Step 1

GPT Training Dynamics Collection

Step 2 **Featu**

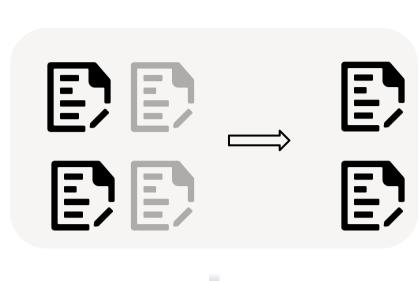
Step 2

Featurized Simulator Training

Step 3

Inference

A curriculum consists of a subset of the training data.



Training and test examples are encoded by the simulator.

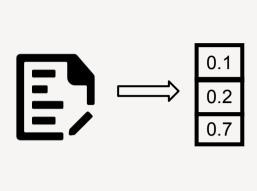
The simulator

assumption.

predicts the test

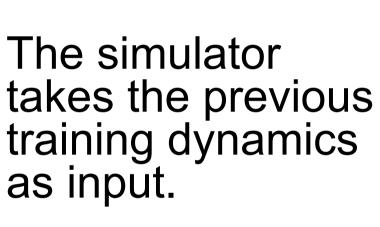
metric under the

n-th order Markov



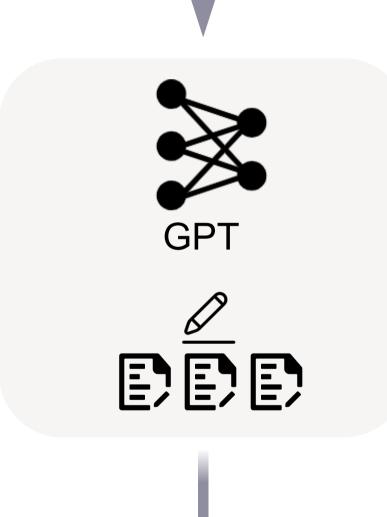




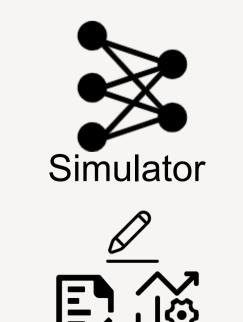




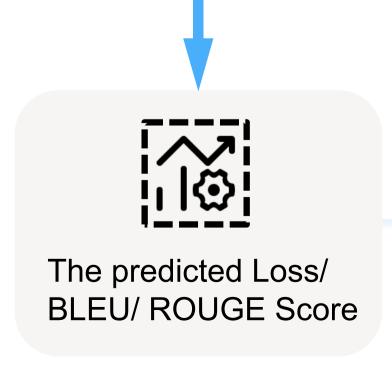
Train the GPT based on the curriculum.



Train the simulator with the MSE loss.



Predict the test metric of the next step.



Compute the test metric during the training.



