

УДК 004.934

ОПТИМИЗАЦИЯ ГИПЕРПАРАМЕТРОВ ТРЕХКЛАССОВОЙ ДИАРИЗАЦИИ

Мешеряков И.Д. (Университет ИТМО)

Научный руководитель – Попов Д.В.

(ГК «ЦРТ»)

В данной работе исследуется задача трехклассовой диаризации и осуществляется применение оптимизатора гиперпараметров для задачи трехклассовой диаризации. В качестве оптимизируемых гиперпараметров, помимо параметров первичной кластеризации, были взяты пороги принятия решения классификации. Также, в работе вводится новая метрика для оценки качества трехклассовой диаризации.

Диаризация – это процесс разделения звукового потока на несколько в соответствии с принадлежностью сегментов тому или иному говорящему. Трехклассовая диаризация, в отличие от классической диаризации, позволяет разделить аудиоданные на 3 потока – оператор, клиент и шумы. Задача трехклассовой диаризации играет значимую роль в решении задачи сегментирования аудиоданных, а актуальность трехклассовой диаризации обусловлена задачами анализа аудиоданных, записанных в зашумленных помещениях с большим количеством дикторов или низким качеством записи. В работе более подробно анализируется трехклассовая диаризация, идет сравнение двухклассовой и трехклассовой диаризаций, выявляются ее проблемы и преимущества. Исследуется оптимизатор гиперпараметров диаризации, основанный на байесовской оптимизации.

В работе рассматривается оптимизатор гиперпараметров классической диаризации, основанный на байесовской оптимизации с гауссовскими процессами в качестве суррогатной функции. При этом, представлена адаптация данного оптимизатора гиперпараметров для решения задачи трехклассовой диаризации. Проверка работоспособности адаптированного оптимизатора была проведена на данных, предоставленных крупной сетью автозаправочных станций и содержащих диалоги «кассир – клиент». Также, для анализа качества трехклассовой диаризации, была введена новая метрика, позволяющая сравнивать между собой результаты классической и трехклассовой диаризации.

Для оптимизации гиперпараметров трехклассовой диаризации был использован адаптированный оптимизатор гиперпараметров классической диаризации. В результате работы адаптированного оптимизатора качество трехклассовой диаризации на исследуемых данных было улучшено. Данное решение может быть использовано для повышения качества трехклассовой диаризации аудиоданных, собранных в людных помещениях.