

Трудности перевода терминов-синонимов

Сабаева Юлия Сергеевна

Студент (бакалавр)

Российский университет транспорта (МИИТ), Москва, Россия

E-mail: sabaevayulia02@gmail.com

В последнее время искусственный интеллект стал важной частью исследований в области лингвистики. Благодаря своей способности обрабатывать огромное количество данных в короткий срок нейронные сети стали незаменимым помощником при анализе, распознавании текстов и в разработке систем машинного перевода.

Однако, ввиду ряда проблем, информация, поступающая от искусственного интеллекта, бывает некорректной. В нашем исследовании мы рассмотрели проблему перевода терминов-синонимов при помощи ИИ. Материал исследования был собран методом сплошной выборки из русских и английских статей в средствах массовой информации. Далее был осуществлен перевод при помощи ресурсов на основе нейросетей и ИИ (ChatGPT, DeepL, GigaChat и др.). Затем был произведен анализ ошибок и их классификация, выявлены основные недоработки на данном этапе.

Во-первых, генеративные нейросети имеют слабое место в области специальных текстов, ввиду чего не справляются с терминологией. Согласно словарю О.С. Ахмановой, термин создан для обозначения точного выражения специализированных понятий [1]. Однако, к примеру, термин «drone» с английского языка переводят, как беспилотник или дрон в качестве полных синонимов. Тем не менее, полные синонимы в языке встречаются сравнительно реже, чем гипонимы или гиперонимы, в результате чего возникает проблема гипо-гиперонимических отношений между терминами. Согласно учебному пособию Мишанкиной Н.А., «гиперо-гипонимические (родо-видовые) отношения — тип отношений между словами одной и той же части речи, разными по форме, имеющими связанные друг с другом значения: одно слово выступает как родовое по отношению к другому — видовому» [3]. При запросе, являются ли беспилотник и дрон синонимами, ИИ утверждает, что это разные термины. Еще одной важной проблемой новых быстроразвивающихся отраслей и не только является отсутствие перевода специализированных терминов в обычных переводных словарях, при этом каждый словарь своей конкретной отрасли дает свое определение конкретному термину. Нейронным сетям неоткуда брать репрезентативные данные по этой проблеме, что влечет ошибки в дальнейшем.

Во-вторых, еще одним проблематичным моментом является область перевода полисемичных терминов. Например, полисемичное слово кисть – кисть руки, кисть рябины и т.д. Если в контексте художник сломал кисточку, то перевод оказывается ошибочным. Недостаток контекста чаще всего наблюдается в диалогичной речи, где фразы короче, нежели в полноценном тексте.

В-третьих, проблемы существуют и в самой терминологии – это терминологизация и детерминологизация слова, что влечет за собой проблемы в синонимии, полисемии и омонимии. Так, например, имя Дрон омонимично термину, а значит вызовет проблему при переводе.

В-четвертых, это проблема в снятии омонимии в корпусе. «На данный момент корпус дезаимбигуирован лишь примерно на 1,5%, а значит предстоит много работы по дезаимбигуированию корпуса» [2].

Таким образом, применение нейросетей, ИИ, в том числе и чат-ботов несомненно облегчает и ускоряет рабочий процесс переводчика, однако машинный перевод до сих пор не достиг уровня адекватности человеческого, несмотря на скорость и качество. Во избежание ошибок в работе ИИ все еще требуется оценка качества и постредактура перевода

лингвистом-переводчиком. Требуется обучение ИИ, составление баз данных (параллельных корпусов текстов), снятие омонимии в корпусе, а также необходимо активное развитие терминологии в сфере беспилотных аппаратов, составление глоссариев и словарей.

Источники и литература

- 1) Ахманова О.С. Словарь омонимов русского языка / О.С. Ахманова. – 3-е изд. – М.: Русский язык, 1986. – 448 с.
- 2) Зобнин А. Как устроен национальный корпус русского языка // YouTube: [сайт]. – 2015. – URL: https://www.youtube.com/watch?v=QMUt9rFY_nA (дата обращения: 25.10.2023).
- 3) Мишанкина Н.А. Лексикология русского языка. Русский язык как иностранный: профессиональная сфера общения: учебное пособие / Н. А. Мишанкина. – изд. 2-е доп. – Томск: Изд-во Томского политехнического университета, 2017. – 124 с.