Метод понижения размерности пространства признаков на основе LASSO-регрессии для задачи   
распознавания эмоций по речи

*маг.**Краснопрошин Д.В.****,*** *проф.**Вашкевич М.И.*

|  |  |
| --- | --- |
|  | *Белорусский государственный университет  информатики и радиоэлектроники (БГУИР)*  *Кафедра электронных вычислительных средств* |
|  | |

**Цель доклада**

Представить метод понижения размерности пространства признаков, основанный на применении LASSO-регрессии, для повышения эффективности распознавания эмоций по речи.

# Разработка системы распознавания эмоций

## **Речевая база**

RAVDES (Ryerson Audio-Visual Database of Emotional Speech and Song), 8 эмоциональных состояний (нейтральность, спокойствие, счастье, грусть, гнев, страх, удивление и отвращение), 12 мужчин и 12 женщин, 1440 аудиофайлов в формате wav (16 бит, 48 кГц)

## **Речевы признаки**

Мел-частотных кепстральных коэффициентов (МЧКК), МЧКК, МЧКК,

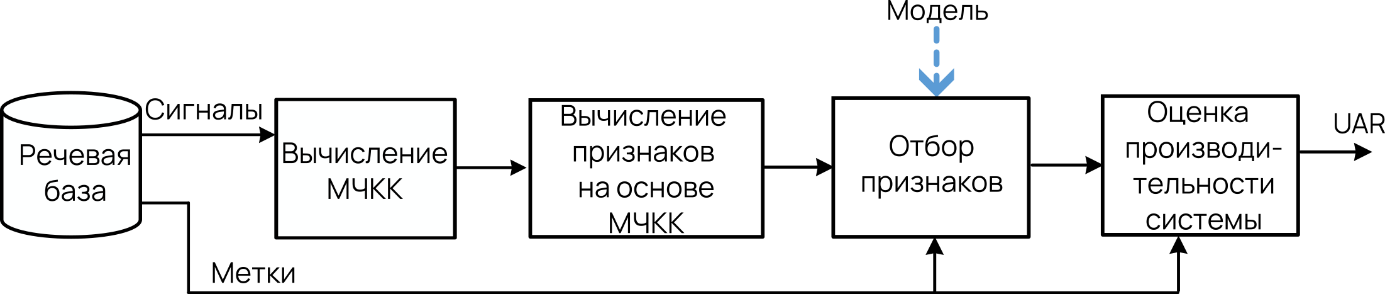
межквартильный размах, коэффициент ассиметрии и эксцесс.

(Изображение)

# Процесс разработки системы распознавания эмоций

Затем был произведен отбор признаков, которые способствуют повышению производительности системы распознавания эмоций. Данный этап выполнялся с учетом классификатора, который будет использоваться в дальнейшем в системе распознавания эмоций.

Качество модели оценивалось с помощью UAR (Unweighted Average Recall) и перекрестной проверки.



# Процедура отбора признаков

Разработан алгоритм отбора признаков для задачи многоклассовой классификации, основанный на использовании LASSO-регрессии:

1) Задача классификации на ***K***-классов **{C1,C2,...Ck}** заменяется на ***K*** задач бинарной классификации, по схеме «один против всех».

2) Для решения задач бинарной классификации применяется метод отбора признаков на основе LASSO. В результате получается **k** подмножеств признаков из их исходного полного набора.

3) Для каждого из K наборов признаков, полученных на шаге 2, применяется метод пошагового исключения переменных (англ. *BSS – backward-stepwise selection*). Этот этап позволяет выбрать оптимальное подмножество наиболее значимых признаков, снижая размерность данных и устраняя избыточность.

4) Полученные на шаге 3 **K** наборов признаков объединяются для формирования итогового набора признаков.

**Применение LASSO для отбора признаков в задаче     бинарной классификации**

Метод LASSO основан на решении задачи линейной регрессии:

|  |  |
| --- | --- |
|  | (1) |

|  |  |
| --- | --- |
| – число примеров в обучающем наборе, – метка класса -го образца, –-й признак -го образца, – -й коэффициент регрессии, – параметр регуляризации. Решение (1) при достаточно большом приводит к тому, что часть коэффициентов становятся в точности нулевыми. |  |

**Применение LASSO для отбора признаков в задаче     бинарной классификации**

После ранжирования признаков была выполнена оценка полноты классификации при использовании возрастающего набора признаков. В ходе этой процедуры определялось число признаков, которое обеспечивало наибольшую точность классификации.

|  |  |
| --- | --- |
| Отбор подмножества признаков по результатам ранжирования с использованием LASSO-регрессии |  |

**Набор данных и извлечение признаков**

В качестве набор данных использовался Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS) [3]. который содержит записи от 24 актеров (12 мужчин и 12 женщин), включая 104 высказывания на каждого актера (60 речевых и 44 песенных). В контексте нашей работы, мы ограничились использованием речевых высказываний, что составило 1440 аудиофайлов в формате wav (16 бит, 48 кГц. RAVDESS содержит эмоциональные состояния: нейтральность, спокойствие, счастье, грусть, гнев, страх, удивление и отвращение.

Речевые признаки рассчитывались на основании мел-частотных кепстральных коэффициентов (МЧКК).

В итоговый набор *исходных признаков* были включены среднее значение МЧКК (34 признака), среднеквадратичное отклонение МЧКК (34 признака), среднее от первой и второй производных от МЧКК (68 признаков), их среднеквадратическое отклонение (68), а также коэффициент асимметрии, эксцесс и межквантильный размах (по 34 признака для каждой характеристики соответственно). Таким образом, для каждого аудиофайла мы получаем 306-компонентный вектор надсегментных признаков МЧКК.

**Тестирование классификатора**

Для тестирования классификатора использовался метод перекрестной проверки по *k*-блокам (*k-fold cross-validation*)

В данной работе использовалась схема разбиения на блоки из работы [1] (в скобках указаны номера актеров):

- блок 0: (2, 5, 14, 15, 16);

- блок 1: (3, 6, 7, 13, 18);

- блок 2: (10, 11, 12, 19, 20);

- блок 3: (8, 17, 21, 23, 24);

- блок 4: (1, 4, 9, 22).

[1] *Multimodal Emotion Recognition on RAVDESS Dataset Using Transfer Learning* / C. Luna-Jiménez, D. Griol, Z. Callejas, R. Kleinlein, J.M. Montero, F. Fernández-Martínez // Sensors. – 2021. – vol. 21. – pp. 1 – 29.

**Результаты экспериментальных исследований**

Исследовалось применение классификаторов на основе

• линейного дискриминантного анализа (ЛДА)

• машины опорных векторов (МОВ).

Оценка производительности выполнялась с использованием невзвешенной средней полноты – UAR (*Unweighted Average Recall*).

**Результаты экспериментальных исследований**

|  |  |  |
| --- | --- | --- |
| Классификатор | Полный набора признаков (306) | Отобранный набор признаков |
| ЛДА | 0,460 | **0,484** (205 признаков) |
| МОВ | 0.456 | **0.473** (208 признаков) |

• После процедуры отбора признаков из изначального набора, состоящего из 306 характеристик, оставлено 205 признаков при повышении точности классификатора в случае с ЛДА и 208 признаков в случае с МОВ.

• Результат демонстрирует значительное уменьшение размерности признакового пространства.

• Одновременно удалось сохранить высокий уровень точности.

**Результаты экспериментальных исследований**

Далее представлена матрица спутанности для лучшей модели.

|  |  |
| --- | --- |
| Матрица спутанности | Сравнение с [1] Привести сравнение |

**Заключение**

В работе *предложен* метод понижения размерности пространства признаков, основанный на применении LASSO-регрессии, для повышения эффективности распознавания эмоций по речи*.*

В работе выполнено сравнение качества класиификации модели с использованием исходного набора признаков, и набора признаков полученных с помощью разработанного метода понижения размерности признакового пространства. Благодаря новому методу **удалось снизить размерность примерно на 30% и одновременно удалось сохранить высокий уровень классификационной точности, что подчеркивает эффективность примененной методологии отбора признаков.**