
Modeling Uncertainty in Vehicle Trajectory Prediction in a Mixed Connected and Autonomous Vehicle Environment using Deep Learning and Kernel Density Estimation

Tao Li

Department of Computer Science
Purdue University
West Lafayette, IN 47906
taoli@purdue.edu

Abstract

The advent of connected and autonomous vehicles (CAVs) will change driving behavior and the travel environment, providing opportunities for safer, smoother, and smarter road transportation. During the transition from the current human-driven vehicles (HDVs) to a fully CAV traffic environment, the road traffic will consist of a “mixed” traffic flow of HDVs and CAVs. Equipped with multiple sensors and vehicle-to-vehicle communications, a CAV can track the trajectories of other CAVs in its vicinity, and ideally, all CAVs in communication range. Such CAV trajectory data can be leveraged with advances in computing and machine learning algorithms to potentially predict trajectory data of HDVs. Based on these predictions, CAVs can react accordingly to avoid or mitigate traffic flow oscillations and accidents. Most previous studies can only generate trajectory predictions with no evaluation of uncertainty, e.g., ignoring the distribution of prediction. In this study, we propose a novel framework that can model the uncertainty in the leading HDV’s trajectory prediction for a CAV by integrating deep learning models with Kernel Density Estimation (KDE) method. Distribution of the HDV’s future trajectory can then be calculated and are critical to develop an optimal CAV control mechanism. Deep learning models have shown powerful prediction performances in many applications. Empowered with large scale neural networks, carefully designed computer architectures, and novel training algorithms, studies show that deep learning-based methods significantly outperform normal vehicle trajectory prediction methods in accuracy. However, deep learning models usually require huge efforts for tuning of hyperparameters in model structure (number of layers, size of each layer) and optimization methods (learning rate and mini-batch size in stochastic gradient descent). They also have issues like prone to overfitting and get stuck at local optimum. According to a recent study, models that have best performance are not necessary to be the one that uses optimal parameters. Our framework trains deep learning models with various hyperparameters in a parallel computing grid of massive nodes and generate trajectory prediction samples, then apply the KDE to estimate the prediction distribution to model the uncertainty. Not like traditional uncertain modeling methods, the KDE method does not require a lot of samples, therefore it is more efficient as fewer runs of deep learning model training are needed; furthermore, the KDE does not rely on prior knowledge or assumption about underlying distributions. This property makes KDE a good fit for the framework, because given the features of deep learning, it’s very hard or almost impossible to get such distribution of predictions through mathematical deductions. This framework is tested on the NGSIM dataset.

References

- [1] Bernard W Silverman. *Density estimation for statistics and data analysis*. Routledge, 2018.
- [2] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *nature*, 521(7553):436, 2015.
- [3] Chiyuan Zhang, Samy Bengio, Moritz Hardt, Benjamin Recht, and Oriol Vinyals. Understanding deep learning requires rethinking generalization. *arXiv preprint arXiv:1611.03530*, 2016.
- [4] Zhengbing He. Research based on high-fidelity ngsim vehicle trajectory datasets: A review. *Research Gate*, pages 1–33, 2017.
- [5] Yeping Hu, Wei Zhan, and Masayoshi Tomizuka. Probabilistic prediction of vehicle semantic intention and motion. *arXiv preprint arXiv:1804.03629*, 2018.
- [6] Simon J Sheather and Michael C Jones. A reliable data-based bandwidth selection method for kernel density estimation. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 683–690, 1991.