870>