



Нейросеть как инструмент изучения процессов познания


Елена Михайловна Бурнаева

Доцент кафедры математики и информационных технологий
Тихоокеанский государственный университет
Хабаровск, Россия
001681@pnu.edu.ru
 0000-0002-6633-0040

Александр Иванович Рудаков

Студент гр.ПСП-21
Тихоокеанский государственный университет
Хабаровск, Россия
2022102932@pnu.edu.ru
 0000-0007-8083-1603

Поступила в редакцию 19.04.2023
Принята 01.05.2023
Опубликована 25.06.2023

 10.25726/z0438-7779-0104-c

Аннотация

В настоящее время требуется такая классификация задач, которые должны решать системы искусственного интеллекта, при чем, задачи при использовании искусственных нейросетей (в виде получения субъективно и объективно новой информации). Нейросети уже столь сильно проникли в нашу повседневную жизнь, что представить наше существование без них уже сейчас – практически невыполнимая задача. Нейросети обучаются выполнять всё больший и больший круг задач – начиная от банальных функций общения, продолжая созданием уникальных и проработанных графических изображений, рассказов и заканчивая даже психологической помощью. Каждый день возможности применения таковой технологии расширяются и облегчают нашу современную жизнь. И оставался открытым лишь вопрос, когда же нейросети будут обучены понимать человека «без слов».

Ключевые слова

нейросеть, искусственный интеллект, семантический анализ, психологическая помощь, информационные технологии.

Введение

Можно смело сказать, что нейросеть овладела необходимыми навыками для того, чтобы начать считывать и мыслимые образы с поразительной точностью, воссоздавая таковые в различных (текст, графика и т.д.) форматах. Именно этот вопрос и рассматриваем, а точнее – ту проблемную область, относящуюся к настоящему, который может быть решен с помощью интегрирования подобной технологии и в ту или иную практическую деятельность.

Исследования учёных из Университета Тахаса, возглавленное студентом-аспирантом Джерри Тангом, было опубликовано 1 мая в журнале Nature Neuroscience. В работе было применено создание декодера человеческих мыслей и образов. Создание такового преследовало за собой два этапа: для начала нейросеть научили понимать, какие слова или фразы связаны сигналами в определённом блоке коры человека и характером этого сигнала. Для этого были применены автобиографические рассказы с известного источника Reddit. Однако, чтение таковых проходило во время прохождения функционального МРТ, что и позволяло зафиксировать эти сигналы. Потом испытуемые попадали в аппарат МРТ, где проходило дополнительное сканирование мозга. В этот раз им запускались рассказы,

которых они ранее не слышали и с которой ранее не встречалась нейросеть. Всего, в общей сумме, было записано 48 часов мозговых сигналов от трёх разных людей (по 16 часов на каждого). Модель GPT же, в свою очередь, смогла декодировать мысли не только из услышанных людьми рассказов, но также из категории беззвучных мультфильмов студии Pixar.

Материалы и методы исследования

Общие результаты эксперимента представлены далее:

- точность «чтения мыслей» 72-84%. Воспринимаемая речь (испытуемые слушали запись);
- точность «чтения мыслей» 41-74%. Мысленная речь (испытуемые про себя рассказывали историю длительностью не более минуты);
- точность «чтения мыслей» 21-45%. Декодинг субъективной интерпретации ранее запечатлённого зрительными анализаторами (ролики Pixar без звука).

Отдельно поражает сознание тот факт, что декодирующей технологией была не только правильно уловлена и интерпретирована суть мыслей, но и воссоздана точная структура словосочетаний: «кричать и плакать», «оставь меня наедине с собой». Однако, без ошибок в данном процессе тоже обходилось, т.к. система работает именно на уровне семантики, а точнее – идей и смысла – получаемый продукт представляет собой не конкретные слова, которые некто услышал или подумал – это основная идея и мысль, единственное – быть может, выраженная иными словами и их связями. Можно сразу же отметить, что распознавание определённых фраз вызывает некоторые сложности. Дело обстоит в том, что большую проблему составляет распознавание естественных, «человеческих» языков. Представлено не столь мало вариантов, которые различными способами решают такую проблему. В статье (Корешкова, 2021) рассмотрены различные способы, но мы рассмотрим лишь некоторые. Одним из наиболее популярных в сообществе вариантов обработки естественного языка является Word2Vec. Данный вариант был представлен компанией Google в 2013 году. Суть заключается в том, что слова преобразуются в вектор, который, в свою очередь, отражает семантические свойства слова. В таком случае слова, близкие по значению, будут находиться рядом в N-мерном пространстве. Это позволяет продемонстрировать семантическую близость различных слов с минимальными потерями. Помимо того, в этом случае слова, даже при рассмотрении в разных языках при простых трансформациях дают подобные друг другу вектора, и следовательно, позволяет без особых проблем производить переводы слов.

RNN (Рекуррентные нейронные сети). Базовый вариант RNN подвержен проблеме исчезающих и взрывающихся градиентов. В связи с этим, в качестве ответвления были разработаны две модернизированные структуры на базе RNN: долгосрочная кратковременная память (LSTM), а также управляемые рекуррентные блоки (GRU). Эта система имеет очень высокую чувствительность к порядку слов, особенно – в близком значении.

CNN. Данная нейросеть хорошо зарекомендовала себя в компьютерном зрении, где распознавались образы лучше, чем человеком. На данный момент, как инструмент распознавания естественного языка – данная модель используется в распознавании речи и её генерации. Однако, также отмечается проблема словообразования. В связи с разницей в таковом процессе в различных языках – сложно создать общую систему наполнения и обучения. Если в английском языке даже при простых правилах образования форм глагола, существуют некоторые исключения из правил (неправильные формы прошедшего времени глаголов). В русском же подобных форм слов могут быть десятки (или даже сотни).

Результаты и обсуждение

Возможно, основываясь именно на основе такового примера, учёная группа Университета Осаки, Япония, смогла выйти на совершенно новый уровень (Стрекопытов, 2023). Синтезировав распознавание отдельных «сообщений» с помощью MPT с нейросетью Stable Diffusion. Эта нейросеть с открытым кодом по структуре не выделяется в перечне прочих LLM (Large language models), таких как DALL-E2 от компании OpenAI или Midjourney. В основе данного метода стоит диффузия, методика машинного

обучения, где какой-либо визуальный образ создаётся путём последовательного приближения, т.е. – всегда новая итерация базируется на дополнительной подсказке (текстовой). Однако, нельзя не отметить, что японцы добавили к стандартной схеме Stable Diffusion ещё один этап обучения, который сопоставлял данные активности мозга четырёх человек с демонстрируемыми им изображениями. В случае Американского опыта было рассмотрено распознавание текста, в случае же с Японским опытом – были получены результаты (и очень высокого уровня) распознавания образов, что, в своём роде, выводит данный вектор исследований на совершенно новый уровень. Как было указано, результаты в большинстве своём соответствовали гипотезе о двух потоках зрительной информации, которая была сформулирована ещё в 1983 году американским нейропсихологом Мортимером Мишкиным. Суть его заключалась в том, что в коре головного мозга есть два анатомически и функционально разных канала для обработки пространственной и предметной информации: «Где?» и «Что?». В результате опытов было выявлено, что затылочный (дорсальный) канал «Где?», в свою очередь, несёт функцию восприятия пространства, в то время как височный (вентральный), также именуемый как канал «Что?» – наиболее прочно взаимосвязанный с памятью, отвечает за узнавание. В данном случае, была объединена зрительная и семантическая информация. Алгоритм был составлен так, чтобы сравнивать наблюдаемые паттерны нейронной активности, которые возникали при просмотре различных фотографий, с имеющимися образцами в наборе обучающих данных. В таком случае, с помощью таковых сигналов становилось возможным построение общего объёма и определённой перспективы. Следом за этим вступали в работу подсказки с декодера семантических сигналов, после чего первичная картинка постепенно приобретала всё более правдоподобные очертания узнаваемых объектов. В результате эксперимента, точность смыслового содержания по сравнению с оригиналом была до 80%. В большинстве случаев также получалось воссоздать даже цветовую гамму исходного изображения. Также отмечается, что семантический декодер был настроен на распознавание конкретных заученных заранее сигналов, что в очередной раз сближает данный эксперимент с опытом американских коллег. Также не следует забывать и про дополнительные подсказки, которые давались испытуемыми, которые мысленно проговаривали картинку, тем самым облегчая работу декодеру. Также интересен опыт с исследователями из Нидерландов (Стрекопытов, 2023). В данном случае, картинки и текст были заменены на лица людей, а, следовательно – исследовался совершенно иной блок мозга. Основная суть сохранилась (MPT + нейросеть, создающая образы и картинки с опорой на текст). В данном случае испытуемые изучали портреты ранее незнакомых людей, в то время как система «попиксельно» отслеживала реакцию нейронов, проводила её в компьютерный код и пересобираала портреты.

Достижения в этом направлении появляются с каждым днём, однако, временно данные технологии находятся ещё довольно далеко от того уровня, при котором их можно было бы применять на практике в различных учреждениях. Однако, Тирза Дадо из Университета Радбауд, являющийся когнитивным психологом, в своём интервью уже упоминал о разработке камеры для имплантации в мозг людей, ослепших в результате болезни или несчастного случая (сохранивших опыт зрительной активности) (Стрекопытов, 2023). Ещё большие проблемы в данном случае вызывает факт уникальности людей. Пускай центры протекания определённых мозговых реакций локализованы в определённых блоках – это не позволяет нам судить об унификации самих алгоритмов мышления. В данном случае, особый интерес, представляет эффект Еськова-Зинченко, выраженный тем, что любая выборка параметров организма человека по своему роду уникальна (Еськов, 2021). Допустимо ли рассматривать тип мышления различных людей в качестве параметра, не допуская унификации – вопрос сложный и неоднозначный. Однако, ранее упомянутые учёные из Японии называют подобный алгоритм универсальным и не столь восприимчивым к различным испытуемым. В таком случае, можно предположить, что ответ кроется в проведении эксперимента на большем количестве испытуемых из различных слоёв населения и различного культурного слоя. Также довольно большую проблему в случае практического применения можно выявить в следующем, что проблема, опять же, кроется в оценке подобного метода исследований со стороны общественности (Сабиров, 2022). В современную эпоху гуманизма, антропоцентризма и человеческой ценности применения подобных практических методик вызовет большой спор, т.к. формально человек лишается права свободно мыслить и хранить свои

секреты. С другой же стороны, также проблемой для внедрения в практическое применение является сложность, а то и вовсе невозможность проверки результатов на достоверность. Вероятнее всего, результаты такого метода будут иметь аналогично рекомендательный характер, подобный тому, который носят результаты исследований на полиграфическом оборудовании.

Заключение


На самом деле, вновь анализируя ранее упомянутые источники, можно выделить уже даже сейчас несколько возможностей, которые даёт данный вектор исследований и какие вопросы он ставит перед научным сообществом. В случае же внедрения таковой технологии в общую структуру, упомянутую как «сильный ИИ» (Дурнев, 2018), особенно с упомянутой там темой применения ИИ в военном деле и национальной безопасности. В данном случае, мы можем увидеть замечательный инструмент противодействия терроризму и поиску лиц, причастных к нему. В качестве возможностей данного метода нельзя не упомянуть о возможности общения с пациентами, утратившими возможности пользования речью (механическое влияние, состояние глубокой комы и т.д.). Помощь ослепшим людям вновь воспринимать образы. Исследования, которые приблизят нас к пониманию процессов, протекающих у животных во взаимодействии с миром. Также нельзя не упомянуть о возможности применения таковой технологии в криминалистике, однако, с учётом множественных улучшений и доработок.

Список литературы


1. Дурнев Р.А., Крюков К.Ю., Жданенко И.В. Искусственный интеллект: иллюзии, проблемы, перспективы // Инноватика и перспектива. 2018. № 4 (25). С. 157-171.
2. Еськов В.М., Филатов М.А., Газя Г.В. Возможности создания искусственного интеллекта на базе искусственных нейросетей // Успехи кибернетики / Russian Journal of Cybernetics. 2021. № 3 (7) 2. С. 44-52.
3. Корешкова Т. Семантический анализ для автоматической обработки естественного языка. https://rdc.grfc.ru/2021/09/semantic_analysis/
4. Сабиров К.А., Петров Е.А., Хасанов Ф.Ф. Нейросеть и предубеждение: фундаментальные проблемы // Инструменты и механизмы современного инновационного развития: сборник статей Международной научно-практической конференции: в 5 частях, Пермь, 05 декабря 2017 года. Том Часть 4. Пермь: Общество с ограниченной ответственностью "Аэтерна", 2017. С. 126-129h
5. Стрекопытов В. «С высокой точностью». Создана первая машина для чтения мыслей. <https://ria.ru/20230315/diffuziya-1857798685.html>

Neural network as a tool for studying cognitive processes


Elena M. Burnaeva

Associate Professor of the Department of Mathematics and Information Technology
Pacific National University
Khabarovsk, Russia
001681@pnu.edu.ru
 0000-0002-6633-0040

Alexander I. Rudakov

Student gr.PSP-21
Pacific National University
Khabarovsk, Russia
2022102932@pnu.edu.ru
 0000-0007-8083-1603

Received 27.03.2023
Accepted 01.04.2023
Published 15.05.2023

 10.25726/z0438-7779-0104-c

Abstract

Currently, such a classification of tasks is required that artificial intelligence systems must solve, moreover, tasks when using artificial neural networks (in the form of obtaining subjectively and objectively new information). Neural networks have already penetrated so much into our daily lives that it is almost an impossible task to imagine our existence without them even now. Neural networks are trained to perform a larger and larger range of tasks - starting from the banal functions of communication, continuing with the creation of unique and elaborate graphic images, stories, and even ending with psychological help. Every day, the possibilities of using such technology are expanding and facilitating our modern life. And only the question remained open, when will neural networks be trained to understand a person "without words".

Keywords

neural network, artificial intelligence, semantic analysis, psycho-logical assistance, information technology.

References

1. Durnev R.A., Krjukov K.Ju., Zhdanenko I.V. Iskusstvennyj intellekt: illjuzii, problemy, perspektivy // *Innovatika i perspektiva*. 2018. № 4 (25). S. 157-171.
2. Es'kov V.M., Filatov M.A., Gazja G.V. Vozmozhnosti sozdaniya iskusstvennogo intellekta na baze iskusstvennyh nejrosetej // *Uspehi kibernetiki / Russian Journal of Cybernetics*. 2021. № 3 (7) 2. S. 44-52.
3. Koreshkova T. Semanticheskij analiz dlja avtomaticheskoi obrabotki estestvennogo jazyka. https://rdc.grfc.ru/2021/09/semantic_analysis/
4. Sabirov K.A., Petrov E.A., Hasanov F.F. Nejroset' i predubezhdenie: fundamental'nye problemy // *Instrumenty i mehanizmy sovremennogo innovacionnogo razvitija: sbornik statej Mezhdunarodnoj nauchno-prakticheskoi konferencii: v 5 chastjah, Perm', 05 dekabrja 2017 goda. Tom Chast' 4. Perm': Obshhestvo s ogranichennoj otvetstvennost'ju "Ajeterna", 2017. S. 126-129h*
5. Strekopytov V. «S vysokoi tochnost'ju». Sozdana pervaja mashina dlja chtenija myslej. <https://ria.ru/20230315/diffuziya-1857798685.html>