Modeling Cycling Route Choice Using Convolutional Neural Networks

Katrin Lubashevsky*[1,2], Iryna Okhrin[1,3], Stefan Huber[4], and Sven Lißner[4]

[1]Center for Interdisciplinary Digital Sciences (CIDS), Technische Universität Dresden, Germany
[2]Center for Information Services and HPC (ZIH), Dresden University of Technology, Germany
[3]Center for Scalable Data Analytics and Artificial Intelligence (ScaDS.AI) Dresden/Leipzig, Germany
[4]Chair of Transport Ecology, Institute of Transport Planning and Road Traffic, Faculty of Transport and Traffic Science "Friedrich List", Dresden University of Technology, Germany

Abstract

Bicycles offer significant potential for sustainable urban transport by reducing emissions and noise when replacing motorized traffic. Understanding cyclists' route choices is crucial for city planners, and traditionally is analyzed using multinomial logit (MNL) models. Recent advancements in machine learning (ML) show capability for greater predictive accuracy. This study explores potential of the ML modeling technique convolutional neural networks (CNN) for bicycle route choice modeling using GPS data from around 183,000 trips across six German cities. By comparing CNN with MNL, we evaluate their strengths, limitations, and implications for cycling infrastructure planning.

Keywords: Route choice, cycling, machine learning, convolutional neural network

1 Introduction

Bicycles offer immense potential for creating eco-friendly urban transport systems by reducing noise and pollutant emissions when replacing motorized transport for daily trips (Koska & Rudolph, 2016; Buehler & Pucher, 2021). Their popularity is growing across Germany, with more people using them for regular travel (Nobis, 2019). However, effective promotion of cycling requires well-developed infrastructure adapted to cyclists' needs. Limited municipal resources necessitate prioritizing measures based on data and informed planning.

Understanding cyclists' route choices is a crucial aspect of transportation planning. Data on cycling routes and the models derived from them can assist network planning (Huber et al., 2019). Traditional approaches rely on discrete choice theory models, such as multinomial logit (MNL), which effectively analyze factors affecting route selection and traffic demand. These models have long been a cornerstone in transport science (Ortúzar & Willumsen, 2011).

However, advancements in machine learning (ML) methods, with their high predictive accuracy, challenge the dominance of traditional models (Zhang & Xie, 2008; Stenneth et al., 2011; Zhao et al., 2020). ML models have demonstrated over 90% prediction accuracy in various transportation studies, outperforming established models like MNL (see Section 2). Despite this, ML's potential for analyzing cyclists' route preferences remains unexplored.

The aim of the presented study is to reveal the potential of machine learning methods for bicycle route choice analysis and modeling. We therefore examine existing research, first. Hence, we apply different an MNL and convolutional neural networks (CNNs), that are tailored to the structure of route choice data, on a large GPS data set containing around 183,000 trips from six different cities in Germany. In the last step, we compare the model results in terms of their predictive qualities, highlighting the strengths and limitations of ML models as well as the implications for bicycle planning.

2 Literature Review

Over the past few decades, numerous international studies leveraged GPS data to analyse bicycle route choice (e.g. Menghini et al., 2010; Broach et al., 2012; Casello & Usyukov, 2014; Khatri et al., 2016; Ton et al., 2017; Peng Chen & Childress, 2018; Scott et al., 2021; Huber et al., 2024).

A synthesis of their findings reveals consistent preference trends: longer distances and steeper slopes tend to deter route choice, while the presence of cycling infrastructure and well-maintained surfaces increases the likelihood of selecting a given route. Additionally, high traffic volumes and fast-moving motorized traffic reduce the probability of a route being chosen, as does a high density of intersections.

With advancements in machine learning (ML), methods from these fields have been increasingly incorporated into a wide range of research areas. Researchers in traffic science have also recognized the potential benefits of ML, leading to a significant increase in its application across various domains over the past decade. Some studies applied ML methods explicitly to route choice. Early research on route choice using ML methods dates back to the 1990s (e.g. Yang et al., 1993). However, most of these investigations focus on analyzing and modeling route choice behavior in motorized traffic. Nevertheless, the studies demonstrate that analyzing and modeling route choice behavior based on various influencing factors is feasible and valuable. Moreover, these studies have achieved high levels of accuracy (e.g. Sun & Park, 2017; Lai et al., 2018; Marra & Corman, 2021). Yang et al. (1993) used neural networks (NN) with stated preference (SP) data, achieving 93% accuracy. Lee et al. (2005) developed a genetic neural network combining NNs and genetic algorithms, achieving 97.3% accuracy, outperforming the MNL (75.2%). Dia & Panwai (2006) showed NNs significantly surpassed traditional models' accuracy (50–73%) with 95–97% accuracy. Sun & Park (2017) found support vector machines offered better efficiency and comparable accuracy to NNs. The results of Lai et al. (2018) indicate that ML models significantly outperformed traditional models in predictive accuracy and performance. Marra & Corman (2021) uses Deep NNs, more specifically CNNs, on GPS data to analyze route choice in public transport and comparing it to Path Size Logit (PSL), where the CNNs outperform the PSL in terms of prediction accuracy (75-97%).

Two studies specifically address bicycle route choice using machine learning methods. Meng & Zheng (2023) developed a bike-ability framework to create personalized cycling route recommendations. Their model incorporates factors like infrastructure, land use, and distance to calculate visual perceptibility and assess road network impedance. However, it does not utilize revealed preference data, such as GPS trajectories, to analyze actual route choices.

Sobhani et al. (2018) investigated the applicability of ML methods for modeling cyclists' route choices using a dataset from a household survey in Toronto, Canada. This dataset included GPS routes, trip purposes, and temporal factors enriched with road network attributes from Toronto GIS. The authors estimated several ML models, including Decision Trees and Random Forests, and compared their performance with that of a traditional Expanded PSL model. Unfortunately, their findings remain incomplete as only a conference abstract has been published, leaving details of their outcomes unavailable.

The literature highlights significant research on bicycle route choice, but most ML applications focus on motorized traffic. The studies of Meng & Zheng (2023) and Sobhani et al. (2018) demonstrate the potential of ML in this domain. Nonetheless, the limited number of studies addressing bicycle-specific route choice suggests that ML applications in cycling remain underexplored, warranting further research to harness their potential for improving cycling infrastructure and planning.

3 Data description

A route choice dataset was created using data from the Germany-wide „CITY CYCLING“ campaign (Lißner et al., 2023), in which participants are tracking their day-to-day cycling trips via a smartphone app for the period of three weeks between May and September. In this study, we consider data from the year 2022 of 6 German cities with different characteristics, see Table 1. For each tracked route, two alternative routes are generated to provide options that could have been chosen instead. This setup allows for a comprehensive analysis of the factors influencing route choice by comparing the selected route to its plausible alternatives for each origin-destination pair (Huber, 2022).

Secondary data, that contains characteristics of every route in the data set was gathered using OpenStreetMap. In this study, we consider a total of thirty variables that describe the route properties. These variables describe the route distance, its slope properties, the existence and type of cycling infrastructure, the share of feeder roads along the route, speed limits for motorized traffic and the properties of surface. Besides the distance variable, all the variables are referring to the share of the route, that possesses these properties. The distance variable can be considered

Attribute	Darmstadt	Dresden	Freiburg	Munich	Offenburg	Weimar
Area in km^2	122.07	328.48	153.04	310.7	78.37	84.47
Population	162,243	563,311	236,140	1,512,491	61,670	65,620
Population class	large	large	large	large	medium	medium
Topography class	flat	hilly	hilly	flat	hilly	hilly
Number of trips	19,825	32,364	41,862	70,995	14,652	3718

Table 1: Characteristics of the six cities, including the number of recorded cycling trips.

in two ways: absolute distance (in meters) or a distance ratio, where the longest route within a route triple gets assigned value one and the other ones are divided by the longest route. In this study, only these distance ratios are included and not the absolute distance, such that all the input values are between zero and one. A full list of the input variables can be found in the appendix in Table 2.
For MNL the variables are pre-selected, such that only nine of the thirty are considered. This preselection is necessary for several reasons, particularly when modeling complex behaviors like route choice. Including too many, especially irrelevant or collinear variables, increases the risk of overfitting the training data, while pre-selection focuses on essential variables to capture true relationships and improve generalization. Moreover, pre-selection simplifies the model, highlighting key variables like route distance or slope for more precise insights.

4 Models

In the following section, two different modeling techniques are described: MNL and CNN. All the models are trained on n data pairs $\left(\{\mathbf{X}^{(1)}, \mathbf{y}^{(1)}\}, \dots, \{\mathbf{X}^{(n)}, \mathbf{y}^{(n)}\}\right)$, where $\mathbf{X}^{(k)}$ is the k-th input matrix and $y^{(j)}$ the respective label, which are structured as follows:

$$\mathbf{X}^{(k)} = \begin{bmatrix} x_{1,1}^{(k)} & x_{1,2}^{(k)} & \dots & x_{1,p}^{(k)} \\ x_{2,1}^{(k)} & x_{2,2}^{(k)} & \dots & x_{2,p}^{(k)} \\ x_{3,1}^{(k)} & x_{3,2}^{(k)} & \dots & x_{3,p}^{(k)} \end{bmatrix}, \quad \mathbf{y}^{(k)} = \begin{pmatrix} y_1^{(k)} \\ y_2^{(k)} \\ y_3^{(k)} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \tag{1}$$

Each input matrix has three rows: the first row contains p variables for the tracked route, while the next two rows represent its generated alternatives. The corresponding output $\mathbf{y}^{(k)}$ is a 3-dimensional label vector, where 1 indicates the tracked route and 0 indicates the generated alternative routes indicating which route was tracked.
The prediction performance of the models is evaluated using accuracy, which is calculated as the proportion of cases where the predicted choice matches the actual choice. Cross-validation is performed to prevent overfitting by splitting the whole dataset into 70% training data and 30% test data. The models are trained using the training set, and their accuracy is evaluated on the test set. Additionally, accuracy can be calculated on the training data to measure the goodness of fitting.

Multinomial Logit

Established route choice models, are based on the random utility maximization theory (see e.g. McFadden & Train, 2000), where it is assumed, that every choice that is made by a person, aims to maximize an individual utility, considering all the information for existing alternatives. The utility of the k-th routes triple associated with i-th possible route is defined as follows:

$$U_i^{(k)} = \sum_{j=1}^{p} \beta_i x_{i,j}^{(k)} + \varepsilon_i^{(k)}, \tag{2}$$

where $x_{i,j}^{(k)}$ is an entry of the input matrix $\mathbf{X}^{(k)}$ from (1), $\beta = (\beta_1, \dots, \beta_p)$ is a vector of regression coefficients, that should be estimated and a random term $\varepsilon_i^{(k)}$ that aims to address the fact, that route choice made by humans, are sometimes influenced by factors, that are not represented within the data on which the model is trained. To obtain the regression coefficient, the utilities of the chosen routes need to be maximized while the utilities of the alternatives are minimized.

The logistic regression aims to model the probabilities of the tracked route and its alternatives. The probability for choosing a route i within the k-th triple of route alternatives is then modeled by the multidimensional logistic function:

$$P(i|X^{(k)}) = \frac{e^{U_i^{(k)}}}{\sum_{j=1}^{3} e^{U_j^{(k)}}} . \tag{3}$$

The route with the largest probability is predicted as the chosen one. The regression coefficients β are estimated using maximum likelihood estimation. Together with respective significance measures, these coefficients provide impact and the importance of each input variable. This study focuses on the comparison of models by their accuracy, and we refer to *Huber et al. 2025* for the investigation of the route variables influence.
For implementation of the MNL the Python package *xlogit* (Arteaga et al., 2022) was used.

Convolutional Neural Networks

CNNs are developed for image classification problem (LeCun et al., 1998, 1989; Broach et al., 2012), which allows matrix input data. This is not the case for simple feed-forward neural networks, and considering that the input of the route choice data is also in matrix form, it is quite intuitive to utilize a CNN for this use case, too.
The first CNN-architecture aims to mimic the MNL model suggested by Marra & Corman (2021). It has the form of Figure 1. Besides the input layer, this CNN has one convolutional layer with

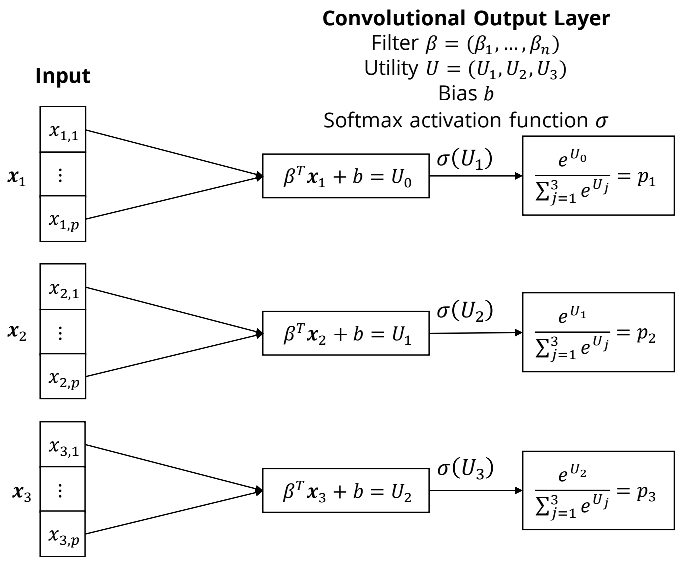


Figure 1: Multinomial logit in the form of a one-layer convolutional neural network (short: CNN-1C).

an one-dimensional filter β, which is equivalent to the regression coefficient in the MNL model, a bias b, and a softmax-activation function σ, which is the same logistic function, that is used for multinomial logistic regression. That way, the objective function, resulting from the CNN, has an equivalent structure as the MNL. The only difference is the calculation of the hyperparameters, which are determined by backpropagation in the CNN model, (see e.g. LeCun et al., 1998).
Using this one-layer CNN (later on referred to as CNN-1C) as a starting point, the neural network structure can be expanded in many different ways, e.g. by adding additional layers or by using more than one filter vector β. In this study, we created three more CNN structures. For the first variation of the CNN a dense layer was added after the convolutional layer (CNN-1C1D) and for the second variation a second convolutional layer with four filters was added between the input and the existing convolutional layer (CNN-2C). In a third version, both of these layers are added, resulting in a total of three layers (CNN-2C1D). The detailed CNN structures can be seen in Figure 4.
So far, the in- and output are ordered such that for every data pair $\{\mathbf{X}^{(k)}, \mathbf{y}^{(k)}\}$ the first row represents the chosen route and the other entries are alternatives that were not chosen. With this data structure, these slightly more complex networks are easily able to learn, that the upper row

always gets label one. This is why the data needs to get shuffled, such that the model does not simply learn the order. To prevent that, the in- and output pair gets randomly shuffled, such that it exemplarily could look like the following:

$$\mathbf{X}^{(k)} = \begin{bmatrix} x_{2,1}^{(k)} & x_{2,2}^{(k)} & \dots & x_{2,p}^{(k)} \\ x_{3,1}^{(k)} & x_{3,2}^{(k)} & \dots & x_{3,p}^{(k)} \\ x_{1,1}^{(k)} & x_{1,2}^{(k)} & \dots & x_{1,p}^{(k)} \end{bmatrix}, \quad \mathbf{y}^{(k)} = \begin{pmatrix} y_2^{(k)} \\ y_3^{(k)} \\ y_1^{(k)} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Such randomly shuffled data pairs are used for training and testing of all the CNN versions. The CNNs are implemented in Python using the TensorFlow interface (Abadi et al., 2015).

5 Empirical Results and Discussion

For each of the six cities, all the models described in Section 4 are trained independently. Due to the randomness in the training process, each CNN architecture is trained one hundred times to account for variability in the resulting models. This approach enables a thorough analysis of accuracy distribution across multiple runs, providing insights into the model's consistency and performance variability. The accuracy empirical distributions are visualized using violin plots in Figure 2, with the red dashed line indicating the accuracy of MNL for comparison.
Two key trends emerge from the violin plots for accuracy in Figure 2. First, the CNN-1C and CNN-2C models perform surprisingly worse than the MNL, whereas the CNN architectures that include a dense layer slightly outperform the MNL up to Weimar, a medium sized city with the smallest number of observed tracked routes. In case of the large cities the accuracy of the CNNs with a dense layer is approximately 1.6% higher than that of the MNL. On the other hand if only the best models from these 100 models per CNN architecture was picked, it can be seen, that these CNN-1C1D and CNN-2C1D models have slightly better accuracies (up to 3.2% improvement). In conclusion, future studies should explore more complex CNN architectures and variations in depth, encompassing both shallow and deep designs. Second, models for CNN-1C1D and CNN-2C1D achieve the best results for Munich, the city with the most collected trips, and worst for Weimar, the city with the smallest number of collected trips. This coincides with the known fact that complex models like CNN perform well for data with many observations.
The results for the medium-sized cities of Offenburg and Weimer are different from those for the large cities and require further analysis. The variation in terms of prediction quality is very large, especially for the models, that have a dense layer (CNN-1C1D and CNN-2C1D), which at the same time also contain the best models. This can be explained by the rather small sample size (14,652 for Offenburg and 3718 for Weimar, whereas the other cities range between 19,800 and almost 71,000, see Table 1). Besides that, in the study of Huber et al. (2025), it was already explained that both cities had lower significance levels for some variables in the MNL models than the other cities, which was reasoned by a higher heterogeneity in the data.
That is why a closer look was taken at the training accuracies of these two model architectures for Offenburg and Weimar. In practice, a model that already has a bad fit or accuracy on the training data would not be considered for further prediction tasks. Following this reasoning, all models with a training accuracy greater than 66% are re-examined. The large cities are not affected by this filtering step, but for the medium cities Offenburg and Weimar, some models get excluded: for Offenburg two models are filtered out for CNN-1C1D, one for CNN-2C1D and for Weimar 53 models for CNN-1C1D and 26 for CNN-1C1. Based on that filtering, Figure 3 shows the prediction accuracy of only the models whose training accuracy was above 66%. The additional filtering step is not meant to distort the results. Instead, the intention behind this was to show how important it might be to inspect the training quality before moving on to testing the model. This showed that for future investigations, it is recommended to conduct such a filtering step.
While this study demonstrates that CNNs can achieve higher prediction accuracy, the improvement is less than expected. Previous studies (see Section 2) have shown that ML techniques have the potential to achieve accuracy exceeding 90%. Nonetheless, there remains significant room for enhancement, particularly through optimizing the number of layers and filters in the models. Future work could delve deeper into these aspects while also addressing the need to optimize training and prediction runtimes, potentially by leveraging highly parallelized algorithms.
This paper primarily focuses on developing novel model architectures and evaluating their predictive capabilities rather than examining the influence of input variables. A comprehensive analysis of the input variables has already been conducted in the study by Huber et al. (2025), which analyzes the same dataset. Therefore, the regression coefficients and significance levels derived

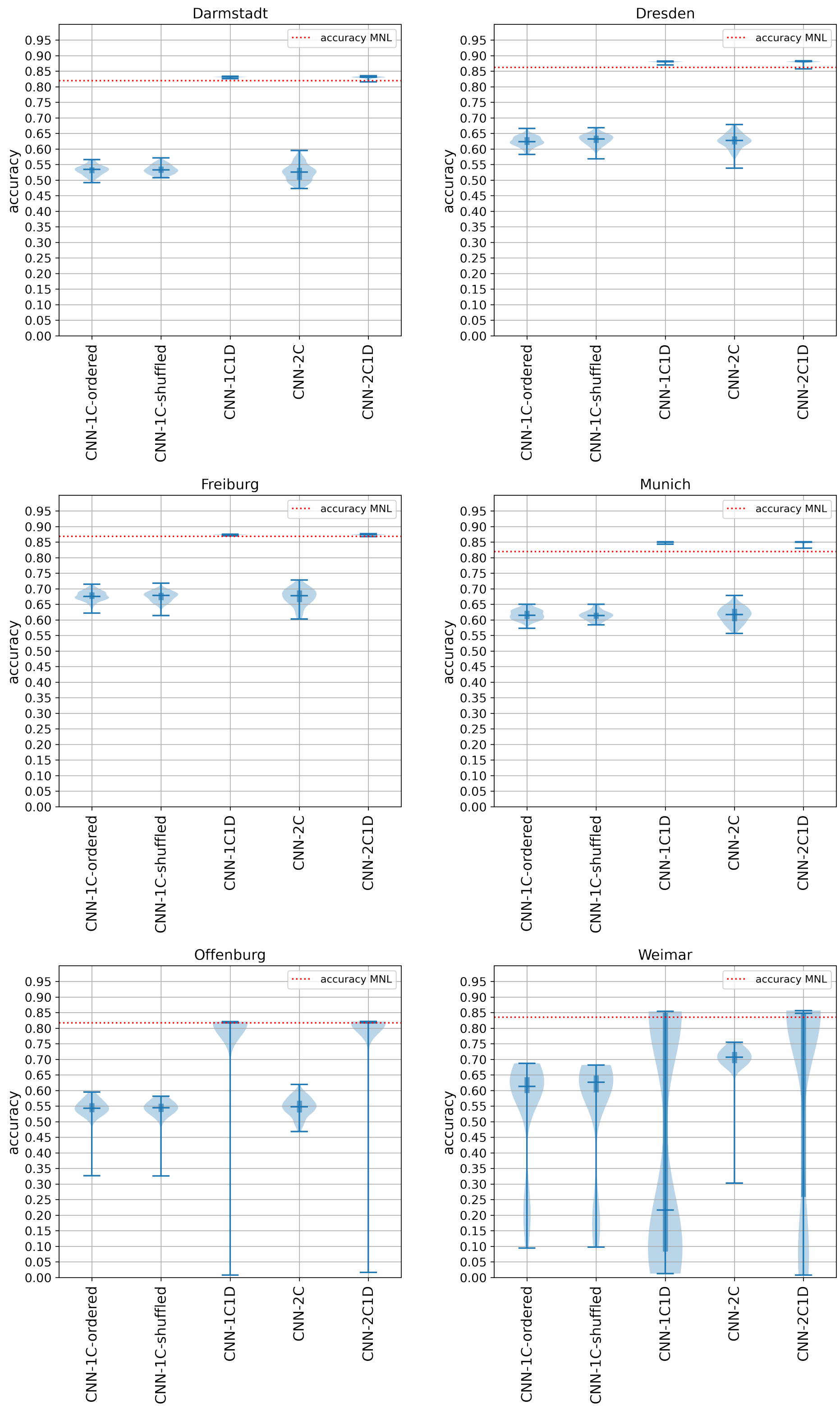


Figure 2: Violin plots of accuracy for different CNN architectures based on 100 fits. The red dashed lines represent the accuracy of MNL.

from the MNL model are not discussed here. However, future research will aim to enhance the explainability of machine learning methods, such as CNNs, enabling more meaningful comparisons with models like MNL in this context. One further idea for future studies is to include personal attributes into the models.

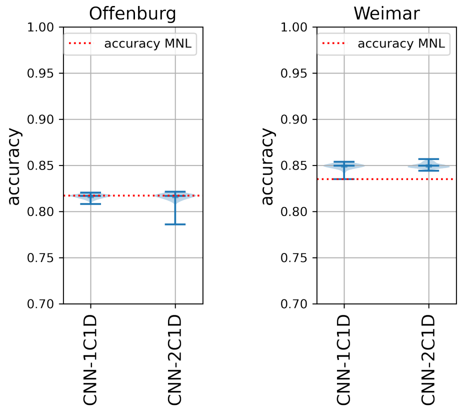


Figure 3: Violin plots for model quality of one filtered CNNs with dense layer

6 Conclusion

In this study multiple CNN architectures for modeling cycling route choice were presented and compared to MNL in terms of their prediction quality. The model architectures that included a dense layer showed the highest accuracy among the compared models. In the future research deeper CNN architectures should be investigated. It could also be deduced, that the models heavily depend on the amount and quality of the data, which was very apparent for the medium city Weimar. In order to prevent unnecessary predictive calculations, we recommend taking a closer look at the goodness of fit to the training data before proceeding to the testing step. Altogether, there is a high potential of CNNs to improve the state of the art in route choice modeling.

Acknowledgements

The authors gratefully acknowledge the willingness of all CITY CYCLING participants to donate their cycling data for the research project „Radverkehrsverhalten in Deutschland - RiD“.

Funding

The authors appreciate the funding by the Deutsche Forschungsgesellschaft DFG (German Research Association) within the project (project number: 514587991).

References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., … Zheng, X. (2015). *TensorFlow: Large-scale machine learning on heterogeneous systems.* (Software available from tensorflow.org)

Arteaga, C., Park, J., Beeramoole, P. B., & Paz, A. (2022). xlogit: An open-source python package for gpu-accelerated estimation of mixed logit models. *Journal of Choice Modelling*, *42*, 100339. doi: https://doi.org/10.1016/j.jocm.2021.100339

Broach, J., Dill, J., & Gliebe, J. (2012). Where do cyclists ride? a route choice model developed with revealed preference gps data. *Transportation Research Part A: Policy and Practice*, *46*(10), 1730-1740. doi: https://doi.org/10.1016/j.tra.2012.07.005

Buehler, R., & Pucher, J. (2021). *Cycling for sustainable cities.* The MIT Press.

Casello, J. M., & Usyukov, V. (2014). Modeling cyclists' route choice based on gps data. *Transportation Research Record*, *2430*(1), 155-161. doi: 10.3141/2430-16

Dia, H., & Panwai, S. (2006, 01). Comparative evaluation of models of route choice and driver compliance with traffic information. *29th Australasian Transport Research Forum, ATRF 06*.

Huber, S. (2022). *Analyse des routenwahlverhaltens von radfahrenden auf grundlage gps-basierter daten zum real beobachteten verkehrsverhalten* (Doctoral dissertation). doi: 10.13140/RG.2.2.13480.14087

Huber, S., Lindemann, P., & Schröter, B. (2024). Safety and bicycle route choice: To what extent do accident risk and perceived safety influence bicycle route choice? *Transportation Engineering*, *18*, 100240. doi: https://doi.org/10.1016/j.treng.2024.100240

Huber, S., Lißner, S., & Francke, A. (2019, 05). Utility of gps data for urban bicycle traffic planning in germany: Potentiality, limitations and prospects. *International Journal of Transport Development and Integration*, *3*, 1-14. doi: 10.2495/TDI-V3-N1-1-14

Huber, S., Lißner, S., Lindemann, P., & Lubashevsky, K. (2025). Different city = different cycling behaviour? a comparative analysis of cycling behaviour in german cities. *European Transport Research Review*. ((in publication))

Khatri, R., Cherry, C. R., Nambisan, S. S., & Han, L. D. (2016). Modeling route choice of utilitarian bikeshare users with gps data. *Transportation Research Record*, *2587*(1), 141-149. doi: 10.3141/2587-17

Koska, T., & Rudolph, F. (2016). *The role of walking and cycling in reducing congestion: A portfolio of measures.* (FLOW Project (2016). The Role of Walking and Cycling in Reducing Congestion: A Portfolio of Measures. Brussels)

Lai, X., Fu, H., Li, J., & Sha, Z. (2018). Understanding drivers' route choice behaviours in the urban network with machine learning models. *IET Intelligent Transport Systems*, *13*. doi: 10.1049/iet-its.2018.5190

LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., & Jackel, L. D. (1989). Backpropagation applied to handwritten zip code recognition. *Neural Computation*, *1(4)*, 541–551. doi: https://doi.org/10.1162/neco.1989.1.4.541

LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, *86(11)*, 2278–2324. doi: https://doi.org/10.1109/5.726791

Lee, S., Kim, Y., Namgung, M., & Kim, J. (2005). Development of route choice behavior model using linkage of neural network and genetic algorithm with trip information. *KSCE Journal of Civil Engineering*. doi: 10.1007/BF02829045

Lißner, S., Huber, S., & von Harten, M. (2023, 07). Influence of the cycling campaign city cycling on cycling behaviour in germany. In *European transport conference (etc) – milano.* doi: 10.26128/2023.65

Marra, A. D., & Corman, F. (2021). A deep learning model for predicting route choice in public transport [Conference Paper]. Ascona: STRC. doi: 10.3929/ethz-b-000504159

McFadden, D., & Train, K. (2000). Mixed mnl models for discrete response. *Journal of Applied Econometrics*, *15*(5), 447–470.

Meng, S., & Zheng, H. (2023). A personalized bikeability-based cycling route recommendation method with machine learning. *International Journal of Applied Earth Observation and Geoinformation*, *121*, 103373. doi: https://doi.org/10.1016/j.jag.2023.103373

Menghini, G., Carrasco, N., Schüssler, N., & Axhausen, K. (2010). Route choice of cyclists in zurich. *Transportation Research Part A: Policy and Practice*, *44*(9), 754-765. doi: https://doi.org/10.1016/j.tra.2010.07.008

Nobis, C. (2019). *Mobilität in deutschland – mid analysen zum radverkehr und fußverkehr.* (Studie von infas, DLR, IVT und infas 360 im Auftrag des Bundesministeriums für Verkehr und digitale Infrastruktur (FE-Nr. 70.904/15). Bonn, Berlin. www.mobilitaet-in-deutschland.de)

Ortúzar, J. d. D., & Willumsen, L. (2011). *Modelling transport, fourth edition.* John Wiley & Sons, Ltd. doi: 10.1002/9781119993308.fmatter

Peng Chen, Q. S., & Childress, S. (2018). A gps data-based analysis of built environment influences on bicyclist route preferences. *International Journal of Sustainable Transportation*, *12*(3), 218–231. doi: 10.1080/15568318.2017.1349222

Scott, D. M., Lu, W., & Brown, M. J. (2021). Route choice of bike share users: Leveraging gps data to derive choice sets. *Journal of Transport Geography*, *90*, 102903. doi: https://doi.org/10.1016/j.jtrangeo.2020.102903

Sobhani, A., Haji Esmaeli, S., & Sobhani, A. (2018). *On the use of machine learning approaches for implicit modeling of cycling route choice: An application of machine learning versus path sampling-logit model framework.* (IATBR 2018: 15th International Conference on Travel Behaviour Research, IATBR 2018 ; Conference date: 15-07-2018 Through 20-07-2018)

Stenneth, L., Wolfson, O., Yu, P. S., & Xu, B. (2011). Transportation mode detection using mobile phones and gis information. In *Proceedings of the 19th acm sigspatial international conference on advances in geographic information systems* (p. 54–63). New York, NY, USA: Association for Computing Machinery. doi: 10.1145/2093973.2093982

Sun, B., & Park, B. B. (2017). Route choice modeling with support vector machine. *Transportation Research Procedia*, *25*, 1806-1814. (World Conference on Transport Research - WCTR 2016 Shanghai. 10-15 July 2016) doi: https://doi.org/10.1016/j.trpro.2017.05.151

Ton, D., Cats, O., Duives, D., & Hoogendoorn, S. (2017). How do people cycle in amsterdam, netherlands?: Estimating cyclists' route choice determinants with gps data from an urban area. *Transportation Research Record*, *2662*(1), 75-82. doi: 10.3141/2662-09

Yang, H., Kitamura, R., Paul P. Jovanis, K. M. V., & Abdel-Aty, M. A. (1993). Exploration of route choice behavior with advanced traveler information using neural network concepts. *Transportation*, *20*. doi: 10.1007/BF01307059

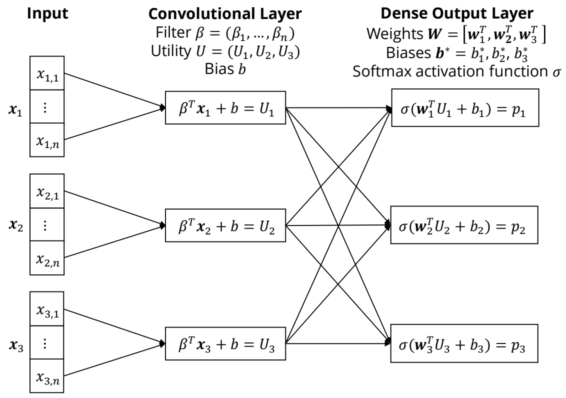
Zhang, Y., & Xie, Y. (2008). Travel mode choice modeling with support vector machines. *Transportation Research Record*, *2076*(1), 141-150. doi: 10.3141/2076-16

Zhao, X., Yan, X., Yu, A., & Van Hentenryck, P. (2020). Prediction and behavioral analysis of travel mode choice: A comparison of machine learning and logit models. *Travel Behaviour and Society*, *20*, 22-35. doi: https://doi.org/10.1016/j.tbs.2020.02.003

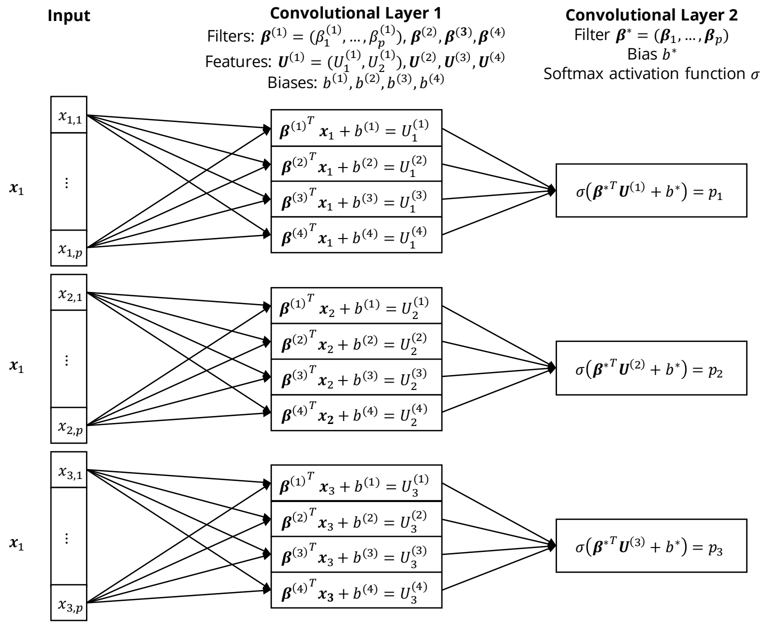
Appendix

Category	Variables
Distance ratio	
Cycle traffic routing	**mixed traffic**, cycle road, cycle path, cycle lane, shared cycle pedestrian path, pedestrian path pedestrian path with bicycle permitted
Road type	**main roads**, secondary roads, not specified
Surface condition	good, medium, **bad**, not specified
Surface material	**asphalt**, stone, **cobble stone**, compact, not specified
Slope	average, maximum, **0-2% slope**, 2-4% slope, >4% slope, **negative slope**
Max. speed of motorized traffic	30 km/h, 50 km/h, > **50** km/h, not specified

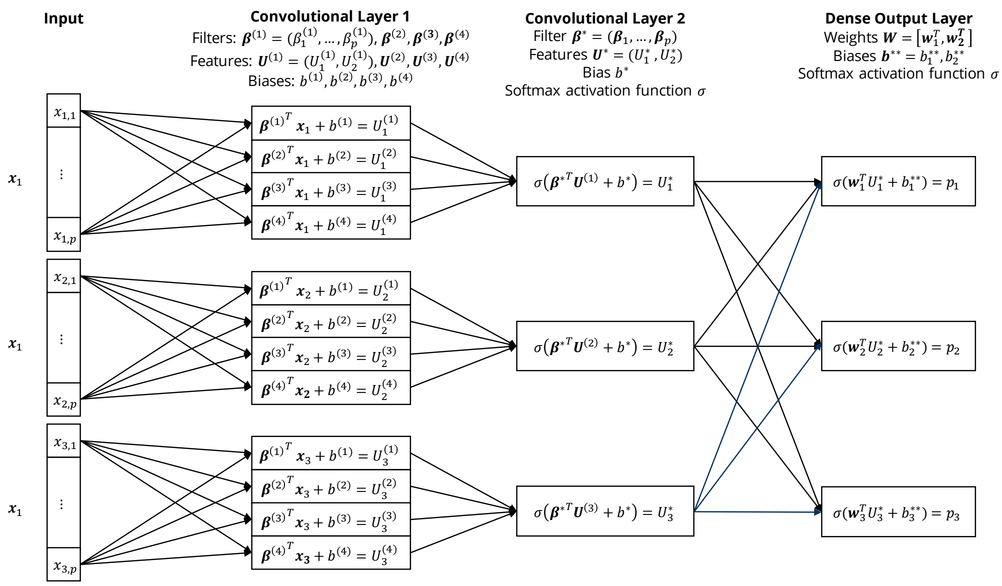
Table 2: Full list of input variables, ordered by categories. All variables, except for the distance ratio describe the share of the route, that has that respective attribute. For the distance ratio, the longest route within a route triple gets assigned value one and the remaining routes are divided by the longest route. The range of all the variables is between zero and one. For MNL only the nine variables are used, that are labelled with bold letters here.



(a) CNN with one convolutional and one dense layer (CNN-1C-1D)



(b) CNN with two convolutional layers (CNN-2C)



(c) CNN with two convolutional and one dense layers (CNN-2C-1D)

Figure 4: Architecture of CNNs.