Understanding fine-tuning

Il y 2 types de fine-tuning pour un modèles:

1. feature extraction
   Dans cette méthode le modèle pré-entrainé a déjà appris des features donc on ne les change pas excépter la dernière couche qui est entraîner sur la tâche spécifique souhaitée
2. full-fine-tuning
   Contrairement au feature extraction, le full fine-tuning entraîne le modèle dans son entièreté. Cette méthode exige plus de mémoire.

Il est possible d'aborder le fine-tuning de différente manière:

a. supervised fine-tuning:
   →basic hyperparameter tuning: manually adjusting the model hyperparameters (learning rate, batch size...)
   → transfer learning:: a model pre-trained on a large, general dataset is used as a starting point.. The model is then fine-tuned on the task-specific data, allowing it to adapt its pre-existing knowledge to the new task.
   → multi-task learning: the model is fine-tuned on multiple related tasks simultaneouslySo that the model can develop a more robust and generalized understanding of the data. Ths works well when the tasks are closely related or when there is limited data for individual tasks.
   → few-shot learning: In this method the model is given a few example during inference to learn a new task. The goal is to guide the model by providing examples in the prompt. This method can be integrated into the reinforcement learning from human feedback (RLHF) if the small amount of data has human feedback
   → task-specific fine-tuning: adapting the model parameters to the requirements of the targeted task thereby enhancing its performance to that particular domain.

b. reinforcement learning from human feedback (RHLF)
   → reward modeling: The model generates multiple output and a human evaluators rank these output based on their quality.
   → proximal policy optimization: Les mises à jour des poids sont minimiser pour éviter qu'elles soient trop éloignés pour éviter les changements abrupts et permettre une meilleur stabilité durant l'entrai^nement
   → comparative ranking: similar to reward modeling but in this method the model learns from multiple raking provided for each output (human evaluators). It focus on the comparison between different outputs.
   → preference learning: The model generate multiple outpu and the human indicate which output it prefers. Then the model adjust it's behavior to align with the human evaluator's preference.

→ parameter efficient fine-tuning: this methode update a fraction of the model parameters by adding or modifying layers.