Analysis on monitor model performance degradation in production.

When deploying a machine learning model into production, it is important to continuously monitor its performance to **detect degradation** over time. This can happen due to **data drift** (changes in the input features such as property types or land sizes) or **concept drift** (changes in the housing market itself, e.g., price inflation). To address this, I would regularly log predictions and compare them with actual house sale prices once available. Metrics such as **MAE, RMSE, and R^2** would be recalculated periodically to check if the error rates are increasing compared to the baseline established during training.

In addition to monitoring prediction quality, I would also track the **integrity of input data** by validating ranges, checking for missing values, and detecting new categorical values that were not present during training. Automated alerts would be set up to flag unusual changes in feature distributions or when error metrics exceed a set threshold. If performance drops significantly, the model would be retrained with the latest data to adapt to new market conditions. This ensures the system remains reliable, accurate, and aligned with real-world changes.

By combining automated monitoring, periodic evaluation, and retraining strategies, the deployed house price prediction model can remain robust and accurate over time, ensuring it continues to deliver reliable insights for real-world decision-making.

Github Repository Link: Wyse10/Npontu