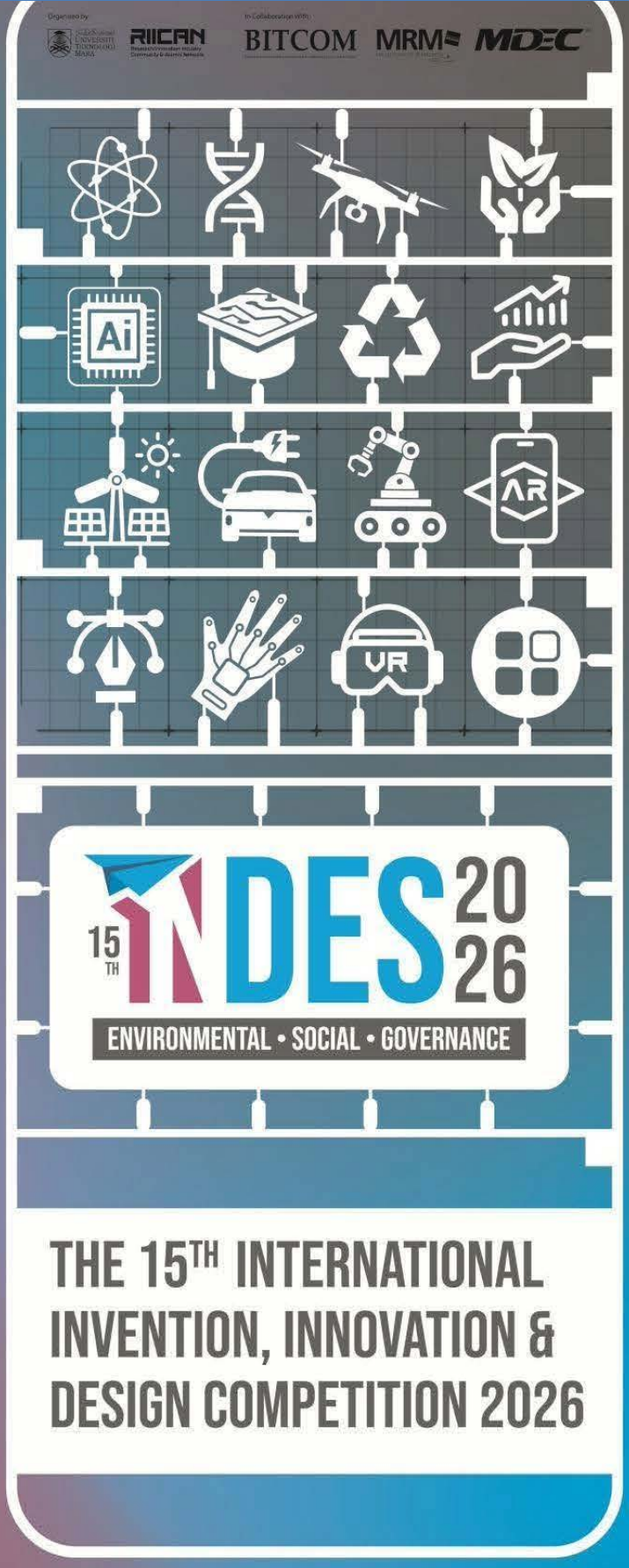


E-BOOK OF

EXTENDED ABSTRACT

E-BOOK OF EXTENDED ABSTRACT

THE 15th INTERNATIONAL
INVENTION, INNOVATION &
DESIGN COMPETITION 2026

Organized by:
Office of Research, Industry,
Community & Alumni Network
UiTM Perak Branch

© Unit Penerbitan UiTM Perak, 2026

All rights reserved. No part of this publication may be reproduced, copied, stored in any retrieval system or transmitted in any form or by any means; electronic, mechanical, photocopying, recording or otherwise; without permission on writing from the director of Unit Penerbitan UiTM Perak, Universiti Teknologi MARA Perak Branch.

Unit Penerbitan UiTM Perak,
Universiti Teknologi MARA,
Perak Branch, 32610
Seri Iskandar Perak, Malaysia.

Perpustakaan Negara Malaysia Cataloguing in Publication Data

No e- ISBN 978-967-2776-67-3

Cover Design: Asrul Nizuan Ariffin

Typesetting : Georgia

EDITORIAL BOARD

Editor-in-Chief

Muhd Syahir Abdul Rani

Managing Editors

Muhammad Wajihuddin Johari
Paul GnanaSelvam A/L Pakirnathan
Nazirul Mubin Mohd Noor
Nurdiyana Mohamad Yusof
Halimatussadiah Iksan

Copy Editors

Sheema Liza Idris
Nurul Munirah Azamri
Zarlina Mohd Zamari
NoorAileen Ibrahim
Farahidatul Akmar Awaludin
Siti Syairah binti Fakhruddin
Noor Fazzrienee JZ Nun Ramlan
Madaha binti Hanafi @ Mohd Ghani
Halimatun Saadiah Abd Mutalib
Hanitha A/P Satthiananthan
Syaza Kamarudin
Iza Faradiba Mohammad Patel
Nuramira Anuar
Noorlinda Alang
Noraziah Azizan
Faridatul Akma Wan Mohd Rashdi
Azurawati Zaidi
Ong Elly
Noran Nur Munira Khalili
Amirah Athirah Amir Yazid
Nurshahirah Azman
Daljeet Singh Sedhu A/L Janah Singh
Budiman Sabri Ahmad
Amirul Farhan Ahmad Tarmizi
Sarah Sakinah Assa'adah Johari

PRICEVISION: THE PIONEERING MACHINE LEARNING FRAMEWORK FOR THE MALAYSIAN SHOP PRICE INDEX (MSPI)

Intan Faiqah Hamizah Mohd Firazan[1], Junainah Mohamad[2*], Najma Azman[3], Mohd Nasurudin Hasbullah[4], Jufiza A. Wahab[5]

[1,2,3,4]Faculty of Built Environment, Universiti Teknologi MARA, Perak Branch, Seri Iskandar Campus, Malaysia
[5]Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, Perak Branch, Tapah Campus, Malaysia

mjunainah@uitm.edu.my

ABSTRACT

The dynamics of shop property prices are critical to Malaysia's commercial real estate performance. However, no comprehensive index currently exists to measure this sector. This research introduces the Malaysian Shop Price Index (MSPI), a pioneering data-driven framework developed in collaboration with the Valuation and Property Services Department (JPPH). The study transitionally shifts from traditional Hedonic models, which face linear constraints, to advanced Machine Learning (ML) analytics. Using a robust dataset of 20,702 transactions (2013–2024) from Kuala Lumpur and Selangor, the research identified 22 key physical, locational, and legal attributes influencing value. After rigorous data cleaning and Z-score outlier removal, 18,199 valid cases were analyzed. Variance Inflation Factor (VIF) analysis ensured the removal of multicollinearity, maintaining a strict threshold (VIF < 5). Six ML algorithms which are Linear Regression, Decision Tree, Random Forest, Support Vector Regressor (SVR), XGBoost, and Multi-Layer Perceptron (MLP) were evaluated. Results revealed that Random Forest and XGBoost achieved near-perfect accuracy (R^2 =0.997), significantly outperforming legacy Laspeyres Hedonic models. Launched in February 2026, the MSPI serves as Malaysia's first intelligent indicator for commercial property trends. By integrating ML into national property data, this innovation enhances market transparency and provides a superior tool for evidence-based policy and valuation decisions.

Keywords: MSPI, machine learning, JPPH, commercial valuation, real estate analytics

1. INTRODUCTION

Property price indices serve as vital economic barometers, functioning as essential benchmarking instruments for financial stability, investment performance, and national policy evaluation. In the banking sector, these indices act as a crucial barometer for monitoring asset prices and mitigating risks associated with collateral valuation. While Malaysia has successfully established the Malaysian House Price Index (MHPI) since 1997 and various rental indices, a significant gap remains in the commercial sector, specifically for shop properties. Commercial real estate transactions are historically thin and irregular, posing significant methodological and data-governance challenges for traditional indexing. Currently, the industry and national property bodies rely heavily on the traditional Hedonic Pricing Method. However, this legacy model faces critical limitations, including linear constraints, model specification complexity, multicollinearity, heteroscedasticity, and an inability to accurately capture complex, non-linear market behaviors (Bourassa et al., 2025; Malpezzi, 2024).

In aligning with the national strategic plans mandating the adoption of Industry 4.0 automation technology, this study pioneers the transition toward digital transformation in property valuation. By integrating advanced Machine Learning (ML) analytics (Yin et al., 2021), this research seeks to overcome the constraints of legacy models through the introduction of PriceVision, a pioneering ML-driven framework used to develop the Malaysian Shop Price Index (MSPI). Created in strategic

collaboration with the Valuation and Property Services Department (JPPH) and funded by the NAPREC grant, this represents Malaysia's first ML-based commercial property index. Officially launched in February 2026, this pioneer index provides a robust, intelligent, and highly accurate framework to drastically enhance market transparency and support evidence-based policy and valuation decisions for stakeholders.

2. METHODOLOGY

The methodology for developing the Malaysian Shop Price Index (MSPI) utilizes a structured quantitative approach to transition from traditional hedonic models to advanced Machine Learning (ML) analytics. The study focuses on Kuala Lumpur and Selangor, utilizing a comprehensive dataset of 22,702 commercial shop transaction records sourced from NAPIC. To ensure the integrity of the data and market authenticity, rigorous cleaning was conducted to remove non-arm's-length sales, missing and incomplete fields, properties that are five stories and above, non-terrace shops, and general outliers. Additionally, Z-score analysis was applied for specific outlier removal, yielding a finalized, robust dataset of 18,199 valid cases for analysis.

Following data preparation, the research identified 22 key variables categorized into physical, locational, and legal attributes that influence property value. Physical factors encompass the position of the shop, type of lot, condition of the shop, build-up size, and land area. Locational attributes consist of the street layer, classification of location, and condition of location. Legal factors include tenure, purchaser status, land use, and first transfer status. To ensure a parsimonious model and mitigate redundancy, Variance Inflation Factor (VIF) analysis was performed to eliminate multicollinearity, strictly maintaining a threshold of $VIF < 5$.

The computational phase involved parallel testing of six machine learning models to identify the best predictive performance. The algorithms evaluated were Linear Regression (OLS), Decision Tree Regressor, Random Forest Regressor, Support Vector Regressor (SVR), XGBoost Regressor, and Multi-Layer Perceptron (MLP). Finally, the index construction was executed using hedonic regression ML predictions, establishing 2018 as the Base Year with an index value of 100 to serve as a stable reference point.

3. FINDINGS

The experimental phase evaluated the transition from traditional indexing to an intelligent, data-driven framework, establishing the Malaysian Shop Price Index (MSPI) as a high-precision tool for commercial real estate monitoring. Comparative analysis revealed significant performance variations between traditional and advanced modelling approaches. Traditional Ordinary Least Squares (OLS) Linear Regression achieved only moderate accuracy, recording an R^2 of 0.4520, a Root Mean Square Error (RMSE) of 0.623, and a Mean Absolute Percentage Error (MAPE) of 12.21%. In contrast, ensemble-based machine learning algorithms demonstrated exceptional predictive power. Specifically, the Random Forest model achieved near-perfect accuracy with an R^2 of 0.9973, an RMSE of 0.031, and a MAPE of 1.12%. The Extreme Gradient Boosting (XGBoost) model performed similarly well, yielding an R^2 of 0.9970, an RMSE of 0.033, and a MAPE of 1.18%. These findings confirm the robust precision of ensemble-based models in predicting shop property prices over legacy linear models. The construction of the MSPI utilized hedonic regression ML predictions with 2018 established as the base year (Index = 100). Both the Random Forest and XGBoost models perfectly replicated the actual market index trends with negligible forecast errors. When compared with conventional indices, legacy Laspeyres Hedonic models exhibited the highest discrepancy due to their rigid fixed-base-year structure and limited ability to account for temporal market adjustments. Figure 1 shows the actual versus predicted price index against chained and Laspeyres hedonic.

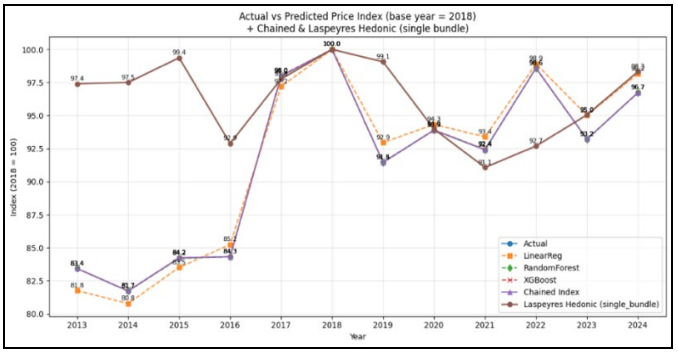


Figure 1 Actual vs predicted price index against chained and Laspeyres Hedonic.

Furthermore, the MSPI provides critical granular market insights as shown in Figure 2 regarding shop property price changes across specific districts. Based on the 2024 index overview against the 2018 base year, the Petaling region demonstrated the most significant commercial growth, recording an index of 140.7 (+40.7%). Other regions indicating strong positive market conditions include Klang at 121.6 (+21.6%) and Sepang at 121.4 (+21.4%). Conversely, Kuala Langat was the only district to indicate a market decline, with its index dropping to 83.9 (-16.1%). Ultimately, these results confirm that integrating Machine Learning analytics into national property data provides a robust, responsive, and intelligent indicator of both spatial and temporal price movements in Malaysia's dynamic commercial landscape.

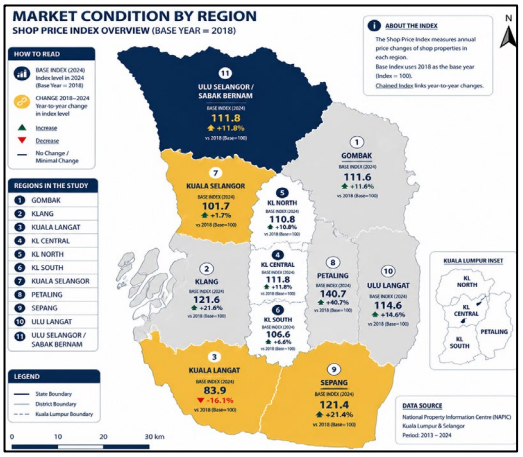


Figure 2 Market condition by region.

4. CONCLUSION

The findings confirm that integrating Machine Learning analytics into national property data provides a more robust and responsive representation of price movements. The established MSPI serves as an essential benchmark for JPPH, property valuers, and policymakers to support evidence-based decisions in Malaysia's dynamic commercial landscape. Beyond its current implementation, the PriceVision framework presents significant commercialisation potential, including seamless integration with the JPPH national property database and utilization across valuation firms, banks, and investment platforms. Future work will focus on expanding the MSPI to nationwide coverage, developing a real-time analytics dashboard, incorporating other commercial property types, and exploring hybrid machine learning and deep learning models to further advance real estate data intelligence.

ACKNOWLEDGEMENT

This research was funded by the National Real Estate Research Coordinator (NAPREC) grant under RMK-12 [Grant No.: 100-TNCPI/GOV 16/6/2 (050/2024)]. We also extend our appreciation to Universiti Teknologi MARA (UiTM) and the Valuation and Property Services Department (JPPH) for their administrative and technical support.

REFERENCES

Bourassa, S. C., Hoesli, M., Mayer, M., & Stalder, N. (2025). Reflections on hedonic price modeling. *Journal of European Real Estate Research*.

Malpezzi, S. (2003). Hedonic pricing models: A selective and applied review. *Housing economics and public policy*, *1*, 67-89.

Yin, L., Recker, W., & Chen, A. (2021). Improving property valuation with machine learning: Evidence from the real estate market. *Computers, Environment and Urban Systems*, *85*, 101549.

E-BOOK OF EXTENDED ABSTRACT THE 15th INTERNATIONAL INVENTION, INNOVATION & DESIGN COMPETITION 2026

eISBN 978-967-2776-67-3

9 789672 776673

Date of Publication: 13 July 2026

UNIT PENERBITAN UITM PERAK

(online)