

УДК 81'37

Виталий Викторович Тур, к. филол. н.

Белорусский государственный экономический университет, Минск, Беларусь
эл. почта: vitaly.tur@gmail.com

Vitalij Viktorovich Tur, Cand. of Sc. (Philology)

Belarussian State Economic University, Minsk, Belarus
e-mail: vitaly.tur@gmail.com

ЧЕЛОВЕК VS. GPT-3: ОСОБЕННОСТИ КОНСТРУИРОВАНИЯ ЛЕКСИЧЕСКОГО ЗНАЧЕНИЯ

Представляемая в докладе работа является продолжением многочисленных исследований, посвященных раскрытию особенностей интерпретации языковых значений генеративной языковой моделью GPT-3. Предлагается эксперимент, оценивающий возможности GPT-3 на основе контекста интерпретировать значения вымышленных слов. Сравниваются особенности конструирования новых значений машиной и человеком. Указывается на невозможность модели GPT-3 разграничивать семантические признаки незнакомых слов по степени их значимости в контексте.

Ключевые слова: генеративная языковая модель GPT-3, структура лексического значения, релевантные семантические признаки, концептуальное прогнозирование

HUMAN VS. GPT-3: PECULIARITIES OF LEXICAL MEANING CONSTRUCTION

The report describes an experiment that evaluates the ability of the generative language model GPT-3 to interpret the meanings of made-up words based on context. Peculiarities of the construction of new meanings by the machine and humans are compared. The GPT-3 model's

impossibility to distinguish between the semantic features of unfamiliar words according to the degree of their significance in the context is pointed out.

Key words: generative language model GPT-3, lexical meaning structure, relevant semantic features, conceptual prediction

Сегодня генеративная языковая модель GPT-3 по праву набирает все большую мировую популярность, привлекая повышенное внимание благодаря своим неординарным языковым возможностям в плане коммуникации с человеком, а также порождения текстов, которые без специальных инструментов уже практически невозможно отличить от текстов, порожденных человеком. Благодаря выдающимся успехам в области развития искусственного интеллекта беспрецедентно высокое внимание со стороны всего общества обращено сегодня на проблемы языковых значений. Неспециалисты в области когнитивной семантики задаются вопросом, действительно ли машина «понимает» значения слов, если может порождать бесконечное количество новых осмысленных текстов.

Не все однозначно и для специалистов. С одной стороны, очевидно, что речи о понимании в «человеческом» смысле быть не может, поскольку данная модель не оперирует на концептуальном уровне. Разработчики подчеркивают, что модель лишь выполняет задачу *имитации* человеческой речи, основываясь на статистическом обучении лексико-семантической сочетаемости на базе корпусов из миллионов существующих текстов, а также на словарях, к которым она имеет доступ в процессе обработки сообщений.

С другой стороны, растущее количество лингвистических исследований модели GPT-3 свидетельствует о том, что машина все больше удивляет лингвистов возможностями интерпретации все новых, нетривиальных оттенков значений, которые, очевидно, открываются благодаря заложенным алгоритмам по обработке и комбинированию словарных и контекстуальных данных. Представляется, что именно возможности по обработке контекстуальных значений, «вытягиванию» смыслов из предложенного контекста, представляет особый интерес в плане семантического исследования языковых возможностей рассматриваемой модели. Если интерпретация известных значений слов действительно является лишь имитацией работы мозга на концептуальном уровне (ведь словари и выделенные в них признаки, которые использует машина, представляют собой результат работы человека), то гораздо больший интерес представляли бы возможности модели в плане работы с неизвестными ей значениями.

В проведенном эксперименте машине предлагались созданные нами короткие тексты с вымышленным словом; при этом давалось задание назвать все признаки в значении этого слова, которые могут быть выведенными на основе контекста. Сравните, например, некоторые отрывки из текстов и признаки в значении вымышленного слова *флигл*, которые выводились моделью на основе контекста: *Одна из книг лежала плашмя поверх других, а на ней лежал флигл => это объект, который можно положить на книгу; Ого! Это ведь флигл? – улыбнулся Ник, взяв предмет двумя пальцами и протерев его от пыли => он достаточно мал, чтобы взять его двумя пальцами; Неужели*

у вас все еще пользуются им перед сном? – спросил он, разглядывая комнату через прозрачную поверхность предмета => 1. Это устаревший предмет, утративший ценность, 2. Имеет прозрачную поверхность, сквозь которую можно смотреть, 3. Как-то связан со сном и т. д.

Далее в ходе эксперимента то же задание было предложено людям. Результаты позволяют говорить о том, что возможности вычленения признаков машиной и человеком в целом сопоставимы. Так, в некоторых случаях люди демонстрировали лучшую способность к выявлению очевидных из контекста признаков, связанных с эмоциями действующих лиц; в других случаях, напротив, модель GPT-3 смогла назвать признаки, которые не замечали люди.

Наиболее интересным, однако, на наш взгляд, являлся следующий этап эксперимента, на котором опрашиваемым предлагалось расположить названные признаки по степени их релевантности и значимости в контексте. Эксперимент показал, что данное задание не вызвало трудностей у людей, однако представило собой непреодолимую сложность для машины. Модель GPT-3 уклонялась от ответа, сообщая, что располагает недостаточными данными об обозначаемом объекте, чтобы судить о его наиболее релевантных признаках. Данные результаты, на наш взгляд, демонстрируют принципиальную разницу в способе интерпретации языковых значений людьми и моделью GPT-3, а также позволяет лучше понять, каким образом человек хранит и использует значения слов.

Мы полагаем, что в отличие от нейросети GPT-3, человек начинает конструирование значения «с конца», т. е. на основе имеющихся сведений (какими незначительными бы они не были) сразу выстраивает теорию о том, что данное слово может обозначать, а в процессе поступления новых данных уточняет, модифицирует, или полностью заменяет имеющиеся теории на новые. В свете данной гипотезы становится объяснимой та легкость, с которой человек приписывает ту или иную степень релевантности любому признаку – теория о значении слова автоматически образует структуру значения, где каждый признак получает ту или иную степень релевантности в контексте выстроенной теории.

Имитация работы по конструированию теорий о семантике незнакомых слов и конструированию новых значений требует от модели GPT-3 решения проблемы оценки релевантности семантических признаков.