

$$\mathcal{L}^{DAS} = - \sum_{(X,Y) \in \mathcal{D}} \log P(Y|\tilde{X}) = - \sum_{(X,Y) \in \mathcal{D}} \log P(y_1, \dots, y_i|\tilde{X}, y_{<i}),$$