

4. Ускорить завершение единой межуровневой образовательной онлайн-платформы.

Список использованных источников

1. Концепция цифровой трансформации процессов в системе образования Республики Беларусь на 2019–2025 годы : утв. Министром образования Респ. Беларусь 15 марта 2019 г. – Мн. : Министерство образования, 2019. – 49 с.

2. Открытое образование : нац. платформа : [сайт]. – URL: <https://openedu.ru> (дата обращения: 10.11.2025).

УДК 004.855.5

ЛЕГКОВЕСНЫЕ МОДЕЛИ АУДИОСЕПАРАЦИИ КАК РЕСУРС И ИНФРАСТРУКТУРА ИННОВАЦИОННОГО РАЗВИТИЯ ОБРАЗОВАНИЯ

Малинчик Н. С., магистрант

*Научный руководитель – канд. физ.-мат. наук,
доцент Статкевич С. Э.*

*Гродненский государственный университет имени Янки Купалы
Гродно, Республика Беларусь*

Аннотация. Обоснование роли легковесных моделей аудиосепарации как ресурса и элемента инфраструктуры инновационного образования, раскрывая их значимость для подготовки специалистов, внедрения инженерных практик и формирования современной цифровой среды учебных заведений.

Ключевые слова: легковесные модели, аудиосепарация, машинное обучение, образовательная инфраструктура, ресурсы образования.

Работа с аудиосигналами на курсе машинного обучения позволяет студентам освоить все ключевые этапы современного ML-процесса (см. рис. 1): от формирования датасетов и feature extraction до проектирования и оптимизации нейросетевых архитектур, выбора метрик и тестирования на разных платформах. Одномерная структура аудиосигнала упрощает знакомство с базовыми концепциями

(convolution, pooling, нормализация, downsampling) и отличиями между распространенными архитектурами (сверточные, рекуррентные, encoder-decoder), минуя сложности, типичные для computer vision [1]. Работа с аудио также наглядно иллюстрирует построение пайплайнов, аугментации, балансировки классов, техник фреймирования и спектрограммирования (STFT, log-mel и др.).

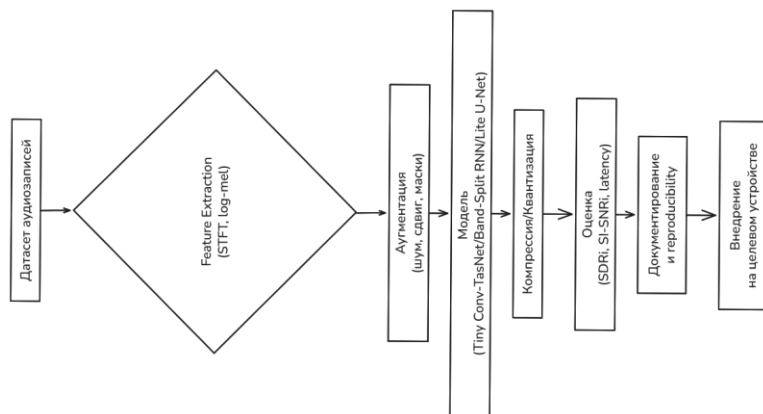


Рисунок 1 – Диаграмма ML-процесса

Основу курса составляют практически проекты: извлечение речи из шума, сегментация аудио, интеллектуальные микшеры. Для этих задач студенты подбирают и настраивают архитектуры, представленные в табл. 1.

Таблица 1 – Сравнение легковесных архитектур для обработки аудио

Модель	Тип Архитектуры	Особенности	Примеры задач
Tiny Conv-TasNet	Сверточная, bottle-neck	Минимализм, temporal conv, skip	Разделение речи
Band-Split RNN	Многополосный RNN	Band splitting, GRU/LSTM по полосам	Сегментация, шумоподавление
Lite U-Net	Encoder-Decoder с skip-links	Variable depth/width, агрессивный downsampling	Интеллектуальный микшер

В лабораторных заданиях студенты осваивают приемы pruning (отсев незначимых весов, сокращение размерности каналов), quantization (замена float32 на int8/float16, анализ деградации качества), связывают компрессию с latency и расходом памяти. По мере продвижения в курс изучаются методы distillation (обучение компактных моделей по выходам teacher), band-splitting, fine-tuning моделей для ARM/MCU, а также влияние FLOPs и mixed-precision на RAM footprint при реальном развертывании. Итоговые этапы посвящены воспроизводимости: студенты тестируют модели (SDRi, SI-SNRi, latency) на разных аудио-сценах и устройствах, сравнивая с базовыми конфигурациями.

Результаты студенческих экспериментов приносят пользу и кафедре: аккумулярованные пайплайны и source-коды становятся источником лучших практик и демонстрационных кейсов, а неудачные архитектуры – исключаются из дальнейших циклов, что экономит ресурсы. Разработка студентами внутренних сервисов (автоматическая фильтрация лекций, улучшение аудио-курсов, автоматическое транскрибирование) расширяет инструментарий для университета и облегчает поддержку образовательного контента.

Легковесные архитектуры обеспечивают максимальную доступность прототипирования – студент буквально за несколько занятий может пройти весь цикл от идеи до рабочего демо. Быстрая сборка, обучение и тестирование компактной модели позволяют реализовать функциональный ML-проект даже в рамках краткого курса, не требующего значительных вычислительных ресурсов. Такой формат способствует развитию инженерного мышления, ускоряет получение обратной связи и мотивирует на самостоятельные эксперименты.

Упрощенная интеграция ML-компонент достигается благодаря компактности и низким требованиям к ресурсам: легковесные модели легко включаются в учебные приложения, лабораторные стенды, модули для обработки аудио на образовательных платформах. Студенты могут внедрять и тестировать свои решения в реальной инфраструктуре, наблюдать результат работы на настоящих устройствах и пользовательских сервисах, что укрепляет связь между теорией и практикой.

Кроме того, легковесные модели служат стартовой базой для последующего усложнения архитектур и applied-исследований. Освоение минималистичных структур и техник оптимизации позволяет студенту

шаг за шагом переходить к экспериментам с более тяжелыми, глубокими сетями, комбинировать разные подходы и создавать гибридные решения. Такой путь способствует развитию устойчивых навыков проектирования ML-систем, открывает перспективы для самостоятельной научной работы и участия в совместных R&D-проектах.

Инженерные навыки, полученные в процессе работы с легковесными моделями, напрямую востребованы в индустрии: здесь критичны не только качество, но и возможность запуска решений на embedded-, edge- и мобильных устройствах. Освоение компрессии, квантизации, анализа trade-off'ов «качество-размер», оформление reproducibility – все это повышает мобильность выпускников и их готовность к командной и самостоятельной applied-работе. Такой подход помогает факультету актуализировать учебные курсы и поддерживать баланс между современной наукой и практикой.

Интеграция задач аудиосепарации и компактных ML-моделей дает синергетический эффект: студенты приобретают значимую экспертизу в работе с «живыми» сигналами [2], учатся критически анализировать результаты, а образовательная система накапливает отработанные решения и инновационные практики. Все это ускоряет переход от теории к практике, формирует платформу для цифровой среды, готовой отвечать вызовам высокой технологичности. Расширение проектных задач (интеграция мультимодальных данных, аудио-видео-поток) знакомит студентов с принципами комплексных интерфейсов и управления потоками разнородных сигналов – особенно актуально для дистанционных и гибридных образовательных форматов. Опыт прототипирования real-time систем и систем обратной связи развивает и технические, и коммуникативные компетенции. В результате полученные знания становятся прочной основой для карьеры в AI, embedded, EdTech и цифровых медиа.

Список использованных источников

1. Luo, Y. Conv-TasNet: Surpassing Ideal Time-Frequency Masking for Speech Separation / Y. Luo, N. Mesgarani // IEEE/ACM Transactions on Audio, Speech, and Language Processing. – 2019. – Vol. 27, № 8. – P. 1256–1266.
2. Ronneberger, O. U-Net: Convolutional Networks for Biomedical Image Segmentation / O. Ronneberger, P. Fischer, T. Brox // Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015 : Lecture Notes in Computer Science. – 2015. – Vol. 9351. – P. 234–241.