В Санкт-Петербурге разработали ИИ-инструмент для анализа человеческого поведения

01.10.2026

Сотрудники и студенты НИУ ВШЭ совместно с экспертами из Санкт-Петербургского Федерального исследовательского центра РАН и Центра практического ИИ Сбера разработали нейросеть, способную одновременно распознавать эмоции, оценивать видимые черты личности и выявлять амбивалентность (неуверенность или противоречивость поведения). Результаты исследования опубликованы в журнале IEEE Access.

Для естественного взаимодействия с человеком искусственному интеллекту недостаточно анализировать только содержание речи и выражения лица. Важно учитывать невербальные особенности поведения, которые помогают цифровым системам точнее интерпретировать контекст общения и адаптироваться под запросы.

Модель, разработанная учеными НИУ ВШЭ совместно с экспертами из СПб из ФИЦ РАН и Сбербанка, получает информацию о поведении человека сразу по нескольким каналам. Она анализирует видео, голос и содержание речи, а также сгенерированное описание мимики, взгляда, позы и жестов. Таким образом нейросеть одновременно распознает эмоции, оценивает воспринимаемые черты личности и выявляет амбивалентность.

По словам доцента Школы информатики, физики и технологий НИУ ВШЭ — Санкт-Петербург

Дмитрия Рюмина, новый подход позволяет анализировать разные проявления человека во взаимосвязи друг с другом.

> *«Существующие модели часто специализируются на конкретном корпусе данных и хуже переносят знания. Чтобы решить эту проблему, мы объединили несколько разнородных корпусов, каждый из которых был размечен для своей задачи, и научили модель использовать всю информацию одновременно. Самое интересное, что это повысило качество работы с данными, которые она читала впервые»,*

— рассказывает **Дмитрий Рюмин**.

Нейросеть обучали на суперкомпьютере НИУ ВШЭ «Харизма». Для трех задач использовали разные корпуса данных — каждый со своей разметкой.

> *«Основная сложность разработки ИИ-инструмента заключалась не только в разной разметке корпусов, но и в том, что информативность модальностей зависела от задачи. Например, для оценки воспринимаемых черт личности особенно важно видео, а для выявления амбивалентности — содержание речи. Поэтому мы объединили в модели общие компоненты, которые формируют связи между разными модальностями, и специализированные, — они выделяют наиболее важную информацию для решения каждой задачи»,*

— объясняет руководитель проекта в отделе прикладных технических решений НИУ ВШЭ и старший научный сотрудник Лаборатории речевых и мультимодальных интерфейсов СПб ФИЦ РАН **Елена Рюмина**.

Исследователи сравнили три режима обучения. В первом модель работала только с размеченными данными, во втором — добирала их из других корпусов, а в третьем обучалась на нескольких задачах одновременно. Дополнительные данные повысили распознавание эмоций примерно на 6 %, а качество оценки воспринимаемых черт личности — на 3,2 %.

> *«Совместные исследования с академическими партнерами позволяют нам решать проблемы на пути к созданию технологий, готовых к реальному применению. В ходе этой работы удалось выяснить, что использование данных о личности и амбивалентности заметно улучшило распознавание эмоций на незнакомых данных. Однако до универсального эмоционального ИИ еще далеко: многозадачное обучение пока не всегда превосходит специализированные модели, не хватает данных для других психологических характеристик, поэтому важно держать под контролем этические вопросы»,*

— отметил научный директор Центра практического искусственного интеллекта Сбера **Андрей Савченко**.

Способность работать с новыми данными проверили на отдельном корпусе, который модель не видела во время обучения. Многозадачность улучшила качество распознавания эмоций на 13,7 %.

Результаты исследования показывают потенциал разработки для цифровых образовательных систем, интеллектуальных интерфейсов и технологий поддержки пользователя, где важно учитывать эмоциональные и поведенческие особенности человека.

> *«Мы показали, что единую мультимодальную модель можно обучать на отличающихся корпусах с разной разметкой и использовать знания, полученные при решении разных задач, совместно. Это большой шаг к созданию устойчивых мультимодальных систем»,*

— резюмирует **Дмитрий Рюмин**.

До конца 2026 года участники проекта планируют повысить адаптивность модели. Следующим шагом станет внедрение нейросетевых экспертов, которые будут по-разному учитывать голос, содержание речи, жесты и другие особенности поведения в зависимости от поставленной задачи. Параллельно специалисты работают над персонализированными цифровыми аватарами, способными адаптировать мимику, речь и поведение к эмоциональному состоянию пользователя.

Материал: пресс-служба Санкт-Петербургского Федерального исследовательского центра РАН