

БОРИС ОРЕХОВ

ЭКСТРАПОЛЯЦИЯ ТЕКСТА
С ПОМОЩЬЮ ИСКУССТВЕННЫХ
НЕЙРОННЫХ СЕТЕЙ.

АРХИВ ГАХН

1. НЕЙРОСЕТИ КАК ИНСТРУМЕНТ

ИСКУССТВЕННЫЕ нейронные сети плотно вошли в современную инженерную практику, подарив человечеству очередную надежду на создание сильного искусственного интеллекта. Такое воодушевление уже охватывало техническое сообщество несколько раз за предыдущие полстолетия, но такие периоды сменялись временем разочарования, называемыми «зимой искусственного интеллекта». Однако успехи последних лет действительно впечатляющие, пусть и несколько преувеличены в кричащих заголовках СМИ: нейросети неплохо умеют помогать компьютерам распознавать людей на фотографиях, переводить звучащую речь в текст, улучшать качество картинки. Скептики, разумеется, правы: нейросети не более (а иногда и менее) «интеллектуальны», чем простейшие организмы, реагирующие на внешние раздражители, но несомненно и то, что реагируют

они на эти раздражители правильно и практически без вмешательства человека.

Искусственными нейронными сетями занимаются преимущественно специалисты технического профиля, разрабатывающие новые архитектуры и подходы, позволяющие улучшить качество их работы. Гуманитарии только подступаются к осмыслению того, как этот новый инструмент может изменить представление об исследовательских практиках, о самом предмете исследования и взаимодействии с ним. Мне кажется, что текстовое наследие ГАХН — это хороший повод поговорить на эту тему. Хотя бы потому что привычный к интеллектуальным новациям коллектив, нам это наследие оставивший, не побоялся бы такой темы¹. Вступление в неизвестное (или по меньшей мере слабо изученное) всегда требует научной смелости, и сегодня именно ее мы можем позаимствовать у академиков ГАХН.

Нейросети (таким сокращенным словом я буду пользоваться, чтобы отличать их от естественных нейронных связей у живых организмов, обладающих нервной системой) — это продолжение и развитие концепции «машин-

1. «Речь шла о поисках нового типа гуманитарного знания, способного установить взаимосвязь философии, исторических наук о культуре, естественных наук о человеке и социальных наук с целью „всестороннего изучения искусства“». Плотников С., Подземская Н., Чубаров И. М. Предисловие, в: *Логос*. 2010. № 2. С. 6.

ного обучения», в центре которой стоит проблема принятия решений. Если мы собираемся доверить какую-то деятельность компьютеру, то он должен каким-то образом самостоятельно давать ответы на многочисленные вопросы, которые неизбежно будут вставать в процессе этой деятельности. Например, в тот момент, когда на наш электронный адрес приходит очередное письмо, почтовая программа должна решить, отнести это отправление к желательным и поместить в папку «Входящие», или пометить как спам. Как машина должна это решить? Приблизительно до 1980-х годов считалось, что наиболее эффективный способ — это привлечь к работе эксперта, он вручную пропишет для компьютера специальную систему правил и та будет на них ориентироваться. Например, если в письме есть кричащие цвета, большие буквы нелепых шрифтов, картинок больше, чем текста, то, вероятнее всего, это назойливая реклама, а не сообщение, которое мы были бы рады прочесть. Программы строились как так называемые экспертные системы, то есть слепо следующие разработанному человеком алгоритму исполнители.

Но чем дальше, тем больше у такого подхода обнаруживалось недостатков. Привлекать эксперта к написанию правил оказалось дорого, ситуация все время менялась (например, спамеры придумывали новые способы обхода фильтров), и эксперты не успевали переписыва-

вать свои алгоритмы, наконец, сами предметы, с которыми имели дело эксперты, оказывались порой настолько сложны, что просто не поддавались описанию в простых терминах вроде «если X, то Y» и содержали неочевидные даже самим специалистам закономерности. Одновременно с этим стремительно развивалась компьютерная техника, дешевле становились вычисления и хранение информации. Все это создало очевидные предпосылки для смены курса: надо было сделать так, чтобы машина сама выявляла закономерности в информации (которой теперь накопилось много), научалась принятию правильных решений так, как учится этому человек в первые годы жизни — просто наблюдая за тем, что происходит вокруг и делая выводы.

Оказалось, что это не только возможно, но и эффективно — экономически и организационно, что и решило дело, потому что принимающие решения роботы были нужны прежде всего коммерческим компаниям. Затратную работу исследования больших массивов данных можно было передоверить машине, которая стала находить решения сама, ориентируясь на сложные связи между ансамблями параметров. Детальное обозрение этих параметров давно выходит за рамки человеческих возможностей, современные модели обучаются на сотнях гигабайт информации, что должно равняться сотням тысяч книг в привычном нам эквиваленте.

2. НЕЙРОСЕТИ И ТЕКСТЫ: ОБУЧЕНИЕ И ПОРОЖДЕНИЕ

Нейросети — это особенное подмножество машинного обучения, они требуют больше данных, учатся медленнее, но выдают гораздо более качественный результат. Особенно стал известен случай с распознаванием образов на картинках специально подготовленного для такой задачи набора ImageNet. В какой-то момент прогресс нейронных сетей в принятии решений о том, кто именно изображен на фотографии, достиг такой отметки, при которой стало ясно, что люди, пытаясь решить ту же задачу, ошибаются чаще роботов².

Но нейросети, как видно из этого экскурса, это прикладной инструмент. Может ли он помочь исследователю, который заинтересован не в прикладных задачах, а в получении нового знания? Что особенно важно — может ли он помочь ученому-гуманитарию, имеющему дело со специальным недискретным и плохо формализуемым исследовательским объектом?

Нейросети существуют разных видов и используются они, разумеется, для разных задач, а число их видов постоянно увеличивается. Но сейчас нас будут интересовать специфиче-

2. He K. et al. Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification in: *Proceedings of the IEEE international conference on computer vision*. 2015. P. 1026–1034.

ски такие, которые умеют строить языковые модели. Для компьютера речь — это просто последовательность единиц языка, поэтому и вся языковая модель представляет собой набор вероятностей — с какой-то за текущим словом идет одно слово, а с какой-то другой вероятностью другое слово. Такие сети научаются на уже существующих текстах и в качестве побочного, но важного для нас следствия, могут, ориентируясь на выученные вероятности, сами порождать последовательности слов, букв — чего угодно, из чего последовательности (например, тексты, но это могут быть и ноты, и программный код, и ДНК) состоят. Удивительной особенностью таких порожденных компьютером новых текстов является то, что для читателя очевидна их зависимость от текста-источника, то есть от того корпуса, на котором сеть обучалась. Такие искусственные тексты в разных отношениях *похожи* на свой исходник³, его узнают по порожденным произведениям, это доказывается экспериментально⁴. Такого эффекта не производили ранние попытки ге-

3. Orekhov B., Fischer F. Neural reading. Insights from the analysis of poetry generated by artificial neural networks in: *Orbis Litterarum*, 2020, vol. 75, no. 5, p. 230–246.
4. См., например, *Peter Potash, Alexey Romanov, and Anna Rumshisky*. Ghostwriter: Using an LSTM for automatic rap lyric generation. In Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, pages 1919–1924, 2015.

нерировать тексты на компьютере⁵: в ходе тех экспериментов десятилетия назад, без нейросетей, произведение составлялось компьютером из случайно выбранных из словаря слов. И хотя эти словари обычно создавались не случайным образом, а тоже выбирались из какого-то текста-источника, конечный результат этот источник никак не напоминал. Для исследователя такая разница в эффектах должна значить, что текст — это все же в первую очередь не набор слов, а их контекстуальные связи. Не парадигматика, а синтагматика (отсюда следует заметная ограниченность, а к настоящему моменту и исчерпанность эвристического потенциала частотных словарей). И это уже значимый научный вывод.

Но не обязательно на этом выводе останавливаться. Можно работать с этими порожденными текстами и дополнять ими наши исследовательские коллекции. Разумеется, речь не идет о том, чтобы наводнить состав оригиналов электронными подделками. Сгенерированные тексты иерархически всегда будут ниже, зани-

5. «...в машину был вложен словарь „Камня“ О. Мандельштама и правила сочетания этих слов в стихи; машина выдала 21 отрывок разных стихотворных размеров (от 4 до 12 строк). Казалось бы, заданный словарь должен был обеспечить хотя бы отдаленное ощущение близости получающихся текстов к поэтике Мандельштама. Этого не произошло». Гаспаров М. Л. *Метр и смысл*. Москва: ФОРТУНА СЛ, 2012. С. 6.

мая подчиненное положение. Но определенный смысл в них все же есть.

3. ЧЕМУ ВЫУЧИЛАСЬ НЕЙРОСЕТЬ НА АРХИВЕ ГАХН?

Дело в том, что изучая систему понятий, мы всегда вынуждены изучать язык как набор контекстов, в которых эти понятия представлены. И текстовые материалы ГАХН не исключение, несмотря на наличие концептуальных словарей, составленных членами академии. Принципиальная недостаточность словарей давно осознана лингвистами: слово не может быть описано в словарном формате во всей полноте нюансов. Зато эти нюансы восстанавливаются из контекстов. Помня об этом, не трудно прийти к заключению, что исследователь может оказаться в ловушке существующего набора текстов, рамки которого не позволят решить его задачу. Например, ограниченность корпуса не позволит корректно истолковать некоторый термин по наличным контекстам, проследить идейную траекторию субъекта речи, реконструировать смысловую наполненность и сочетаемость. Так, исследователи Бахтина страдают от того, что философ не давал в своих работах ясных определений используемым терминам, а существующие словоупотребления оказываются текучими и слабо поддаются формализации (что, по всей видимости, пред-

ставляет проблему только для исследователей, но не для самого Бахтина). Нейросети могут добавить материала, расширить контекст сказанного, следуя при этом траектории обучающих исходных текстов. Нейросети не способны порождать новые смыслы, вся их продукция — это усреднение и суммирование обучающего корпуса. По сути, мы имеем дело с экстраполяцией исходных данных⁶, которая может быть полезна для изучения языка научного авангарда 1920-х годов (вспомним, что именно проблеме языка ГАХНовцы уделяли особое внимание⁷).

У меня уже есть такой опыт для корпуса текстов М. М. Бахтина: порожденные искусственными нейронными сетями «псевдобахтинские» отрывки опубликованы в последнем номере журнала «Диалог. Карнавал. Хронотоп»⁸. Такого же рода эксперимент позволяют поставить и оцифрованные ныне архивы ГАХН.

6. Андрусенко В. А. *Экстраполяция в научном познании: специфика и проблема правильности*. Иркутск: Издательство Иркутского университета, 1986; Попова Н. Л. *Экстраполяция как средство научного познания и интегративный фактор в науке*. Киев : Наукова думка, 1985.
7. «...стала выработка нового языка, способного выразить основные измерения понимания искусства, — формирование языка искусства или, вернее, интерпретация искусства как специфического языка, «грамматика» которого отличает его от языка повседневной коммуникации и языка науки». Плотников и др. *Предисловие*. С. 7.
8. Орехов Б. В. Нейронная сеть переписывает М. М. Бахтина, в: *Диалог. Карнавал. Хронотоп*. 2018. № 1. С. 181–186.

Материалом для обучения нейросети стали оцифрованные тексты архива ГАХН⁹ общим объемом 4,7 мегабайт. На них была обучена символьная (то есть опирающаяся на символ, а не на слово как на единицу данных) четырехслойная рекуррентная сеть с долгой краткосрочной памятью (LSTM) размером 512 нейронов. К концу обучения значение потерь (loss) составляло 1,242.

Посмотрим на получившиеся у нейросети контексты:

С этой стороны сознание не дает нам воззрений на произведение, которые вынесены понятием о кино, т. е. искусство — параллельные к своей особенности пространства, которые подразумеваются поставленными как в произведении и на странице. Его начала для соответствующих отношений проблемы и образующего ответа, и других соответствующих, это безусловно сознательная деятельность как индивидуальной театральной формы. Последние выражения в том, что не оставляет общее в себе. Конкретное становится как нечто в области вопроса о том, что пример построения (проблема постановки и в искусствоведениях) вещей и проблем, и оснований конкретнее метода выражения субъективного начала последнего завершен-

9. «Искусство как язык — языки искусства. Государственная академия художественных наук и эстетическая теория 1920-х годов». Москва: Новое литературное обозрение, 2017 и др.

ного содержания. Такой эстетический отдел, где предметом служит разная форма, остается под восприятием. Другие образы составляются в интересе вопроса и подходят с анализа общественного. Его антракт позволяет увидеть дело не в восприятии научного явления, а лишь как другого на своем бытии, и проанализировать образ с развитием эстетического момента.

Поэтика отношения образа к действительности и принципиального понятия совершенно искусство. В конце Академии, не совпадаясь взглядами должного искусства, противоположность выражения почти предполагает строить возможности развития и образования. Последний момент значимости в нем в связи с противопоставлением формы того, что не своей литературы и которые вещи не совпадает создание произведения с помощью работы в слове.

Чтение такого текста (если относиться к нему с исследовательской серьезностью) может представлять собой изнурительную работу. Дело в том, что наше читательское сознание привыкло видеть в предложении не только грамматически оформленное высказывание, но и семантически завершенную структуру. Однако порожденные нейросетью тексты такого свойства лишены. В результате читатель постоянно держит себя в напряжении, пытаясь найти смысловую вершину фразы, что и приводит к усталости от чтения, или даже чувству растерянности и разочарования.

В то же время в представленном тексте внимательный взгляд обнаруживает многие ключевые для эстетических дискуссий 1920-х годов термины и аспекты. Во-первых, искусство мыслится как синтетический объект, вклад в который вносят и кино, и литература («и на странице», «работы в слове»), и театр. Во-вторых, здесь проявляет себя проблема сознательности/бессознательности творчества («сознание не дает нам», «сознательная деятельность»), являющаяся частным проявлением осмысления роли бессознательного в структуре человеческой личности, разумеется, весьма актуальная для 1920-х. В-третьих, затрагивается вопрос социальном преломлении эстетического («подходят с анализа общественного»). Обращает на себя внимание и многократное появление даже в таком коротком отрывке слова «форма», безусловно, одного из ключевых понятий для дискуссий внутри ГАХН, а конструкция «сознательная деятельность <...> формы» как раз преломляет идею Г. Г. Шпета об интеллектуальном содержании внутренней формы¹⁰.

Любопытно и сконструированное нейросетью слово «призначимость». Механизм по-

10. Грюбель Р. «Красноречивей слов иных / Немые разговоры» Понятие формы в сборнике ГАХН «Художественная форма» (1927) в контексте концепций Густава Шпета, русских формалистов и Михаила Бахтина, в: *Логос*. 2010. № 2. С. 21.

рождения текста у той нейросети, продукция которой представлена в этом материале, таков, что не предполагает наличия заранее заготовленного словаря. В отличие от тех экспериментов по текстовой генерации компьютерами, которые проводились с 1950-х до 2000-х годов, нейросеть ориентируется только на распределения букв, на буквенные последовательности, и не оперирует понятиями слова, словосочетания или предложения. Слова, фразы, абзацы формируются из выученных сетью вероятностей следования символов в тексте, а корни, суффиксы и другие морфемы являются своего рода функцией для этих вероятностей. Это означает, что нейросеть «не знает», какие слова существуют в языке, а какие нет, а следовательно, может породить свои собственные лексические единицы, хотя они всегда будут похожи на настоящие. В данном случае мы видим гибрид «признака» и «значимости», что является по-своему концептуально глубоким сближением на новом уровне двух понятий, этимологически родственных, но к настоящему моменту далеко разошедшихся по содержанию. К такого же рода любопытным неологизмам, не вошедших в этот отрывок, можно отнести: «дальнейство» («дальнейшее действие?»), «инуемость» (синоним литературного «инаковость»), «рукоственность», «буквопресссионизм» («буква + экспрессионизм») и др.

Еще пример:

Общее художественное созерцание отдельных жизней принято различать теми же сторонами в конечном смысле и интерпретации методологическим фактором. В другом представлении особенности философии искусств в противоположность этическому составу в постановке науки, в противополжении самосознания образованию театра (потому что в нем представление о значении и языковой проблеме докладчик не открывает само понятие «положение». В создании продукта представление о современной разногурному произведению – представления или работы созерцания между действительностями.

Разумеется, приведенные примеры не дают нам новых стенограмм заседаний ГАХН, законченных сентенций и идеологем. Нам нужно самим в ходе медленного чтения извлекать из этих отрывков и собирать значимые элементы интеллектуальных конструкций. Но примечательно, что в такие отрывки все же попадают наиболее значимые фрагменты теории.

Новые поколения нейросетей на основе технологии Transformer способны порождать более убедительные текстовые примеры, но обучаться они должны на грандиозных массивах текста, далеко выходящих за пределы того, что способны предоставить нам архивы ГАХН. Следовательно — такая текстовая продукция всегда бу-

дет отталкиваться не столько от наследия ГАХН, сколько от распыленного по интернету коллективного бессознательного. Такой пример мы можем получить, если инициировать генерацию текста в нейросети ruGPT¹¹, точкой отсчета будет типичная для обсуждений в академии реплика:

Искусство, собственно, обучаемо по пространству — как эстетикой, так и философией, — это и есть то самое искусство, которое есть, по определению Ницше, «сверхчеловек». «Сверхчеловек — искусство» (в этом, собственно говоря, вся суть). Это и есть искусство «полюмя», «полное искусство» — «сверхчеловек». Это искусство, которое не имеет ни начала, ни конца.

11. <https://github.com/sberbank-ai/ru-gpts>