

Специалисты Сбера разработали метод HuBERT-CTC, который повышает точность автоматического распознавания русской речи искусственным интеллектом (ИИ). Технология была представлена на международной конференции Interspeech 2025.

Метод использует целевые переменные из CTC-модели распознавания, что позволяет формировать более семантические представления данных. Существующие модели обычно опираются на низкоуровневые акустические переменные.

Новый подход показал значительное улучшение в обработке русского языка. Он снижает количество ошибок распознавания на 50% по сравнению с моделью Whisper-large-v3 от OpenAI.

Важное преимущество технологии — возможность обучения на неразмеченных аудиоданных. Это решает проблему дефицита качественных размеченных данных для обучения ИИ-моделей.

Метод масштабируется по размеру модели и объёму данных. Одна архитектура работает как в онлайн-, так и в офлайн-режиме без необходимости переобучения.

Разработка может быть применена в голосовых помощниках, контакт-центрах, системах аналитики телефонных звонков и мультимодальных чат-ботах.