

Real Estate Pro: Website For Property Search

Abstract:

By connecting buyers and sellers (or agents) in a smooth and effective manner, our website makes real estate purchases and sales easier than ever. Setting up viewing schedules, uploading images and videos, and updating property listings are all easy tasks for agents and sellers. Customers may now plan visits that suit their needs and view the most recent information at all times. Buyers can explore properties with detailed media, such as images and videos of the house and the neighbourhood, thanks to a seamless, user-friendly experience. They can use the platform to compare houses according to location and pricing, enabling them to make informed, cost-effective choices. The website provides you with all the information you require to feel secure in your decisions, whether you're searching for a new residence or an investment. One of the most notable features is the integrated chat box, which facilitates quick and simple contact between buyers and sellers. Customers don't have to leave the platform to arrange viewings, ask questions, or obtain additional information. By bringing the parties together in one location, this feature helps to eliminate the typical back and forth. Together, these capabilities make the website an easy-to-use and practical way to handle real estate transactions.

By facilitating communication, property viewing, and speedy decision-making, it reduces the inconvenience of conventional procedures in real estate. This website offers everything you need to make the transaction easy, quick, and stress-free, whether you're buying or selling.

Introduction:

Both buyers and sellers frequently encounter difficulties in the current real estate market, which can cause the process of property transfers to move more slowly. Keeping property listings current is a challenge for sellers and agents, and

purchasers struggle to sort through the many postings, arrange viewings, and make well-informed judgments with scant information.

Furthermore, there may be gaps in communication between buyers and sellers, which frequently results in missed opportunities or delays. Inefficient and time-consuming, the typical real estate purchasing and selling procedure frequently entails numerous platforms, phone conversations, and in-person visits. The requirement for buyers to evaluate houses across various areas and price ranges, as well as for sellers to handle numerous inquiries and viewings, exacerbates this issue even further. A centralized platform that simplifies the whole real estate process—from listing properties to viewing, comparing, and communicating—is required to address these issues and guarantee a seamless, effective experience for all involved.

This platform must be user-friendly and accessible, offer real-time communication capabilities, make it simple to access the most recent property listings, and provide facilities for scheduling viewings and comparing costs.

Paper 1:

The research article "A Real Estate Valuation Model Using Boosted Feature Selection" by Kankawee Chanasit and others suggests a better approach to real estate value estimation when data is incomplete or high-dimensional. The simple Artificial Neural Networks (ANNs) are excellent predictors of complex, non-linear interactions but may be un-interpretable and get worse as the number of useless features becomes excessively large. To overcome this, authors propose a new ****Boosted Feature Selection**** approach by combining AdaBoost.R2 (boosting algorithm) with an ****Improved Garson's Algorithm (IGA)**** as feature sensitivity analysis. It selects the most significant features one at a time by tuning their weights in terms of prior prediction errors.

It was tested on the synthetic data sets (Friedman data set) as well as on actual data sets (Boston Housing, Home2nd of Thailand, and Zillow Prize data). Experimental outcome proved that proposed technique was far better than conventional feature selection methods like Recursive Feature Elimination (RFE), RReliefF, and Mutual Information (MI). It indeed did simplify the models, improved the prediction

quality, but with fewer errors. Most notably, the approach allowed the most dominant traits in setting housing prices to be established, providing insight into different property markets.

In practice, the model performed quite adequately by placing in the bottom 24% of the Zillow Prize competition, which is a sign that it could be employed practically. The study also illustrated that the ANN model, with appropriate feature selection, can outperform conventional tree-based models such as XGBoost, particularly under time-evolving data patterns.

Overall, this paper emphasizes the significance of choosing proper features in real estate valuation models to prevent issues such as the curse of dimensionality. The future will enhance stopping criteria for feature selection, minimize computational time, and investigate ensemble feature selection algorithms to make improved models stronger and scalable on large datasets.

Paper 2:

The paper "Developing a Mobile Application for Smart Real Estate Information" by A. C. Aydinoglu and R. Bovkir outlines the process of creating a mobile GIS-based application for the management of urban real estate information. Land administration will be based on robust, complete, and interoperable land and property information. Geographic Information Systems (GIS) offer effective means for spatial analysis in the form of enhanced land assessment, urban planning, and investment.

Researchers created a mobile application, Smart Real Estate, by combining multiple datasets of environmental characteristics, public facilities, cultural characteristics, transport, land use, and socio-economic variables. Six theme groups contain twenty parameters like schools, health centers, transport stops, parks, and socio-economic variables like people and education. GIS analysis like slope, aspect, density, and Euclidean distance calculation was utilized to create thematic surfaces. Fuzzy logic was treated similarly for dealing with uncertainty in everyday life.

It was built with Java and Android Studio and has support for PostgreSQL/PostGIS databases and RESTful web services. Users can have address search, real estate

search with thematic filtering on many factors, and land evaluation with thematic scoring. Property listings can be seen by users and real estate can be evaluated based on specific geographic factors using Google Maps API for visualization.

Unlike historical real estate application, Smart Real Estate offers thematically ranked dynamic ratings in bar graphs. It enhances the ability of residents to use urban land data reliably and enable smart city goals through openness and knowledge-based decision-making. Big data integration, mobile GIS, and open data services are considered essential by the study for urban planning. The app is a shining light to the promise of next-generation smart city technology leveraged toward sustainable, human-centric city management.

Paper 3:

The article "Public Real Estate Management System in the Procedural Approach – A Case Study of Poland and Slovakia" by Marta Gross, Ryszard Żróbek, and Daniela Špírková acknowledges research on public real estate management in Poland and Slovakia from the procedural approach. The research seeks to establish how the concepts of good governance can help public real estate management systems. Although specific to individual countries, both are comparable in nature because both are post-socialist.

The research compares three best practices: sale of public assets, registration of rights, and public-private partnerships (PPPs). Sale of assets has to be executed by open procedures in Poland, whereas in Slovakia empty state buildings have to be disposed off first to public institutions. The Polish system was more open as well as anti-corruption policy. However, the Slovakian system performs better with more online services and faster procedures despite its faster registration process in Poland.

In the case of PPPs, Poland precedes Slovakia because it has more dynamic projects, more liberal legislation, and more entrenched political support, primarily for the construction of infrastructure. It follows secondly by Slovakia because it has delayed implementation of its projects, along with weak political backing. The 0–1 system of scoring was utilized with the goal of comparison of both nations

based on indices of efficiency worth 13 and 8, respectively, for Poland and Slovakia.

The research finds that Poland is a leader in public disposal of assets and application of PPPs, but Slovakia is a leader in real property registration systems. The two countries can be improved by studying the best international practices and learning from one another's do's and don'ts. The research is aimed at the efficiency culture within real estate management and promotes broader aspiration for the achievement of sustainable, open, and people-focused government in the context of public real estate. Good governance is able to impact significant success in management in most contexts across the globe..

Paper 4:

The working paper "Searching for Property with a Smartphone – Trends, Promises and Perspectives" by Veronika Lang and Peter Sittler discusses the increased relevance of smartphone apps in searching property in the German-speaking nations. As there has been increased use of smartphones and apps, property business has undergone a sudden shift to mobile technology. The study compares the availability, content, functionality, and capability of Austrian, Swiss, and German real estate applications and contrasts the two Austrian applications--[immobilien.net](#) and [wohnnet.at](#)--in complete detail. The study determines that iOS is largest in size with more real estate applications to use than Android systems. None of the applications have Augmented Reality (AR), and geo-coding of the properties is absent in most, and only 10% of Austrian listings of properties are geo-coded. The applications were checked on operating systems, number and quality of property information, degree of information, ease of use, and property types available.

Testing identified [immobilien.net](#) as targeting Vienna inner suburbs, with AR capability but only with very limited Austro-wide coverage. [Wohnnet.at](#) has wider Austro coverage with no AR functionality. Both suffered from partial address and restricted search functionality customisation.

Across the German-speaking region, Germany boasts the largest number of apps and functionalities like route planners and geo-coding of social media of networks.

Swiss apps are coded geographically more and carry multiple languages, particularly through Comparis and ImmoScout24.ch. Augmented Reality is good to look at but underdeveloped because of low geo-coding rates and technical limitations.

The study keeps smartphone application potential high in order to support seamless searching of properties but depends on prolonged geo-coding, vast lists of property listings, and constant calibration with technology. Future innovation, client need, and the future of smartphones will still drive how property mobile searches evolve into the future with prospective possibilities but key development obstacles to overcome.

Paper 5:

The article "A Systematic Review of Smart Real Estate Technology: Drivers of, and Barriers to, the Use of Digital Disruptive Technologies and Online Platforms" authored by Fahim Ullah, Samad M. E. Sepasgozar, and Changxin Wang provides insights about how digital disruptive technologies influence the real estate sector. The research tackles 213 papers comprehensively to explore the application of nine mega technologies—drones, IoT, cloud computing, SaaS, big data, 3D scanning, wearable technology, VR/AR, and AI/robotics—the Big9 for Smart Real Estate (SRE) construction.

SRE is explained along the Big9 dimensions of technology innovation, user-centricity, and sustainability. The analysis selects key stakeholders—consumers, agents and associations, regulators and government, and support industries—and follows the path of how digital technologies address their needs to avert post-purchase or rental regret. Technology Adoption Model (TAM) is adapted in order to evolve a conceptual model that links Big9 technologies, online dissemination platforms, and needs of stakeholders.

The research acknowledges that although property has a gigantic market, it trails all companies in exploiting technology. It focuses on developing sophisticated distribution platforms such as websites, mobile applications, and social networking platforms to get to as many consumers as possible with sophisticated property information. Effective management has huge potential to trigger more informed

decision-making, minimize risk, and improve market transparency through big data analysis and artificial intelligence

Implications reveal humongous potential for Big9 technologies with minimal current levels of adoption due to talent shortages, technology tight paradigms, and resistance to change. Recommendation requires recommendations for improvement of better ways through which technologies must be re-architected to empower intelligent, green, and people-centric property management. Finally, digital transformation must be undertaken with a vision to revolutionize the obsolete property market into smart and customer-centric business..

Paper 6:

Mingzhao Li, Zhifeng Bao, Timos Sellis, Shi Yan, and Rui Zhang's "HomeSeeker: A Visual Analytics System of Real Estate Data" presents HomeSeeker, a visual analytics system that can help homebuyers to search for and locate real estate markets. HomeSeeker is proposed as a reaction to counteract some typical drawbacks of current commercial systems like a deficiency of geo-related information, insufficient learning support on local markets, and insufficient multidimensional comparative support.

HomeSeeker integrates heterogeneous data sources such as property attributes, distance to transit, school districts, amenities, and populations. HomeSeeker can learn incrementally for new users and handles natural, coherent visualizations. The most significant visualization tools in HomeSeeker are choropleth maps, dot maps, glyphs on maps, several coordinated views, parallel coordinates, geo-coded scatter plots, boolean tables, image cards, word clouds, and spider charts.

System architecture facilitates navigation within regions, suburbs, and properties at profile-based, suburb-level, and property-level views. Properties may be filtered against and compared against user-specified requirements such as price, amenity proximity, and school grade. Actual case studies illustrate that users, either investor or first-home buyer, may choose suitable properties by iteratively filtering candidates against specific requirements.

HomeSeeker was developed using HTML5, JavaScript (D3.js), Google Maps API, MySQL, and PHP. System testing and expert evaluation made the system easy and

user-friendly, but scalability complexity and visualization complexity needed to be enhanced. HomeSeeker extension to other cities, dynamic updation of data, and optimization of large datasets are upcoming extensions.

In all, HomeSeeker is superior to conventional property search sites in the way that it provides location-based, multi-dimensional data searching for seeking with a consumer-centric methodology in attempting to help smarter, more participatory, and more efficient house seeking by various parties

.Paper 7:

Article "Real Estate Politik: Democracy and the Financialization of Social Networks" by Joanne Cheung describes how social networking site business models erode the ideal of democracy. Cheung draws a parallel between commodification of human land for money and of human attention for corporate profit through such sites as Facebook. Spaces are construed to refer to privately owned public space where private interest rather than public concern prevails, i.e., property development for pecuniary gain maximisation.

Accounts explain how platforms monetize social interaction (liking, sharing) to generate user engagement through mechanisms concealing moral responsibility and stripping users of agency. Platform-based business models commodify mundane conversation, amplifying system polarization and inequalities. Algorithmic reward systems encourage content that supports confirmation bias and limiting users' exposure to dissent, eroding the basis for democratic debate.

Cheung discloses contrasted histories of extractive platform business model and colonial land use, and city design codes, land trusts, and Indigenous land control as templates to reconfigure digital space. The expert boundaries are forced to be reworked on democratic lines, the communal ownership models negotiated out of corporate monopolies, and the extraction stewardship.

Redesign of incentives and ownership is structurally necessary, but not on their own, without values. Platforms must move away from the poles of laboring for shareholders' profit interests to laboring for common good and democratic. Colonial and finance origins of platform structure must be understood to reform.

In short, Cheung urges the creation of new technical, legal, and governance infrastructures to reclaim social networks for the public good. In this manner, society can rebuild online communication spaces from tools of economic exploitation to spaces of authentic democratic participation, inclusiveness, and counter-action to structural injustices..

Paper 8:

The article "Real Estate Recommendation Approach for Solving the Item Cold-Start Problem" by Jirut Polohakul, Ekapol Chuangsuwanich, Atiwong Suchato, and Proadpran Punyabukkana deals with new real estate unit recommendation in web applications. The item cold-start problem arises when there are no new units that have interaction history and thus are less suitable for using usual recommendation techniques. The authors suggest a new solution using a content-based model and a session-based recommendation algorithm to address this problem with good performance for warm-start items.

Their approach is trained in a session-based manner altered by accounting for sequential user behavior and context to predict user profiles. They utilize NARM's attention mechanism and the latent cross (LC) method to effectively process contextual features. Following the prediction of a user profile, an appropriate filtering module utilizing a weighted cosine similarity nearest-neighbors algorithm suggests suitable properties. The used dataset is comprised of more than 13 million interactions of Thai real estate search engine.

Experimental findings indicate that the system proposed works amazingly well in cold-start and warm-start situations. While it will not be a best performer in either case individually, it is a good compromise by outperforming most of the conventional methods which operate on a single dimension. Comparisons with baseline models of Item-KNN, NARM, STAMP, and content-based models illustrate the challenge of doing extremely well on both cold- and warm-start recommendation at the same time.

The conclusion is that besides sequential patterns, contextual information improves recommendation quality, particularly in real estate websites where activities of users vary and listings refresh dynamically. A fine-tuning of the weight of

importance towards further personalizing recommendation is a potential direction of future optimization. Overall, the solution offers an effective solution to real estate websites with the dynamic issue of user and item cold-start.

Paper 9:

The article "A Real Estate Valuation Model Using Boosted Feature Selection" by Kankawee Chanasit et al. presents a novel method of real estate price estimation using an Artificial Neural Network (ANN) with boosted feature selection.

Conventional real estate information such as Thailand data usually are sparse but high-dimensional data, and therefore feature selection is very crucial. Authors blend the boosting algorithms with the enhanced Garson's algorithm (IGA) for the sensitivity analysis to incrementally choose significant features. The strategy is designed to preserve interpretability with improved prediction ability and lower computational cost.

Friedman dataset real and synthetic datasets (Boston Housing, Home2nd, and data for the Zillow Prize competition) were utilized during analysis in the research. Experiments indicate that their model always performs better compared to other standard feature selection methods such as RFE, RReliefF, and Mutual Information (MI) to identify crucial features with more accurate prediction errors. Their model ranked in the top 24% of the Zillow Prize competition.

The model integrates IGA and AdaBoost.R2 in a way that ANN dynamically adjusts feature selection parameters based on sample error. BIGA technique suggested learns intrinsic informative features which are capable of dealing with local variation and data anomaly. Findings indicate feature selection improves the performance of ANNs over tree models, particularly in noisy market trends.

The article goes on to say that BIGA is computationally more expensive but better in feature selection and dimensionality reduction for house prices. Future work is keen to enhance the stopping criterion in feature selection and extend the approach to ensemble systems. Overall, the research proves that the integration of boosting methods with sensitivity analysis can be a useful tool to enhance automated real estate appraisal systems with computational cost-effectiveness and prediction accuracy.

Paper 10:

Real Estate Recommendation Method for Solving the Item Cold-Start Issue is a paper that explains how to solve one of the most common issues in recommendation systems when there are new items and sparse interaction data. For the topic of real estate, the research explains how traditional systems perform poorly when users start developing interest in new and old properties at the same time. The authors recommend a hybrid approach using a session-based recommendation model, content filtering, and context to solve cold-start and warm-start scenarios quite effectively.

The method predicts a session-based to next-interacted item encoded feature prediction for user profile enabling without identifiers. The model uses sequential click streams and capitalizes on context features like location or device type via the latent cross (LC) approach. The model proposes the application of a nearest-neighbors strategy and weighted cosine similarity to predicted user profiles. Performance is measured in terms of recall metrics Recall@K and MRR@K against a collection of reasonable cold-start-warm-start trade-off baselines. Experiments are performed on real data on a Thai property portal, with over 13 million interactions, 3 million users, and nearly 7,000 items.

The new approach is more adaptive with a superior compound case but not necessarily cold-start or warm-start case per se. The paper suggests that simple methods such as STAMP and Item-KNN are optimal for warm-start cases but not cold-start cases, and simple content-based methods are optimal for cold-starts but worst suited to warm-starts. More discussion suggests that the activity in the future would be focused on feature optimization to user-specific match quality targets.

Quality in research means that balance of performance on both sets of items' attributes is needed to optimize tasks like real estate, which involves user demand from new homes through to years-old homes. Their hybrid approach optimizes balance, optimizing reach and relevance of recommendations in a manner to enable high item turnover and short-term user contexts

.Paper 11:

The article "An End-to-End Named Entity Recognition Platform for Vietnamese Real Estate Advertisement Posts and Analytical Applications" targets problems of getting structured information from Vietnamese real estate ads. The article proposes an end-to-end platform that can harvest, clean, process, and analyze real estate data. The design consists of two primary modules: Noise Filtering and Named Entity Recognition (NER).

The Noise Filtering module removes low-quality posts using a highly optimized PhoBERTlarge model with an F1 score of 0.8697. The NER module retrieves effective features like property type, location, and price from PhoBERT-based transformer high-performance posts. The above modules combined help to produce a high-quality dataset, addressing noisy, inconsistent, or incomplete listings problems of real estates.

The used training data are 24,695 Vietnamese real estate listings on large websites, labeled based on professional guidelines using Doccano. Approximately 58.3% of samples were applied for NER training in preprocessing. Some of the challenges in the Vietnamese language such as the utilization of abbreviations, free-text style, and the entity boundary challenge are explained in the paper.

Experimental outcomes show that transformer-based models perform better than conventional methods like BiLSTM-CRF for this problem. PhoBERT models with Vietnamese data only attains the best overall performance. Ablation analysis also shows that the noise filtering module added further enhances the NER performance up to 13.81%.

Production deployment of the platform includes auto-pipelines in Apache Airflow and hosting over PostgreSQL and Redshift databases. Superset dashboards are implemented within the platform to display trends like property distribution per region, price bands, demand types, and property types. The system processed more than 400,000 records within production and support data-driven decisions within real estate firms.

Finally, the paper also mentions some limitations in convenience of acquisition of exact address details and potential system weaknesses which are expected to emerge in the future. Future research aims to maximize module performance as well as expand the amount of information that is recovered.

Paper 12:

Paper "Explainable and Fair AI: Balancing Performance in Financial and Real Estate Machine Learning Models" forms a framework by combining fairness and transparency with ML models such as LightGBM and XGBoost for loan decisions and prediction of house prices. The study identifies towards the absence of practices for fairness and explainability in finance and property sectors where discriminatory models have profound societal consequences.

The study uses fairness techniques such as Calibrated Equalized Odds and Intersectional Fairness coupled with explainability tools such as SHAP (SHapley Additive exPlanations) and Partial Dependence Plots (PDPs). The authors conducted a large number of experiments on publicly available datasets with fairness constraints while not significantly harming model accuracy. LightGBM was found to be slightly better than XGBoost in the trade-off between fairness and accuracy.

Data preprocessing involved cleaning, feature engineering, and normalization. The models were tested on accuracy, precision, recall, F1-score, and fairness measures such as Disparate Impact and Equal Opportunity Difference. SHAP and LIME methods gave global and local explanations of predictions that pointed out major features such as Credit History and Applicant Income for loan approval and Overall Quality and Living Area for house prices.

The findings highlight that the addition of fairness constraints marginally impacts performance but yields humongous ethical gains. The article also illustrates how SHAP and LIME complement each other, where SHAP provides global information and LIME provides local interpretability.

New studies are applying fairness-aware learning in real estate, comparing fairness-led LightGBM and XGBoost models, preventing the compromise between fairness and accuracy, and proving the scalability of fairness-aware AI systems. It

also suggests directions for future research in the form of adversarial fairness techniques and fairness-aware explanation tools.

The study emphasizes the additional importance of responsible AI regulation, regulation compliance, and stakeholder trust building. It promotes using fairness-aware models for actual financial and real estate systems with no compromise in scalability or operation performance.

Paper 13:

The paper "Foundational AI in Insurance and Real Estate: A Survey of Applications, Challenges, and Future Directions" offers an extensive overview of applications, challenges, and future directions of Artificial Intelligence (AI) in insurance and real estate. It points to the intensifying influence of AI through technological developments like machine learning, deep learning, natural language processing (NLP), and computer vision. Applications include risk assessment, fraud detection, property valuation, dynamic pricing, to smart building management, all promoting efficiency and decision-making.

Among the most crucial issues that are realized are interpretability of the model, ethics and compliance, scalability, and data quality. The article presents the argument that while AI is forecasted to revolutionize sectors, sectors will have to grapple with privacy threats, data bias used to train the system, and challenges faced when trying to integrate AI in current systems. It points to regulatory measures such as the Fair Housing Act and GDPR that affect the implementation of AI.

Future directions are dynamic pricing with reinforcement learning, explainable AI for regulatory compliance, and smart sustainable buildings with AI. Interfacing with technologies such as IoT, AR/VR, and blockchain will further drive these industries. Future research directions the paper suggests are standardizing datasets, generalizing models to markets better, making algorithms fairness-aware, and increasing interdisciplinary collaboration between technologists, ethicists, and lawyers.

It generates value primarily by concentrating attention onto insurance and real estate markets, plugging gaps such as absence of standardized benchmarks, domain

explainability needs, and transferability across markets. The paper suggests that responsible AI should be developed on the pillars of transparency, ethical behaviour, good governance, and frequent skill refreshes.

Overall, this survey emphasizes that sustainable, explainable, and fair AI solutions have the potential to bring tremendous power to insurance and property competitiveness and innovation while ensuring societal and moral norms. It provides a guide to scholars, researchers, and lawmakers on the way AI's future value added to the industry can affect.

Paper 14:

The article "Optimal Real Estate Pricing and Offer Acceptance Strategy" by Gonen Singer and Eugene Khmelnitsky discusses the issue of the optimal expected revenue of the seller, under a series of offers acceptance in real estate markets. Contrary to static models with stable offers distributions, the authors propose a dynamic model in which listing price adjusts the quantity and offer distribution.

The seller at every level of selection posts a listing price, receives bids, and chooses to take the highest bid or persist, incurring holding and communicating costs. Optimal stopping is applied in the model in listing price and holding-back price strategies. Increasing the listing price statistically achieves the impact of increasing amounts of offers but decreasing offer rates, placing a price-demand trade-off.

The authors construct a threshold-based optimal policy for both finite and infinite time horizons. For the infinite time horizon, a stationary policy is constructed in which the reservation price is calculated by maximizing expected revenue with equalization of marketing costs. For the finite time horizon case, policies are repeatedly updated at each stage as the deadline for sales draws near.

Numerical value calculations of market value estimates imply that with higher offer spread (i.e., with more distinct assets), listing and reservation prices should be greater from the perspective of the sellers. Sellers can charge more as there is

greater market demand, and sellers' holding costs being greater imply selling faster at lower prices.

Validation of the model is done through an actual case study of Ames house prices in Iowa. Linear regression is applied to predict market values and determine the best reservation and listing prices as per them. Results show sellers' gains due to dynamic adjustment of strategies with market feedback and time constraints.

The study shows that a theory-of-optimal-stopping-based dynamic pricing and acceptance policy yields significantly better performance for the seller. Future research can investigate non-stationarity in the offer distribution and learning-based adaptive selling policies to achieve optimal real estate selling policies.

Paper 15:

The paper "Can Web Search Queries Predict Price Change on the Real Estate Market?" by Anna Baj-Rogowska and Nina Rizun deals with the possibility of utilizing web search volume, as discovered with the help of Google Trends Search Volume Index (SVI), in order to predict changes in real estate prices, or, at least, on the Polish market. It combines cross-correlation analysis and machine learning (ML) classification techniques to predict house price rises, falls, or none. Unlike other approaches, the research uses SVI statistics in "credit to purchase property" and "property" to train ML models and shows it is possible by utilizing online search behaviours as a single predictor variable that has adequate time-lag to inform decision-making.

The approach involves gathering Google Trends and Polish House Price Index (HPI) data from 2011-2019, timestamp difference adjustment, and expert selection and cross-correlation methods for cleaning keywords. Cross-correlation analysis identifies that there exists a four-month lag window under which search activity precedes price change. Machine learning models Random Forest, Bagged CART, and GLMNET confirm that SVI can predict direction of the market with high accuracy.

It is revealed that housing price volatility has very high positive correlation with real estate search volume, but lower correlation with credit searches. Cross-correlation and Granger causality tests also established that search volumes lead housing prices and not the other way around. Machine learning classification results categorize that SVI alone can be extremely accurate in prediction, especially through methods such as Random Forest ensemble.

Theoretical contribution of the study is to incorporate predictive modeling with unbounded real-time online search behaviour data. In practice, it promotes utilization of SVI in real estate predictive models and makes market participants take decisions based on adequate information. Some of the limitations are reflected by authors: generalizability to other nations and utilization of a bounded dataset. Future research can utilize merging SVI with traditional predictors and employing analysis on multiple real estate markets.