Prompt tuning- soft prompts learned through backpropagation and can be tuned to incorporate signals from any number of labeled examples

- Condition frozen language models to perform specific tasks
    - Prepend additional tunable tokens to input text
    - Trained end- to end and can condense signal from full labeled dataset
- Prompt tuning becomes more competitive at scale (language model capacity): with larger models, this method is able to match the performance of model tuning (far better than prompt design)
- Robustness to domain transfer and efficient “prompt ensembling”
- Can use same model for many tasks

Model tuning “fine tuning”- all model parameters are tuned during adaptation

Prompt design “priming”-

- Prompts: task description and/ or several examples
- Single model can serve many different tasks
- Error- prone and requires human involvement and the effectiveness is limited by how much conditioning text can fit into model input
- Performance lags behind tuned models

Explicitly separating task-specific parameters from the “generalist” parameters needed for general language-understanding has a range of additional benefits.

- By capturing task definition in the prompt while keeping the generalist parameters fixed, we are able to achieve better resilience to domain shifts
- “Prompt ensembling”: learning multiple prompts for the same task can boost quality and is more efficient than classic model ensembling
- Interpretable

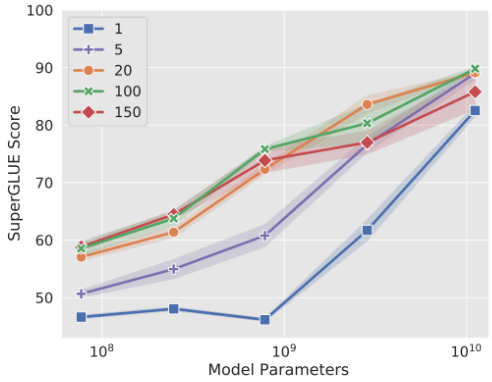
Prompt tuning removes restriction that the prompt be parameterized by the model since it has its own parameters which can be updated

- Prompt design selects tokens from fixed vocabulary of frozen embedding

Initialize prompt tokens with embeddings of the valid target tokens

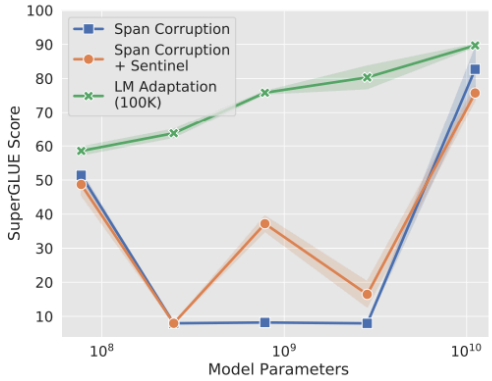
Span corruption is not that good for training frozen models to be later conditioned by prompts

Prompt tuning prevents model from modifying its general understanding of language, modulates representation of input, reduces ability to overfit by memorizing specific lexical cues and spurious correlations

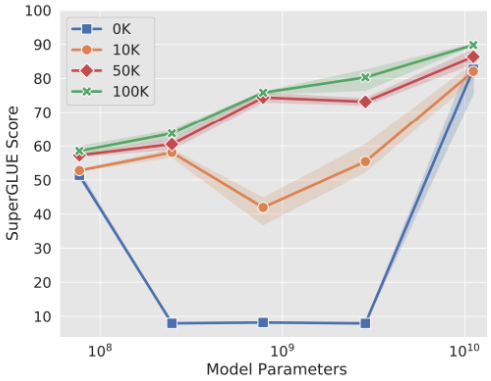


(a) Prompt length

(b) Prompt initialization



(c) Pre-training method



(d) LM adaptation steps