**Model Performance**

Bảng dưới đây so sánh hiệu suất của một số mô hình học máy chính trong dự đoán giá nhà:

| **Mô hình** | **MAE** | **RMSE** | **R²** |
| --- | --- | --- | --- |
| Linear Regression | ~11,500 | ~14,500 | ~0.72 |
| Random Forest | ~9,000 | ~11,500 | ~0.81 |
| XGBoost | ~8,000 | ~10,200 | 0.88 |
| ANN | ~8,500 | ~11,000 | ~0.85 |

Theo các nghiên cứu, các mô hình boosting như XGBoost thường có hiệu suất vượt trội so với mô hình cơ bản như Linear Regression. Cụ thể, Ho et al. (2021) [1] và Pai et al. (2020) [2] đã chỉ ra rằng các mô hình học máy hiện đại có khả năng xử lý dữ liệu phức tạp và mối quan hệ phi tuyến tốt hơn.

**Research Gaps and Future Directions**

Dù đã đạt được nhiều tiến bộ, một số thách thức vẫn còn tồn tại:

* **Thiếu dữ liệu phi cấu trúc:** Chưa khai thác đầy đủ hình ảnh, văn bản hay dữ liệu địa lý.
* **Khó giải thích mô hình:** Một số thuật toán (như Random Forest, XGBoost) cần tích hợp các kỹ thuật giải thích như SHAP hoặc LIME để minh bạch hóa quyết định của mô hình.
* **Chưa tích hợp yếu tố thời gian:** Giá nhà có xu hướng biến động theo thời gian, tuy nhiên hầu hết các nghiên cứu chưa xem xét yếu tố này.

**Conclusion**

Tổng hợp các nghiên cứu cho thấy rằng các mô hình học máy, đặc biệt là các thuật toán boosting, mang lại hiệu quả dự đoán cao trong bài toán định giá nhà. Tuy nhiên, cần có thêm các nghiên cứu để khai thác dữ liệu phi cấu trúc, tăng cường khả năng giải thích và tích hợp yếu tố thời gian vào mô hình.

**References (Part 3)**

**[1]. Ho, W.K.O., Tang, B.S., Wong, S.W. (2021) Predicting property prices with machine learning algorithms. *Journal of Property Research*, 38(1), 48-70.  
[2]. Pai, P.F., Wang, W.C. (2020) Using machine learning models and actual transaction data for predicting real estate prices. *Applied Sciences*, 10(17), 5832.**