

**ПРОЕКТИРОВАНИЕ И РАЗРАБОТКА С ПРИМЕНЕНИЕМ МЕТОДОВ
МАШИННОГО ОБУЧЕНИЯ СЕРВИСА ДЛЯ АНАЛИЗА ЗАПИСЕЙ ДЕЛОВЫХ
ТЕЛЕФОННЫХ РАЗГОВОРОВ С ЦЕЛЬЮ РАСПОЗНАВАНИЯ ЭМОЦИЙ**

Кузьмичева К.Е. (Университет ИТМО)

Научный руководитель – д.т.н., доцент Лямин А.В.

(Университет ИТМО)

Аннотация. В докладе поднимается проблема определения тональности речи. В ходе исследования моделей для решения данной проблемы строится алгоритм автоматического анализа записей телефонных разговоров с целью идентификации тональности речи, на основе которого проектируется и разрабатывается программный модуль.

Введение. Одним из направлений аффективных вычислений является обнаружение и распознавание эмоций по данным, собранным из языка тела человека, выражения лица, голоса и других сигналов, аналогичных тем, что используют люди, чтобы воспринимать эмоции друг друга. Распознавание эмоциональной информации требует извлечения закономерностей из собранных данных, на основании которых система будет выдавать ответ. Например, может быть выведена закономерность, что явно высокий тон голоса проявляется у радостного и заинтересованного человека, а, например, нахмуренные брови могут говорить о сосредоточенности либо о негативном настрое. В данной работе рассматриваются технологии обработки эмоциональной речи. Анализ речи является эффективным методом выявления аффективных состояний, средняя точность которого, согласно последним исследованиям, составляет 70-80%. Для сравнения, среднестатистический человек способен определять эмоции в речи с точностью, примерно, 60%. Однако более точными считаются системы, использующие другие методы обнаружения эмоций, такие как физиологические состояния или мимика.

Основная часть. Для повышения точности распознавания эмоций в докладе приводится ряд предложений, которые в дальнейшем будут использованы при построении классификатора.

Во-первых, успешность работы классификатора во многом зависит от качества обучающей выборки. Необходимо уделить внимание специфике обрабатываемой информации. В данном случае это будет анализ деловых телефонных разговоров. Обучать программу нужно на выборке именно такой тематики, ведь эмоции, которые выражаются в спонтанной речи, труднее поддаются распознаванию, чем, например, эмоции в речи человека, выступающего на сцене. При этом возникает сложность балансировки распределения данных по классам, так как в спонтанной речи, тем более деловой, проявление сильных эмоций, таких как счастье и гнев, маловероятно. Однако для составления качественной обучающей базы данных необходимо собрать приблизительно равное количество примеров каждого класса.

Во-вторых, технологии обработки эмоциональной речи распознают эмоциональное состояние человека с помощью вычислительного анализа характеристик речевого сигнала, т. е. используют только акустические параметры. В докладе же предлагается строить составной классификатор, который будет объединять результаты анализа, сделанного на основе характеристик голоса, и анализа тональности произнесенных человеком слов. Это усложнит систему, но повысит качество распознавания.

В-третьих, сосредоточимся на выделении только позитивного и негативного настроения говорящего. Большое количество типов эмоций резко снижает качество распознавания,

поэтому будем давать оценку в интервале $[-1; 1]$, где -1 – абсолютно негативная экспертная оценка, а 1 – абсолютно позитивная.

Предполагается следующий порядок действий при загрузке записи телефонного разговора в сервис на анализ. В начале происходит обработка аудиозаписи, в ходе которой производится очистка от шумов, удаление пустых участков и разделение каналов. После этого запись отправляется в модуль распознавания эмоций по голосу и в модуль аннотирования аудиоданных в текст, из которого текст передается в модуль распознавания эмоций в тексте. Модули распознавания эмоций анализируют речь и возвращают дискретизированные по времени оценки тональности для каждого канала. На последнем шаге происходит объединение результатов анализа, полученных из обоих модулей распознавания эмоций, в итоговый отчет.

Выводы. В результате работы была предложена модель классификатора для получения наиболее высокого процента точности в задаче идентификации тональности речи, построен алгоритм работы сервиса, решающего задачу распознавания эмоций в записях деловых телефонных разговоров. Разработанный программный модуль может быть внедрен, например, в систему call-центра для мониторинга работы сотрудников при общении с клиентами.

Кузьмичева К.Е. (автор)

Лямин А.В. (научный руководитель)