A Decision Support System for Property Valuation in Australia

Mehrdad Ziaee Nejad

Ph.D. Thesis
This thesis is submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

University of Technology Sydney
Faculty of Engineering and Information Technology
January 2020

Certificate of Original Authorship

I, Mehrdad Ziaee Nejad declare that this thesis, submitted in fulfilment of the requirements for the award of PhD, in the Computer Science School at the University of Technology Sydney.
This thesis is wholly my own work unless otherwise reference or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.
This document has not been submitted for qualifications at any other academic institution.
This research is supported by an Australian Government Research Training Program Scholarship.

Signature:

Production Note:
Signature removed prior to publication.

Date: 08/07/2020

Dedication

I dedicate this thesis to my darling wife for her passion and patience and to my parents who passed away.

Acknowledgement

I would like to express my profound appreciation to Professor Jie Lu, who supervised this Ph.D. program, for all of her knowledgeable comments, precious support, right guidance at hard times, and great suggestions along the way. I want to thank Jie for her willingness to let my research follow my interests. I am also grateful to Dr. Vahid Behbood and Dr. Mohsen Naderpour, my co-supervisors, for knowledgeable suggestions and valuable advices through my PhD study.

I would like to thank my wife, Sepideh, for her endless support during my Ph.D. journey. She truly has been shoulder-to-shoulder with me during this journey. Thank you very much for being so supportive in all circumstances throughout the four years of this Ph.D., for understanding the stress I was subject to, and for giving me the freedom to follow my scientific interests unconditionally. I could not have accomplished this without your constant love and support.

I express my gratitude to all my friends and colleagues in the decision systems and e-service intelligence (DeSI) lab in the Centre for Artificial Intelligence (CAI) for their help and valuable comments during my study.

I also appreciate the Faculty of Engineering and Information Technology, the School of Computer Science and the CAI at the University of Technology Sydney for conference registration and travel funds provided during this research.

Abstract

Residential real estate price is one of the key components of our economy developments and also has been a major concern of the public, bank industry, government and investors. The accurate estimation of sale price and its changes have an important role in decision-making of related departments and organisations. In Australia, one of the biggest investments for people is in residential real estate. Therefore, many studies and research works have been carried out to build an automated valuation model to predict sale prices of residential properties accurately as much as possible.

The complexity of estimating residential real estate price is related to its main characters, such as immovability, durability and it is highly dependent on location and property structure. In traditional Automated Valuation Models (AVMs), the standard information related to property is used for predicting the sale price of properties including structural features (number of bedrooms, bathrooms, age, floor area, etc.), locational features (distance to Central Business District, amenities and etc.) and demographic features (reigning profiles). But, in a real-world situation, a valuer visually inspects the property to estimate its sale price instead of using numbers on paper. People browse real estate online websites to trade their properties. Real estate agents or sellers provide images of each property to help buyers in decision-making when purchasing property. A reason to use images and videos is that they act like universal languages for people with different backgrounds. The main content of an image or video can be easily understood.

However, images, which are probably the most important factors on a buyer's initial decision-making process, have been ignored in the process of prediction of sale

price in existing AVM models due to the difficulty of engaging visual features in a regression problem. This study uses property images to propose a new advanced property valuation decision support system (DSS) to predict residential real estate sale prices.

The proposed DSS includes these components: data management, pre-processing, classification, prediction and user interface. The data management component takes care of data flows in the system. Pre-processing prepares the inputs for the prediction model through cleansing, scaling and feature enrichment. The prediction component takes into account the visual features and incorporates them into the regression problem. The convolutional deep neural network is used to extract the visual features from property images. These include bedroom, bathroom, living room, kitchen, balcony and front images. In regards to labelling images in a property valuation DSS environment automatically, an image classification model is developed whereby arrival images are labelled and stored in a real estate database.

Also, in this study, two separate areas of real estate property prediction – mass appraisal and housing indices – are combined to improve the accuracy of prediction of an individual property. Inspired by the stratified median price method, this study shows how a median price index derived from K-mean clustering is created to extract the trend and historical features related to segment (cluster) to which that property belongs. To create a proper index based on the median price, the structural, locational and demographic features of properties are considered as aggregated features. The adjusted median sale price index is used as input of a Long-Short Time Memory (LSTM) model to create and extract sale price trends and historical features to be add to the deep neural network regression model to improve the accuracy of predicting sale prices.

Two main prediction models to handle real-world situations regarding availability of property images in the proposed valuation DSS are developed: the visual prediction model and the non-visual prediction model. The visual prediction model considers the standard meta-data, visual features extracted and sale price market features. The non-visual prediction model considers meta-data and extracted housing market features.

This study presents the improvement of accuracy in sale price prediction based on visual features and housing sale price index features. The improvement of prediction by adding adjusted median sale price features compared to the baseline model which uses meta-data features only, is 8.23% decreasing in the mean absolute error. By adding visual features, the mean absolute error is dropped by 9.16% compared to the non-visual prediction model. In other words, a mainly visual-based prediction model has 16.65% lower error rate than a baseline prediction model.

Table of Contents

List of Figures

List of Tables