

Technique pour l'inférence de scores

12 juillet 2024

1 Introduction

Dans ce document, nous présentons une technique raffinée d'ajustement pour inférer des scores basée sur les scores de similarité fournis par un modèle de langage (LLM). Cette approche utilise les scores de similarité des deux meilleures réponses de référence comme poids pour ajuster le score inféré.

2 Méthode

Étant donné un ensemble de réponses de référence avec des scores connus et leurs scores de similarité correspondants à une nouvelle réponse, nous suivons ces étapes :

2.1 Étape 1 : Normalisation des Scores de Similarité

Normaliser les scores de similarité sim_i en sim'_i de manière à ce que $\sum_{i=1}^n \text{sim}'_i = 1$:

$$\text{sim}'_i = \frac{\text{sim}_i}{\text{sim}_2 + \text{sim}_3}$$

où $\text{sim}_2 = 0.8400$ et $\text{sim}_3 = 0.7000$ sont les scores de similarité des deux meilleures réponses de référence.

$$\text{sim}'_2 = \frac{0.8400}{0.8400 + 0.7000} = \frac{0.8400}{1.5400} \approx 0.5455$$

$$\text{sim}'_3 = \frac{0.7000}{0.8400 + 0.7000} = \frac{0.7000}{1.5400} \approx 0.4545$$

2.2 Étape 2 : Ajustement Pondéré

Ajuster les scores de référence s_i en utilisant les scores de similarité normalisés sim'_2 et sim'_3 :

$$\text{score_inféré} = \text{sim}'_2 \cdot s_2 + \text{sim}'_3 \cdot s_3$$

Cette méthode assure que le score inféré est une moyenne pondérée des deux meilleurs scores de référence, où les poids sont déterminés par les scores de similarité normalisés fournis par le LLM.

3 Exemple

Considérons un exemple avec les scores de référence et les scores de similarité :

Score de Référence	Score de Similarité
1	0.2000
2	0.8400
3	0.7000

En suivant les étapes :

— **Étape 1** : Normaliser les deux meilleurs scores de similarité :

$$\text{sim}'_2 \approx 0.5455$$

$$\text{sim}'_3 \approx 0.4545$$

— **Étape 2** : Ajustement pondéré :

$$\text{score_inféré} = 0.5455 \cdot 2 + 0.4545 \cdot 3$$

$$\text{score_inféré} \approx 1.091 + 1.3635$$

$$\text{score_inféré} \approx 2.4545$$

Ainsi, le score inféré est d'environ 2.4545.

4 Conclusion

Cette technique raffinée d'ajustement offre une approche plus nuancée pour inférer des scores basée sur les scores de similarité des deux meilleures réponses de référence. En utilisant des poids normalisés, la méthode assure que le score inféré reflète l'importance relative de chaque score de référence en fonction de sa similarité avec la nouvelle réponse.