Housing Price Forecasting

BY Yurii Spizhovyi

MIT Emerging Talent

Why This Project?

- Housing affordability is a national issue in Canada
- Accurate price forecasts can help people plan better
- Time-series forecasting is a core Data Science skill I want to master
- Real housing datasets (CREA, StatCan, CMHC) make this a meaningful project

The Problem

Which forecasting approach predicts housing prices most accurately?

Traditional vs modern:

- ARIMA
- Prophet
- LSTM

The Hypothesis

- LSTM will perform best because it captures nonlinear patterns
- Prophet will perform well for seasonality
- ARIMA will be stable but less accurate for long horizons

Research Included

- Reviewing DS literature on time-series forecasting
- Studying housing market dynamics
- Learning models architectures
- Exploring StatCan, CMHC, BoC, CREA datasets

Testing Method

All models were evaluated fairly using the same time windows and metrics.

Methods used:

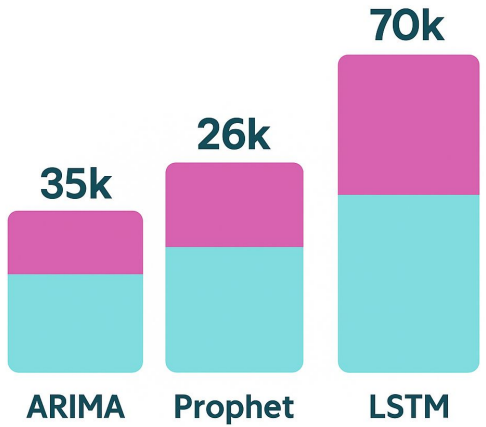
- Built a unified ETL pipeline
- Extracted features from 2005–2025 housing datasets
- Split data: Train (2005–2020) → Test (2021–2025)
- Trained ARIMA, Prophet, and LSTM on identical inputs
- Saved predictions to Postgres
- Compared MAE, RMSE, MAPE, R^2

Experiment data

Home price forecasts Vancouver

- ARIMA produced moderate error (~35k)
- Prophet achieved lower error (~26k)
- LSTM showed very high error (~70k)

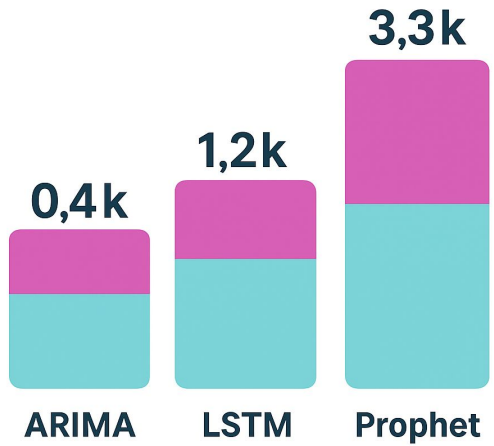
These values are based on Mean Absolute Error (MAE) computed over all horizons for Vancouver home price forecasts.



Experiment data

Rent price forecasts Vancouver

- ARIMA ≈ 0.4k — best performance
- LSTM ≈ 1.2k — moderate accuracy
- Prophet ≈ 3.3k — highest error

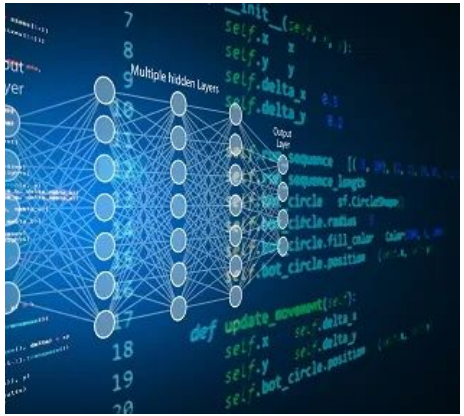


These values are based on Mean Absolute Error (MAE) computed over all horizons for Vancouver rent price forecasts.

Discoveries

What I learned:

1. Prophet performed best for home price forecasts.
2. ARIMA delivered the best rent forecasts and worked well for short horizons.
3. LSTM struggled because the dataset was too small for deep learning.
4. Best model differs slightly by city and target (price vs rent)



Key Takeaway

Prophet currently provides the best balance of accuracy and stability for Canadian housing forecasts.

Conclusion

- The experiment validated that different models behave differently depending on horizon
- **Prophet** + **ARIMA** outperform **LSTM** with current data size
- Results support using **Prophet** for early product prototypes

Next Steps:

- Expand dataset (more cities, more features)
- Extend dataset with rental listings, mortgage cohorts, and construction data
- Improve LSTM with architecture tuning
- Develop Hybrid-LSTM model combining macro indicators and micro market signals

Thank you!