

.LMs

en working in the industry, participating in user groups, and participating in other certificate programs.

ə of the input variable that yields the highest/lowest output value.

e-trained causal language models like GPT-2 for text generation tasks.

for reinforcement learning, including a value head used for estimating the value function, crucial for models like

okens at each step in sequence generation, used to improve the quality of outputs in language models by

le between the current policy and the reference model, impacting the exploration and exploitation trade-off in
ə parameter for the DPO loss.

s or choices, often used to model pairwise preferences, where the probability of one item being preferred over

des not require iterative or numerical methods to solve.

le for processing by machine learning models, especially in reinforcement learning scenarios.

parameters in an optimization problem. It is used to guide models toward better performance by minimizing the

ence datasets, for use in training models like those using direct preference optimization (DPO).

ə learning models. In this context, it refers specifically to the IMDB dataset used for sentiment analysis.

rocessed by machine learning models, particularly language models.