

Title of the Project ➤

AI-Driven Agents for Fair and Efficient Property Valuation

Prepared by: Fahad Abubaker Bahashwan

Literature Review

➤ Research Paper OVERVIEW

Research Paper Title

DoRA: Domain-Based Self-Supervised Learning Framework for Low-Resource Real Estate Appraisal

<https://arxiv.org/pdf/2309.00855>

Conference

ACM International Conference on Information and Knowledge Management (CIKM 2023)

Authors

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

University

National Yang Ming Chiao Tung University, Taiwan

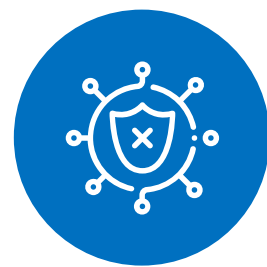
A photograph of a large, modern, multi-story university building with a grid-like facade and many windows. In the foreground, there is a red heart-shaped sculpture and large white letters spelling 'NYCU'.

National Yang Ming Chiao
Tung University, Taiwan

KEY IDEA

self-supervised learning (SSL)

Use self-supervised learning (SSL) to learn property patterns from unlabeled data, then fine-tune with minimal labeled examples. (Training Model)



HOW IT WORKS:

1

**Pre-Training Stage
(Unlabeled Data)**

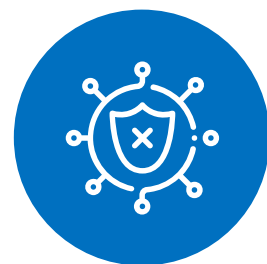
2

**Fine-Tuning Stage
(Few Labeled Examples)**

KEY IDEA

self-supervised learning (SSL)

Use self-supervised learning (SSL) to learn property patterns from unlabeled data, then fine-tune with minimal labeled examples. (Training Model)



Baselines Compared to DoRA:

XGBoost

Traditional supervised
method
(fails with minimal data)

Graph Models (MugRep)

Struggle in rural areas with
sparse neighbors



Numerical Features

e.g., square footage, age, and rooms



Categorical Features

Property type: apartment, house

Parking availability: yes, no



Pol (Point of interest) Features

YIMBY (schools, parks) vs NIMBY (factories) counted within 1km



Economic Features

e.g., land price, unemployment rate, GDP



> Data Description



Source

Taiwan Real Estate Transaction Platform (2015–2021)



Size of Data

- Unlabeled
434,243 properties (used for pre-training)
- Labeled
64,266 transactions (24k test, 40k validation)



Features

- 39 Real Estate
23 numerical (e.g., age, rooms), 16 categorical (e.g., property type)
- 16 Pol (Point of Interest)
Nearby facilities (e.g., schools, factories)
- 7 Economic
Macro-level indicators (e.g., GDP)

Evaluate: DoRA vs XGBoost

The results table comparing DoRA and XGBoost for houses in **Wugu District** (industrial/rural area with extremely limited transactions)

Metric	DoRA	XGBoost	Improvement
MAPE	2.55%	162.85%	160.3% lower
HR10	100%	0%	100% higher

MAPE = Mean Absolute Percentage Error

HR10 = Hit Rate within 10%

> Advantages & Limitations



Advantages

- Works with 1–5 labeled examples (traditional methods need 100+)
- Integrates domain knowledge
- Deployed at E.SUN Bank as a mortgage calculator



Limitations

- Tested only on Taiwan data (may not generalize globally)
- Requires Pol data (may not be available everywhere)



Thank You