

Optimization des Hyperparamètres appliquée au Fine Tuning de LLM

Basé sur l'article : *Bayesian and Partition-Based Optimization for Hyperparameter Optimization of LLM Fine-Tuning*

Nathan Davouse

Sommaire

1. Introduction

2. Design et Implémentation

3. Résultats et Analysis

4. Conclusion

Large Language Models

Summary

- ▶ State-of-the-art of Natural Language Processing (NLP) problems
- ▶ Architecture : Transformers^a block, mixed with classical layers (MLP, Conv)

^aVaswani et al, Attention is all you need, 2017

Self Attention

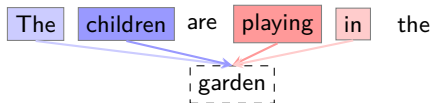


Figure: Self Attention mecanism illustration

Self attention is the key of LLM, used to compute the context of each token.

Fine Tuning

Hyperparameter Optimization

Problem Formulation

Related Works

Sommaire

1. Introduction

2. Design et Implémentation

3. Résultats et Analysis

4. Conclusion

Search Space

Search Strategy : B0

Search Strategy : S00

Search Strategy : BaMSOO

Performance Estimation Strategy

Implémentation

Sommaire

1. Introduction

2. Design et Implémentation

3. Résultats et Analysis

4. Conclusion

Expérimentation

LHS : Résultats

Résultats des 3 algorithmes

Analyse

Prospectives

Sommaire

1. Introduction

2. Design et Implémentation

3. Résultats et Analysis

4. Conclusion

Conclusion

Une conclusion

Merci.

Annexes 1 : Architecture d'un LLM

MHA, Transformers

Annexes 2

Pour du contenu supplémentaire, une deuxième fois