Фото: Freepik

Тренажер на основе нейронных сетей поможет улучшить произношение на английском языке

***Студенты Санкт-Петербургского государственного университета аэрокосмического приборостроения разработали тренажер для корректировки произношения на английском языке с использованием современных нейронных сетей.***

Тренажер позволяет корректировать произношение с использованием современных моделей нейронных сетей, таких как Wav2Vec2, предварительно дообученных на задачу распознавания фонем. Такой подход обеспечивает высокую точность распознавания речи и чёткую транскрипцию. Пользователь может записывать своё произношение, получать его транскрипцию и сравнивать её с эталонной версией. Программа анализирует ошибки, дает рекомендации и отслеживает прогресс в обучении.

Система преобразует устную речь пользователя в цифровой формат, анализируя акустические сигналы. Для этого используются технологии автоматического распознавания речи – такие, как модели, обученные на больших объемах данных для распознавания фонем, слов и фраз.

– Мы собрали датасет, включающий аудио и соответствующие фонемные транскрипции. Далее был разработан прототип системы CAPT, позволяющий точно распознавать произношение на уровне фонем. Система анализирует полученные акустические данные и сравнивает их с эталонной моделью речи, затем генерирует обратную связь для пользователя. Она может указать на ошибки в произношении конкретных звуков, выделить проблемные слова или части фраз, повторно воспроизвести эталонную речь для подражания, – рассказал один из разработчиков проекта Михаил Дорохин, студент ГУАП направления «Прикладная информатика».

Применение модели Wav2Vec2 позволяет точно распознавать речь и эффективно анализировать фонетическую транскрипцию. Тренажер поможет изучающим английский язык совершенствовать разговорные навыки независимо от наличия преподавателя или носителя языка. В дальнейшем разработчики планируют расширить функционал системы, интеграцию с другими инструментами обучения и улучшить алгоритмы обратной связи для более детального анализа ошибок произношения.