УДК 004.82
DOI: 10.18523/2617-3808.2024.7.98-101

Андрощук М. В.

СУЧАСНІ ПІДХОДИ ДО ВИКОРИСТАННЯ БАЗ ЗНАНЬ ДЛЯ ВИРІШЕННЯ ПРОБЛЕМ ВЕЛИКИХ МОВНИХ МОДЕЛЕЙ

Розглянуто можливості інтеграції великих мовних моделей (ВММ) із базами знань для підвищення точності й надійності їхніх відповідей. Визначено переваги такого поєднання, зокрема зменшення ризику галюцинацій — генерування неправильної або вигаданої інформації. Проаналізовано різні підходи до об'єднання ВММ із базами знань, їхні сильні та слабкі сторони. Обговорено перспективи та виклики у застосуванні цієї технології в різних галузях, зокрема інформаційний пошук, підтримку прийняття рішень та автоматизоване створення контенту. Стаття надає огляд сучасного стану досліджень у цій сфері та окреслює напрями для подальшого вивчення.

Ключові слова: ВММ, база знань, галюцинації, RAG.

Вступ

Великі мовні моделі (ВММ) пройшли значну еволюцію від простих статистичних підходів до складних нейронних мереж, здатних розуміти й генерувати людську мову на високому рівні. Незважаючи на їхній потенціал у різних сферах застосування, ВММ стикаються з низкою викликів, як-от проблеми з точністю та достовірністю генерованої інформації, потенційні упередження, етичні питання та технічні обмеження. Особливу увагу привертає проблема «галюцинацій» — генерування неправдивої, але правдоподібної інформації. Для подолання цих викликів розробляють різні підходи, зокрема техніки самоперевірки, навчання на якісніших даних і використання баз знань як достовірних джерел інформації.

Розвиток ВММ

Великі мовні моделі мають довгу історію розвитку, яка сягає кількох десятиліть. Їхня еволюція тісно пов'язана з прогресом у галузі оброблення природної мови та штучного інтелекту. Ці моделі пройшли шлях від простих статистичних підходів до складних нейронних мереж, що здатні розуміти та генерувати людську мову на високому рівні.

На початку свого шляху ці моделі базувалися на простих статистичних підходах, які давали змогу аналізувати текст на базовому рівні. Вони могли виконувати такі завдання, як підрахунок частоти слів або визначення найбільш імовірного наступного слова в реченні. Однак ці ранні моделі мали обмежені можливості й не могли впоратися зі складними лінгвістичними структурами. З розвитком комп'ютерних технологій і появою більш потужних алгоритмів машинного навчання мовні моделі почали ставати все більш складними. Вони перейшли від простих статистичних методів до використання нейронних мереж, що дало їм змогу краще розуміти контекст та семантику мови.

Сучасні великі мовні моделі являють собою складні нейронні мережі, які навчаються на величезних обсягах текстових даних. Вони здатні не лише розуміти й аналізувати людську мову, а й генерувати власний текст, який часто важко відрізнити від написаного людиною. Ці моделі можуть виконувати широкий спектр завдань, зокрема переклад, відповіді на запитання, написання есе та навіть створення програмного коду [1].

Важливо зазначити, що розвиток великих мовних моделей триває. Дослідники постійно працюють над удосконаленням цих технологій, прагнучи зробити їх ще більш ефективними, точними та етичними у використанні. Майбутнє цих моделей обіцяє ще більш вражальні можливості у розумінні та генерації людської мови, що може мати значний вплив на різні сфери нашого життя — від освіти до бізнесу та наукових досліджень. Попри переваги і перспективи великі мовні моделі мають і проблеми.

Отже, великі мовні моделі демонструють значний потенціал і переваги, які можуть революціонізувати багато галузей та аспектів нашого життя. Ці передові технології здатні обробляти та генерувати людську мову на рівні, який раніше вважали неможливим. Їх застосовують у різноманітних сферах, як-от автоматизований переклад, створення контенту, аналіз даних та підтримка клієнтів. Крім того, великі мовні моделі відкривають нові можливості для наукових досліджень, освіти і творчості.

Однак, поряд із цими надзвичайними можливостями, великі мовні моделі стикаються з певними викликами та обмеженнями, які не можна ігнорувати. Серед них — проблеми з точністю та достовірністю генерованої інформації, потенційні упередження в навчальних даних, які можуть призвести до несправедливих або дискримінаційних результатів, а також етичні питання, пов'язані з конфіденційністю даних і можливим зловживанням технологією. Крім того, існують технічні обмеження, як-от високі обчислювальні вимоги та енергоспоживання, які можуть обмежувати широке впровадження цих моделей.

Хоча ці технології відкривають нові можливості в різних галузях, від медицини до фінансів, від освіти до розваг, важливо критично оцінювати їх застосування та продовжувати інтенсивні дослідження для подолання наявних проблем. Це передбачає розробку більш прозорих та інтерпретованих моделей, удосконалення методів навчання для зменшення упереджень, а також створення надійних механізмів контролю та верифікації генерованого контенту.

Подальший розвиток і вдосконалення великих мовних моделей потребуватиме збалансованого підходу, який враховуватиме як їхні незаперечні переваги, так і потенційні ризики. Це вимагатиме тісної співпраці між дослідниками, розробниками, політиками та громадськістю для створення етичних рамок і регуляторних механізмів, які забезпечать відповідальне використання цих потужних технологій. Важливо також інвестувати в освіту та підвищувати обізнаність суспільства щодо можливостей та обмежень великих мовних моделей, щоб сприяти їх ефективному та безпечному використанню. У кінцевому підсумку, майбутнє великих мовних моделей залежатиме від здатності максимізувати їхній потенціал, одночасно мінімізуючи пов'язані з ними ризики.

Проблеми LLM

Попри досягнення, актуальними залишаються й проблеми: галюцинації, етичність, прозорість, обмеженість предметних областей. Причина більшості проблем лежить у самій природі великих мовних моделей і їхнього оброблення знань [6].

Проблема галюцинацій у великих мовних моделей є однією з найбільш актуальних. Окрім цього, упередження даних призводять до аналогічної упередженості з боку великих мовних моделей [4].

Це явище виникає, коли модель генерує неправдиву або неточну інформацію, яка виглядає правдоподібно. Галюцинації можуть призвести до поширення дезінформації та зниження довіри до AI-систем. Дослідники активно працюють над розробленням методів для виявлення та зменшення галюцинацій у мовних моделях.

Про що знають мовні моделі, важливо розвивати розуміння для можливості гарантування точності замість повноти [7]. Надійність їхніх знань повністю залежить від навчальних даних. Якщо таких даних не вистачає для відповіді, потрібно залучати інструменти, що дають можливість отримувати актуальні дані, незалежні від тренувальних наборів. Наприклад, для медицини це є принциповим моментом, щоб забезпечити правильність лікування і прийняття рішень [8].

Одним із перспективних підходів є використання техніки «самоперевірки», коли модель оцінює власні відповіді на достовірність. Інший метод полягає у навчанні моделей на більш якісних і перевірених даних, щоб зменшити ймовірність генерації неправдивої інформації. Крім того, розробляють системи зовнішньої верифікації, які можуть автоматично перевіряти вихідні дані моделі на відповідність фактам із надійних джерел.

Пропозиції щодо покращення через бази знань

Для розгляду частини проблем, пов'язаних із точністю відповідей, запропоновано підхід із використанням баз знань як достовірних джерел інформації.

Графи знань є потужним інструментом для представлення та аналізу складних взаємозв'язків між різними концепціями та фактами [2]. Вони дозволяють візуалізувати інформацію у вигляді мережі взаємопов'язаних вузлів, що полегшує розуміння складних систем і процесів. Завдяки наявності

інфраструктури для роботи зі знаннями в такому випадку з'являється передбачуваність, визначеність і актуальність даних при інтеграції з ВММ. У галузі охорони здоров'я бази знань відіграють критичну роль у забезпеченні точності діагностики та лікування [5]. Вони містять структуровану інформацію про симптоми, захворювання, методи лікування та взаємодії лікарських засобів, що допомагає медичним працівникам приймати обґрунтовані рішення.

Інтеграція великих мовних моделей із базами знань відкриває нові можливості для автоматичного збагачення та оновлення інформації. Дослідження демонструють потенціал таких рішень для отримання точних і повних результатів [3]. Це дає змогу підтримувати актуальність баз знань і розширювати їх охоплення, використовуючи останні досягнення у сфері оброблення природної мови.

Підхід із використанням баз знань як достовірних джерел інформації є перспективним напрямом у галузі оброблення та аналізу даних. Цей метод передбачає створення та використання структурованих сховищ інформації, які містять верифіковані факти, концепції та взаємозв'язки між ними. Бази знань можуть охоплювати різноманітні галузі, від медицини до інженерії, і слугувати надійним фундаментом для прийняття рішень і проведення досліджень.

Використання баз знань дозволяє підвищити точність і надійність інформаційних систем, мінімізуючи ризик помилок та неточностей, які можуть виникнути при використанні неверифікованих джерел. Крім того, цей підхід сприяє ефективному обміну знаннями між експертами та полегшує процес навчання для новачків у певній галузі. Інтеграція баз знань із сучасними технологіями штучного інтелекту та машинного навчання відкриває нові можливості для створення інтелектуальних систем, здатних аналізувати складні ситуації та надавати обґрунтовані рекомендації на основі достовірної інформації.

Особливо корисними такі рішення є для RAG (Retrieval Augmented Generation) систем, бо поєднує в собі потужність великих мовних моделей і точність фактуальної інформації. Прикладами таких систем є: питально-відповідальні, пошукові, підтримки прийняття рішень та інші.

Висновки

Великі мовні моделі демонструють значний прогрес у розумінні та генеруванні людської мови, але їхній розвиток супроводжується низкою викликів. Ці моделі здатні виконувати широкий спектр завдань, від перекладу та узагальнення тексту до створення оригінального контенту, що робить їх потужним інструментом у різних галузях. Однак, незважаючи на можливості, ВММ стикаються з серйозними проблемами, такими як упередженість у даних, на яких вони навчаються, що може призвести до відтворення та посилення соціальних стереотипів.

Крім того, ВММ часто мають труднощі в забезпеченні послідовності та достовірності генерованої інформації, особливо при роботі зі складними або спеціалізованими темами. Етичні питання, пов'язані з конфіденційністю даних та потенційним зловживанням технологією, також залишаються актуальними. Дослідники і розробники активно працюють над вирішенням цих проблем, прагнучи створити більш надійні, прозорі та етично відповідальні моделі. Незважаючи на ці виклики, потенціал ВММ для революціонізації взаємодії людини з машиною та розширення можливостей у галузі оброблення природної мови залишається величезним.

Використання баз знань дає змогу моделям отримувати доступ до актуальної та перевіреної інформації. Це допомагає зменшити кількість помилок і неточностей у відповідях моделей. Крім того, бази знань можуть регулярно оновлюватися, що забезпечує моделям доступ до найсвіжіших даних.

Список літератури

1. Cooper N. Perplexed: Understanding When Large Language Models are Confused [Electronic resource] / N. Cooper, T. Scholak // arXiv.org. — Mode of access: https://doi.org/10.48550/arxiv.2404.06634 (date of access: 21.09.2024).
2. Min B. Recent Advances in Natural Language Processing via Large Pre-Trained Language Models: A Survey ACM Computing Surveys [Electronic resource] / B. Min et al. — 2023. — Mode of access: https://doi.org/10.1145/3605943 (date of access: 21.09.2024).
3. Nayak A., Timmapathini H. P. LLM2KB: Constructing Knowledge Bases using instruction tuned context aware Large Language Models [Electronic resource] / A. Nayak, H. P. Timmapathini // arXiv.org. — Mode of access: https://doi.org/10.48550/arxiv.2308.13207 (date of access: 21.09.2024).
4. Ranaldi L. A Trip Towards Fairness: Bias and De-Biasing in Large Language Models [Electronic resource] / L. Ranaldi et al. // arXiv.org. — Mode of access: https://doi.org/10.48550/arxiv.2305.13862 (date of access: 21.09.2024).
5. Wang H. Knowledge-tuning Large Language Models with Structured Medical Knowledge Bases for Trustworthy Response Generation in Chinese ACM Transactions on Knowledge Discovery from Data [Electronic resource] / H. Wang et al. — 2024. — Mode of access: https://doi.org/10.1145/3686807 (date of access: 21.09.2024).

6. Yang L. Give Us the Facts: Enhancing Large Language Models with Knowledge Graphs for Fact-aware Language Modeling IEEE Transactions on Knowledge and Data Engineering [Electronic resource] / L. Yang et al. — 2024. — Pp. 1–20. — Mode of access: https://doi.org/10.1109/tkde.2024.3360454 (date of access: 21.09.2024).
7. Yin Z. Do Large Language Models Know What They Don't Know? [Electronic resource] / Z. Yin et al. // arXiv.org. — Mode of access: https://doi.org/10.48550/arxiv.2305.18153 (date of access: 21.09.2024).
8. Zheng D. Large Language Models as Reliable Knowledge Bases? [Electronic resource] / D. Zheng, J. Z. Pan, M. Lapata // arXiv.org. — Mode of access: https://doi.org/10.48550/arxiv.2407.13578 (date of access: 21.09.2024).

References

Cooper, N., & Scholak, T. (n. d.). Perplexed: Understanding When Large Language Models are Confused. arXiv.org. https://doi.org/10.48550/arxiv.2404.06634.

Min, B., Ross, H., Sulem, E., Veyseh, A. P. B., Nguyen, T. H., Sainz, O., Agirre, E., Heintz, I., & Roth, D. (2023). Recent Advances in Natural Language Processing via Large Pre-Trained Language Models: A Survey. ACM Computing Surveys. https://doi.org/10.1145/3605943.

Nayak, A., & Timmapathini, H. P. (2023). LLM2KB: Constructing Knowledge Bases using instruction tuned context aware Large Language Models. arXiv.org. https://doi.org/10.48550/arxiv.2308.13207.

Ranaldi, L., Zanzotto, F. M., Onorati, D., Venditti, D., & Ruzzetti, E. S. (2023). A Trip Towards Fairness: Bias and De-Biasing in Large Language Models. arXiv.org. https://doi.org/10.48550/arxiv.2305.13862.

Wang, H., Zhao, S., Qiang, Z., Li, Z., Liu, C., Xi, N., Du, Y., Qin, B., & Liu, T. (2024). Knowledge-tuning Large Language Models with Structured Medical Knowledge Bases for Trustworthy Response Generation in Chinese. ACM Transactions on Knowledge Discovery from Data. https://doi.org/10.1145/3686807.

Yang, L., Chen, H., Li, Z., Ding, X., & Wu, X. (2024). Give Us the Facts: Enhancing Large Language Models with Knowledge Graphs for Fact-aware Language Modeling. *IEEE Transactions on Knowledge and Data Engineering*, 1–20. https://doi.org/10.1109/tkde.2024.3360454.

Yin, Z., Huang, X., Qiu, X., Wu, J., Guo, Q., & Sun, Q. (2023). Do Large Language Models Know What They Don't Know? arXiv.org. https://doi.org/10.48550/arxiv.2305.18153.

Zheng, D., Pan, J. Z., & Lapata, M. (2024). Large Language Models as Reliable Knowledge Bases? arXiv.org. https://doi.org/10.48550/arxiv.2407.13578.

M. Androshchuk

MODERN APPROACHES TO USING KNOWLEDGE BASES TO ADDRESS THE CHALLENGES OF LARGE LANGUAGE MODELS

This paper examines the potential of integrating Large Language Models (LLMs) with knowledge bases to imrpove the accuracy and reliability of their responses. The advantages of such a combination are evaluated, particularly in reducing the risk of hallucinations – the phenomenon where models generate erroneous or fabricated information. Various methodologies for combining LLMs with knowledge bases are analyzed, along with their respective advantages and limitations. The prospects and challenges of implementing this technology in diverse fields—such as information retrieval, decision support, and automated content creation—are discussed. The paper presents an overview of the current state of research in this domain and delineates directions for future investigation.The integration of LLMs with knowledge bases represents a significant advancement in artificial intelligence, aiming one of the key concerns regarding LLMs—their tendency to generate inaccurate or fabricated information, commonly referred to as hallucinations. This approach leverages the vast language understanding and generation capabilities of LLMs while grounding their outputs in structured and verified information from knowledge bases. The synergy between these two technologies has the potential to significantly enhance the reliability and factual accuracy of AI-generated responses across a wide range of applications. The methodologies for combining LLMs with knowledge bases differ in their implementation and effectiveness. Some approaches involve pre-training LLMs on curated knowledge bases, while others reference knowledge bases externally during the inference process. Each method presents its own set of advantages and challenges, such as balancing computational efficiency against accuracy and maintaining the fluency of LLM outputs while adhering strictly to factual information. The application of this integrated technology extends beyond mere information retrieval, showing promise in complex decision support systems, automated content creation for specialized domains, and contributing to the advancement of explainable AI by providing traceable sources for generated information. As research in this area progresses, it is expected to open new avenues for developing more trustworthy and capable AI systems across various industries and academic disciplines.

Keywords: LLM, knowledge base, hallucinations, RAG.

Матеріал надійшов 23.09.2024