

## Отчёт о проделанной работе

**Тема работы:** Исследование методов многозадачного обучения для автоматического распознавания речи

**Состав коллектива:** Легченко Антон Евгеньевич – студент 4 курса бакалавриата ММФ НГУ (a.legchenko@g.nsu.ru), научный руководитель – Павловский Евгений Николаевич к.ф. – м.н., в.н.с. ЛАПДИМО ММФ НГУ (e.pavlovskiy@g.nsu.ru)

**Информация о работе:** Основная часть работы выполнялась в рамках дипломной работы бакалавриата.

### Научное содержание работы:

- 1. Постановка задачи:** Обучение интегрированных моделей (сепарация речи вместе с ее последующей транскрипцией), изучение и сравнение полученных результатов, изучения влияния дополнительной задачи транскрипции для улучшения моделей сепарации голосов, изучение влияния такого обучения на устойчивость модели транскрипции, аналогичный набор задач, в случае использования методики дистилляции знаний.
- 2. Современное состояние проблемы:** Нейросетевые модели крайне требовательны к вычислительным ресурсам, особенно на этапе обучения. В связи с этим крайне актуальной всегда остается задача получения более компактных моделей, повышения их качества. Популярным подходом к решению данной задачи является обучение малой модели воспроизводить распределение модели учителя, иногда такие модели достигают качества не хуже учителя. Одним из подходов к улучшению качества моделей является многозадачное обучение, когда при обучении модели необходимо решать разные задачи из одной предметной области, благодаря чему повышается качество решения всех задач и получаются более верные внутренние представления о предметной области. На данный момент существуют работы по использованию двух этих подходов к распознаванию текстов, когда одна общая модель обучается решать большое кол-во задач от специализированных моделей учителей, в итоге получается общая модель, решающая задачи не хуже учителей, а по некоторым параметрам даже лучше. Распознавание речи является также ресурсно затратной задачей, но совмещение данных подходов не использовалось, хотя есть работы показывающие их эффективность по отдельности, также в существующих работах использовались небольшие сочетания задач (обычно пары), актуальной задачей является выявление зависимостей и влияния различных задач на внутренние представления модели, поскольку данный вопрос опускается в существующих работах.
- 3. Подробное описание работы:** В рамках работы были составлены и синтезированы обучающие наборы данных с помощью библиотеки torchaudio. Построены архитектуры изучаемых моделей и реализованы алгоритмы их обучения и тестирования, на основе библиотеки Pytorch. Было произведено обучение и сравнение изучаемых решений и решения не использующего исследуемого метода для оценки его эффективности. После получения сравнительных результатов, было исследованы изменения в признаковом пространстве, а также их визуализация с помощью алгоритма понижения размерности tSNE.
- 4. Полученные результаты:** Исследуемый метод дал прирост качества по различным оценкам качества порядка 1%, а также обнаружены вносимые задачей транскрибирования изменения в признаковое пространства модели разделения дикторов.

**Эффект от использования кластера:** Обучение всех моделей происходило на гри ускорителях кластера и было бы крайне затруднительно в связи с высокой вычислительное трудоемкостью задач.

**Перечень публикаций, содержащих результаты работы:** Работа основанная на части результатов была опубликована в материалах студенческой сессии конференции по компьютерной лингвистике ДИАЛОГ 2022 (ссылка на статью на сайте конференции <https://www.dialog-21.ru/media/5799/dialog2022.pdf>) Также все полученные результаты вошли в выпускную квалификационную работу.