

DeepL, новый компьютерный переводчик

Перевод – это искусство, столь же древнее, как и сама цивилизация. Однако древние корни не являются препятствием для инноваций, а нейронные сети и появление чат-ботов на основе искусственного интеллекта (ШИ), таких как ChatGPT от OpenAI, повлекли за собой стремительное развитие переводческих технологий.

Мир становится все более взаимосвязанным, способность общаться на разных языках становится все более важной, а спрос на качественный перевод еще никогда не был столь высоким. Недавние опросы свидетельствуют о том, что более 70% переводчиков в Европе уже используют в работе машинный перевод.

Одним из лидеров на этом рынке является немецкая компания DeepL, которая предлагает переводы с 32 языков, имеет более 100 тыс. корпоративных клиентов, миллиард пользователей через приложения, расширения для браузеров и веб-сайт.

В чем ее уникальность? Компания использует свою модель нейронных сетей, специально обученную для лучшего уровня точности в переводе для бизнеса. Базой для сервиса и обучения ШИ-модели стал ее обширный онлайн-словарь Linguee, который помогает легко конкурировать с такими платформами, как Google Translate и Grammarly.

Как небольшой стартап из Кельна бросил вызов Google Translate и Grammarly? Боится ли он конкуренции с Google, Meta и OpenAI и как планирует развиваться дальше?

Рынок машинного перевода

Мировой рынок перевода на основе ИИ, по прогнозам Statista, вырастет с 5 млрд долл. в 2023 году до 12 млрд долл. в 2030 году. Опрос американских бизнес-лидеров, проведенный KPMG, показывает, что 97% компаний планируют в этом году инвестировать в ИИ-проекты по меньшей мере 250 млн долл., а 6% - вдвое больше.

Способствует этому, что спрос на ИИ-функции, в том числе и в переводе, растет: исследование IBM показало, что 42% компаний активно внедряют ИИ, а еще 40% - изучают его потенциал.

Впоследствии начальные данные Linguee дополнили другими источниками – веб-сайтами компаний и соцсетями, которые помогают технологии DeepL овладевать обычными идиомами.

В чем особенность сервиса DeepL? Ее онлайн переводчик использует архитектуру глубокой нейронной сети Transformer, которая разработана для преодоления ограничений предыдущих нейронных моделей машинного перевода (в частности, RNN) и стала высокоэффективной в создании качественных переводов.

Дело в том, что RNN обрабатывает входные последовательности по одному слову за раз, и влияние предыдущих слов на следующие уменьшается, когда последовательность становится длиннее. Transformer же использует механизм самоуважения, который позволяет ей одновременно обращать внимание на все слова в предложении, фиксируя все зависимости.

"В то время как Google Translate часто делает дословный перевод, пропускающий некоторые нюансы и идиомы (или делает перевод этих идиом совершенно неправильным), DeepL обеспечивает естественный перевод, который ближе к переводу опытного переводчика", – писал сразу после запуска DeepL TechCrunch.

Стартап тогда заявлял, что во время тестирования, проведенного компанией, профессиональные переводчики предпочли сервис DeepL в соотношении 3:1 по сравнению с конкурирующими инструментами от Google, Microsoft и Facebook. Версия DeepL показала лучшие результаты и в тесте на точность перевода по системе BLEU, чем конкуренты.

Нейронные сети DeepL способны улавливать даже самые маленькие нюансы и воспроизводить их в переводе лучше любого другого сервиса. Чтобы оценить качество наших моделей машинного перевода, мы регулярно проводим слепые тесты, во время которых профессиональные переводчики выбирают самый точный перевод, не зная, кто его сделал. DeepL превосходит конкурентов в три раза", – уверяют в DeepL.

Сейчас компания предлагает бесплатные и платные услуги, поддерживая более 30 языков – в том числе украинский (с сентября 2022). Бесплатный уровень ограничивает количество символов, которые пользователь может перевести за раз, безопасность пользовательских данных, и количество документов, которые он может переводить в месяц. Уровень API DeepL позволяет получить доступ к большему количеству глоссариев и переводов.

А DeepL Pro предназначен для корпоративного пользования и имеет три уровня, которые оплачиваются за каждого пользователя в год: начальный – от \$100, расширенный – от \$300, а лучший – почти \$700.

За годы существования DeepL сделала ставку на переводческие услуги для бизнеса в таких сферах, как производство, юриспруденция, розничная торговля, технологии, здравоохранение. Специализированная платформа помогает решать различные задачи компаний – от внутренних коммуникаций до поддержки клиентов и выхода на международные рынки.

Еще одно преимущество переводчика DeepL – можно загружать и переводить документы в форматах PDF, Word, HTML или PowerPoint с сохранением их оригинального формата.

Ее платформа для бизнеса также обеспечивает значительное сокращение издержек и повышение эффективности. Исследование компании Forrester, проведенное в 2024 году, показало, что использование DeepL обеспечило 345% рентабельности инвестиций для глобальных компаний, сократив время перевода на 90% и снизив рабочую нагрузку на 50%.

Такая эффективность позволила компании собрать уже более 100 тыс. корпоративных клиентов, наиболее известные среди которых – японский медиахолдинг Nikkei, провайдер онлайн-курсов Coursera, немецкий железнодорожный холдинг Deutsche Bahn и компания Zendesk, занимающаяся разработкой программного обеспечения.

"У Zendesk мы на собственном опыте убедились, насколько эффективно можно внедрять инструменты искусственного интеллекта в клиентский опыт, и лучший в области перевод сайта DeepL является ярким примером", – говорит технический директор Zendesk Адриан Макдермотт.

По словам инвестировавшего в DeepL представителя Index Ventures Дэнни Римера требования к точности, конфиденциальности и безопасности переводов в корпоративном мире намного выше, чем в потребительском, и качество модели стартапа дает ему преимущество.

Кутиловский признает, что DeepL делает акцент именно на качественном бизнес-переводе. "Когда вы находитесь в отпуске в Испании, очевидно, что перевести меню в ресторане важно, но качество перевода не столь важно. Это

огромная разница, когда вы хотите отправить электронное письмо клиенту, и мы очень быстро это заметили", – говорит он.

Конкуренция и риски

"Технологические гиганты типа Google, Microsoft и Facebook используют машинное обучение для перевода, но маленькая компания под названием DeepL превзошла их всех и подняла планку в этой области. Ее переводческий инструмент такой же быстрый, как и у огромных конкурентов, но более точный и детализированный в сравнении с другими протестированными нами инструментами", – писал TechCrunch после запуска DeepL.

Стартап заявляет, что разрабатывает лучшее в мире программное обеспечение машинного перевода. Это утверждение основывается на "слепых тестах", во время которых профессиональные переводчики выбирают точный вариант перевода текста, не зная, какая компания его создала, объясняют в DeepL.

"Все сводится к тому, чтобы иметь лучшую технологию", – говорит Кутиловский. Он уверяет, что нейронные сети компании могут анализировать огромные объемы данных, собранных из Интернета, чтобы узнать наиболее вероятный перевод для любого предложения.

Но это делают и конкуренты DeepL, также используя нейронные сети. В частности, материнская компания Facebook Meta выпустила программное обеспечение на базе ИИ, которое может автоматически переводить с около 100 языков как устную речь, так и текст. Компания заявила, что ее технология SeamlessM4T является шагом к универсальному переводчику.

Не стоит забывать и о Google и Microsoft. Кроме привычного уже Google Translate, первая развивает также Google Cloud, позволяющий компаниям динамически переводить текст на своих вебсайтах и в приложениях с помощью программного интерфейса API. Он использует модель нейронного машинного перевода, которую корпорация регулярно обновляет.

Microsoft Translator же позволяет предприятиям, разработчикам приложений и поставщикам языковых услуг создавать индивидуальные системы нейронного машинного перевода.

Работает в секторе перевода и Amazon: сервис Amazon Translate поддерживает более 70 языков.

Сможет ли амбициозный стартап поддерживать нынешние темпы роста в условиях жесткой борьбы с гораздо более крупными технологическими компаниями? Кутиловский, похоже, не опасается конкуренции. По его словам, DeepL и дальше будет сосредотачиваться на исследованиях нейронных сетей для перевода и получения наилучших данных для обучения.

"На самом деле мы конкурируем с Google Translate, а значит, и со всем Google, с момента основания компании. Поэтому для нас ландшафт не слишком изменился", – говорит основатель DeepL.

Однако он признает, что в последнее время темп перемен ускорился. "В сфере ИИ проходит много разработок, а скорость исследований выросла. Это, безусловно, вызов для наших команд, ведь им нужно немного ускориться", – говорит он.

Компания также начала выходить за пределы своего основного рынка, запустив уже упомянутый ШИ-инструмент DeepL Write для конкуренции с Grammarly. Стартап также стремится выйти за пределы своей европейской базы, особенно Азии и США.

Несмотря на все, DeepL сохраняет оптимизм по отношению к своему будущему – там убеждены, что сосредоточенность на качестве перевода откроет для нее новые возможности.

"Когда вы улучшаете качество перевода, шаг за шагом, для машинного перевода открывается все больше и больше сфер применения", – говорит Кутиловский.

Но риск для DeepL, по словам аналитиков, в том, что бизнесы уже и так используют один из облачных сервисов Google, Microsoft или Amazon. И как только с развитием ШИ-сервисов этих техногигантов преимущество DeepL в переводах станет менее заметным, он может начать терять корпоративных клиентов, поскольку компаниям проще воспользоваться предложением на платформе, на которой они уже работают.

Также на горизонте у DeepL потенциально может появиться еще один конкурент – OpenAI, выпускающий целый ряд громких инструментов на основе ИИ.

Пока неясно, будет ли OpenAI создавать отдельные переводческие сервисы, или объединится в этом с Microsoft. Кутиловский говорит, что пока нет конкретных планов относительно того, как и будет ли DeepL когда-нибудь сотрудничать с OpenAI, и какую форму это может приобрести. Основатель стартапа признает, что языковые модели, которые использует DeepL, похожи на те, что есть в OpenAI, и что у обеих компаний много общих клиентов.

"Мы всегда находимся в режиме гонки с точки зрения перевода. Мы привыкли к большим противникам, и часть нашей культуры состоит в том, чтобы продвигаться вперед, несмотря на это", – резюмировал он.