

ЭФФЕКТИВНОСТЬ ИСПОЛЬЗОВАНИЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В СОЦИАЛЬНЫХ ЭКСПЕРИМЕНТАХ

Земляник Р. В.

Кафедра информационных технологий автоматизированных систем,
Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь
E-mail: rzemlanik@gmail.com

Рассматривается использование искусственного интеллекта в социальных исследованиях, с акцентом на влияние объема выборки на точность результатов. Анализируется роль больших языковых моделей в создании экспериментальных условий и генерации данных.

ВВЕДЕНИЕ

В последние десятилетия искусственный интеллект (ИИ) стал неотъемлемым компонентом научных исследований, включая такие дисциплины, как социология, психология и другие социальные науки. Его внедрение открыло широкие перспективы для улучшения качества исследований, ускорения анализа данных и разработки более точных прогнозов. Тем не менее, наряду с позитивными аспектами, возникает ряд этических и методологических проблем. Одной из наиболее обсуждаемых тем является возможность использования ИИ для манипуляции данными и подделки результатов социальных экспериментов. Если ИИ используется некорректно или без должного контроля, это может привести к искажению результатов, что потенциально негативно сказывается на науке и общественных мнениях.

Цель данной статьи заключается в анализе того, как ИИ может воздействовать на результаты социальных исследований и опросов, как использование больших языковых моделей (LLM) может повлиять на точность моделирования, а также каким образом объем выборки и другие параметры могут оказывать влияние на результаты. Помимо этого, будут обсуждаться возможные последствия для научного сообщества и общества в целом, а также предложены пути смягчения потенциальных рисков, связанных с внедрением ИИ в социальные исследования.

I. ПРИМЕНЕНИЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В АНАЛИЗЕ ДАННЫХ

Machine learning (ML) является подмножеством искусственного интеллекта и представляет собой ключевой элемент анализа больших данных. С его помощью можно строить модели, которые позволяют не только анализировать имеющиеся данные, но и делать прогнозы. При этом важно отметить, что не все решения на основе ИИ зависят от машинного обучения. Существует множество примеров, когда ИИ применяется без использования МО, особенно в задачах, связанных с автоматизацией процессов или решениями, основанными на заранее заданных правилах.

Тем не менее, применение ИИ в анализе данных открывает широкие возможности для ученых. Например, в медицине ИИ может ускорить процесс диагностики, анализируя огромные объемы медицинских данных и выявляя паттерны, которые могут ускользнуть от внимания врачей. Это позволяет специалистам сосредоточиться на более сложных и индивидуальных аспектах лечения пациентов. Таким образом, ИИ не заменяет человека, а дополняет его, предоставляя инструменты для более глубокого анализа и лучшего понимания проблемы.

Large language models (LLM) являются классом ИИ, который обучен на огромных объемах текстовых данных и способен имитировать естественную человеческую речь, формулируя осмысленные ответы на широкий спектр вопросов. Такие модели обладают потенциалом для применения в социальных науках, где они могут выступать в роли инструментов для проведения опросов, анализа текстов и прогнозирования социальных явлений. LLM способны имитировать сложные процессы, такие как социальное взаимодействие, переговоры, принятие решений и эмоциональная поддержка, что делает их полезными в исследовательских целях.

II. СОЦИАЛЬНЫЕ НАУКИ И ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ

Искусственный интеллект оказывает серьезное влияние на социальные науки, в частности на проведение и интерпретацию социальных экспериментов. Один из главных вопросов заключается в том, каким образом ИИ может изменять или искажать результаты экспериментов. В ряде случаев, если ИИ используется для обработки данных без должного контроля, это может привести к манипуляции результатами. Например, если модели обучены на неполных или предвзятых данных, они могут выдавать неточные или искаженные прогнозы.

Подделка данных с помощью ИИ может иметь значительные последствия на уровне общества. Если люди начинают полагаться на ИИ в процессе принятия решений, особенно в условиях неопределенности или недостатка информации,

это может привести к распространению дезинформации. В таком случае результаты научных исследований могут быть искажены, что подрывает доверие не только к отдельным исследованиям, но и к научному методу в целом. Это особенно опасно в социально значимых областях, таких как политика, медицина и экономика.

С другой стороны, ИИ может быть полезным инструментом для социологов и других исследователей. Например, в социальных науках часто проводится множество экспериментов, связанных с влиянием текста, изображений или других стимулов на поведение участников. Генеративные модели ИИ могут использоваться для создания таких стимулов, что делает эксперименты более точными и достоверными. Более того, они помогают исследователям соблюдать конфиденциальность, заменяя реальные данные синтетическими, что особенно важно в контексте этических норм.

Тем не менее, применение ИИ в исследованиях сопряжено с рядом рисков. Во-первых, слишком сильная зависимость от ИИ может привести к иллюзии понимания, когда исследователи начинают считать, что их модели точно отражают реальность, даже если это не так. Во-вторых, существует опасность создания научной монокультуры, когда одни методы и подходы начинают доминировать над другими, что снижает инновационность науки и увеличивает вероятность ошибок.

III. ВЛИЯНИЕ ОБЪЕМА ВЫБОРКИ НА ТОЧНОСТЬ ПРОГНОЗОВ

Один из ключевых факторов, определяющих точность прогнозов и моделей, основанных на ИИ, – это объем выборки. В исследованиях, основанных на больших языковых моделях, объем данных, на которых обучена модель, напрямую влияет на ее способность точно предсказывать ответы на вопросы или моделировать поведение участников экспериментов.

Увеличение объема обучающих данных приводит к улучшению точности моделей. Например, при моделировании данных опросов использование большего количества вопросов для обучения модели позволяет значительно повысить ее способность корректно предсказывать ответы на другие вопросы. Это особенно важно в условиях, когда часть данных отсутствует или когда респонденты не отвечают на все вопросы. В таких случаях ИИ может использоваться для заполнения отсутствующих данных, для предсказания

ответов на те вопросы, которые не были заданы респондентам напрямую.

Одно из исследований показало, что даже при небольшом объеме данных (например, при наличии только 10 вопросов для обучения) модели могут достигать достаточно высокой точности ($AUC > 0,8$). Однако увеличение числа вопросов до 100 позволяет добиться еще более высокой точности ($AUC = 0,83$). Это означает, что количество данных, на которых обучена модель, напрямую связано с ее производительностью.

Применение таких моделей может быть полезным при разработке опросов и экспериментов, где необходимо учитывать возможность неявки респондентов на опрос или неполные ответы. Например, вместо того чтобы задавать каждому респонденту один и тот же набор вопросов, исследователи могут использовать ИИ для предсказания ответов на дополнительные вопросы, основываясь на уже полученных данных.

IV. ЗАКЛЮЧЕНИЕ

Интеграция ИИ в социальные исследования открывает новые горизонты, но также ставит перед учеными и обществом ряд вызовов. Несмотря на то, что ИИ может значительно улучшить точность прогнозов и моделирование данных, его использование сопряжено с рисками, такими как искажение данных, подмена реальных данных синтетическими и иллюзия понимания. Важно разработать механизмы, которые позволят минимизировать эти риски и обеспечить, чтобы ИИ использовался в исследованиях ответственно и с максимальной пользой для науки и общества.

1. Can large language models help predict results from a complex behavioural science study? [Electronic resource] / S. Lippert, A. Dreber, M. Johannesson, W. Tierney, W. Cyrus-Lai, E. L. Uhlmann, Emotion Expression Collaboration, T. Pfeiffer –Royal society open science, 2024. – Mode of access: <https://royalsocietypublishing.org/doi/10.1098/rsos.240682>. – Date of access: 11.10.2024.
2. AI Should Augment Human Intelligence, Not Replace It [Electronic resource] / D. De Cremer, G. Kasparov–Harvard Business Review, 2021. – Mode of access: <https://hbr.org/2021/03/ai-should-augment-human-intelligence-not-replace-it>. – Date of access: 13.10.2024.
3. Artificial intelligence and illusions of understanding in scientific research [Electronic resource] / L. Messeri, M. J. Crockett–Nature 627, 2024. – Mode of access: <https://www.nature.com/articles/s41586-024-07146-0> – Date of access: 13.10.2024.