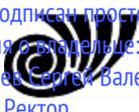


Документ подписан простой электронной подписью Информация о владельце: ФИО: Таскаев Сергей Валерьевич Должность: Ректор	 МИНОБРАЗОВАНИЯ И НАУКИ РОССИИ Федеральное государственное бюджетное образовательное учреждение высшего образования «Челябинский государственный университет» (ФГБОУ ВО «ЧелГУ»)	
Дата подписания: 18.03.2025 14:53:16 Уникальный прозрачный ключ: 04c19ed8bfb98f3b6cb77a486b9a8788b8323233	Рабочая программа дисциплины "Распознавание и синтез речи" по направлению подготовки (специальности) 01.03.02 Прикладная математика и информатика направленности (профилю) Прикладная математика и искусственный интеллект ФГБОУ ВО «ЧелГУ»	стр. 1

Рабочая программа дисциплины (модуля)*

Распознавание и синтез речи

Направление подготовки (специальность)

01.03.02 Прикладная математика и информатика

Направленность (профиль)

Прикладная математика и искусственный интеллект

Присваиваемая квалификация (степень)

бакалавр

Форма обучения

очная

Год(ы) набора 2024

*Рабочая программа дисциплины (модуля) адаптирована для инклюзивного обучения инвалидов и лиц с ограниченными возможностями здоровья

Челябинск 2023 г.



Содержание

1. Цели освоения дисциплины
2. Место дисциплины в структуре ОПОП
3. Компетенции обучающегося, формируемые в результате освоения дисциплины (модуля)
4. Объем дисциплины (модуля)
5. Структура и содержание дисциплины (модуля)
6. Фонд оценочных средств
 - 6.1. Перечень видов оценочных средств
 - 6.2. Типовые контрольные задания и иные материалы для текущей аттестации
 - 6.3. Типовые контрольные вопросы и задания для промежуточной аттестации
 - 6.4. Критерии оценивания
7. Учебно-методическое и информационное обеспечение дисциплины (модуля)
 - 7.1. Рекомендуемая литература
 - 7.2. Перечень ресурсов информационно-телекоммуникационной сети "Интернет"
 - 7.3. Перечень информационных технологий
8. Материально-техническое обеспечение дисциплины (модуля)
9. Методические указания для обучающихся по освоению дисциплины (модуля)
10. Специальные условия освоения дисциплины обучающимися с инвалидностью и ограниченными возможностями здоровья



1. ЦЕЛИ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Цели дисциплины «Распознавание и синтез речи» состоят в выработке у студентов знаний теоретических и практических основ обработки цифровых сигналов, их применению в анализе речи, современных нейросетевых методов распознавания речи и ее синтеза. К задачам дисциплины относятся: получения навыков обработки цифровых сигналов; формирование навыков постановки задач распознавания и синтеза речи; получение практических навыков решения задач распознавания и синтеза речи.

ПК-12 (ПК-9 модели): Результаты обучения:

Знает: [ПК-9.4. 3-1.] принципы построения систем распознавания и синтеза речи, методы и подходы к планированию и реализации проектов по созданию и поддержке систем искусственного интеллекта на основе сквозной цифровой субтехнологии

"Распознавание и синтез речи"

Умеет: [ПК-9.4. У-1.] применять методы и подходы к планированию и реализации проектов по созданию и поддержке системы искусственного интеллекта на основе сквозной цифровой субтехнологии "Распознавание и синтез речи"

2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОПОП

Цикл (раздел) ОПОП: Б1.В.1.11

2.1 Требования к предварительной подготовке обучающегося:

Основы компьютерного зрения

2.2 Дисциплины и практики, для которых освоение данной дисциплины (модуля) необходимо как предшествующее:

Выполнение и защита выпускной квалификационной работы

3. КОМПЕТЕНЦИИ ОБУЧАЮЩЕГОСЯ, ФОРМИРУЕМЫЕ В РЕЗУЛЬТАТЕ ОСВОЕНИЯ ДИСЦИПЛИНЫ (МОДУЛЯ)

ПК-12: Способен создавать и внедрять одну или несколько сквозных цифровых субтехнологий искусственного интеллекта

Знать:

принципы построения систем распознавания и синтеза речи, методы и подходы к планированию и реализации проектов по созданию и поддержке систем искусственного интеллекта на основе сквозной цифровой субтехнологии "Распознавание и синтез речи"

Уметь:

применять методы и подходы к планированию и реализации проектов по созданию и поддержке системы искусственного интеллекта на основе сквозной цифровой субтехнологии "Распознавание и синтез речи"

Владеть:

навыками, методами и подходами в реализации проектов на основе технологии "Распознавание и синтез речи"

В результате освоения дисциплины обучающийся должен

3.1	Знать:
3.1.1	принципы построения систем распознавания и синтеза речи, методы и подходы к планированию и реализации проектов по созданию и поддержке систем искусственного интеллекта на основе сквозной цифровой субтехнологии "Распознавание и синтез речи"
3.2	Уметь:



3.2.1 применять методы и подходы к планированию и реализации проектов по созданию и поддержке системы искусственного интеллекта на основе сквозной цифровой субтехнологии "Распознавание и синтез речи"

3.3 Владеть:

3.3.1 в реализации проектов на основе технологии "Распознавание и синтез речи"

4. ОБЪЕМ ДИСЦИПЛИНЫ (МОДУЛЯ)

Общая трудоемкость	2 ЗЕТ
Часов по учебному плану : 72	Виды контроля в семестрах: зачеты 8
в том числе :	
аудиторные занятия : 48	
самостоятельная работа : 19,75	
контактная работа: 52,25 ИКР: 4,25	

5. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

Код занятия	Наименование разделов и тем /вид занятия/	Семестр / Курс	Часов	Литература
Раздел 1. Цифровая обработка сигналов				
1.1	Введение в курс. Понятия гильбертова пространства и тригонометрического базиса. Неравенство Бесселя и тождество Парсеваля. Дискретные преобразования Фурье. /Лек/	8	2	Л1.1 Л1.2Л2.1
1.2	Преобразование Фурье /Лаб/	8	2	Л1.1 Л1.2Л2.1
1.3	Линейные стационарные системы. Цифровые фильтры. Стабильность и импульсная характеристика. Z-transform. Подходы к построению фильтров. /Лек/	8	2	Л1.1 Л1.2Л2.1
1.4	Обработка сигналов цифровыми фильтрами /Лаб/	8	2	Л1.1 Л1.2Л2.1
1.5	Кепстральный анализ и цифровые фильтры. Сложный кепстр и кепстр. Кепстральный анализ речи. Кепстральное представление Mel-Frequency. /Лек/	8	2	Л1.1 Л1.2Л2.1
1.6	Расчет Mel-Frequency коэффициентов /Лаб/	8	2	Л1.1 Л1.2Л2.1
1.7	Оконное преобразование Фурье. Вейвлет преобразование. Feature engineering аудио-сигнала /Лек/	8	2	Л1.1 Л1.2Л2.1
1.8	Аналогово-цифровые преобразователи и сжатие сигналов /Лек/	8	2	Л1.1 Л1.2Л2.1
1.9	Цифровая обработка сигналов /Ср/	8	6	Л1.1 Л1.2Л2.1
Раздел 2. Распознавание речи				
2.1	Биологические аспекты речи. Выделение признаков, нормализация и аугментации речи. /Лек/	8	2	Л1.1 Л1.2Л2.1
2.2	Извлечение признаков из аудиосигналов. Часть 1 /Лаб/	8	2	Л1.1 Л1.2Л2.1
2.3	Извлечение признаков из аудиосигналов. Часть 2 /Лаб/	8	2	Л1.1 Л1.2Л2.1
2.4	Кластеризация и квантизация в распознавании речи. Скрытые марковские модели /Лек/	8	2	Л1.1 Л1.2Л2.1
2.5	Кластеризация в задачах распознавания речи /Лаб/	8	4	Л1.1 Л1.2Л2.1
2.6	Нейронные сети для распознавания речи. Рекуррентные нейронные сети. Connectionist Temporal Classification /Лек/	8	2	Л1.1 Л1.2Л2.1
2.7	Скрытые марковские модели /Лаб/	8	2	Л1.1 Л1.2Л2.1
2.8	Методы работы с последовательностями. Механизм внимания /Лек/	8	2	Л1.1 Л1.2Л2.1
2.9	Задачи распознавания: идентификация голоса; определение конца предложения; определение активности; распознавание ключевой фразы /Лек/	8	2	Л1.1 Л1.2Л2.1
2.10	ASR модель /Лаб/	8	4	Л1.1 Л1.2Л2.1



2.11	Распознавание речи /Ср/	8	7,75	Л1.1 Л1.2Л2.1
	Раздел 3. Синтез речи			
3.1	Алгоритмы синтеза звука /Лек/	8	2	Л1.1 Л1.2Л2.1
3.2	Использование сверток для обработки последовательностей. WaveNet, DeepVoice и их улучшения. Multi-speaker synthesis. Tacotron /Лек/	8	2	Л1.1 Л1.2Л2.1
3.3	Синтез речи /Лаб/	8	4	Л1.1 Л1.2Л2.1
3.4	Синтез речи /Ср/	8	6	Л1.1 Л1.2Л2.1
	Раздел 4. Курсовая работа			
4.1	Курсовая работа /КурсР/	8	4,25	Л1.1 Л1.2Л2.1

6. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

6.1. Перечень видов оценочных средств

Задания к лабораторным работам 1--9.
Вопросы к зачету.

См. Приложения

6.2. Типовые контрольные задания и иные материалы для текущей аттестации

См. Приложения

6.3. Типовые контрольные вопросы и задания для промежуточной аттестации

См. Приложения

6.4. Критерии оценивания

Каждая лабораторная работа оценивается следующим образом:

Код запускается без ошибок - 2 балла;

выполнены все задания - 3 балла;

выполнено более половины заданий - 1 балл;

студент может пояснить порядок получения результатов и расчетов – 2 балла.

В остальных случаях баллы не начисляются.

Максимальный балл - 8.

Все ЛР, кроме шестой, имеют вес 1. Шестая - вес 2.

Промежуточная аттестация - зачет. Число баллов 6.

На зачете студент отвечает на билет, который содержит 2 теоретических вопроса. При необходимости студенту могут быть заданы дополнительные вопросы по заданиям. Продолжительность зачёта – 60 минут. Максимальный балл за ответ на теоретический вопрос 3 балла. 3 балла - ответ структурирован, приведен анализ положений существующих теорий по вопросу билета, студент логично и доказательно раскрывает проблему, предложенную в билете, ответ не содержит фактических ошибок и характеризуется глубиной, полнотой; 2 балла - ответ имеет достаточный содержательный уровень, однако отличается слабой структурированностью, раскрыто содержание билета, имеются неточности при ответе; 1 балл - ответ имеет фрагментарный характер, отличается поверхностностью и малой содержательностью, имеются неточности при ответе на основные вопросы билета, материал в основном излагается, но допущены фактические ошибки; 0 баллов - допускаются существенные фактические ошибки при ответе, на большую часть дополнительных вопросов по содержанию экзамена; студент затрудняется дать ответ или не дает верных ответов.

Процедура проведения зачета, критерии оценивания.

Оценка за дисциплину формируется на основе полученных оценок за контрольно-рейтинговые мероприятия текущего контроля. Студент может повысить свой рейтинг, пройдя контрольное мероприятие промежуточной аттестации – ответ по билету. Студент выбирает случайный билет, содержащий два теоретических вопроса. Студенту предоставляется не более 60 минут на подготовку ответа. По истечении этого времени студент отвечает экзаменатору вопросы билета. Фиксация результатов учебной деятельности по дисциплине проводится в день зачета при личном присутствии студента.



7. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

7.1. Рекомендуемая литература

7.1.1. Основная литература

	Авторы,	Заглавие	Издательство,	Ресурс
Л1.1	Тампель И. Б., Карпов А. А.	Автоматическое распознавание речи: учебное пособие (https://e.lanbook.com/book/91466)	Санкт-Петербург : НИУ ИТМО, 2016	ЭБС
Л1.2	Столв Е. Л.	Цифровая обработка сигналов. Водяные знаки в аудиофайлах (https://e.lanbook.com/book/212891)	Санкт-Петербург : Лань, 2022	ЭБС

7.1.2. Дополнительная литература

	Авторы,	Заглавие	Издательство,	Ресурс
Л2.1	Гудфеллоу Я., Бенджио И., Курвилль А.	Глубокое обучение (https://e.lanbook.com/book/107901)	Москва : ДМК Пресс, 2018	ЭБС

7.3 Перечень информационных технологий

7.3.1 Программное обеспечение

OpenOffice

Python

Visual Studio

7.3.2 Профессиональные базы данных и информационно-справочные системы

8. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

Лекции: ПК, проектор, экран.

Лабораторные

занятия: ПК с установленным ПО: MS Office, Microsoft Visual Studio, Python, Jupyter notebook.

9. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ (МОДУЛЯ)

От студента требуется систематическая работа в течение семестра, выполнение лабораторных работ и освоение теоретического материала.

10. СПЕЦИАЛЬНЫЕ УСЛОВИЯ ОСВОЕНИЯ ДИСЦИПЛИНЫ ОБУЧАЮЩИМИСЯ С ИНВАЛИДНОСТЬЮ И ОГРАНИЧЕННЫМИ ВОЗМОЖНОСТЯМИ ЗДОРОВЬЯ

Освоение дисциплины инвалидами и лицами с ограниченными возможностями здоровья осуществляется с использованием специальных технических средств и голо информационных технологий, предоставляемых Ресурсным учебно-методическим центром по обучению инвалидов и лиц с ограниченными возможностями здоровья ЧелГУ по запросу обучающегося.

1. Мобильные специальные технические средства для лиц с нарушениями зрения: портативный компьютер с вводом/выводом шрифтом Брайля с синтезатором речи «EIBraile-W14J G2»; ноутбуки с программной экранного доступа NVDA; электронные увеличители для удаленного просмотра; видеоувеличители портативные; тифлоплеер; цифровые диктофоны.

2. Мобильные специальные технические средства для лиц с нарушениями слуха: система свободного звукового поля со встроенной совместимостью с FM-устройствами; радиоклассы «Сонет-PCM» с передатчиком, заушным индуктором и индукционной петлей; система информационная для слабослышащих переносная «Исток» А2 со встроенным плеером – звуковым информатором; документ-камера; программируемые слуховые аппараты индивидуального пользования.

3. Ассистивные информационные технологии: программное обеспечение экранного доступа с синтезом речи NVDA; программы экранного увеличения; программы речевого синтеза для компьютеров и ноутбуков; программы речевого синтеза для мобильных устройств; экранная клавиатура; экранная лупа.

При необходимости для обучающихся с нарушениями зрения на рабочих местах для проведения практических или



лабораторных занятий устанавливается специальное программное обеспечение (программа речевой навигации NVDA, речевые синтезаторы, экранные лупы).

В учебные аудитории обеспечивается беспрепятственный доступ для обучающихся инвалидов и обучающихся с ограниченными возможностями здоровья. В каждой аудитории, где обучаются инвалиды и лица с ограниченными возможностями здоровья, предусматривается соответствующее количество мест для обучающихся с учетом нарушений их здоровья.

Для освоения дисциплины инвалидам и лицам с ограниченными возможностями здоровья предоставляется доступ к печатным источникам, имеющимся в научной библиотеке ЧелГУ, с помощью специальных технических средств; доступ к электронным источникам, представленным в форме электронного документа в фонде научной библиотеки ЧелГУ или электронно-библиотечных системах, с помощью специальных технических и программных средств (рабочее место для незрячего пользователя с программным обеспечением экранного доступа с синтезом речи NVDA, рабочее место с компьютерным роллером и клавиатурой CleVu с большими кнопками и с разделяющей клавиши накладкой).

Учебно-методические материалы для обучающихся из числа инвалидов и лиц с ограниченными возможностями здоровья предоставляются в формах, адаптированных к ограничениям их здоровья и восприятия информации:

Для лиц с нарушениями зрения:

- в печатной форме увеличенным шрифтом,
- в форме электронного документа,
- в форме аудиофайла,
- в печатной форме шрифтом Брайля.

Для лиц с нарушениями слуха:

- в печатной форме,
- в форме электронного документа.

Для лиц с нарушениями опорно-двигательного аппарата:

- в печатной форме,
- в форме электронного документа,
- в форме аудиофайла.

Данный перечень может быть конкретизирован в зависимости от контингента обучающихся.

Для инвалидов и лиц с ограниченными возможностями здоровья освоение дисциплины может быть частично или полностью осуществлено с использованием дистанционных образовательных технологий (Moodle, Adobe Connect Pro и пр.).

В освоении дисциплины инвалидами и лицами с ограниченными возможностями здоровья используется индивидуальная работа. Под индивидуальной работой подразумевается две формы взаимодействия с преподавателем: индивидуальная учебная работа (консультации), т.е. дополнительное разъяснение учебного материала и углубленное изучение материала с теми обучающимися, которые в этом заинтересованы, и индивидуальная воспитательная работа. Индивидуальные консультации направлены на индивидуализацию обучения и установлению воспитательного контакта между преподавателем и обучающимся инвалидом или обучающимся с ограниченными возможностями здоровья.

При проведении процедуры оценивания результатов обучения инвалидов и лиц с ограниченными возможностями здоровья по дисциплине обеспечивается выполнение следующих дополнительных требований в зависимости от индивидуальных особенностей, обучающихся:

- а) инструкция по порядку проведения процедуры оценивания предоставляется в доступной форме (устно, в письменной форме, в письменной форме шрифтом Брайля, устно с использованием услуг сурдопереводчика);
- б) доступная форма предоставления заданий оценочных средств (в печатной форме, в печатной форме увеличенным шрифтом, в печатной форме шрифтом Брайля, в форме электронного документа, задания зачитываются ассистентом, задания предоставляются с использованием сурдоперевода);
- в) доступная форма предоставления ответов на задания (письменно на бумаге, набор ответов на компьютере, письменно шрифтом Брайля, с использованием услуг ассистента, устно).

При проведении процедуры оценивания результатов обучения инвалидов и лиц с ограниченными возможностями здоровья предусматривается использование технических средств, необходимых им в связи с их индивидуальными особенностями. Эти средства могут быть предоставлены ЧелГУ или могут использоваться собственные технические средства. При необходимости инвалидам и лицам с ограниченными возможностями здоровья предоставляется дополнительное время для подготовки ответа на задания, процедура оценивания результатов обучения по дисциплине может проводиться в несколько этапов.

Проведение процедуры оценивания результатов обучения инвалидов и лиц с ограниченными возможностями здоровья допускается с использованием дистанционных образовательных технологий.

Приложение

Фонд оценочных средств ООП «Прикладная математика и искусственный интеллект» по направлению 01.03.02 – Прикладная математика и информатика
Дисциплина «Распознавание и синтез речи»

№ КМ	Вид КМ	Наименование КМ	Оценочные средства
1.	Текущий контроль	Лабораторная работа № 1. Преобразования Фурье	<p>Задание:</p> <ol style="list-style-type: none"> 1. Сгенерировать сигнал заданный формулой 2. Написать функцию прямого преобразования Фурье 3. Написать функцию обратного преобразования Фурье 4. Написать функцию быстрого преобразования Фурье* 5. Применить преобразования к сигналу 6. Сравнить результаты работы функций с <code>scipy.fft</code> и <code>scipy.ifft</code>
2.	Текущий контроль	Лабораторная работа № 2. Обработка сигналов цифровыми фильтрами	<p>Задание:</p> <ol style="list-style-type: none"> 1. Реализовать КИХ-фильтр 2. Реализовать БИХ-фильтр 3. Определите единичный скачок длиной 20 отсчетов 4. Рассчитать импульсную характеристику фильтров 5. Рассчитать амплитудно-частотную характеристику фильтров для единичного импульса 6. Рассчитать фазово-частотную характеристику фильтров для единичного импульса
3.	Текущий контроль	Лабораторная работа № 3. Расчет Mel-Frequency коэффициентов	<p>Задание:</p> <ol style="list-style-type: none"> 1. Загрузить датасет с аудиозаписями, сделанными людьми разного пола 2. Реализовать функцию расчета MFCC 3. Для случайных 10 файлов из папки каждого пола рассчитать MFCC
4.	Текущий контроль	Лабораторная работа № 4. Извлечение признаков из аудиосигналов. Часть 1	<p>Используя датасет из Лабораторной работы №3 выполнить следующие задания:</p> <ol style="list-style-type: none"> 1. Установить библиотеку librosa 2. Разделить гармонические (тональные) и ударные (переходные) сигналы на две формы волны <code>librosa.effects.hpss()</code> 3. Извлечь темп и биты <code>librosa.beat.beat_track(r)</code> 4. Построить нормализованную энергию цветности - Chroma Energy Normalized (CENS) <code>chroma=librosa.feature.chroma_cens()</code> 5. Рассчитать MFCC <code>librosa.feature.mfcc(y=y_harmonic, sr=sr, n_mfcc=20)</code> 6. Рассчитать спектрограмму <code>librosa.stft(y)</code> 7. Рассчитать спектральный центроид <code>librosa.feature.spectral_centroid(y=y, sr=sr)</code>

№ КМ	Вид КМ	Наименование КМ	Оценочные средства
			<ol style="list-style-type: none"> 8. Рассчитать спектральный контраст feature.spectral_contrast(y=y_harmonic,sr=sr) 9. Рассчитать спектральный спад librosa.feature.spectral_rolloff(y=y, sr=sr) 10. Рассчитать спектральную ширину librosa.feature.(y+0.01, sr=sr)[0] 11. Рассчитать скорость пересечения нуля feature.zero_crossing_rate(y_harmonic)
5.	Текущий контроль	Лабораторная работа № 5. Извлечение признаков из аудиосигналов. Часть 2	<p>Используя датасет из Лабораторной работы №3 для каждой аудиозаписи вычислить:</p> <ol style="list-style-type: none"> 1. Средние значения и стандартные отклонения Мел-кепстральных коэффициентов 2. Среднее значение, стандартное отклонение и наклон (skew) спектрального центроида 3. Среднее значение и стандартное отклонения спектрального спада 4. Собрать все в единый .csv файл с помощью pandas.DataFrame; таблица должны содержать: 5. столбец 'filename', где указан номер и название файла 6. столбцы 'mfcc_mean{i}' и 'mfcc_std{i}' - средние значения и стандартные отклонения (по 20 значений) из Мел-кепстральных коэффициентов; 7. столбцы 'cent_mean', 'cent_std', 'cent_skew' - среднее значение, стандартное отклонение и skew (наклон) из Спектрального центроида; 8. столбцы 'rolloff_mean', 'rolloff_std' - среднее значение и стандартное отклонение из Спектрального спада ; 9. столбец 'label' метка (male, female\0,1)
6.		Лабораторная работа № 6. Скрытые марковские модели	<p>Используя датасет из Лабораторной работы №3 и результат лабораторной работы №6 (.csv файл с признаками и разметкой) выполнить следующие задания:</p> <ol style="list-style-type: none"> 1. Разбить данные на train (60% от датасета), val(20%) и test(20%) 2. С помощью метода SVM построить классификатор, распознающий пол говорящего на аудиозаписи 3. Рассчитать accuracy обученной модели 4. Рассчитать precision и recall обученной модели 5. Рассчитать F1 score обученной модели
7.	Текущий контроль	Лабораторная работа № 7. Скрытые марковские модели	<p>Задание: используя MFCC (20 значений) из таблицы, полученной в лабораторной работе №6, как исходные состояния:</p> <ol style="list-style-type: none"> 1. обучить скрытую марковскую модель для классификации пола говорящего 2. рассчитать accuracy 3. рассчитать precision 4. рассчитать recall 5. рассчитать F1 score 6. составить таблицу сравнения метрик с моделью полученной в лабораторной работе №6 7. сравнить скорость работы скрытой марковской модели и модели SVM из лабораторной работы №6
8.	Текущий	Лабораторная работа № 8. ASR	<p>Задание:</p> <ol style="list-style-type: none"> 1. скачать датасет LibriSpeech из torchaudio

№ КМ	Вид КМ	Наименование КМ	Оценочные средства
	конт роль	модель	<ol style="list-style-type: none"> 2. имплементировать архитектуру convnet acoustic model из https://arxiv.org/abs/1609.03193 3. с помощью имплементации https://github.com/pytorch/audio/tree/release/0.12/examples/pipeline_wav2letter 4. обучить asr модель на датасете LibriSpeech 5. заменить convnet acoustic model на свою имплементацию 6. обучить wav2letter со своей имплементацией convnet acoustic model 7. сравнить метрики двух имплементаций 8. выбрать по 50 случайных записей мужского и женского голоса из датасета лабораторной работы №3 9. с помощью модели с наилучшими метриками получить текст который произносят на этих записях
9.	Текущий конт роль	Лабораторная работа № 9. Синтез речи	<p>Задание:</p> <ol style="list-style-type: none"> 1. изучить пример использования Tacotron 2 для синтеза речи 2. для текста полученного для 100 файлов в лабораторной работе №9 с помощью Tacotron 2 сгенерировать аудиофайлы 3. рассчитать для сгенерированных аудиофайлов признаки указанные в лабораторной работе №6 4. сравнить вычисленные признаки с признаками оригинальных аудиофайлов 5. применить для сгенерированных данных модель SVM полученную в лабораторной работе №7 6. Рассчитать accuracy обученной модели 7. Рассчитать precision и recall обученной модели 8. Рассчитать F1 score обученной модели
10	Промежуточный конт роль	Зачет	<p>Перечень вопросов на зачет:</p> <ol style="list-style-type: none"> 1. Дискретное преобразование Фурье 2. Быстрое преобразование Фурье 3. Кепстральные коэффициенты, их применение в анализе аудиосигналов 4. Вейвлет преобразования 5. Линейные фильтры, их типы и свойства 6. Методы извлечения признаков из аудиосигналов 7. Метод SVM для классификации аудиосигналов 8. Скрытые марковские модели. Принцип работы 9. Сверточные сети для анализа аудиосигналов 10. Рекуррентные сети для анализа аудиосигналов 11. Connectionist Temporal Classification 12. Принцип работы механизма внимания 13. Алгоритмы синтеза речи: WaveNet 14. Алгоритмы синтеза речи: DeepVoic 15. Алгоритмы синтеза речи: Tacotron 2

