

# AirBnB Price Recommender

**An Interpretable Deep Learning  
Approach to Price Recommendation**

The Erdős institute Deep Learning Bootcamp Project (Fall 2025)

**Arvind Suresh, Junyu Ma, Reilly McDowell**



[Github](#)



[WebApp](#)

# The host's pricing dilemma



1. Recommendations are “black boxes” (*why this price?*)
2. Comparable listings are “black boxes” (*why these listings?*)
3. External tools are expensive!
4. Airbnb’s internal tool tends to lowball prices

# Our solution: An Intuitive Price Formula

**Recommended price = Neighborhood average x Feature Multipliers**

Feature	Multiplier	Description	Price
Neighborhood base price	-	Average for Bedford-Stuyvesant	\$95.00
Location	x 1.15	+ 15% premium for being near a park	\$109.25
Capacity	x 1.08	+ 8% premium for an extra bed	\$118.00
Quality	x 1.20	+ 20% premium for Superhost & great reviews	\$141.60
Amenities	x 0.95	- 5% discount for lacking central A/C	\$134.52
Description	x 1.02	+ 2% premium for a well-written title	\$137.21
Seasonality	x 0.90	- 10% discount for being an off-peak month	\$123.49
Final Recommended Price			\$123.50

# The Feature Pipeline

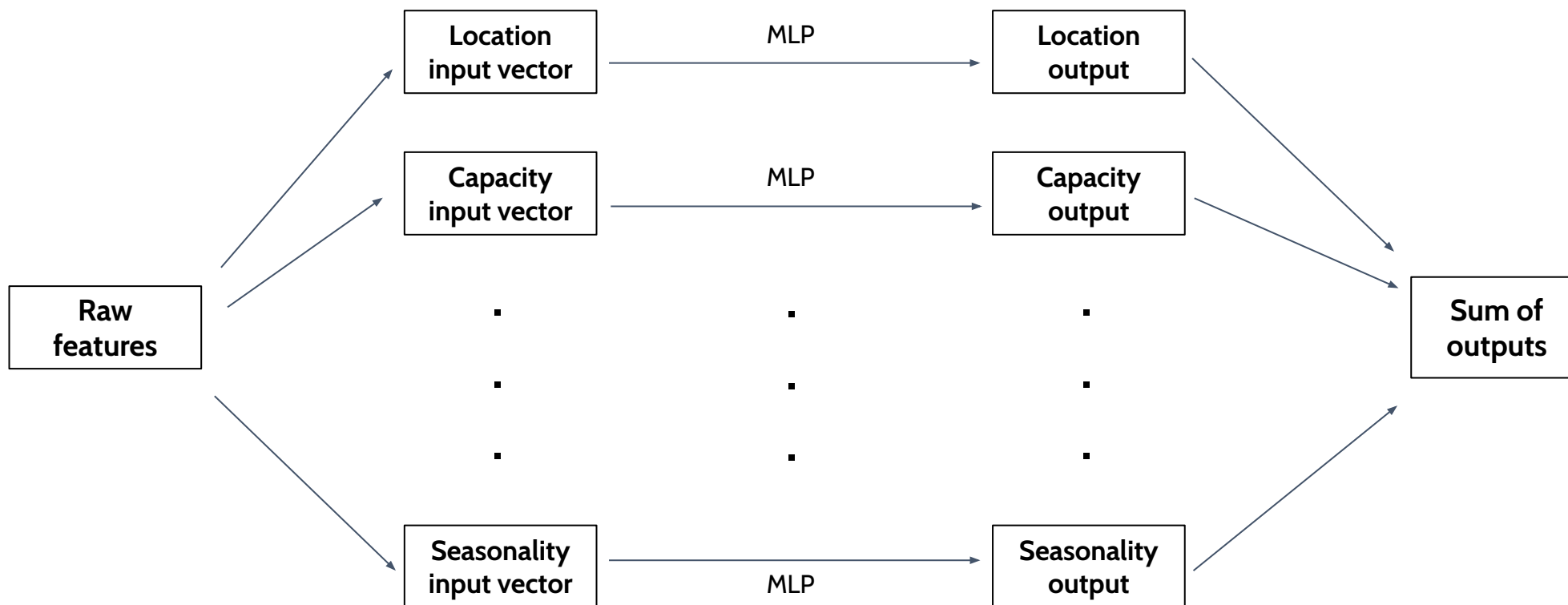
## From Raw Data to Model Inputs

Feature group	Features	Processing technique	Input vector size
Neighborhood base price	-	Pre-computed	-
Location	Latitude & Longitude	Cyclical Positional Encoding	32
Capacity	Accommodates, Bedrooms, Beds, Room/Property Type	One-hot encoding, Standardization	16
Quality	Review Scores, Superhost Status, Total Reviews	Standardization	8
Amenities	Cleaned Amenity Text (e.g., "wifi, pool, kitchen, ...")	Sentence Transformer (pre-trained, fine-tuned head)	384
Description	Raw Listing Description Text	Sentence Transformer (pre-trained, fine-tuned head)	384
Seasonality	Month (1-12)	Cyclical Sine/Cosine Transform	2

# Interpretable Additive NN Architecture

Target: Deviation from log of neighborhood average

Loss function: Mean-squared error



# Data Size and KPI

City	Data Size
NYC	~120k
Toronto	~80k

Metric	Definition	Purpose
MAPE (Price)	% error between true and predicted price	Host-friendly, intuitive accuracy
RMSE (Log Deviation)	Error on log(price) – neighborhood mean	Measures accuracy of multiplicative model

# Explainability Without Compromise

NYC results	Random Forest Baseline	Deep Learning Baseline	Interpretable Additive Model
Validation MAPE (Price)	29.40%	27.89%	27.30%
Validation RMSE (log deviation)	0.3625	0.3305	0.3449

Toronto results	Random Forest Baseline	Deep Learning Baseline	Interpretable Additive Model
Validation MAPE (Price)	31.08%	26.71%	28.91%
Validation RMSE (log deviation)	0.3676	0.3319	0.3440

## Conclusion & Future Work

- **Main Statement:** We created a tool that is not just accurate, but also transparent and trustworthy.
- **Future Work:**
  - Live Airbnb Integration
  - Incorporate Review Text Analysis



# THANK YOU

We gratefully acknowledge:

- **Inside Airbnb**

For providing the open-access dataset that made this project possible.

- **Lindsay Warrenburg**

- **Marcos Ortiz**

For leading the Quant ML Bootcamp and providing foundational instruction that enabled this work.