

Российские ИТ-специалисты представили нейросетевую технологию, позволяющую виртуальным помощникам точнее распознавать голосовые команды на фоне посторонних звуков. Разработка уже внедрена в продуктах «Яндекса», включая умные колонки и ТВ-станции.

Главная особенность подхода — применение механизма внимания, который получает два сигнала одновременно: с эхоподавлением и шумоподавлением. Система в реальном времени выбирает наиболее чёткий вариант, что позволяет повысить точность распознавания речи даже в сложных звуковых условиях.

Разработчики обучили алгоритм на базе записей, сделанных как в лаборатории, так и на пользовательских устройствах. Решение адаптировано под семь микрофонов и показало эффективность как на мощных вычислительных системах, так и на малопроизводительном оборудовании с ограниченными ресурсами.

До этого момента не существовало единого алгоритма, одинаково хорошо работающего и в тестовой среде, и в реальности. В новой технологии удалось добиться устойчивой работы в условиях фонового шума — будь то музыка, звук воды или работающий пылесос.