

Учёные из МГУ, при поддержке НОШ «Мозг, когнитивные системы, искусственный интеллект», разработали новый метод, который эффективно решает две важные задачи: определение спикеров на аудиозаписях и их активности. Этот подход хорошо работает даже при потоковой обработке и перекрытии речи.

Диаризация спикеров — это процесс определения, кто и когда говорит на аудиозаписи. Это важно для многих приложений, включая транскрибирование речи и анализ эмоций. Новый метод эффективен и универсален, способен работать в различных условиях, включая шум и перекрытие речи.

Исследователи МГУ создали метод, который преодолевает ограничения существующих подходов. Он показывает превосходное качество даже в сценариях с перекрытием речи и может применяться в режиме реального времени, что открывает новые перспективы для различных приложений, включая распознавание речи и системы наблюдения.