

Utilisation des LLM pour le raisonnement (mathématique)

MASHS 2024

J. Rynkiewicz

Université Paris 1

2024

Exemple de raisonnement par des LLM

Utilisation des LLM pour le raisonnement (mathématique)

J. Rynkiewicz

Introduction

Les LLM

Chaîne de Markov
Formatage du texte
Les transformers

LLM pour les mathématiques

Stratégies d'inférence
Affinement du modèle
Limites des LLM

Conclusion

- Les modèles de langage de grande taille (LLM) tels que GPT-4 ont révolutionné le domaine de l'intelligence artificielle en démontrant une capacité impressionnante à effectuer de nombreuses tâches.
- Pourtant, nous verrons qu'on peut décrire un LLM comme un "Perroquet stochastique".
- Le 28 septembre 2023, lors d'une présentation, Xavier Leroy montre les limites de raisonnement des LLM :
 - L'IA et ses défis : Xavier Leroy – Quelle confiance en les résultats d'une IA ?, \simeq 2ème minute à la 4ème minute.
 - <https://benchmarks.llmonitor.com/sally>
- Faisons l'expérience nous même : "Sally (une fille) a trois frères. Chaque frère a deux soeurs. Combien de soeurs a Sally ?".
 - ChatGTP (GPT4) (nombre de paramètres inconnu, mais certainement > 200 milliards).
 - Llama 3.1 70B (70 milliards de paramètres).
 - On peut changer le nombre de frères et de soeurs.

Un problème plus compliqué

Utilisation des LLM pour le raisonnement (mathématique)

J. Rynkiewicz

Introduction

Les LLM

- Chaîne de Markov
- Formatage du texte
- Les transformers

LLM pour les mathématiques

- Stratégies d'inférence
- Affinement du modèle
- Limites des LLM

Conclusion

What is the minimum value of $5x^2 + 5y^2 - 8xy$ when x and y range over all real numbers such that $|x - 2y| + |y - 2x| = 40$?

Un problème plus compliqué

What is the minimum value of $5x^2 + 5y^2 - 8xy$ when x and y range over all real numbers such that $|x - 2y| + |y - 2x| = 40$?

- En remarquant que le problème est invariant par $(x, y) \mapsto (-x, -y)$ et $(x, y) \mapsto (y, x)$, on peut se limiter au cas :

- $x \geq 0, y \leq 0$, où la contrainte devient

$$x - 2y + 2x - y = 3x - 3y = 40 \Rightarrow y = x - \frac{40}{3} \text{ pour } 0 \leq x \leq \frac{40}{3}.$$

$$5x^2 + 5y^2 - 8xy = 5x^2 + 5\left(x - \frac{40}{3}\right)^2 - 8x\left(x - \frac{40}{3}\right) = 2x^2 - \frac{80}{3}x + \frac{8000}{9}$$

dont la dérivée en x vaut $4x - \frac{80}{3}$, qui sera négative pour $0 \leq x \leq \frac{20}{3}$ et positive pour $\frac{20}{3} \leq x \leq \frac{40}{3}$ et le minimum sera atteint en

$$2\left(\frac{20}{3}\right)^2 - \frac{80}{3} \frac{20}{3} + \frac{8000}{9} = \frac{7200}{9} = 800$$

Un problème plus compliqué

Introduction

Les LLM

Chaîne de Markov
Formatage du texte
Les transformers

LLM pour les mathématiques

Stratégies d'inférence
Affinement du modèle
Limites des LLM

Conclusion

What is the minimum value of $5x^2 + 5y^2 - 8xy$ when x and y range over all real numbers such that $|x - 2y| + |y - 2x| = 40$?

- En remarquant que le problème est invariant par $(x, y) \mapsto (-x, -y)$ et $(x, y) \mapsto (y, x)$, on peut se limiter au cas :

- $x \geq 0, y \leq 0$, où la contrainte devient

$$x - 2y + 2x - y = 3x - 3y = 40 \Rightarrow y = x - \frac{40}{3} \text{ pour } 0 \leq x \leq \frac{40}{3}.$$

$$5x^2 + 5y^2 - 8xy = 5x^2 + 5\left(x - \frac{40}{3}\right)^2 - 8x\left(x - \frac{40}{3}\right) = 2x^2 - \frac{80}{3}x + \frac{8000}{9}$$

dont la dérivée en x vaut $4x - \frac{80}{3}$, qui sera négative pour $0 \leq x \leq \frac{20}{3}$ et positive pour $\frac{20}{3} \leq x \leq \frac{40}{3}$ et le minimum sera atteint en

$$2\left(\frac{20}{3}\right)^2 - \frac{80}{3}\frac{20}{3} + \frac{8000}{9} = \frac{7200}{9} = 800$$

- Si $x \geq y \geq 0$ alors la contrainte peut rester $x - 2y + 2x - y = 40$ mais alors $\frac{40}{3} \leq x \leq \frac{80}{3}$ et $2x^2 - \frac{80}{3}x + \frac{8000}{9}$ sera minimum en $x = \frac{40}{3}$ et

$$\text{vaudra } 2\left(\frac{40}{3}\right)^2 - \frac{80}{3}\frac{40}{3} + \frac{8000}{9} = \frac{8000}{9} \simeq 888.889 > 800$$

Un problème plus compliqué

What is the minimum value of $5x^2 + 5y^2 - 8xy$ when x and y range over all real numbers such that $|x - 2y| + |y - 2x| = 40$?

- En remarquant que le problème est invariant par $(x, y) \mapsto (-x, -y)$ et $(x, y) \mapsto (y, x)$, on peut se limiter au cas :
- $x \geq 0, y \leq 0$, où la contrainte devient $x - 2y + 2x - y = 3x - 3y = 40 \Rightarrow y = x - \frac{40}{3}$ pour $0 \leq x \leq \frac{40}{3}$.
 $5x^2 + 5y^2 - 8xy = 5x^2 + 5(x - \frac{40}{3})^2 - 8x(x - \frac{40}{3}) = 2x^2 - \frac{80}{3}x + \frac{8000}{9}$
dont la dérivée en x vaut $4x - \frac{80}{3}$, qui sera négative pour $0 \leq x \leq \frac{20}{3}$ et positive pour $\frac{20}{3} \leq x \leq \frac{40}{3}$ et le minimum sera atteint en $2 \left(\frac{20}{3}\right)^2 - \frac{80}{3} \frac{20}{3} + \frac{8000}{9} = \frac{7200}{9} = 800$
- Si $x \geq y \geq 0$ alors la contrainte peut rester $x - 2y + 2x - y = 40$ mais alors $\frac{40}{3} \leq x \leq \frac{80}{3}$ et $2x^2 - \frac{80}{3}x + \frac{8000}{9}$ sera minimum en $x = \frac{40}{3}$ et vaudra $2 \left(\frac{40}{3}\right)^2 - \frac{80}{3} \frac{40}{3} + \frac{8000}{9} = \frac{8000}{9} \simeq 888.889 > 800$
- Finalement, la contrainte pourra être $2y - x + 2x - y = x + y = 40 \Rightarrow y = 40 - x$ et $5x^2 + 5y^2 - 8xy = 5x^2 + 5(40 - x)^2 - 8x(40 - x) = 18x^2 - 720x + 8000$ qui sera minimale en $x = \frac{320}{36} = 20$ et vaudra **800**.

Solution informatique du problème

What is the minimum value of $5x^2 + 5y^2 - 8xy$ when x and y range over all real numbers such that $|x - 2y| + |y - 2x| = 40$? On peut résoudre ce problème à l'aide d'une minimisation sous contrainte (avec le risque de converger vers un minimum local).

```
from scipy.optimize import minimize

def min_value_expression():
    # Define the expression and the constraint
    expression = lambda x: 5*x[0]**2+5*x[1]**2-8*x[0]*x[1]
    constraint = lambda x: abs(x[0]-2*x[1])+abs(x[1]-2*x[0])-40

    # Define the initial guess for the optimizer
    x0 = [0, 0]

    # Use scipy's minimize function to find the minimum
    # Using Sequential Least Squares Programming (SLSQP).
    result = minimize(expression, x0, method='SLSQP',
                      constraints={'type': 'eq', 'fun': constraint})

    # Return the minimum value of the expression
    return result

solution = min_value_expression()
print("Valeur_minimum_:", solution.fun, "en_", solution.x)
```

Un LLM aurait-il pu trouver une des deux solutions?

Un LLM est une Chaîne de Markov

Utilisation des LLM pour le raisonnement (mathématique)

J. Rynkiewicz

Introduction

Les LLM

Chaîne de Markov

Formatage du texte

Les transformers

LLM pour les
mathématiques

Stratégies d'inférence

Affinement du modèle

Limites des LLM

Conclusion

- Un LLM ne sait faire qu'une seule tâche : Prévoir le prochain "mot" connaissant les p "mots" précédents.
- Pour cela, il calcule la probabilité conditionnelle de tous les mots sachant les p "mots" précédents :

$$P(X_{t+1} | X_t, \dots, X_{t-p})$$

- C'est une chaîne de Markov, homogène, d'ordre p .
- Dans une langue, il y a plusieurs millions de "mots" différents, donc on simplifie le dictionnaire de mots possibles : Les "tokens".
- Le plus simple exemple de "tokens" sera les lettres de l'alphabet et les signes de ponctuation.
- On peut implémenter à la main un LLM sur ce petit dictionnaire. C'est un LLM pour les caractères.

Deux exemples de trajectoires de Chaînes de Markov

On a estimé les paramètres d'une chaîne de Markov pour modéliser la suite des caractères de l'oeuvre de Shakespeare.

- Chaîne de Markov d'ordre 1 ($p = 1$).

$$\hat{P}(X_{t+1} = e_k | X_t = e_l) = \frac{\sum_{t=1}^{n-1} \mathbf{1}_{\{e_l, e_k\}}(X_t, X_{t+1})}{\sum_{t=1}^{n-1} \mathbf{1}_{\{e_l\}}(X_t)}$$

- Après l'estimation, on génère une trajectoire :
<https://www.kaggle.com/code/josephrynkiewicz/markov>

Deux exemples de trajectoires de Chaînes de Markov

On a estimé les paramètres d'une chaîne de Markov pour modéliser la suite des caractères de l'oeuvre de Shakespeare.

- Chaîne de Markov d'ordre 1 ($p = 1$).

$$\hat{P}(X_{t+1} = e_k | X_t = e_l) = \frac{\sum_{t=1}^{n-1} \mathbf{1}_{\{e_l, e_k\}}(X_t, X_{t+1})}{\sum_{t=1}^{n-1} \mathbf{1}_{\{e_l\}}(X_t)}$$

- Après l'estimation, on génère une trajectoire :
<https://www.kaggle.com/code/josephrynkiewicz/markov>
- Chaîne de Markov d'ordre 200 ($p = 200$).

$$\hat{P}(X_{t+1} = e_k | X_t = e_l, \dots, | X_{t-199} = e_{l'}) = T_{\hat{\theta}}(X_t = e_l, \dots, | X_{t-199} = e_{l'}),$$

où $T_{\hat{\theta}}(\cdot)$ est un "Transformer".

Deux exemples de trajectoires de Chaînes de Markov

On a estimé les paramètres d'une chaîne de Markov pour modéliser la suite des caractères de l'oeuvre de Shakespeare.

- Chaîne de Markov d'ordre 1 ($p = 1$).

$$\hat{P}(X_{t+1} = e_k | X_t = e_l) = \frac{\sum_{t=1}^{n-1} \mathbf{1}_{\{e_l, e_k\}}(X_t, X_{t+1})}{\sum_{t=1}^{n-1} \mathbf{1}_{\{e_l\}}(X_t)}$$

- Après l'estimation, on génère une trajectoire :
<https://www.kaggle.com/code/josephrynkiewicz/markov>
- Chaîne de Markov d'ordre 200 ($p = 200$).

$$\hat{P}(X_{t+1} = e_k | X_t = e_l, \dots, | X_{t-199} = e_{l'}) = T_{\hat{\theta}}(X_t = e_l, \dots, | X_{t-199} = e_{l'}),$$

où $T_{\hat{\theta}}(\cdot)$ est un "Transformer".

- Après l'estimation on génère une trajectoire :
<https://www.kaggle.com/code/josephrynkiewicz/generative-transformer>

Deux exemples de trajectoires de Chaînes de Markov

On a estimé les paramètres d'une chaîne de Markov pour modéliser la suite des caractères de l'oeuvre de Shakespeare.

- Chaîne de Markov d'ordre 1 ($p = 1$).

$$\hat{P}(X_{t+1} = e_k | X_t = e_l) = \frac{\sum_{t=1}^{n-1} \mathbf{1}_{\{e_l, e_k\}}(X_t, X_{t+1})}{\sum_{t=1}^{n-1} \mathbf{1}_{\{e_l\}}(X_t)}$$

- Après l'estimation, on génère une trajectoire :
<https://www.kaggle.com/code/josephrynkiewicz/markov>
- Chaîne de Markov d'ordre 200 ($p = 200$).

$$\hat{P}(X_{t+1} = e_k | X_t = e_l, \dots, | X_{t-199} = e_{l'}) = T_{\hat{\theta}}(X_t = e_l, \dots, | X_{t-199} = e_{l'}),$$

où $T_{\hat{\theta}}(\cdot)$ est un "Transformer".

- Après l'estimation on génère une trajectoire :
<https://www.kaggle.com/code/josephrynkiewicz/generative-transformer>
- Le changement d'échelle du modèle fait émerger de nouveaux comportements !

- Découper un texte en caractères ne permet pas d'obtenir des résultats proches de l'état de l'art.
- "Tokenizer" un texte consiste en le découper en mots, parties de mot ou caractères. Il sont ensuite encodés en nombres (ids).
- Les deux techniques les plus employées sont : Byte-Pair Encoding (BPE) et WordPiece.
- L'idée de garder entier les mots très utilisés, mais de découper les mots rares ou transformés par la grammaire (adverbe etc...)
- La "tokenization" permet de garder une taille raisonnable pour le vocabulaire (quelques dizaines de milliers de tokens).
- Les meilleurs ensembles de "tokens" sont contruits en gardant les mots (resp. parties de mot) les plus fréquents (BPE) ou bien ceux qui maximisent la vraisemblance du texte (WordPiece).

Les Transformers (Attention is all you need (2017))

Utilisation des LLM
pour le raisonnement
(mathématique)

J. Rynkiewicz

Introduction

Les LLM

Chaîne de Markov

Formatage du texte

Les transformers

LLM pour les
mathématiques

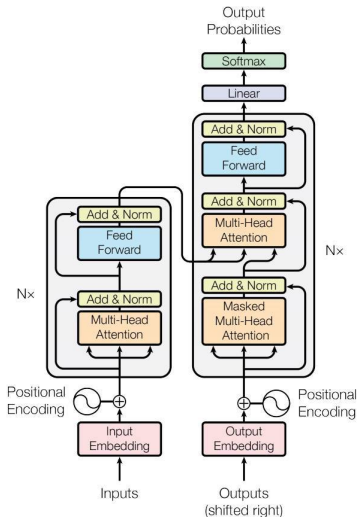
Stratégies d'inférence

Affinement du modèle

Limites des LLM

Conclusion

Ces réseaux ont des modules d'attention qui permettent de tenir compte du contexte des mots.



Le mécanisme d'attention

Utilisation des LLM pour le raisonnement (mathématique)

J. Rynkiewicz

Introduction

Les LLM

Chaîne de Markov

Formatage du texte

Les transformers

LLM pour les mathématiques

Stratégies d'inférence

Affinement du modèle

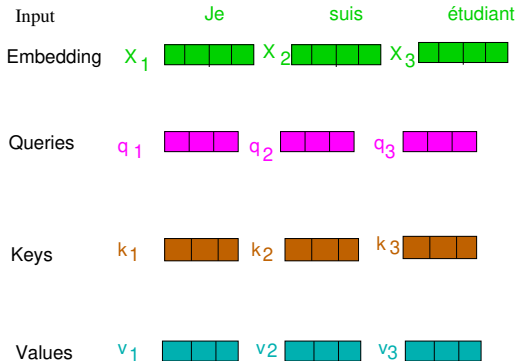
Limites des LLM

Conclusion

- L'attention permet de tenir compte du contexte d'un mot.
- Elle est calculée sur toute la phrase (suite de tokens) de l'entrée.
- Pour traduire la phrase "La brebis n'a pas traversé la rue parce qu'elle était trop fatiguée"
- À quoi fait référence "elle" dans le texte ? La brebis ou bien la rue ?
- C'est une question facile pour un être humain, mais difficile pour une machine.
- La machine doit donc estimer si le mot "elle" est plus lié au mot "brebis" ou au mot "rue".
- La couche de "Self-attention" des transformers proposent une méthode pour permettre cette estimation.
- L'attention remplace chacun des mots de la phrase par une moyenne pondérée des mots de la phrase.

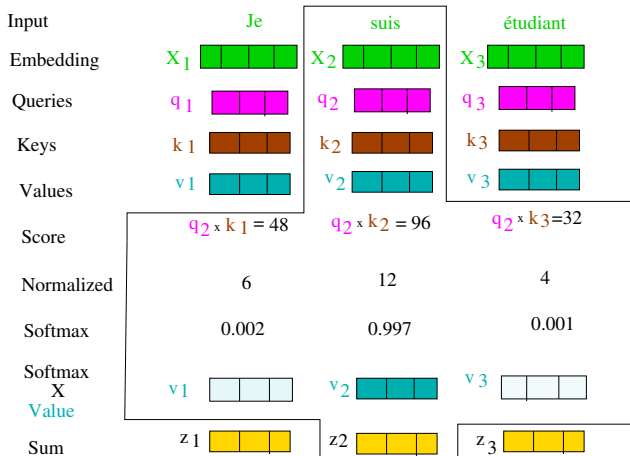
Détail de la couche de “Self-attention” (1)

- On commence par définir trois projections de l'entrée : Le vecteur “Query”, le vecteur “Key” et le vecteur “Value”.
- Ces vecteurs sont créés en multipliant l'entrée de la couche par des matrices de poids.



Détail de la couche de "Self-attention" (2)

- L'embedding est transformé (par des multiplication matricelles) en "query", "key" et "value" puis chaque "value" est pondérée par un softmax du score induit par toutes les keys.
- La somme pondérés des "values" est la sortie "z" de la couche d'attention.



Le transformer a été entraîné à prévoir le “token” suivant. Il calcule seulement la loi de probabilité du prochain “token”, il faut une méthode pour choisir le “token” suivant. Les principales méthodes sont :

- Greedy search : On choisit simplement le “token” le plus probable. Cela a tendance à créer des séquences répétitives.
- Beam search : On calcule les probabilités des suites de “tokens” de longueur fixée NB , et on choisit la suite de “tokens” la plus probable. NB est un hyperparamètre à régler. Cette méthode a toujours tendance à créer des répétitions.
- Top K sampling : On tire aléatoirement parmi les K “tokens” les plus probables. Il faut choisir l'hyperparamètre K .
- Top P sampling : On choisit K telle que la somme des probabilités de K “tokens” soit (à peu près) égale à P .
- Dans la pratique, on mélange souvent les deux méthodes précédentes avec un K maximum.
- On peut modifier la loi conditionnelle grâce à une “température” qui va rendre plus probables (ou moins) les “tokens” les plus probables.

Utilisation de transformer pour les mathématiques

Utilisation des LLM pour le raisonnement (mathématique)

J. Rynkiewicz

Introduction

Les LLM

Chaîne de Markov
Formatage du texte
Les transformers

LLM pour les mathématiques

Stratégies d'inférence
Affinement du modèle
Limites des LLM

Conclusion

- Le Transformer est une chaîne de Markov, on tire au hasard le prochain "token" suivant la probabilité conditionnelle de la chaîne de Markov.
- On peut l'appeler "Perroquet stochastique".
- Un "Perroquet stochastique" peut-il raisonner ?
- Ici par raisonnement on entend : Résoudre un problème de mathématiques (qui n'a jamais été vu par le modèle pendant l'entraînement).
- Le raisonnement serait un phénomène émergent : Il apparait sans avoir entraîné le modèle à raisonner.
- Il ne faut pas perdre de vue que la génération du texte reste stochastique.
- Si on lui pose deux fois la même question, on aura deux réponses différentes.
- Peut-on favoriser (ou défavoriser) le raisonnement ?

What is the minimum value of $5x^2 + 5y^2 - 8xy$ when x and y range over all real numbers such that $|x - 2y| + |y - 2x| = 40$?

- Demandons à ChatGPT
- Demandons à LLaMa 3.1 70b

Stratégie général pour le raisonnement mathématiques

Utilisation des LLM pour le raisonnement (mathématique)

J. Rynkiewicz

Introduction

Les LLM

Chaîne de Markov
Formatage du texte
Les transformers

LLM pour les mathématiques

Stratégies d'inférence
Affinement du modèle
Limites des LLM

Conclusion

- Il faut faire “réfléchir” le modèle : Chaîne de Raisonnement (chain of thought : CoT).
 - Par exemple, on rajoute la phrase : “Raisonne pas à pas” dans le prompt.
 - Cela permet de vérifier les éventuelles erreurs de raisonnement.
 - Le modèle va dépenser plus d'énergie pour les questions difficiles.
- Le modèle peut utiliser des outils pour les calculs compliqués. C'est à dire qu'il peut écrire un programme en python pour résoudre la question.
- Mais le LLM ne peut pas deviner qu'elle va être le résultat du programme, il va donc falloir lui donner.
- Pour lui donner le résultats, on ajoute ce résultat au texte généré et on lui redonne un nouveau prompt qui comprend : l'ancien prompt, ce qu'il a généré, et la réponse du programme.
- On recommence ce processus jusqu'à ce que le programme donne la réponse.
- ⇒ ToRA : A Tool-Integrated Reasoning Agent for Mathematical Problem Solving (ICLR 2024)

Modèle spécialisé pour les mathématiques

Utilisation des LLM pour le raisonnement (mathématique)

J. Rynkiewicz

Introduction

Les LLM

Chaîne de Markov
Formatage du texte
Les transformers

LLM pour les mathématiques

Stratégies d'inférence
Affinement du modèle
Limites des LLM

Conclusion

- On va utiliser le LLM DeepSeekMath 7B-RL
- C'est un modèle LLaMa2, open source, entraîné par une entreprise chinoise qui comprends 7 milliards de paramètres (petit LLM).
- Le prompt utilisé est : "Below is a math problem you are to solve : "What is the minimum value of $5x^2 + 5y^2 - 8xy$ when x and y range over all real numbers such that $|x - 2y| + |y - 2x| = 40$?" Analyze this problem and think step by step to come to a solution with programs. After solving the problem, output the final numerical answer within boxed."
- Le modèle est utilisé dans un programme python, pour l'"aider" à résoudre le problème grâce à ToRA.

Modèle spécialisé pour les mathématiques

Utilisation des LLM pour le raisonnement (mathématique)

J. Rynkiewicz

Introduction

Les LLM

Chaîne de Markov
Formatage du texte
Les transformers

LLM pour les mathématiques

Stratégies d'inférence
Affinement du modèle
Limites des LLM

Conclusion

- On va utiliser le LLM DeepSeekMath 7B-RL
- C'est un modèle LLaMa2, open source, entraîné par une entreprise chinoise qui comprends 7 milliards de paramètres (petit LLM).
- Le prompt utilisé est : "Below is a math problem you are to solve : "What is the minimum value of $5x^2 + 5y^2 - 8xy$ when x and y range over all real numbers such that $|x - 2y| + |y - 2x| = 40$?" Analyze this problem and think step by step to come to a solution with programs. After solving the problem, output the final numerical answer within boxed."
- Le modèle est utilisé dans un programme python, pour l'"aider" à résoudre le problème grâce à ToRA.
- Comme la génération est stochastique, les résultats dépendront de la graine.

Modèle spécialisé pour les mathématiques

Utilisation des LLM pour le raisonnement (mathématique)

J. Rynkiewicz

Introduction

Les LLM

Chaîne de Markov
Formatage du texte
Les transformers

LLM pour les mathématiques

Stratégies d'inférence
Affinement du modèle
Limites des LLM

Conclusion

- On va utiliser le LLM DeepSeekMath 7B-RL
- C'est un modèle LLaMa2, open source, entraîné par une entreprise chinoise qui comprends 7 milliards de paramètres (petit LLM).
- Le prompt utilisé est : "Below is a math problem you are to solve : "What is the minimum value of $5x^2 + 5y^2 - 8xy$ when x and y range over all real numbers such that $|x - 2y| + |y - 2x| = 40$?" Analyze this problem and think step by step to come to a solution with programs. After solving the problem, output the final numerical answer within boxed."
- Le modèle est utilisé dans un programme python, pour l'"aider" à résoudre le problème grâce à ToRA.
- Comme la génération est stochastique, les résultats dépendront de la graine.
- Comment fiabiliser les résultats ?

Modèle spécialisé pour les mathématiques

Utilisation des LLM pour le raisonnement (mathématique)

J. Rynkiewicz

Introduction

Les LLM

Chaîne de Markov
Formatage du texte
Les transformers

LLM pour les mathématiques

Stratégies d'inférence
Affinement du modèle
Limites des LLM

Conclusion

- On va utiliser le LLM DeepSeekMath 7B-RL
- C'est un modèle LLaMa2, open source, entraîné par une entreprise chinoise qui comprends 7 milliards de paramètres (petit LLM).
- Le prompt utilisé est : "Below is a math problem you are to solve : "What is the minimum value of $5x^2 + 5y^2 - 8xy$ when x and y range over all real numbers such that $|x - 2y| + |y - 2x| = 40$?" Analyze this problem and think step by step to come to a solution with programs. After solving the problem, output the final numerical answer within boxed."
- Le modèle est utilisé dans un programme python, pour l'"aider" à résoudre le problème grâce à ToRA.
- Comme la génération est stochastique, les résultats dépendront de la graine.
- Comment fiabiliser les résultats ?
- On recommence plusieurs fois la procédure et on garde la solution majoritaire : "Consistency".
- La solution majoritaire est la solution du code confirmée par le texte du LLM.

Modèle spécialisé pour les mathématiques

Utilisation des LLM pour le raisonnement (mathématique)

J. Rynkiewicz

Introduction

Les LLM

Chaîne de Markov
Formatage du texte
Les transformers

LLM pour les mathématiques

Stratégies d'inférence
Affinement du modèle
Limites des LLM

Conclusion

- On va utiliser le LLM DeepSeekMath 7B-RL
- C'est un modèle LLaMa2, open source, entraîné par une entreprise chinoise qui comprends 7 milliards de paramètres (petit LLM).
- Le prompt utilisé est : "Below is a math problem you are to solve : "What is the minimum value of $5x^2 + 5y^2 - 8xy$ when x and y range over all real numbers such that $|x - 2y| + |y - 2x| = 40$?" Analyze this problem and think step by step to come to a solution with programs. After solving the problem, output the final numerical answer within boxed."
- Le modèle est utilisé dans un programme python, pour l'"aider" à résoudre le problème grâce à ToRA.
- Comme la génération est stochastique, les résultats dépendront de la graine.
- Comment fiabiliser les résultats ?
- On recommence plusieurs fois la procédure et on garde la solution majoritaire : "Consistency".
- La solution majoritaire est la solution du code confirmée par le texte du LLM.
- "CoT+ToRA+Consistency" est l'état de l'art pour l'inférence.

Comment améliorer le modèle ?

- Pour améliorer le modèle, il faut modifier ses poids sur de nouvelles données.
- Affinement du modèle sur un corpus sélectionné.
 - Le modèle de DeepSeek est, à l'origine, un modèle qui écrit du code informatique (DeepSeek coder) et dont les poids ont été affinés sur un ensemble de pages web spécialisées en math.
 - Les auteurs ont essayé d'utiliser les articles de mathématiques sur Arxiv mais cela a dégradé les résultats.
 - \Rightarrow DeepSeek base
- Affinage supervisé.
 - Construction d'une base au bon format sur des bases connus.
 - Le bon format du prompt est de poser le problème en invitant à le résoudre par un raisonnement et des programmes.
 - Le bon format des solutions (CoT+ToRA).
 - \Rightarrow DeepSeek instruct
- Pour ces deux étapes il faut des données de très grande qualités.
- Cela nécessite de sélectionner et de créer les données à la main. C'est très coûteux en main d'oeuvre humaine.
- Il faut utiliser pendant des heures (jours, semaines) du matériel performant pour faire apprendre le modèle.

Continuer à améliorer le modèle.

Utilisation des LLM pour le raisonnement (mathématique)

J. Rynkiewicz

Introduction

Les LLM

Chaîne de Markov

Formatage du texte

Les transformers

LLM pour les
mathématiques

Stratégies d'inférence

Affinement du modèle

Limites des LLM

Conclusion

On peut continuer à affiner le modèle avec l'apprentissage par renforcement.

- L'apprentissage par renforcement est une méthode d'apprentissage automatique dans laquelle un agent (un programme) apprend à prendre des décisions en interagissant avec un environnement. Il reçoit une récompense ou une punition en fonction des actions qu'il effectue.
- L'objectif de l'agent est de maximiser la récompense (bonne réponse à la question). Pour cela, il explore différentes actions et apprend à choisir celles qui lui rapportent le plus de bénéfices à long terme.
- Ce processus est souvent comparé à un jeu où l'agent apprend en essayant, en se trompant et en ajustant son comportement pour s'améliorer progressivement.
- Un algorithme d'apprentissage par renforcement populaire est "Proximal Policy Optimization (PPO)".
- Au lieu d'être optimisé pour prévoir le token suivant, ils sont optimisés pour prévoir la bonne réponse finale.
- Cette méthode est extrêmement coûteuse en calcul.
- ⇒ DeepSeek rl

Retour sur le problème de Sally

Utilisation des LLM pour le raisonnement (mathématique)

J. Rynkiewicz

Introduction

Les LLM

Chaîne de Markov
Formatage du texte
Les transformers

LLM pour les mathématiques

Stratégies d'inférence
Affinement du modèle

Limites des LLM

Conclusion

- Le modèle DeepSeekMath 7B-RL est basé sur le modèle LLaMa 2, qui échouait sur le problème du nombre de soeurs de Sally.
- On va utiliser la méthode “CoT+ToRA+Consistency” pour ce problème.

Retour sur le problème de Sally

Utilisation des LLM pour le raisonnement (mathématique)

J. Rynkiewicz

Introduction

Les LLM

Chaîne de Markov
Formatage du texte
Les transformers

LLM pour les mathématiques

Stratégies d'inférence
Affinement du modèle

Limites des LLM

Conclusion

- Le modèle DeepSeekMath 7B-RL est basé sur le modèle LLaMa 2, qui échouait sur le problème du nombre de soeurs de Sally.
- On va utiliser la méthode “CoT+ToRA+Consistency” pour ce problème.
- Dans ce cas, le modèle n'a jamais été capable d'écrire un code valide.
- La solution majoritaire est donc celle du texte généré par le LLM.

Retour sur le problème de Sally

Utilisation des LLM pour le raisonnement (mathématique)

J. Rynkiewicz

Introduction

Les LLM

Chaîne de Markov
Formatage du texte
Les transformers

LLM pour les mathématiques

Stratégies d'inférence
Affinement du modèle
Limites des LLM

Conclusion

- Le modèle DeepSeekMath 7B-RL est basé sur le modèle LLaMa 2, qui échouait sur le problème du nombre de soeurs de Sally.
- On va utiliser la méthode “CoT+ToRA+Consistency” pour ce problème.
- Dans ce cas, le modèle n'a jamais été capable d'écrire un code valide.
- La solution majoritaire est donc celle du texte généré par le LLM.
- Cette méthode n'est pas parfaite, il fait encore des erreurs sur des problèmes simples.

Conclusion

Utilisation des LLM pour le raisonnement (mathématique)

J. Rynkiewicz

Introduction

Les LLM

Chaîne de Markov
Formatage du texte
Les transformers

LLM pour les mathématiques

Stratégies d'inférence
Affinement du modèle
Limites des LLM

Conclusion

- Les LLM sont des “Perroquets stochastiques” (Chaînes de Markov qui tirent au hasard les tokens suivants).
- Pourtant, ils peuvent résoudre des problèmes mathématiques.
- Les modèles généralistes (ChatGPT) sont encore limités pour certains problèmes (notamment ceux qui demandent des calculs).
- Les petits modèles spécialisés sur un domaine restreint peuvent avoir de meilleures performances que les modèles généralistes sur ce domaine.
- Il faut les aider : “CoT+ToRA+Consistency”, car ils restent peu fiables.
- On peut imaginer d'utiliser ces modèles en sciences humaines pour la classification.
- Mieux comprendre ce qu'est le raisonnement (philosophie, sciences cognitives) permettra certainement de mieux cerner les limites de ce modèle.