

Project Title: AI-Driven Agents for Fair and Efficient Property Valuation

Prepared by: Fahad Abubaker Bahashwan

Advisor: Dr. Ashraf Elsayed

Abstract

This project aims to build an AI-powered system that automatically analyzes maps and estimates property values. The motivation is to make property valuation faster, more accurate, and less reliant on manual work. The approach combines deep learning for land use classification, AI agents for collecting data, machine learning for predicting values, and natural language generation for creating reports. Key datasets include satellite images, real estate listings, and public records. The expected outcome is a tool that helps investors, planners, and homeowners understand property values and the factors influencing them.

1. Introduction

Property valuation is essential for real estate, urban planning, and investment. Traditional methods are slow and require experts to gather and analyze lots of data. With advances in AI, it's possible to automate much of this process. However, challenges remain in combining different data sources, making accurate predictions, and explaining the results. This report presents a structured plan to build an AI system that addresses these challenges, with sections covering the problem, motivations, objectives, related work, methodology, datasets, timeline, and references.

2. Problem Specification

The main problem is to automatically analyze maps and estimate property values using AI. This includes:

- Classify land types (e.g., buildings, roads, Industrial zones) from satellite images using image segmentation.
- Collecting relevant data (e.g., prices, amenities) from various sources (data aggregation).
- Predicting property values (regression).
- Explaining the predictions and generating readable reports.

The project focuses on urban and suburban areas with available data. It assumes access to satellite images, real estate listings, and public information.

3. Motivations

- **Real-world application:** Accurate, fast valuations help buyers, sellers, investors, and city planners.
- **Research significance:** Combines computer vision, NLP, and machine learning in a real-world setting.
- **Industry need:** Current solutions are slow, manual, or fragmented.
- **Personal motivation:** Interest in applying AI to solve practical problems in real estate.

4. Objectives

- Build an AI system that can analyze maps and predict property values automatically.
- Use deep learning to classify land use from images.
- Develop AI agents to collect and organize data.
- Create models to predict values and explain key factors.
- Generate reports that are easy to understand.
- As summary make property valuation more transparent, fair, and accessible.

5. Literature Review

- **Title of research:** Autonomous GIS - the next-generation AI-powered GIS ¹

Authors: Zhenlong Li, Huan Ning

Affiliation: Geoinformation and Big Data Research Laboratory (GiBD), Department of Geography, University of South Carolina.

Submitted: May 9, 2023 | **Last Updated:** May 28, 2023

Summary: This research introduces Autonomous GIS, a new type of AI-driven geographic information system that uses Large Language Models (LLMs) like GPT-4 to automate spatial data analysis. The system aims to perform tasks such as data collection, processing, and visualization without human intervention. A prototype called LLM-Geo was developed to demonstrate its capabilities. In three case studies (e.g., analyzing population near hazardous waste sites, tracking COVID-19 mobility trends), LLM-Geo generated accurate maps, charts, and numerical results, significantly reducing manual work. The goal is to make GIS tools faster, easier, and more accessible to non-experts.

Methodology: Core AI Model: GPT-4 (via API), Tested on real-world spatial problems, including population analysis, COVID-19 trends, and death rate visualization.

Findings: LLM-Geo successfully automated spatial analysis tasks, producing correct results (e.g., population counts, maps) in all case studies. Reduced manual effort by generating code and workflows automatically. Demonstrated feasibility of integrating LLMs into GIS for deterministic, data-driven tasks.

- **Title of research:** DoRA: Domain-Based Self-Supervised Learning Framework for Low-Resource Real Estate Appraisal ²

Authors: Wei-Wei Du, Wei-Yao Wang, Wen-Chih Peng

Affiliation: National Yang Ming Chiao Tung University, Taiwan | ACM

Submitted: Sep 14, 2023

Summary: DoRA introduces an innovative approach to real estate appraisals by leveraging domain-based self-supervised learning. This framework targets low-resource scenarios, effectively combining geographic prediction and contrastive learning to create robust, generalizable models. Extensive benchmarks validate its superiority, suggesting wide applicability in real-world financial markets. Implementations such as E.SUN Bank's system indicate practical benefits and enhanced efficiency in property valuation tasks.

Methodology: The thesis introduces DoRA, a domain-based self-supervised learning framework specifically developed for real estate appraisal in low-resource settings. Its methodology integrates both intra-sample geographic prediction and inter-sample contrastive learning to efficiently process unlabeled real estate data. DoRA was rigorously tested on transaction data from the Taiwan Real Estate Transaction Platform between 2015 and 2021, incorporating features like real estate metadata, points of interest, and economic and geographical data. This framework demonstrated significant performance improvements over existing self-supervised, graph-based, and supervised approaches, especially in scenarios with limited data availability. Notably, DoRA's success in enhancing property valuation precision has led to its adoption by E.SUN Bank for their online mortgage calculator, underscoring its practical value in financial services. The tool helps institutions manage high-cost domain expert tasks more efficiently, especially in regions lacking sufficient transactional data, thereby boosting appraisal objectivity and accuracy.

Findings: DoRA is a domain-based self-supervised learning framework for real estate appraisal in low-resource settings. It uses an intra-sample geographic prediction task to equip representations with domain knowledge and inter-sample contrastive learning to generalize representations for limited downstream data. Tested on real-world Taiwan real estate data from 2015-2021, DoRA significantly outperformed existing self-supervised, graph-based and supervised approaches in few-shot scenarios by at least 7.6% for MAPE, 11.59% for MAE and 3.34% for HR10%. DoRA demonstrates the effectiveness of geographical knowledge and contrastive learning for low-resource property valuation. It has been deployed by E.SUN Bank to improve efficiency and objectivity in their mortgage appraisal process.

- **Title of research:** A Comprehensive Overview Regarding the Impact of GIS on Property Valuation³

Authors: Gabriela Droj, Anita Kwartnik-Pruc, Laurentiu Droj

Affiliation: Faculty of Economic Sciences, University of Oradea, Romania | International Journal of Geo-Information

Submitted: March 29, 2024

Summary: This academic paper provides a comprehensive review of the impact of Geographic Information Systems (GIS) on property valuation. It highlights how GIS technology enhances traditional valuation methods by incorporating spatial data analysis, addressing the subjectivity and inefficiency often found in manual appraisals. The study focuses on the integration of GIS with hedonic models, artificial intelligence, and mathematical appraisal models to improve accuracy and support mass appraisal techniques. Ultimately, the research demonstrates the strategic role of GIS-based valuation for various stakeholders, including governments, banks, and developers, especially in navigating complex urban landscapes and promoting sustainable development.

Methodology: The researchers focused on three main areas. First, they looked at hedonic models, which estimate property value based on features like location, size, and surroundings. Second, they explored how Artificial Intelligence (AI), including machine learning and neural networks, can help analyze complex data. Third, they examined mathematical models, such as statistics and fuzzy logic. They only included studies that combined Geographic Information Systems (GIS) with real estate valuation to ensure the research was relevant to automated and location-based analysis.

Findings: GIS helps reduce guesswork in valuing properties by adding detailed location-based information, such as how close a property is to parks or areas that might flood. When GIS is added to traditional models, it gives a better picture of how things like the environment and natural risks affect property value. AI tools, like neural networks, improve the accuracy of property valuation by processing large and varied data sources, such as satellite images and even social media. Mathematical models, especially when paired with GIS, help speed up the process of valuing many properties at once, which is useful for taxes or mortgage planning. The study also found that collaboration between experts in different fields is important to handle big challenges like climate change, city growth, and fairness in access to property. Overall, GIS supports smarter planning by showing how things like green spaces, and infrastructure affect property values.

6. Proposed Methodology

- **Land Use Classification:** Use deep learning to segment satellite images and classify land types.
- **Data Collection Agents:** Build AI agents using reinforcement learning and NLP (e.g., BERT) to scrape and structure data from websites and APIs.
- **Valuation Prediction:** Use machine learning models (e.g., XGBoost, Graph Neural Networks) to predict property values, incorporating spatial and economic features.
- **Explainability:** Apply SHAP (Shapley Additive Explanations) to show which factors influence each prediction.
- **Report Generation:** Use template-based and transformer models to generate human-readable reports.
- **Tools:** PyTorch, Scikit-learn, Hugging Face, FastAPI.

7. Datasets Description

- **Satellite Images:** OpenStreetMap, SpaceNet for land use.
- **Real Estate Listings:** Scraped from property websites (price, area, features).
- **Public Records:** Government data on sales, permits, and demographics.
- **Preprocessing:** Clean and normalize data, handle missing values, augment images.

8. Timeline of Second Semester Project (10 Weeks)

Week	Planned Activity
1	Literature review, environment setup, start data collection
2	Data preprocessing, initial analysis
3	Implement and test land use segmentation model
4	Refine segmentation, start building data collection agents
5	Integrate multi-modal data, implement valuation models
6	Fine-tune models, develop explainability tools
7	System integration, build web app prototype
8	Evaluate system, collect user feedback
9	Write report, prepare visualizations
10	Final review, thesis submission, presentation prep

9. References

1. Li, Z., & Ning, H. (2023). Autonomous GIS: the next-generation AI-powered GIS. arXiv preprint arXiv:2305.06453. <https://arxiv.org/abs/2305.06453>
2. Du, W.-W. (2023). DoRA: Domain-Based Self-Supervised Learning Framework for Low-Resource Real Estate Appraisal. arXiv preprint arXiv:2309.00855. <https://arxiv.org/abs/2309.00855>
3. Droj, G.; Kwartnik-Pruc, A.; Droj, L. A Comprehensive Overview Regarding the Impact of GIS on Property Valuation. ISPRS Int. J. Geo-Inf. 2024, 13, 175. <https://doi.org/10.3390/ijgi13060175>