

Группа «Т-Технологии» опубликовала в открытом доступе свою разработку — потоковую модель распознавания русской речи T-one. Решение уже используется в собственных сервисах компании, включая службу поддержки Т-Банка и мобильного оператора Т-Мобайл.

Модель размером 70 млн параметров показывает высокую точность даже при работе с зашумленными записями из колл-центров. По внутренним тестам компании, T-one превосходит по качеству распознавания более крупные открытые аналоги, включая Whisper Large-v3 с 1,5 млрд параметров.

Ключевое преимущество разработки — возможность обработки аудиопотоков в реальном времени с минимальной задержкой. Это позволяет использовать модель для автоматизации звонков, создания голосовых ассистентов и анализа разговоров.

Метрики качества

T-one по качеству конкурирует с более крупными открытыми оффлайн-моделями в телефонии

Домен	T-one + LM, 70M	GigaAM-RNNNT v2, 243M	GigaAM-CTC v2, 242M	Whisper Large-v3, 1.5B
Колл-центр	8.63	10.22	10.57	19.39
Прочая телефония	6.20	7.88	8.15	17.29
Именованные сущности	5.83	9.55	9.81	17.87
Common Voice 19	5.32	2.68	3.14	5.78

Т-Банк

«Т-Технологии» выложили в открытый доступ не только саму модель, но и код для ее запуска в производственных средах. Решение распространяется по лицензии Apache 2.0, что позволяет свободно использовать и модифицировать его для коммерческих проектов.

«Т-Технологии» открыли доступ к модели распознавания речи

Разработчики отмечают, что компактный размер T-one дает возможность развернуть систему без дорогостоящего оборудования. Это снижает затраты на внедрение и делает технологию доступной даже для небольших компаний.