Residual Neural Network Learning Approach for Behaviour Choice Modelling

Melvin Wong[1], Bilal Farooq[1,2]

[1]Laboratory of Innovations in Transportation (LITrans), Department of Civil Engineering, Ryerson University, Toronto, Canada

[2]Canada Research Chair in Disruptive Transportation Technologies and Services

Introduction

- Machine learning and DNNs have emerged as powerful modelling strategies (Wong and Farooq, 2020), understanding how we can leverage on these strategies effectively is one of the primary motivations in choice modelling.
- We present a Residual Logit (ResLogit) model for integrating a data-driven Deep Neural Network (DNN) architecture into a travel choice model.
- Our proposed method can be formulated as a Generalized Extreme Value (GEV) model – This is based on a deep learning strategy known as Residual Neural Networks.

Hypothesis

- Residual skipped connections between layers enables unobserved random heterogeneity in decision process to be modelled effectively and resolves the shortcomings in machine learning for behaviour choice modelling.
- Our proposed ResLogit model allows for richer set of higher order substitution patterns than a standard logit might be able to achieve.

Background

- Enhancing discrete choice models with DNNs and deep learning algorithms is one of the areas of research that have shown promising results (Wong et al., 2018, Sifringer et al., 2018; Borysov et al., 2019).
- It has been observed that increasing the number of layers beyond a certain threshold will degrade the model due to overfitting, unreachable optimal solutions, and model identification problems (Glorot et al., 2011).
- Model interpretability, identification tests and the applicability of machine learning algorithms has not yet been clearly justified in travel behaviour modelling applications.
- Recent work has shown that the Residual Neural Network strategy significantly improves optimization in deep neural network architectures with marginal or no loss in performance.

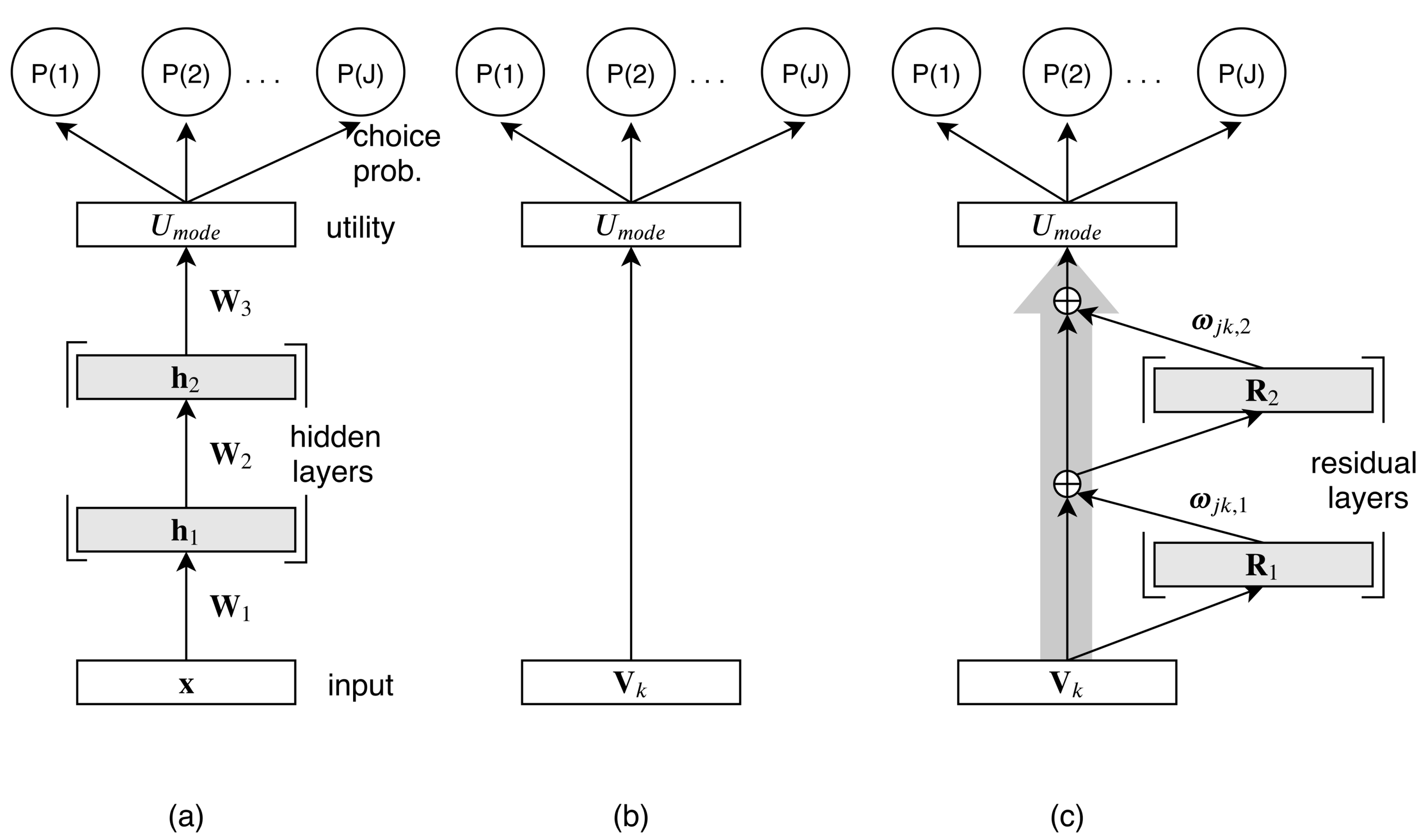

Figure 1: Model framework (a) MLP-DNN model with 2 hidden layers, (b) MNL model, (c) our proposed ResLogit model

Objectives

- Propose a practical extension of discrete choice travel behaviour modelling using a ResNet architecture.
- Define the GEV structure and generating function for our ResLogit model.
- Analyze the properties of our proposed approach using two SP/RP dataset, and evaluate the performance on mode choice prediction and value of time estimation.
- Show how a conventional feedforward DNN model fails to capture the underlying random heterogeneity in the data, and why a residual model structure works better in practice.

Methodology

ResLogit GEV model specification:

$$G(y_1,\dots,y_j) = \sum_{j\in\mathcal{C}} \Big(\Big[\prod_{m=1}^{M} R_{j,m} \Big] y_j^{\mu} \Big),$$

Derivative of the generating function G:

$$\frac{\partial G}{\partial y_j} = G_j(y_1,\dots,y_j) = \mu \Big[\prod_{m} R_{j,m} \Big] y_j^{\mu-1}.$$

Choice probability equation:

$$P(j|\mathcal{C}) = \frac{y_j G_i}{\mu G},$$

$$= \frac{e^{\mu V_j + \sum_{m=1}^{M} \ln R_{j,m}}}{\sum_{j'\in\mathcal{C}} e^{\mu V_{j'} + \sum_{m=1}^{M} \ln R_{j',m}}} \quad \forall \quad M \geq 1.$$

Residual function:

$$R_{j,m} = \begin{cases} \dfrac{1}{1+\exp\Big(\sum_{k\in\mathcal{C}} \omega_{jk,m} V_k\Big)} & \text{if } m=1, \\ \dfrac{1}{1+\exp\Big(\sum_{k\in\mathcal{C}} \omega_{jk,m} (V_k + \ln R_{j,m-1})\Big)} & \text{if } m>1 \end{cases}$$

Case Study 1: 2016 Train Hotel SP travel survey dataset

Value of Time (VoT) analysis on a stated preference dataset. We compared our proposed model with a Mixed Logit and MNL model.

Utility equations:

Logit $U_{mode} = \beta_{cost} X_{mode\_cost} + \beta_{TT} X_{mode\_TT} + \varepsilon_{mode}$

Mixed Logit $U_{mode} = \bar{\beta}_{cost} X_{mode\_cost} + \bar{\beta}_{TT} X_{mode\_TT} + \sigma_{cost}\mathcal{N}(\bar{\beta}_{cost},1) + \sigma_{TT}\mathcal{N}(\bar{\beta}_{TT},1) + \varepsilon^{*}_{mode}$

ResLogit $U_{mode} = \beta_{cost} X_{mode\_cost} + \beta_{TT} X_{mode\_TT} - \ln(1+\exp(\boldsymbol{\omega}_{j,1}\cdot V_0)) + \varepsilon^{*}_{mode}$

	mode share	mean travel cost ($)	mean travel time (h)
mode			
car	0.239	115	7.26
car rental	0.021	179	6.29
bus	0.057	84	11.72
plane	0.056	178	4.04
train	0.060	48	6.12
TrH	0.567	164	13.78

Table 1: Train Hotel survey characteristics

Parameters	Logit	Mixed Logit	ResLogit
ASC_{car}	-2.53 (0.25)	-2.58 (0.26)	2.19 (2.84)
$ASC_{car\ rental}$	-1.7 (0.12)	-1.71 (0.12)	2.48 (4.61)
ASC_{bus}	-3.46 (0.23)	-3.6 (0.25)	6.18 (25.1)
ASC_{plane}	-2.36 (0.25)	-2.42 (0.27)	-0.62 (1.32)
ASC_{train}	-2.27 (0.21)	-2.38 (0.22)	1.3 (0.77)
ASC_{TrH}	ref.	ref.	ref.
β_{cost}	-0.58 (0.11)	-0.64 (0.12)	-0.48 (0.13)
β_{TT}	-0.072 (0.02)	-0.064 (0.024)	0.05 (0.01)
σ_{cost}		0.71 (0.22)	
σ_{TT}		0.2 (0.06)	
VoT (CAD$/hour)	**12.41**	**9.98**	**-9.75**
log-likelihood	-2060.1	-2058.3	-2036.6
sample size	1788	1788	1788
$\bar{\rho}^2$	0.29	0.295	0.302

Table 2: Estimated model parameters and VoT calculation, std. error in parenthesis

Parameters	Logit	Mixed Logit	ResLogit
ASC_{car}	-0.71 (0.37)	-0.67 (0.41)	0.18 (0.37)
$ASC_{car\ rental}$	-0.92 (0.19)	-0.89 (0.21)	3.81 (10.7)
ASC_{bus}	-1.26 (0.30)	-1.27 (0.31)	-1.26 (0.37)
ASC_{plane}	-0.91 (0.42)	-0.93 (0.53)	-1.01 (0.65)
ASC_{train}	ref.	ref.	ref.
β_{cost}	-2.62 (1.56)	-0.40 (0.22)	-0.10 (0.06)
β_{TT}	-0.134 (0.04)	-0.15 (0.06)	-0.02 (0.01)
σ_{cost}		0.64 (0.34)	
σ_{TT}		0.09 (0.10)	
VoT (CAD$/hour)	**5.11**	**38.20**	**20.59**
log-likelihood	-884.1	-843.3	-821.0
sample size	775	775	775
$\bar{\rho}^2$	0.214	0.213	0.216

Table 3: Estimated model parameters and VoT calculation **without TrH** mode, std. error in parenthesis

Case Study 2: 2016 Mtl Trajet RP travel survey dataset

Effects of model depth on performace metrics – ResLogit vs MLP-DNN

- Sample size: Training=42,256; Validation=18,109
- Accuracy: MNL (72.01%); MLP-DNN (70.95±1.62%); ResLogit (76.09%±0.61%)

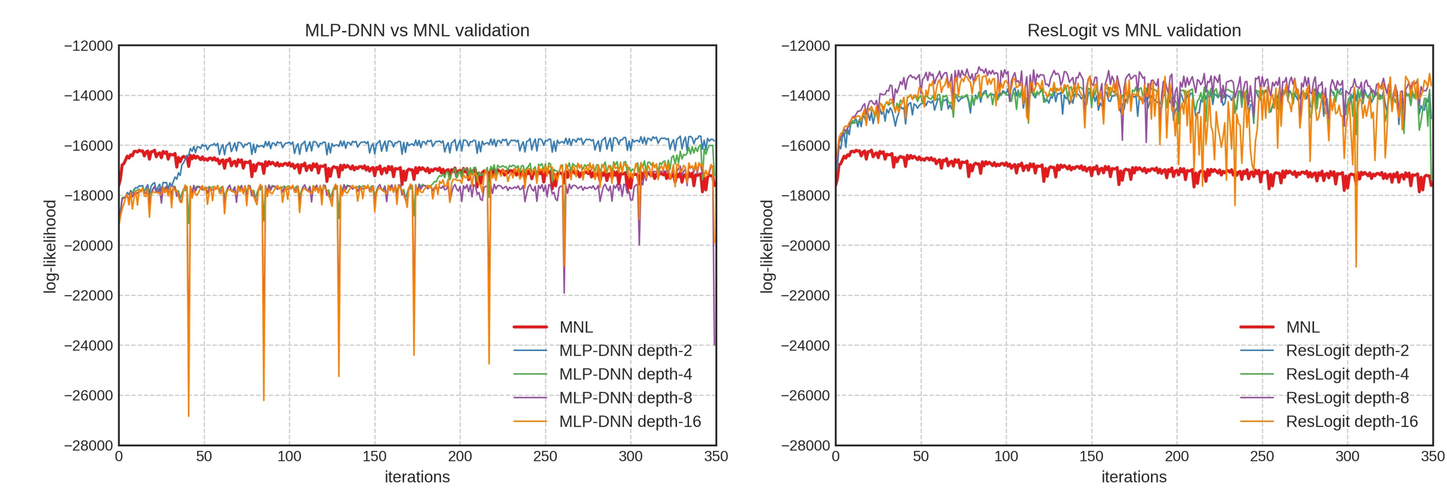

Figure 2: Model framework (a) MLP-DNN model with 2 hidden layers, (b) MNL model, (c) our proposed ResLogit model

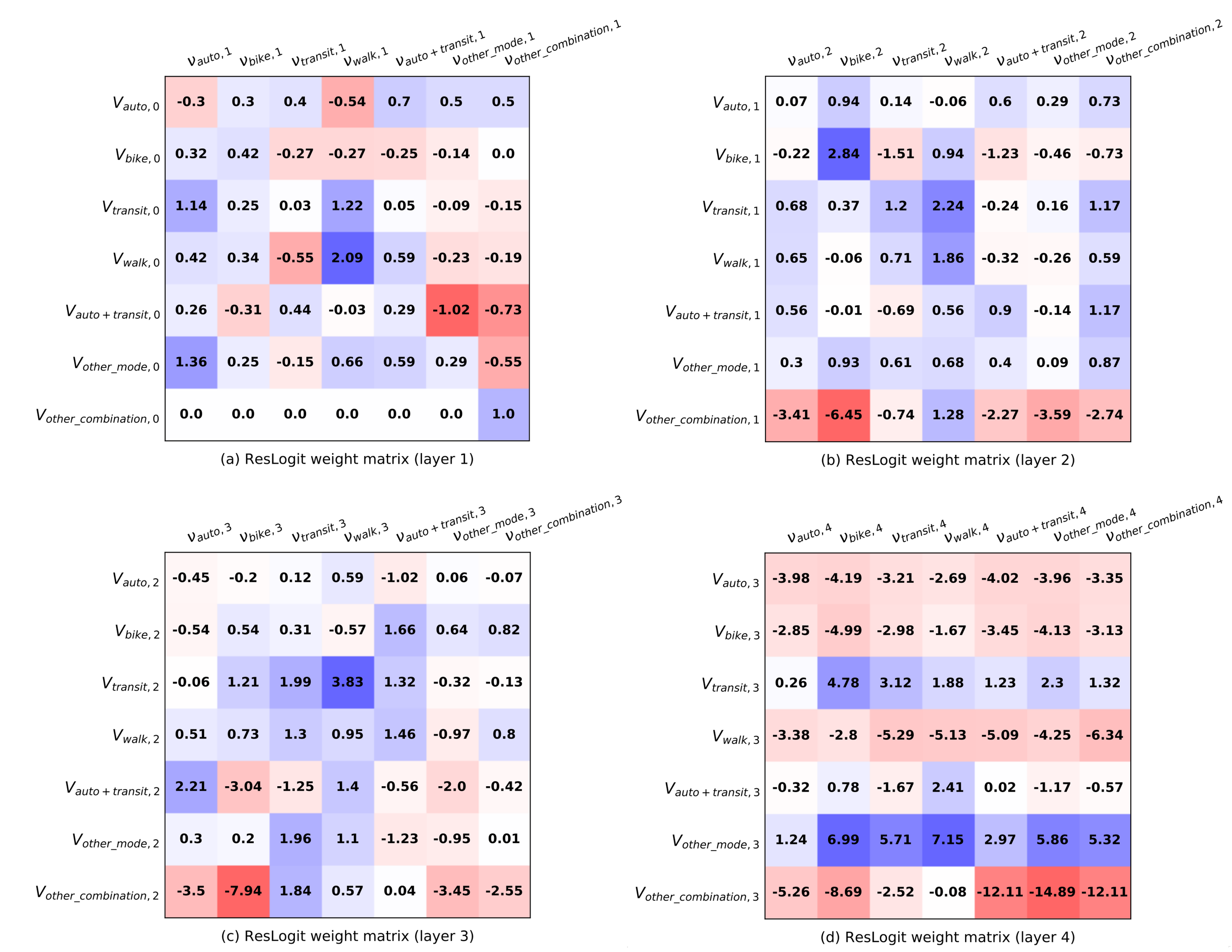

Figure 3: First 4 layers of weight matrix parameters from the ResLogit model (ResLogit depth-16)

Conclusion

- We proposed a ResLogit model using a variant of residual neural network and machine learning for discrete choice analysis.
- We are able to account for random heterogeneity while maintaining economic interpretability and without loss of model performance.
- Our method can be formulated as a GEV random utility maximization model.

References

[1] Wong, M., Farooq, B., 2020. A bi-partite generative model framework for analyzing and simulating large scale multiple discrete-continuous travel behaviour data. Transportation Research Part C: Emerging Technologies 110, p. 247—268.

[2] Wong, M., Farooq, B., Bilodeau, G.A., 2018. Discriminative conditional restricted boltzmann machine for discrete choice and latent variable modelling. Journal of Choice Modelling 29, p. 152—168.

[3] Borysov, S.S., Rich, J., Pereira, F.C., 2019. How to generate micro-agents? A deep generative modeling approach to population synthesis. Transportation Research Part C: Emerging Technologies 106, p. 73—97.

[4] Sifringer, B., Lurkin, V., Alahi, A., 2018. Enhancing discrete choice models with neural networks, in: Proceedings of the 18th Swiss Transportation Research Conference.

[5] Glorot, X., Bordes, A., Bengio, Y., 2011. Deep sparse rectifier neural networks, in: Proceedings of the 14th International Conference on Artificial Intelligence and Statistics, p. 315—323.