

Transformer train loss on random tasks with different regular alphabet sizes

