A Deep Learning Approach to Better Urban Mobility Planning

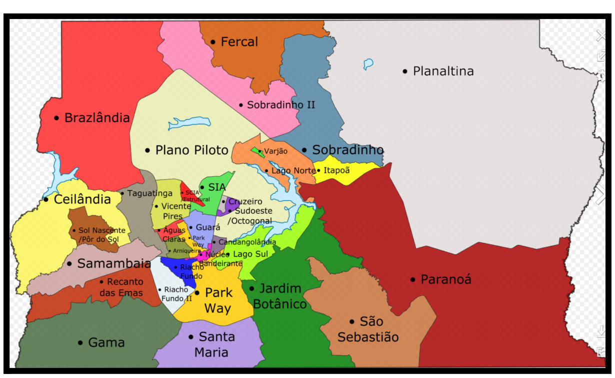
Project Team: Zach Clayton (zclayton@stanford.edu), Kei Tomozawa (keitomo@stanford.edu)

Abstract

Urban centers are becoming increasingly populated and will continue to do so for the foreseeable future. Understanding how people travel and by which means is a crucial aspect for urban planners as this influx occurs. One important question for planners is determining the mode residents choose to travel. Transportation analysis has previously relied on discrete choice models to predict mode, and now machine learning and deep learning have made strides in helping planners. With this in mind, we explore how a deep learning model performs in comparison to the status quo in predicting mode of travel. The goal of this paper is to further validate the use of deep learning in urban planning circles.

1 Introduction

Adaptation to slow the effects of climate change will require decisive and extensive action in numerous sectors, particularly transportation. Greenhouse gas emissions from transportation account for approximately 28% of the total for the US, making this sector the largest contributor[1]. Reducing the Vehicle Miles Traveled (VMT), a quantity that aggregates all the distance traveled for all vehicles in a particular area, is crucial and can be achieved through more robust public transit options. However, any improvements in urban mobility alternatives must prioritize the needs of especially vulnerable members of a community, and this starts by understanding the socioeconomic and demographic factors affecting people's transit habits. We will investigate a dataset from the Federal District of Brasil (a region surrounding the nation's capital of Brasilia) to better understand the necessary steps for an equitable and sustainable improvement in public transit. The main objective will be to predict transportation mode so urban planners can design better mobility solutions through an equity lens.



Map of Federal District of Brasil (analogous to Washington, D.C.) and its subdistricts

2 Related Work

Deep learning approaches have increasingly been applied to urban mobility problems. One such study analyzed 86 classifiers and found that deep learning models performed the best of any non-ensemble method (ensemble methods performed better) in predicting travel mode. Another study compared the results of a deep learning approach to discrete choice models like the random utility model. Again, the deep learning model performed better at predicting travel mode. Thus, evidence suggests that deep neural networks have much to offer in the urban planning space. We would like to build off of this research and validate the performance of a deep learning approach.

3 Dataset and Features

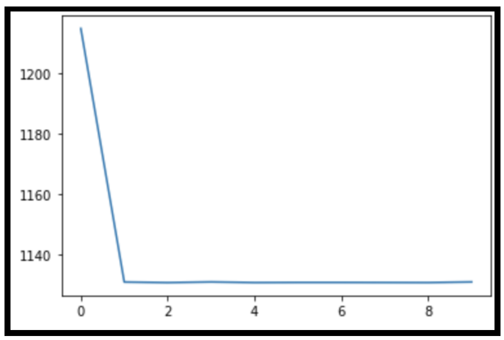
The dataset, a mobility survey conducted by the metro company of the Federal District of Brasil, was obtained from Kaggle[2]. The main objective of this survey was to study mobility patterns and socioeconomic characteristics of the Federal District population. This insight would allow urban planners to assess the impact of future transportation planning models.

Temporally, the survey was conducted over a span of nine months from March 12, 2016 to December 15, 2016. Residents from a set of households across all Administrative regions of the Federal District were randomly selected to participate in the survey. Stratified random sampling based on spatial distribution and socioeconomic indicators were used to ensure that there was representation. Overall, a total of 19,252 responses were obtained that spanned over 61,358 residents and 113,398 trips performed during a work day. This can be more easily summarized as:

Data	
household	19252 obs. of 20 variables
person	61358 obs. of 11 variables
trips	113398 obs. of 49 variables

4 Methods

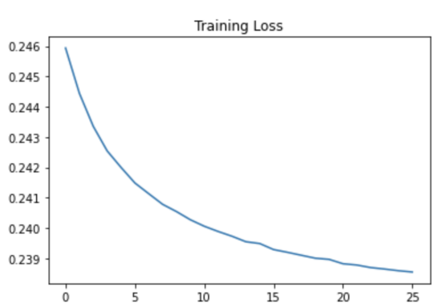
Though we wanted to predict mode of transportation, we started out by building a simple primitive model (Keras sequential model, Adam Optimizer, ten epochs) to predict trip duration based on three input features, which were common modes of transportation: cars, bicycles, and walking. This was actually our original plan, but we pivoted later to predict mode. Intuitively, we expect private vehicles to indicate longer trip durations and bicycles/walking would be indicative of shorter trip durations. Using this model, we aimed to predict duration of travel. We expected that this simple model would give us a baseline estimate, and then we could improve performance by including more features in a more advanced deep neural network model. The mean squared error loss function of the simple model is shown below.

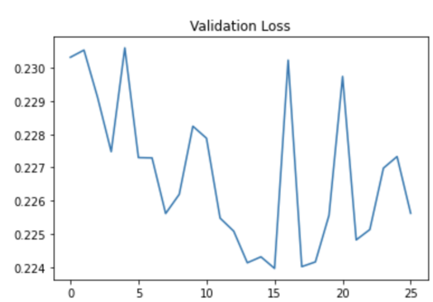


Description of deep neural network including plots

5 Experiments/Results/Discussion

With the simple sequential model sketched out, we developed a more complex deep neural network. While we had hoped that this more robust DNN would be based on a paper that explored DNN architectures[3], we did not quite achieve this. Instead, we focused on building a multi classification DNN from the Keras library. This involved identifying a set of 8 input features: age, income, education, # of people in household, # of bedrooms, # of cars, # of bikes, # of motorcycles and mapping them to a set of 8 output labels: individual transport, walking, collective public transport, collective private transport, bicycle, individual public transport, combined and other. We used a neural network with 10 hidden layers, RELU activation and a softmax classifier for 30 epochs. We used an 80/20 split for training and validation. Our loss graphs are shown below:

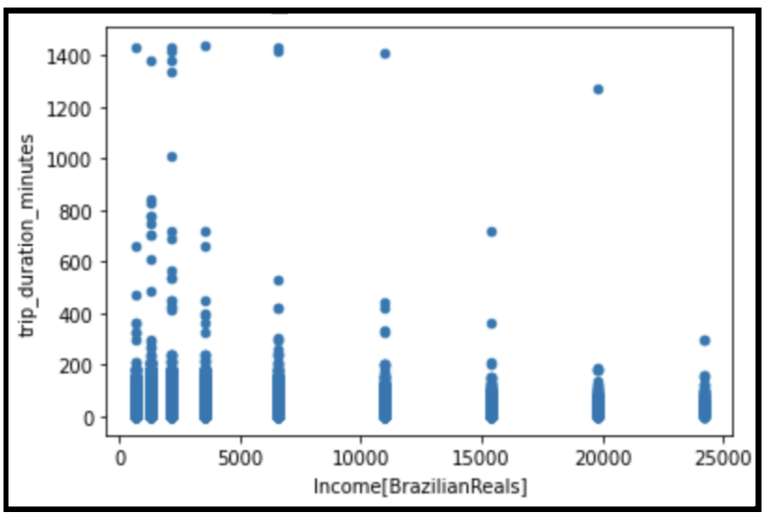




Clearly, our model did not perform well. Though loss slightly decreased during training, it was marginal. We did not achieve good prediction accuracy (~45%). Much work need to be done to finetune our model moving forward.

6 Conclusions

Deep learning models and data analysis can offer valuable insights into urban transportation planning, but these models must prioritize equity and sustainability to ensure that cities are meeting both their climate goals and supporting underserved communities. During our process, we talked with Professor Michael Kahan, Co-Director of the Urban Studies Department and Stanford, who urged us to consider several key factors. First, what is the relationship between income and travel duration, and are jobs and centers of commerce located near residential areas? These questions are key to urban planning. Brasilia, designed in the mid 20th-Century, was actually aimed to prioritize transit by car, and this decision still has impacts on citizens' lives today. The following figure illustrates the relationship between income and transit duration.



Relationship between Monthly Income and Trip Duration
(for reference, 1 USD is equivalent to 5.08 Brazilian Reals)

We ultimately determined that one subdistrict within Brasilia, called Plano Pilato, was where most citizens were commuting for work, university and school studies, and commerce. However, lower-income residents--which we defined as citizens making less than the Brazilian minimum monthly wage of 1,100 BRL (US$209)--lived in districts much farther from the economic center of Brasilia. On the other hand, citizens in the dataset's highest income bracket--above 11,000 BRL per month (US$2,166 per month)--were much more likely to live in Plano Pilato. As shown below, Plano Pilato was clearly designed as a commerce center with few residential options, so only wealthier residents could afford to live there.

View of Plano Pilato, Brasilia's economic center

Consequently, lower-income residents who lived in distant districts but still needed to commute daily to Plano Pilato would need to spend more money to travel the longer distance. At the same time, these community members would have less disposable income. Furthermore, longer travel times would mean lower-income residents would have less time available to spend with friends and family.

7 Contributions

Kei and Zach drafted the Project Proposal together. For the first milestone, Zach created the initial code for the data input and simple model, and Kei created the milestone report to position the project's work within the urban planning field and describe upcoming steps. Lastly, both Kei and Zach expanded the program with development of the deep neural network model and data analysis to answer urban planning questions; both Zach and Kei also created the final report and final video.

References

[1] EPA. *Carbon Pollution from Transportation.* United States Environmental Protection Agency, 2021. https://www.epa.gov/transportation-air-pollution-and-climate-change/carbon-pollution-transportation

[2] Miranda, D. *Urban Mobility Survey (Federal District, Brazil)*. Kaggle, 2020. https://www.kaggle.com/danielefm/urban-mobility-survey-federal-district-brazil?select=Stage.csv

[3] Wang, Shenhao. *Deep Neural Networks for Choice Analysis: Architecture Design with Alternative-Specific Utility Functions*. Transportation Research Part C Emerging Technologies, 2020.

[4]Shenhao Wang & Baichuan Mo & Stephane Hess & Jinhua Zhao, 2021. "Comparing hundreds of machine learning classifiers and discrete choice models in predicting travel behavior: an empirical benchmark," Papers2102.01130, arXiv.org