

УДК 004.522

ПРОБЛЕМЫ ИСПОЛЬЗОВАНИЯ ТЕХНОЛОГИЙ РАСПОЗНАВАНИЯ РЕЧИ

Степанов Д.А., ст. преп., Тарасенков Д.А., студ.

*Витебский государственный технологический университет,
г. Витебск, Республика Беларусь*

Технология распознавания речи представляет собой возможность, которая позволяет программе обрабатывать человеческую речь в письменном формате. Хотя его обычно путают с распознаванием голоса, распознавание речи фокусируется на переводе речи из вербального формата в текстовый, тогда как распознавание голоса просто стремится идентифицировать голос отдельного пользователя.

Различные алгоритмы и вычислительные методы используются для распознавания речи в тексте и повышения точности транскрипции. Ниже приведены примеры некоторых из наиболее часто используемых методов: обработка естественного языка, скрытые марковские модели, N-граммы, нейронные сети, диаризация говорящего.

Сегодня множество отраслей используют различные приложения речевых технологий, помогая предприятиям и потребителям экономить время и даже жизни. Вот некоторые примеры: автомобилестроение, технологии, здравоохранение, продажи, безопасность.

Современные проблемы распознавания речи вызваны двумя основными факторами – охватом и громкой средой. Это требует еще более точных систем, которые могут справиться с самыми амбициозными вариантами использования ASR: живые интервью, распознавание речи на громком семейном ужине или встречах с разными людьми. Это предстоящие задачи, которые необходимо решить для распознавания голоса следующего поколения.

Помимо этого, распознавание речи должно быть доступно для большего количества языков и охватывать широкие темы. Потому что на данный момент ASR требуется много данных для хорошей работы, и некоторые из них просто не были собраны для определенных языков и тем. Не добавляя их, системы ASR останутся заметно недостоверными. Также к проблемам использования технологий распознавания речи можно отнести: неточность и ложные толкования, время и недостаточную эффективность, акценты и местные различия, фоновый шум и громкую среду, конфиденциальность и безопасность данных.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Чекмарёв, А. Речевые технологии – проблемы и перспективы / А. Чекмарёв // Компьютерра, № 49, с. 26-43, 1997 г.
2. Davies, K.H., Biddulph, R. and Balashek, S. (1952) Automatic Speech Recognition of Spoken Digits, J. Acoust. Soc. Am. 24 (6), pp. 637-642.