Polysona: Modular Driving Styles in Trajectory Prediction

Laura Zheng[1], Hamidreza Yaghoubi Araghi[1], Tony Wu[1], Tianyi Zhou[2], and Ming C. Lin[1]

[1]University of Maryland, College Park, [2]Mohamed bin Zayed University of Artificial Intelligence

Introduction

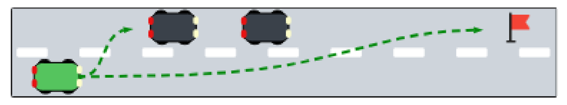


Fig. 1: **Motivating Example: Lane change occupied by other vehicles.** The ego vehicle has the goal of changing lanes and reaching a particular waypoint. At its current velocity, a direct lane change would result in a collision with the other vehicles. Here, the ego vehicle is faced with a natural stressor to make a decision: either decelerate and merge behind the other vehicles, or accelerate and pass the other vehicles first, then merge. Both outcomes are plausible, yet lead to drastically different trajectories by average displacement error metrics, which are computed per-timestep.

Trajectory prediction objectives and metrics often reward predictions that are common modes of behavior; average lane following, average car following, and average notions of social space on the road.

More out-of-distribution behavior types are difficult to massage from existing objectives in trajectory prediction (not to be confused with rarer maneuvers like U-turns).

In this paper, we present **Polysona, a parameter-efficient Mixture of Experts framework for modeling latent driving styles in trajectory prediction models**. This is one of the first works exploring latent behavior modeling without any additional annotations or specialized datasets.

Background

An assumption we make is that *driving* style, *or persona,* is strongly correlated with second-order kinematics such as acceleration and steering.

Task: Given a pre-trained encoder-decoder trajectory prediction model, we want to learn a set of LoRA experts such that each expert encodes a different latent driving persona.

When learning such experts jointly, this can be considered as a *Mixture-of-LoRA (MoL)* problem.

In our experiments, we attempt to model a latent driving behavior variable with N=3 LoRA experts at prior probabilities (0.3, 0.6, 0.1) respectively; this prior is motivated by prior work in behavioral long-term human subject studies conducted by the National Highway Transportation Safety Administration (NHTSA) [1].

Methodology

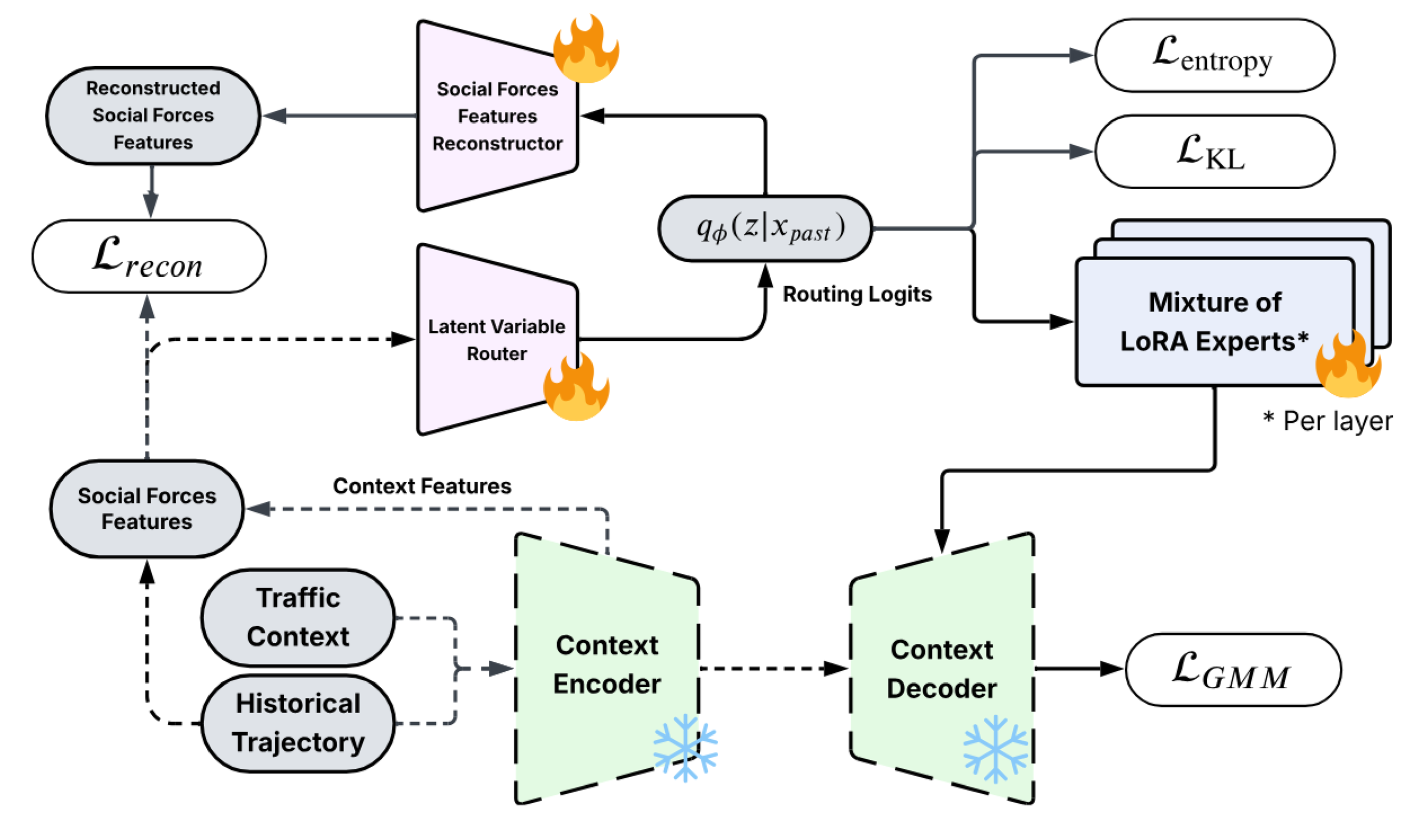
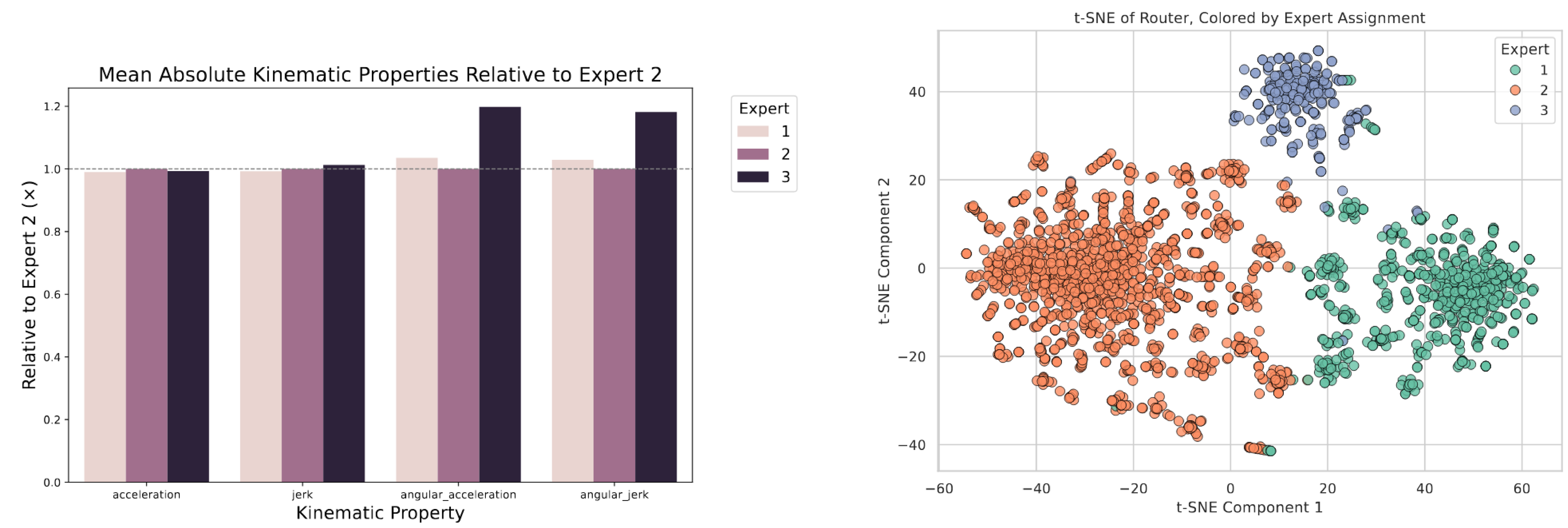


Fig. 2: **Training Overview.** We use a Mixture-of-LoRA (MoL) approach to learn a driving style latent variable for trajectory prediction models. Solid lines indicate gradient flow in backpropagation. In summary, we construct social forces features from historical agent trajectories and train a routing network in a Variational Autoencoder-like fashion; in other words, the router predicts logits for discrete latent classes, and a reconstructor network then learns to re-construct social forces features. The predicted routing logits and then used to select experts in the MoL layers, which are attached to certain decoder layers of the base model. All modules except the base model are trained end-to-end.



(a) Relative magnitudes of second-order kinematics.

(b) t-SNE visualization of router embeddings.

Fig. 3: **Visualizations of expert representations in Ours+CAT.** In (a), we show the mean magnitudes of acceleration, jerk, angular acceleration, and angular jerk between experts, relative to Expert 2. Expert separability is more pronounced in lateral (angular) kinematics than in longitudinal kinematics (acceleration, jerk), suggesting that *differences between experts are better captured by steering behavior than by speed profiles alone*. In (b), we project router embeddings to 2D with t-SNE and color each point by the router's expert assignment. The clear three-cluster separation confirms that our discrete routing does not suffer from expert collapse, and that the router learns meaningful groupings in the social forces feature space.

Results

Table 1: Trajectory Prediction Benchmark Performance Comparisons. We compare existing trajectory prediction models, TAE (a behavior-aware trajectory predictor), and a CVAE baseline against our best-performing variant (Ours+MoV). Bolded values indicate best performance in respective columns.

Method	# Trainable Params	brierFDE↓	minADE↓	minFDE↓	MissRate↓
Autobot [28]	1.5M	2.4439	0.8892	1.7817	0.2803
Wayformer [29]	15.2M	2.5747	0.9348	1.9718	0.3574
MTR [30]	65.2M	2.1702	0.8645	1.7094	0.3107
MTR+Actions [26]	65.2M	2.1605	0.8658	1.7102	0.3208
TAE [27]	5.7M	2.9430±.070	1.0421±.040	2.2721±.069	0.4014±.017
CVAE [31]	**99.5K**	2.1603±.003	**0.8617**±.001	1.7032±.003	0.3164±.000
Ours+MoV [15]	195K	**2.1588**±.001	0.8619±.001	**1.7016**±.001	**0.3158**±.000

Table 2: Comparison across MoL mixing strategies. Interestingly, the best all-around results are achieved with the MoV approach, which uses IA3 for finetuning, instead of LoRA.

Method	# Trainable Params	∇Router	brierFDE↓	minADE↓	minFDE↓	MissRate↓
MTR+Actions [26]	65.2M	-	2.1605	0.8658	1.7102	0.3208
Ours+Polytropon [16]	527K	✗	2.1602±.000	0.8623±.000	1.7037±.000	0.3164±.001
Ours+C-Poly [17]	969K	✗	2.1613±.002	0.8625±.001	1.7048±.002	0.3169±.000
Ours+HyperFormer [18]	527K	✗	2.1601±.001	0.8623±.000	1.7038±.001	0.3159±.001
Ours+CAT [34]	526K	✓	2.1607±.002	0.8624±.001	1.7041±.002	0.3171±.001
Ours+MoV [15]	195K	✓	**2.1588**±.001	**0.8619**±.001	**1.7016**±.001	**0.3158**±.000
Best % Improvement			0.066%	0.423%	0.497%	1.645%

Conclusion

- We present Polysona, one of the first approaches proposed for latent driving style modeling without specialized annotations of datasets.
- Polysona learns distinct LoRA experts using a VAE-like objective, such that latent social forces features must be classified/reconstructed by three distinct scalars, clustering behavior into three distinct bins.
- Results show that, even without true real-world evaluation datasets focused on aggressive driving, finetuning under Polysona still improves performance on trajectory prediction metrics.
- A limitation of this work is that the proper metrics and tooling for evaluating driving personas is under-explored for behavior prediction and simulation; thus, the results in this paper focus on existing benchmarks as a proxy for generalization to rarer personas. Valuable future work includes curating a labeled dataset of long-tail driving behaviors for improved evaluation.

Acknowledgements

This project is supported in part by Dr. Barry Mersky and Capital One E-Nnovate Endowed Professorships and UMD-ARL Cooperate Agreement.

[1] Klauer, S. G., Dingus, T. A., Neale, V. L., Sudweeks, J. D., & Ramsey, D. J. (2009). *Comparing real-world behaviors of drivers with high versus low rates of crashes and near crashes* (No. DOT-HS-811-091).