Extended Abstract - Theory-based residual neural networks: A synergy of discrete choice models and deep neural networks

Shenhao Wang*, Baichuan Mo, Jinhua Zhao

Massachusetts Institute of Technology
Cambridge, MA 02139

Keywords: demand modeling, travel behavior, deep neural networks.

1. Introduction

As machine learning (ML) is increasingly used in the transportation field, we observe a tension between data-driven ML methods and classical theory-driven methods. Take travel behavior research as an example: researchers can analyze travel mode choice by using discrete choice models (DCMs) under the framework of random utility maximization (RUM), or using data-driven methods such as ML classifiers without any substantial behavioral understanding. This tension creates practical difficulty in choosing one method over the other, and prevents scholars from tackling travel behavior problems under a unified framework. However, a closer examination reveals that the two methods are complementary in terms of prediction, interpretation, and robustness, prompting us to ask how to synergize them rather than treating them as disparate or even conflicting methods. Deep neural networks (DNNs) and DCMs can be complementary because the former are more predictive [6, 5, 7, 8, 3], but less interpretable and robust [2, 9, 6, 9, 1], while the latter are less predictive, but more interpretable and robust.

2. Methodology

To address the aforementioned challenge, this study designs a theory-based residual neural network (TB-ResNet) that synergizes DNNs and DCMs, demonstrating that this synergy is not only feasible but also desirable, leading to a simultaneous improvement in prediction, interpretation, and robustness. We first demonstrate that DNNs align with the RUM framework by briefly recounting McFadden (1974) and Wang et al. (2020) [10, 11]. Second, we present the TB-ResNet framework, which augments DNNs to DCMs to fit the utility residuals with a $(\delta, 1-\delta)$ formulation, resembling the essence of the standard residual network (ResNet) [4], as shown in Figure 1. The TB-ResNet

*Corresponding author: shenhao@mit.edu

framework consists of a DCM utility function $V_{T,k}(z_i, \tilde{x}_i)$ and a DNN utility function $V_{DNN,k}(z_i, \tilde{x}_i)$ weighted by δ and 1-δ:

$$v_{TB-ResNet,ik} = (1-\delta)v_{T,ik} + \delta v_{DNN,ik} = (1-\delta)V_{T,k}(z_i, \tilde{x}_i) + \delta V_{DNN,k}(z_i, \tilde{x}_i) \quad (1)$$

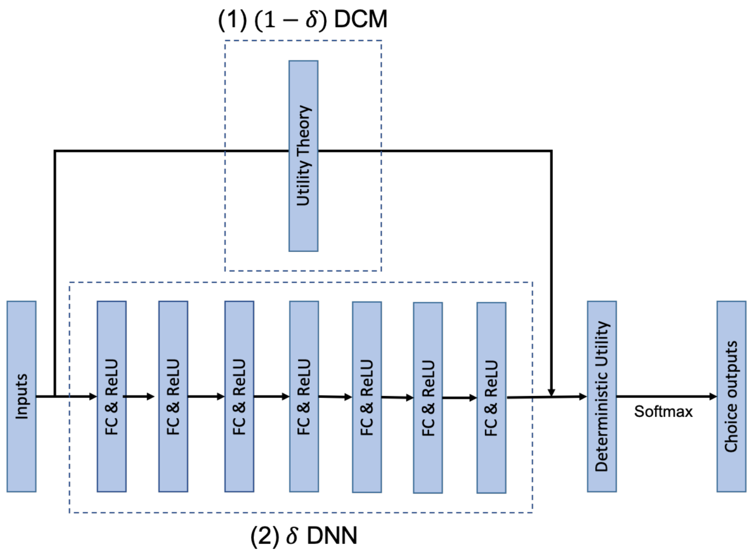


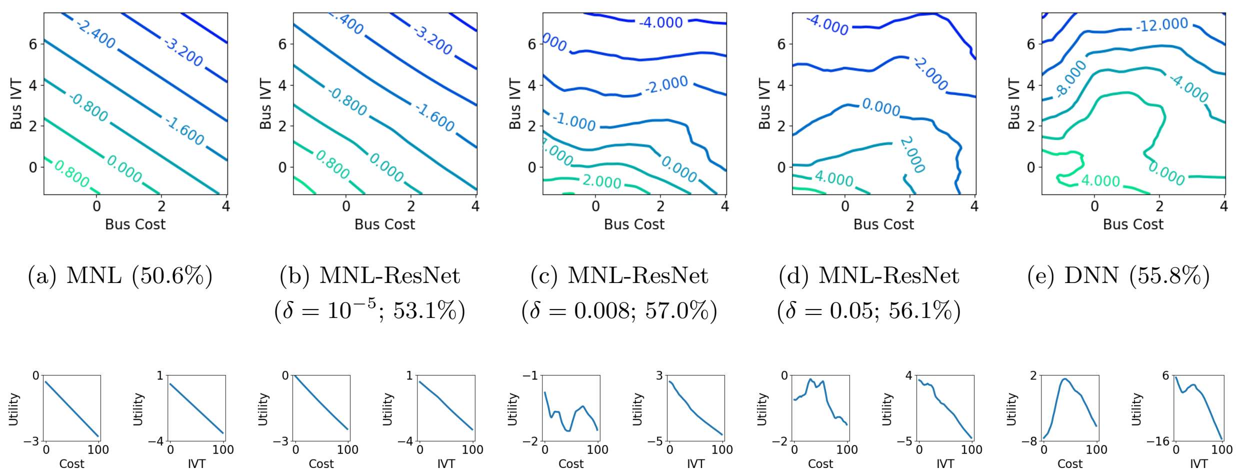
Fig. 1. Architecture of TB-ResNet. Both DCM and DNN are flexible: the DNN block uses seven layers as an example, but it can be any depth or width; the DCM block can take any utility specification under the RUM framework.

This TB-ResNet framework can be understood from six interwoven perspectives: architecture design, model ensemble, gradient boosting, regularization, flexible function approximation, and theory diagnosis. The regularization perspective can be formally demonstrated by using the state-of-the-art statistical learning theory to illustrate the intuition that DNNs tend to be too complex to capture the reality and DCMs tend to be too simple to do so. Then we design three instances of TB-ResNets using multinomial logit models (MNL-ResNets), prospect theory (PT-ResNets) for risk preference, and hyperbolic discounting (HD-ResNets) for time preference, showing that the simple TB-ResNet framework can incorporate a wide range of DCMs that are part of the utility maximization framework.

3. Results

3.1. Utility functions of TB-ResNets as combination of DCMs and DNNs

Figures 2 visualizes how utilities vary with input values. We can observe the complementary nature of DNNs and DCMs by comparing only the two graphs of DNNs and DCMs on the right and left ends of Figures 2. On one hand, the utility functions of the MNL model are very regular and intuitive, as shown by subfigures 2a. In subfigures 2f and 2g, the utility values of choosing the bus mode linearly decrease as bus costs and in-vehicle travel time increase. These highly regular utility functions in DCMs are interpretable, although it is also likely that the true utility functions can be much more complex than the smooth and regular MNL, leading to their misspecification errors and underfitting. However, on the other hand, the utility functions of DNNs for the MNL scenario are very irregular and highly counter-intuitive, although they have higher prediction accuracy, as shown by subfigure 2e. For example, in Figure 2n, the DNN predicts that the utility of using buses first increases as the travel cost increases, violating the basic principle of economics theory. Overall, it is critical to observe the complementary nature of DNNs and DCMs: DCMs might be too simple and regular to capture reality, while DNNs might be too complex and irregular to do so. TB-ResNets achieve a flexible compromise between DCMs and DNNs, the degree of which is controlled by δ, as shown in Figures 2b, 2c, and 2d.



(a) MNL (50.6%) (b) MNL-ResNet ($\delta = 10^{-5}$; 53.1%) (c) MNL-ResNet ($\delta = 0.008$; 57.0%) (d) MNL-ResNet ($\delta = 0.05$; 56.1%) (e) DNN (55.8%)

(f) Cost (g) IVT (h) Cost (i) IVT (j) Cost (k) IVT (l) Cost (m) IVT (n) Cost (o) IVT

Fig. 2. Utility functions of MNL-ResNets, MNL, and DNNs. Upper row: visualization of 2D utility functions, and percentages in the parentheses represent the prediction accuracy. Lower row: visualization of 1D utility functions, and every pair of figures on the lower row correspond to the figure directly above on the upper row.

3.2. *Prediction, Interpretation, and robustness of TB-ResNets over DCMs and DNNs*

As summarized in Table 1, our empirical results can also demonstrate that TB-ResNets are more predictive, interpretable, and robust overall than pure DCMs and DNNs, although several exceptions exist. Compared to DNNs, TB-ResNets are more predictive, interpretable, and robust because the DCM component in TB-ResNets can stabilize the utility functions and regularize the DNN component. Compared to DCMs, TB-ResNets are more predictive and interpretable because richer utility functions are augmented to the skeleton DCM by the DNN component in TB-ResNets. The TB-ResNets are formulated with a flexible $(\delta, 1-\delta)$ weighting, thus taking advantage of both the simplicity of the DCMs and the richness of the DNNs, preventing the underfitting of the DCMs and the overfitting of the DNNs, and providing insights into the completeness of the DCM theories. Our findings are consistent across the three scenarios (MNL, PT, and HD). While exceptions exist in the PT scenario and the comparison to DCMs for robustness evaluation, our main findings hold from both theoretical and empirical perspectives.

Models	**Prediction**	**Interpretability**	**Robustness**
Compared to DNNs	Marginal improvement (by stabilization and regularization)	Significant improvement (by stabilization and regularization)	Significant improvement (by stabilization and regularization)
Compared to DCMs	Significant improvement (by augmenting and enriching utility function)	Significant improvement (by augmenting and enriching utility function)	No improvement

Table 1: Comparison of TB-ResNets to DCMs and DNNs

4. Conclusion

Researchers often treat data-driven and theory-driven models as two disparate or even conflicting methods in travel behavior analysis. However, the two methods are highly complementary because data-driven methods are more predictive but less interpretable and robust, while theory-driven methods are more interpretable and robust but less predictive. Using their complementary nature, this study designs a theory-based residual neural network (TB-ResNet) framework, which synergizes discrete choice models (DCMs) and deep neural networks (DNNs) based on their shared utility interpretation. The TB-ResNet framework is simple, as it uses a (δ, 1-δ) weighting to take advantage of DCMs' simplicity and DNNs' richness, and to prevent underfitting from the DCMs and overfitting from the DNNs. This framework is also flexible: three instances of TB-ResNets are designed based on multinomial logit model (MNL-ResNets), prospect theory (PT-ResNets), and hyperbolic discounting (HD-ResNets), which are tested on three data sets. Compared to pure DCMs, the TB-ResNets provide greater prediction accuracy and reveal a richer set of behavioral mechanisms owing to the utility function augmented by the DNN component in the TB-ResNets. Compared to pure DNNs, the TB-ResNets can modestly improve prediction and significantly improve interpretation and robustness, because the DCM component in the TB-ResNets stabilizes the utility functions and input gradients.

Overall, this study demonstrates that it is both feasible and desirable to synergize DCMs and DNNs by combining their utility specifications under a TB-ResNet framework. This study highlights a synergetic perspective and a comprehensive model evaluation based on three criteria, so future studies could take the complementarity of the data-driven and theory-driven methods beyond simple prediction comparison. We hope that this work can pave the way for future studies to create more links between the data-driven and theory-driven methods, because their complementary nature provides immense opportunities, their underlying perspectives are interwoven, and their synergy can overcome their respective weaknesses.

References

[1] Finale Doshi-Velez and Been Kim. "Towards a rigorous science of interpretable machine learning". In: (2017).

[2] Alex A Freitas. "Comprehensible classification models: a position paper". In: *ACM SIGKDD explorations newsletter* 15.1 (2014), pp. 1–10.

[3] Edward L Glaeser et al. "Big data and big cities: The promises and limitations of improved measures of urban life". In: *Economic Inquiry* 56.1 (2018), pp. 114–137.

[4] Kaiming He et al. "Deep residual learning for image recognition". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016, pp. 770–778.

[5] Matthew G Karlaftis and Eleni I Vlahogianni. "Statistical methods versus neural networks in transportation research: Differences, similarities and some insights". In: *Transportation Research Part C: Emerging Technologies* 19.3 (2011), pp. 387–399.

[6] Sotiris B Kotsiantis, I Zaharakis, and P Pintelas. "Supervised machine learning: A review of classification techniques". In: *Emerging artificial intelligence applications in computer engineering* 160 (2007), pp. 3–24.

[7] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. "Imagenet classification with deep convolutional neural networks". In: *Advances in neural information processing systems*. 2012, pp. 1097–1105.

[8] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. "Deep learning". In: *Nature* 521.7553 (2015), pp. 436–444.

[9] Zachary C Lipton. "The mythos of model interpretability". In: *arXiv preprint arXiv:1606.03490* (2016).

[10] Daniel McFadden. "Conditional logit analysis of qualitative choice behavior". In: (1974).

[11] Shenhao Wang, Baichuan Mo, and Jinhua Zhao. "Deep neural networks for choice analysis: Architecture design with alternative-specific utility functions". In: *Transportation Research Part C: Emerging Technologies* 112 (2020), pp. 234–251. ISSN: 0968-090X.