

402 A Algorithm for MPF

403 The algorithm for MPF is provided in Alg. 1. We use the agent-centric dataloader generated by
 404 trajdata [19] to access various prediction datasets ([31, 18, 17]) in a unified representation. In
 405 lines 15-16, trajectories from the standalone predictors are sampled, and in lines 8-13, the belief
 406 update discussed in Section 3.3.2 is performed.

Algorithm 1 Multi-Predictor Fusion (MPF)

```

1: Input: learn_predictor, rule_predictor, dataloader (agent-centric)
2: Output: pred_MPF  $\leftarrow []$  (list of MPF predictions)
3: Hyperparameters: number of trajectory samples per scene  $N$ , prediction horizon  $T$ , Bayes learning rate
    $\eta$ , convex combination factor with prior  $\gamma$  in (8), belief prior  $b_0$ .
4: for episode in dataloader do
5:    $t \leftarrow 0, b \leftarrow b_0$ 
6:   learn_predictor_history  $\leftarrow \emptyset$ , rule_predictor_history  $\leftarrow \emptyset$ ,
7:   for scene in episode do
8:     if  $t > 0$  then
9:        $x_t \leftarrow \text{observe\_current\_agent\_state}()$ 
10:      likelihood_learn  $\leftarrow \text{get\_likelihood}(\text{learn\_predictor\_history}, x_t)$ 
11:      likelihood_rule  $\leftarrow \text{get\_likelihood}(\text{rule\_predictor\_history}, x_t)$ 
12:       $\alpha \leftarrow (\text{likelihood\_learn} * b^l) / (\text{likelihood\_rule} * b^r)$ 
13:      new_belief  $\leftarrow \left[ \frac{\frac{\alpha^\eta}{1+\alpha^\eta}}{\frac{1}{1+\alpha^\eta}} \right]$ 
14:       $b \leftarrow (1 - \gamma)\text{new\_belief} + \gamma b_0$ 
15:     end if
16:      $\{x_{t+1:t+T,i}^l\}_{i=1}^N \leftarrow \text{learn\_predictor.sample}(\text{scene}, N, T)$ 
17:      $\{x_{t+1:t+T,i}^r\}_{i=1}^N \leftarrow \text{rule\_predictor.sample}(\text{scene}, N, T)$ 
18:     learn_predictor_history  $\leftarrow \{x_{t+1,i}^l\}_{i=1}^N$ 
19:     rule_predictor_history  $\leftarrow \{x_{t+1,i}^r\}_{i=1}^N$ 
20:     pred_MPF.append(belief_sampler( $\{x_{t+1:t+T,i}^l\}_{i=1}^N, \{x_{t+1:t+T,i}^r\}_{i=1}^N, b$ ))
21:      $t++ = 1$ 
22:   end for
23: end for
24: return pred_MPF

```

407 B Further Details on Experimental Evaluation

408 B.1 Hyperparameters

409 In our experimental evaluation we set the hyperparameters in line 3 of Alg. 1 as $N = 20, T = 8$,
 410 $\eta = 0.1, \gamma = 0.02$, and an uninformed prior $b_0 = [0.5, 0.5]$. The time dt between two timesteps is
 411 0.5 seconds.

412 B.2 Dataset Details

The number of scenes and episodes per dataset are provided in Table 2 below:

Dataset	Scenes	Episodes
nuPlan-mini val	118030	6814
nuScenes val	66620	4140
Lyft val 10%	744448	49196

Table 2: Dataset Details

414 B.3 Learning rate study

415 We studied the effect of varying the learning rate η in the Bayes belief update (Section 3.3.2). The
 416 mean difference from best (MDB) for MPF for different choices of η are reported in Table 3. The
 417 performance of MPF is not very sensitive to the choice of η , however, a clear trend does emerge
 418 which points towards smaller η performing better than larger η on an average across the datasets.

Dataset	MDB (%)			
	@ $\eta = 0.1$	@ $\eta = 0.4$	@ $\eta = 0.7$	@ $\eta = 1.0$
NuPlan-mini	4.40	4.32	4.21	3.95
NuScenes	5.30	5.25	5.55	6.00
Lyft	4.48	4.80	5.36	6.10
Avg. MDB on Datasets	4.73	4.79	5.04	5.35

Table 3: MDB for different η in Bayes update.