Existing evaluation metrics

In the framework, the following metrics are currently implemented (notations below the table):

Metric	Input	Formula
ADE (marginal)	Trajectories	$\frac{1}{\lvert P_{N_p}\rvert \sum_{i=1}^{N_S} N_{A,i}} \sum_{i=1}^{N_S} \frac{1}{N_{O,i}} \sum_{p \in P_{N_p}} \sum_{t \in T_{O,i}} \sum_{j=1}^{N_{A,i}} \sqrt{D_{i,p,j}(t)}$
ADE (joint)	Trajectories	$\frac{1}{\lvert P_{N_p}\rvert N_S} \sum_{i=1}^{N_S} \frac{1}{N_{O,i}} \sum_{p \in P_{N_p}} \sum_{t \in T_{O,i}} \sqrt{\frac{1}{N_{A,i}} \sum_{j=1}^{N_{A,i}} D_{i,p,j}(t)}$
min ADE (marginal)	Trajectories	$\frac{1}{\sum_{i=1}^{N_S} N_{A,i}} \sum_{i=1}^{N_S} \frac{1}{N_{O,i}} \sum_{j=1}^{N_{A,i}} \min_{p \in P_{N_p}} \sum_{t \in T_{O,i}} \sqrt{D_{i,p,j}(t)}$
min ADE (joint)	Trajectories	$\frac{1}{N_S} \sum_{i=1}^{N_S} \frac{1}{N_{O,i}} \min_{p \in P_{N_p}} \sum_{t \in T_{O,i}} \sqrt{\frac{1}{N_{A,i}} \sum_{j=1}^{N_{A,i}} D_{i,p,j}(t)}$
most likely ADE (marginal)	Trajectories	$\frac{1}{\sum_{i=1}^{N_S} N_{A,i}} \sum_{i=1}^{N_S} \frac{1}{N_{O,i}} \sum_{j=1}^{N_{A,i}} \sum_{t \in T_{O,i}} \sqrt{D_{i,p^*_{i,j},j}(t)} \, , \quad \text{with } p^*_{i,j} = arg \min_{p \in P_{N_p}} \hat{L}_{i,p,j}$
most likely ADE (joint)	Trajectories	$\frac{1}{N_S} \sum_{i=1}^{N_S} \frac{1}{N_{O,i}} \sum_{t \in T_{O,i}} \sqrt{\frac{1}{N_{A,i}} \sum_{j=1}^{N_{A,i}} D_{i,p^*_i,j}(t)} \, , \quad \text{with } p^*_i = \arg\min_{p \in P_{N_p}} \hat{L}_{i,p}$
FDE (marginal)	Trajectories	$\frac{1}{\lvert P_{N_p}\rvert \sum_{i=1}^{N_S} N_{A,i}} \sum_{i=1}^{N_S} \sum_{p \in P_{N_p}} \sum_{j=1}^{N_{A,i}} \sqrt{D_{i,p,j}(\max T_{O,i})}$
FDE (joint)	Trajectories	$\frac{1}{\lvert P_{N_p}\rvert N_S} \sum_{i=1}^{N_S} \sum_{p \in P_{N_p}} \sqrt{\frac{1}{N_{A,i}} \sum_{j=1}^{N_{A,i}} D_{i,p,j}(\max T_{O,i})}$
min FDE (marginal)	Trajectories	$\frac{1}{\sum_{i=1}^{N_S} N_{A,i}} \sum_{i=1}^{N_S} \sum_{j=1}^{N_{A,i}} \min_{p \in P_{N_p}} \sqrt{D_{i,p,j}(\max T_{O,i})}$
min FDE (joint)	Trajectories	$\frac{1}{N_S} \sum_{i=1}^{N_S} \min_{p \in P_{N_p}} \sqrt{\frac{1}{N_{A,i}} \sum_{j=1}^{N_{A,i}} D_{i,p,j}(\max T_{O,i})}$
Brier min FDE (marginal)	Trajectories	$\frac{1}{\sum_{i=1}^{N_S} N_{A,i}} \sum_{i=1}^{N_S} \sum_{j=1}^{N_{A,i}} \left(1 - \frac{\exp(\hat{L}_{i,p^*_{i,j},j})}{\sum_{p=1}^{\lvert P_{N_p}\rvert} \exp(\hat{L}_{i,p,j})}\right)^2 \sqrt{D_{i,p^*_{i,j},j}(\max T_{O,i})} \, , \quad \text{with } p^*_{i,j} = arg \min_{p \in P_{N_p}}$

Brier min FDE (joint)	Trajectories	$\frac{1}{N_S}\sum_{i=1}^{N_S}\left(1-\frac{\exp(\hat{L}_{i,p_i^*})}{\sum_{p=1}^{\lvert P_{N_p}\rvert}\exp(\hat{L}_{i,p})}\right)^2 \sqrt{\frac{1}{N_{A,i}}\sum_{j=1}^{N_{A,i}} D_{i,p_i^*,j}(\max T_{O,i})}\ , \ \text{with } p_i^* = \arg\min_{p\in P_{N_p}} \frac{1}{N_{A,i}} \ldots$
most likely FDE (marginal)	Trajectories	$\frac{1}{\sum_{i=1}^{N_S} N_{A,i}}\sum_{i=1}^{N_S}\sum_{j=1}^{N_{A,i}}\sqrt{D_{i,p^*_{i,j},j}(\max T_{O,i})}\ , \ \text{with } p^*_{i,j} = arg\min_{p\in P_{N_p}} \hat{L}_{i,p,j}$
most likely FDE (joint)	Trajectories	$\frac{1}{N_S}\sum_{i=1}^{N_S}\sqrt{\frac{1}{N_{A,i}}\sum_{j=1}^{N_{A,i}} D_{i,p_i^*,j}(\max T_{O,i})}\ , \ \text{with } p_i^* = \arg\min_{p\in P_{N_p}}\hat{L}_{i,p}$
NLL (marginal)	Trajectories	$-\frac{1}{\sum_{i=1}^{N_S} N_{A,i}}\sum_{i=1}^{N_S}\sum_{j=1}^{N_{A,i}} L_{i,j}$
NLL (joint)	Trajectories	$-\frac{1}{N_S}\sum_{i=1}^{N_S} L_i$
ECE (marginal)	Trajectories	$\frac{1}{201}\sum_{k=0}^{200}\left\lvert\left(\frac{1}{\sum_{i=1}^{N_S}N_{A,i}}\left\lvert\left\{i,j \mid \frac{1}{\lvert P_{N_p}\rvert}\left\lvert\left\{p\in P_{N_p}\mid \hat{L}_{i,p,j} > L_{i,j}\right\}\right\rvert > \frac{k}{200}\right\}\right\rvert\right) + \frac{k}{20\ldots}\right.$
ECE (joint)	Trajectories	$\frac{1}{201}\sum_{k=0}^{200}\left\lvert\left(\frac{1}{N_S}\left\lvert\left\{i \mid \frac{1}{\lvert P_{N_p}\rvert}\left\lvert\left\{p\in P_{N_p}\mid \hat{L}_{i,p} > L_i\right\}\right\rvert > \frac{k}{200}\right\}\right\rvert\right) + \frac{k}{200} - \right.$
Miss rate (marginal)	Trajectories	$\frac{1}{\sum_{i=1}^{N_S} N_{A,i}}\sum_{i=1}^{N_S}\sum_{j=1}^{N_{A,i}}\begin{cases}1 & \sqrt{D_{i,p,j}(\max T_{O,i})} > 2 \forall p\in P_{N_p}\\ 0 & \text{otherwise}\end{cases}$
Miss rate (joint)	Trajectories	$\frac{1}{N_S}\sum_{i=1}^{N_S}\max_{j\in\{1,\ldots,N_{A,i}\}}\begin{cases}1 & \sqrt{D_{i,p,j}(\max T_{O,i})} > 2 \forall p\in P_{N_p}\\ 0 & \text{otherwise}\end{cases}$
Collision rate (marginal)	Trajectories	Needs to be rewritten, but should evaluate collision rate against GT of other agents
Collision rate (joint)	Trajectories	Needs to be rewritten, but should evaluate collision rate against predicted futures of oth available, otherwise against GT
Area under Curve (AUC)	Classifications	$\frac{1}{N_S}\sum_k \frac{\left(\sum_{i=1}^{N_S} r_{i,k}p_{i,k}\right) - \frac{1}{2}N_k(N_k+1)}{N_S - N_k}\ ,$ where likelihood rank $r_{i,k}$ fulfills $r_{i_1,k} > r_{i_2,k} \Rightarrow \frac{\hat{p}_{i_1,k}}{\hat{p}_{i_1,k} + \max_{\hat{k}\neq k}\hat{p}_{i_1,\hat{k}}} \geq \frac{\hat{p}_{i_2,k}}{\hat{p}_{i_2,k}+\max_{\hat{k}\neq k}\hat{p}_{i_2,\hat{k}}}\ , \ \text{with } N_k = \sum_{i=1}^{N_S} p_i \ldots$

ECE	Classifications	Look at the corresponding python file for a definition.
TNR-PR	Gap acceptance classifications (for '*t0_type*': '*crit*')	$$\frac{1}{\lvert\{i \,\vert\, p_{i,k=accepted} = 0\}\rvert} \sum_{i} \begin{cases} 1 & p_{i,k=accepted} = 0 \land \hat{p}_{i,k=accepted} > \underset{\hat{i} \in \{\hat{i} \vert p_{\hat{i},k=accepted} = 1\}}{\min} \hat{p} \\ 0 & otherwise \end{cases}$$

Here, the following notation is used:

- N_S: Number of samples in the dataset.
- i: Index for those samples.
- For trajectory based metrics, the following notation is used:
  - $N_{A,i}$: The number of evaluated agents in each sample.
  - j: Index for those agents.
  - P: Set of indicators to different, randomly predicted trajectories (with $P =$ **self.data_set.num_samples_path_pred**).
  - P_{N_p}: A random subset of P with $|P_{N_p}| = N_p$. In the respective metrics, N_p is set by using '*num_preds*' = <N_p> when defining the respective kwargs dictionary in the *simulations.py* file.
  - p: Index for those predictions.
  - $T_{O,i}$: The output timesteps for trajectories, with length $N_{O,i}$.
  - t: Value of those timesteps.
  - $x_{i,j}(t)$, $y_{i,j}(t)$: Recorded positions of an agent.
  - $\hat{x}_{i,p,j}(t)$, $\hat{y}_{i,p,j}(t)$: Predicted positions of an agent. We can then define the following shortcuts for squared distances: $D_{i,p,j}(t) = (x_{i,j}(t) - \hat{x}_{i,p,j}(t))^2 + (y_{i,j}(t) - \hat{y}_{i,p,j}(t))^2$
  - P_{KDE}: A probability density function trained over predicted samples (from P, not \$P_{N_p}). We can then define the following log likelihoods:
    - For marginal metrics, we have:

      $$L_{i,j} = \ln\left(P_{KDE,i,j}\left(\{\{x_{i,j}(t), y_{i,j}(t)\} \,|\, \forall\, t \in T_{O,i}\}\right)\right)$$
      $$\hat{L}_{i,p,j} = \ln\left(P_{KDE,i,j}\left(\{\{\hat{x}_{i,p,j}(t), \hat{y}_{i,p,j}(t)\} \,|\, \forall\, t \in T_{O,i}\}\right)\right)$$
    - For joint metrics, we have:

      $$L_{i} = \ln\left(P_{KDE,i}\left(\{\{\{x_{i,j}(t), y_{i,j}(t)\} \,|\, \forall\, t \in T_{O,i}\} \,|\, \forall\, j \in \{1, ..., N_{A,i}\}\}\right)\right)$$
      $$\hat{L}_{i,p} = \ln\left(P_{KDE,i}\left(\{\{\{\hat{x}_{i,p,j}(t), \hat{y}_{i,p,j}(t)\} \,|\, \forall\, t \in T_{O,i}\} \,|\, \forall\, j \in \{1, ..., N_{A,i}\}\}\right)\right)$$
- For classification based metrics, the following notation is used:
  - N_B: Number of potentially classifiable behaviors.
  - k: Index for those behaviors.
  - $p_{i,k}$, $\hat{p}_{i,k}$: The actual and predicted likelihood of behavior k in smaple i.