

## ДЕТЕКЦИЯ ЭМОЦИЙ В ДЕТСКОЙ РЕЧИ

Матвеев А.Ю. (Университет ИТМО)

Научный руководитель – к.т.н. Махныткина О.В.

(Университет ИТМО)

Исследования выполнены за счет финансирования университета ИТМО в рамках НИР №621296 «Разработка технологий для персонификации разговорного искусственного интеллекта».

В работе рассматривается задача автоматической классификации эмоций в детской речи. Производится анализ подходов и алгоритмов, а также демонстрируются результаты тестирования на образцах речи Русских и Индийских детей.

**Введение.** Задаче автоматического определения эмоций в речи сейчас уделяется существенное внимание, однако, в отношении некоторых специфичных подзадач, таких как детекция эмоций в детской речи или в кросс-культурном окружении, доступно лишь ограниченное количество опубликованных исследований и результатов. В современной литературе по направлению детекции эмоций на данный момент нет устоявшегося мнения в какой степени выражение эмоций в речи и возможности их автоматической классификации зависит от культурных и языковых различий.

Высокая акустическая и лингвистическая вариативность детской речи создает уникальные особенности применения методов автоматической классификации эмоций и психоневрологических состояний.

Автоматическая детекция эмоций в детской речи в кросс-культурном окружении еще сильнее усложняет ситуацию и в современной литературе можно найти лишь единичные примеры подобных исследований, в то время как значимость задачи подтверждается множеством конкурсов и грантов по данному направлению.

**Основная часть.** Одной из наиболее важных проблем в задачах классификации речи является выбор акустических признаков. Акустические признаки должны быть достаточно содержательными для того, чтобы презервировать информацию, необходимую для дифференциации между эмоциональными состояниями, но при этом достаточно компактными для возможности эффективного обучения системы. В результате анализа, в работе были выбраны наборы признаков DisVoice, Geneva Minimalistic Acoustic Parameter Set (GeMAPS) и extended Geneva Minimalistic Acoustic Parameter Set (eGeMAPS).

Для классификации образцов речи на основе выделенных признаков в задачах детекции эмоций зачастую используются такие алгоритмы как метод опорных векторов, скрытые марковские модели, смешанные Гауссовы модели, нейронные сети. В данной работе рассматриваются особенности применения метода опорных векторов и многослойного перцептрона для классификации.

Для тестирования используется 2505 образцов Русской речи и 417 образцов Тамильской речи. В интра-культурном окружении при тестировании применяется кросс-валидация, а в кросс-культурном также применяется схема leave-one-culture-out.

**Выводы.** Для оценки результатов применяются матрицы ошибок, а также такие характеристики как точность, полнота, невзвешенная средняя точность, и взвешенная и невзвешенная средняя полнота для мультиклассовой классификации. Демонстрируется, что результаты автоматической детекции сравнимы с классификацией экспертами, достигая точности свыше 80%. Производительность предложенных подходов позволяет рассматривать

возможность их практического применения, например, в рамках НИР №621296 «Разработка технологий для персонификации разговорного искусственного интеллекта».