ON THE EFFECT OF PRE-TRAINING FOR TRANSFORMER IN DIFFERENT MODALITY ON OFFLINE REINFORCEMENT LEARNING

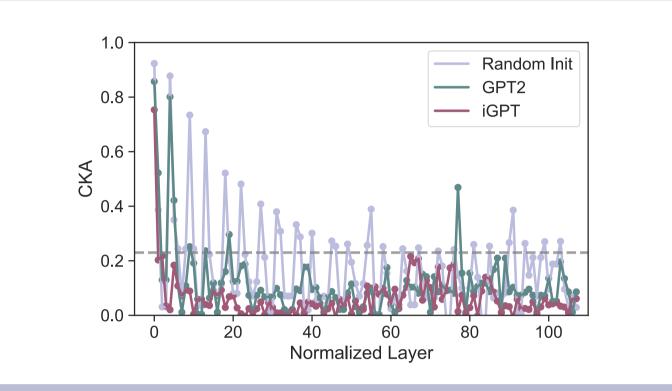
Shiro Takagi (Independent Researcher)



Q: Why can pre-trained transformers with text data be efficiently fine-tuned to offline RL tasks? & What is happening inside the Transformers?

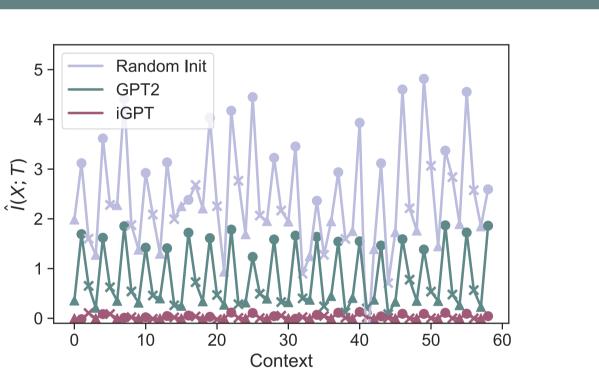
1. Feature reuse may not be the cause.

- We compared different models' CKAs between pre- and post-fine-tuning.
- We found that pre-trained models changed their representations more largely.



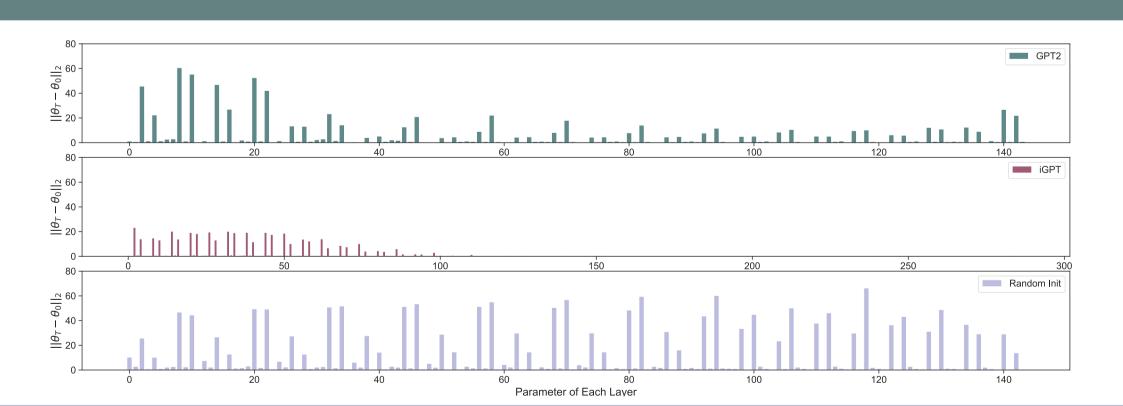
2. Better fitting to data may not be the cause.

- We computed the estimated mutual information between data and representations.
- We found that pre-trained models contain less information of data.



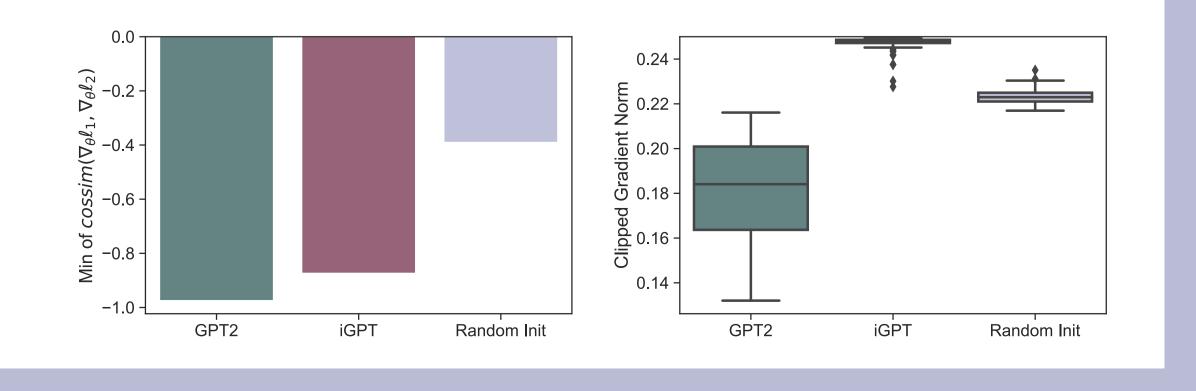
3. Some unchanged parameters may be the cause.

- We analyzed the change in weights after training.
- We found that pre-trained models don't change parameters so much.



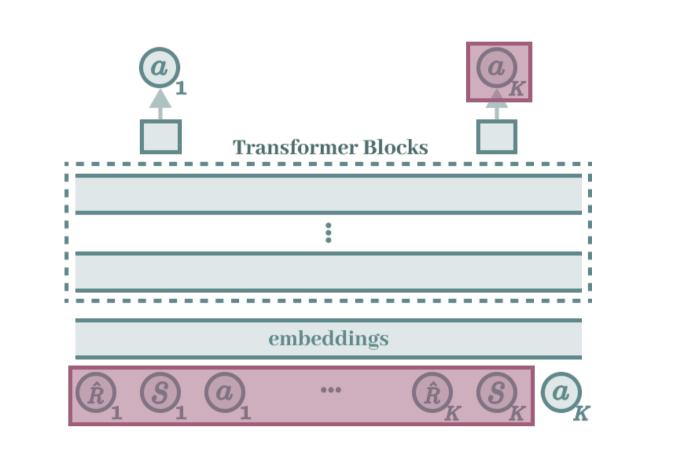
4. Large gradient w/ gradient clipping may degrade the performance.

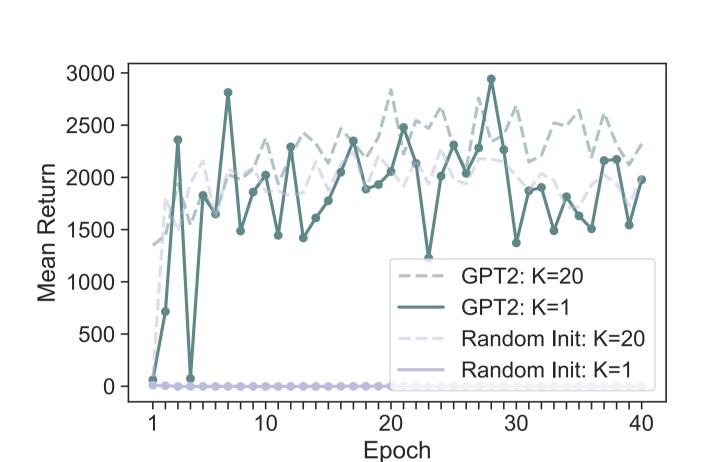
- We studied the gradient norm and confusion at an early phase.
- We found that gradient norms of iGPT collapse by gradient clipping.



5. Pre-trained models learn efficiently even without context.

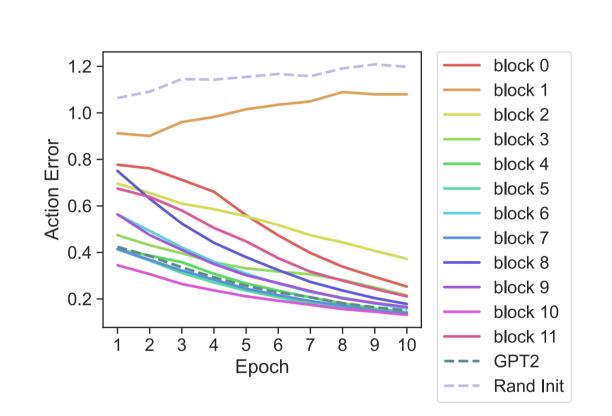
- We trained language-pre-trained and randomly initialized models with no access to the context.
- We found that removing the context information doesn't slow down the convergence of language-pre-trained models but slows down that of randomly initialized models.

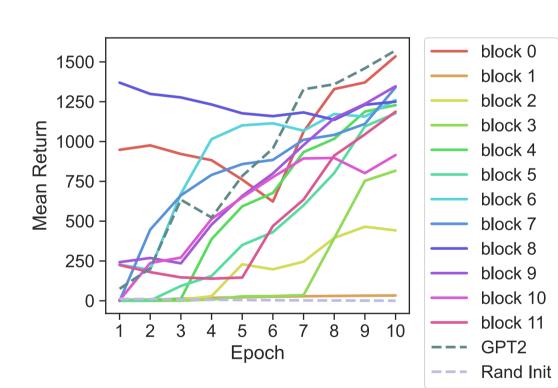




6. Even a single pre-trained block improves training efficiency.

- We replaced a Transformer block of randomly initialized models with that of language-pre-trained models and fine-tuned them.
- We found that replacing a block helps the models decrease the action error and increase the mean return more quickly.





7. Pre-acquired way to use context is a cause of the success.

- We compared the change in attention distance of the models during fine-tuning.
- Attention distance is the average distance b/w keys and queries weighted by attention.
- We found that the attention distance of the pre-trained models changes less than that of the randomly initialized models.

