Technical Report: Uncertainty vs. Capacity in Spatial Housing Price Modeling

A Rigorous Comparison of Hierarchical Bayesian Models and Deep Neural Networks

Pan Zhichao | January 2026

1. ABSTRACT

This study investigates the trade-off between model capacity and interpretability in the context of spatial housing price prediction. We compare a **Multi-Slope Hierarchical Bayesian Model** against a **Deep Neural Network (MLP)** baseline. Contrary to common intuition, our rigorous cross-validation (5-Fold, 3 Seeds) reveals that simple Linear Regression models (RMSE: 0.499 ± 0.017) marginally outperform over-parameterized Neural Networks (RMSE: 0.531 ± 0.021) on this dataset. This finding underscores the importance of Occam's Razor in **Small Tabular Regimes** and highlights the unique value of Bayesian methods for uncovering spatial heterogeneity.

2. RELATED WORK

Recent benchmark studies have consistently shown that deep learning struggles on small-to-medium tabular datasets. **Grinsztajn et al. (2022)** and **Shwartz-Ziv & Armon (2022)** demonstrated that tree-based methods often outperform neural networks due to locality inductive biases. While existing literature compares deep learning to tree-based methods, few studies directly compare neural networks to Bayesian hierarchical models on spatial data. Our work fills this gap by validating the "Small Tabular Regime" hypothesis in a spatial context.

3. METHODOLOGY

3.1 Data & Schema

The dataset consists of California housing prices with spatial coordinates (N=20,640, subsampled to 2,000 to simulate a strict small-data regime and evaluate algorithmic data efficiency). We employ a strict **Schema-Driven Architecture** to ensure all models consume identical features. The target is standardized Median House Value.

3.2 Models

- **Hierarchical Bayesian Model:** Multi-slope partial pooling using NUTS sampler. Varies slopes for income, age, and rooms by spatial cluster.
- **PyTorch MLP:** 3-Layer Perceptron (64 → 32 → 1) with ReLU activations and Dropout.
- **Linear Regression:** Standard OLS baseline.

3.3 Experimental Protocol

To rule out "lucky seeds," we implemented a rigorous evaluation framework: Stratified 5-Fold Cross-Validation repeated with 3 distinct random seeds (n=15 independent runs).

4. RESULTS

Table 1: Cross-Validated Performance Comparison (Lower is better).

Rank	Model	RMSE (Mean ± Std)	95% CI	Notes
1	**Linear Regression**	**0.499 ± 0.017**	[0.490, 0.508]	**Best Generalization**
2	PyTorch MLP	0.531 ± 0.021	[0.520, 0.542]	Signs of Overfitting
3	Spatial Emb NN	0.566 ± 0.025	[0.553, 0.579]	Over-parameterized

4.1 Statistical Significance

The performance gap between Linear Regression and MLP is $\Delta \approx 0.032$. Given the standard deviations, this difference is statistically significant (Cohen's d ≈ 1.6). In this specific data regime, the extra "capacity" of the Neural Network proved to be a liability rather than an asset.

5. DISCUSSION & CONCLUSION

The Triumph of Simplicity: The fact that Linear Regression outperforms the MLP suggests that the underlying relationship is largely linear within the data range. The Neural Network likely overfits to noise in the small training set.

The Value of Bayesian Inference: While the predictive performance of linear models is superior, the Hierarchical Bayesian model provides unique insights. It identifies distinct spatial clusters where housing prices decouple from income, offering causal-adjacent interpretability that black-box models miss.

> ***Implication:*** *For high-stakes spatial decision making, we advocate for a "Bayesian-First" workflow. Establish a transparent probabilistic baseline before resorting to black-box methods.*

REFERENCES

- Grinsztajn, L., Oyallon, E., & Varoquaux, G. (2022). Why do tree-based models still outperform deep learning on tabular data? *arXiv preprint arXiv:2207.08815*.
- Shwartz-Ziv, R., & Armon, A. (2022). Tabular data: Deep learning is not all you need. *Information Fusion*, 81, 84-90.
- Gelman, A., & Hill, J. (2006). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press.