

УТВЕРЖДЕНА

Решением Ученого совета
АНО ВО «Центральный университет»
«07» марта 2024 г.
Протокол № 1

**Рабочая программа дисциплины (модуля)
«Речевые технологии»**

Направление подготовки: 02.04.01 Математика и компьютерные науки

Направленность (профиль) подготовки: Продуктовая аналитика

Квалификация (степень) выпускника: магистр

Форма обучения: очная

Срок освоения программы: 2 года

Год набора: 2024

**Москва
2024**

Содержание

1. Краткая характеристика дисциплины (модуля)	3
2. Перечень планируемых результатов обучения.....	4
3. Тематический план.....	6
4. Содержание дисциплины (модуля).....	6
5. Учебно-методическое обеспечение	7
6. Материально-техническое обеспечение	7
7. Методические и оценочные материалы	9

1. Краткая характеристика дисциплины (модуля)

Рабочая программа дисциплины (модуля) «Речевые технологии» составлена в соответствии с федеральным государственным образовательным стандартом высшего образования – магистратура по специальности 02.04.01 Математика и компьютерные науки, профиль Продуктовая аналитика, утвержденный приказом Министерства науки и высшего образования Российской Федерации № 810 от 23.08.2017 года.

Изучение дисциплины (модуля) «Речевые технологии» позволяет освоить современные подходы к созданию интеллектуальных аудиосистем, востребованных в голосовых ассистентах, системах распознавания и синтеза речи, мультимодальных приложениях. Полученные знания и навыки обеспечивают основу для разработки конкурентоспособных решений в области обработки и генерации аудио и речи.

Место дисциплины (модуля) в структуре образовательной программы

Настоящая дисциплина (модуль) включена в учебный план по программе подготовки магистратуры по направлению 02.04.01 Математика и компьютерные науки, профиль Продуктовая аналитика и входит в вариативную часть Блока 1, формируемую участниками образовательных отношений.

Дисциплина (модуль) изучается на 2 курсе в 4 семестре, доступна для прохождения при условии успешного завершения дисциплин (модулей) «Основы Python», «Machine Learning (Машинное обучение)», «Deep Learning (Глубокое обучение)» и общеуниверситетского факультатива «Математика в DS».

Цель изучения дисциплины (модуля): формирование системного понимания современных методов обработки и генерации речи, а также практических навыков разработки, обучения и оценки речевых моделей с использованием современных инструментов глубокого обучения.

Задачи изучения дисциплины (модуля):

— формирование знаний по темам: основные задачи и приложения в области речевых технологий; способы представления и анализа аудиосигнала в цифровом виде; архитектура и принципы работы классических и современных нейросетевых моделей, решающих задачи, связанные с речью; основные метрики для оценки качества речевых систем; концепции обучения на неразмеченных или слабо размеченных данных; аудиокодеки и подходы языкового моделирования в задачах генерации аудио; передовые исследования мультимодальных и многозадачных больших аудиомоделей, методы их адаптации; научные статьи в области обработки речи, ключевые идеи и предлагаемые подходы

— освоение умений: применять стандартные библиотеки для загрузки, предобработки и анализа аудиосигналов, извлекать и визуализировать различные акустические признаки в среде Python; реализовывать и обучать базовые модели для задач анализа и генерации звука и речи с использованием фреймворка PyTorch; работать с многозадачными end-to-end моделями и строить каскады из множества компонент; анализировать и интерпретировать результаты работы речевых моделей, рассчитывать и сравнивать ключевые метрики качества;

— формирование навыков решения типовых задач в области речевых технологий, включая выбор архитектуры модели, функции потерь и метрик качества; навыков проектирования и реализации комплексных пайплайнов, состоящих из нескольких моделей; навыков практического использования современных фреймворков и библиотек для глубокого обучения (PyTorch, Hugging Face, SpeechBrain) применительно к аудиоданным.

2. Перечень планируемых результатов обучения

Компетенции, формируемые в результате освоения дисциплины (модуля) при проведении учебных занятий в форме контактной работы обучающихся с педагогическими работниками Университета и в форме самостоятельной работы обучающихся:

Компетенция	Содержание компетенции	Индикатор компетенции	Перечень планируемых результатов обучения по дисциплине (модулю)
УК-6.	Способен определять и реализовывать приоритеты собственной деятельности и способы ее совершенствования на основе самооценки	УК-6.1.	Знает основные методы самооценки и анализа своей деятельности, а также принципы управления временем и целеполагания
		УК-6.2.	Умеет ставить реалистичные и достижимые цели, определять приоритеты в своей деятельности, а также разрабатывать и внедрять планы по совершенствованию своих навыков и компетенций на основе полученной самооценки
		УК-6.3.	Имеет практический опыт применения методов самооценки в своей профессиональной деятельности, включая участие в тренингах, семинарах и проектах, направленных на развитие личной эффективности и профессионального роста
ОПК-2.	Способен создавать и исследовать новые математические модели в естественных науках, совершенствовать и разрабатывать концепции, теории и методы	ОПК-2.1.	Знает основные математические модели и методы, используемые в естественных науках, включая статистическое моделирование, дифференциальные уравнения и численные методы, а также современные подходы к исследованию и анализу данных
		ОПК-2.2.	Умеет разрабатывать и адаптировать математические модели для решения конкретных проблем в естественных науках, проводить их анализ и верификацию, а также интерпретировать полученные результаты в контексте научных исследований
		ОПК-2.3.	Имеет практический опыт создания и исследования математических моделей в рамках научных проектов или исследований, включая участие в публикациях, конференциях или коллаборациях, где были разработаны и апробированы новые концепции и методы
ПК-3.	Способен решать задачи	ПК-3.1.	Знает методы и инструменты

	<p>профессиональной деятельности в области продуктовой аналитики, формулировать результаты анализа и выявлять последствия полученных данных для принятия обоснованных решений и оптимизации продуктов</p>		продуктовой аналитики
		ПК-3.2.	Умеет применять аналитические инструменты и программное обеспечение для обработки и визуализации данных, а также формулировать выводы на основе проведенного анализа
		ПК-3.3.	Имеет опыт работы над реальными проектами в области продуктовой аналитики, включая анализ пользовательского поведения и оптимизацию продуктов на основе полученных данных
ПК-4.	<p>Способен публично представлять собственные и известные научные результаты</p>	ПК-4.1.	Знает основные принципы эффективного публичного выступления, методы визуализации данных и основные требования к научным презентациям, включая структуру и содержание
		ПК-4.2.	Умеет четко и логично формулировать свои научные результаты, адаптируя их для различных аудиторий, а также использовать визуальные средства для улучшения восприятия информации
		ПК-4.3.	Имеет практический опыт участия в научных конференциях, семинарах или других мероприятиях, где успешно представлял свои и известные научные результаты, получая обратную связь и взаимодействуя с аудиторией

3. Тематический план

№ п/п	Наименование раздела дисциплины (модуля)	Трудоемкость, академические часы				ТКУ (текущий контроль успеваемости)
		<i>Очная форма</i>				
		Аудиторная работа		Контроль	Самостоятельная работа	
Лекции	Семинары (практические занятия)					
1	Анализ сигнала	2	2		10	Домашние задания
2	Распознавание речи	4	4		19	Домашние задания
3	Биометрия	2	2		10	Домашние задания
4	Синтез речи	6	6		27	Домашние задания
5	Большие аудио модели	2	2		10	Домашние задания
	<i>Зачет</i>			6		
	<i>Итого:</i>	<i>16</i>	<i>16</i>	<i>6</i>	<i>76</i>	
	<i>Объем дисциплины (модуля) (в ак. ч.)</i>	<i>114</i>				
	<i>Объем дисциплины (модуля) (в зач. ед.)</i>	<i>3</i>				

4. Содержание дисциплины (модуля)

№п/п	Наименование раздела дисциплины (модуля)	Содержание дисциплины (модуля) по темам
1	Анализ сигнала	Введение в речевые технологии и цифровая обработка сигналов
2	Распознавание речи	Распознавание речи: классические подходы Распознавание речи: предобучение
3	Биометрия	Голосовая биометрия и диаризация
4	Синтез речи	Синтез речи: классические подходы Синтез речи: вокодеры, конверсия и клонирование голоса Синтез речи: токенизация аудио, нейрокодеки, языковое моделирование
5	Большие аудио модели	Большие аудио модели и мультимодальность

5. Учебно-методическое обеспечение

Университет располагает полным набором лицензионного и свободно распространяемого программного обеспечения, включая продукты отечественного производства.

Каждый студент в течение всего периода обучения получает индивидуальный неограниченный доступ к электронно-библиотечной системе и электронной информационно-образовательной среде университета. Эти системы предоставляют возможность доступа к ресурсам из любой точки, где есть подключение к сети Интернет, как на территории университета, так и за его пределами.

Студентам обеспечен удаленный доступ к современным профессиональным базам данных и информационным справочным системам.

Основная литература:

1. Бенгфорт, Б. Прикладной анализ текстовых данных на Python. Машинное обучение и создание приложений обработки естественного языка : практическое руководство / Б. Бенгфорт, Р. Билбро, Т. Охеда. - Санкт-Петербург : Питер, 2020. - 368 с. - (Серия «Бестселлеры O'Reilly»). - ISBN 978-5-4461-1153-4. - Текст : электронный. - URL: <https://znanium.com/catalog/product/1733693>.

2. Николенко, С. И. Глубокое обучение : практическое руководство / С. И. Николенко, А. Кадури, Е. Архангельская. - Санкт-Петербург : Питер, 2019. - 480 с. - (Серия «Библиотека программиста»). - ISBN 978-5-496-02536-2. - Текст : электронный. - URL: <https://znanium.ru/catalog/product/1760785>

Дополнительная литература:

1. Шолле, Ф. Глубокое обучение на Python : практическое руководство / Ф. Шолле. - Санкт-Петербург : Питер, 2023. - 576 с. - (Библиотека программиста). - ISBN 978-5-4461-1909-7. - Текст : электронный. - URL: <https://znanium.com/catalog/product/2123389>.

6. Материально-техническое обеспечение

Университет располагает материально-технической базой, соответствующей действующим противопожарным правилам и нормам и обеспечивающей проведение всех видов дисциплинарной и междисциплинарной подготовки, практической и научно-исследовательской работ обучающихся, предусмотренных учебным планом.

Помещения, которые представляют собой учебные аудитории для проведения занятий лекционного типа, занятий семинарского (практического) типа, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, а также помещения для самостоятельной работы и помещения для хранения и профилактического обслуживания учебного оборудования. Помещения укомплектованы специализированной мебелью и техническими средствами обучения, служащими для представления учебной информации большой аудитории.

Изучение дисциплины (модуля) обеспечивается в учебных аудиториях, оснащенных:

- столами и стульями;
- компьютерной техникой;
- механическими калькуляторами;
- специализированным оборудованием, включая демонстрационное оборудование.

Помещения для самостоятельной работы обучающихся, в том числе приспособленные для использования инвалидами и лицами с ограниченными возможностями здоровья, оснащены компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду Университета.

Обучающимся предоставляется доступ (в том числе удаленный) к ресурсам информационно-телекоммуникационной сети «Интернет», электронным ресурсам (в том числе электронным библиотечным системам, современным профессиональным базам данных и информационным справочным системам):

№	Наименование портала (издания, курса, документа)	Ссылка
1.	Научная электронная библиотека eLibrary.ru библиотека	https://elibrary.ru/defaultx.asp
2.	База данных для IT-специалистов	https://habr.com
3.	База данных ScienceDirect	https://www.sciencedirect.com
4.	Официальный сайт Министерства науки и высшего образования Российской Федерации	https://minobrnauki.gov.ru/
5.	Федеральный портал «Российское образование»	https://www.edu.ru/
6.	Информационная система "Единое окно доступа к образовательным ресурсам"	http://window.edu.ru/
7.	Единая коллекция цифровых образовательных ресурсов	http://school-collection.edu.ru/
8.	Федеральный центр информационно - образовательных ресурсов	http://fcior.edu.ru/

Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), в том числе комплект лицензионного программного обеспечения, современные профессиональные базы данных и информационные справочные системы:

Наименование ПО	Производство	Лицензионное / свободно распространяемое
Операционные системы:		
Microsoft Imagine (Windows Client, Server)	зарубежное	лицензионное
Браузеры:		
Яндекс.Браузер	отечественное	свободно распространяемое
Google Chrome	зарубежное	свободно распространяемое
Офисные приложения:		
Microsoft Imagine (Visio, OneNote)	зарубежное	лицензионное
TeXstudio	зарубежное	свободно распространяемое
Adobe Acrobat Reader	зарубежное	свободно распространяемое
Программное обеспечение для планирования и учета времени:		
Toggle app	зарубежное	свободно распространяемое
Системы управления проектами:		
Microsoft Imagine (Project)	зарубежное	лицензионное
Системы управления базами данных:		
Microsoft Imagine (SQL Server)	зарубежное	лицензионное
Системы резервного копирования (backup):		
Acronis Backup Advanced for HyperV	зарубежное	лицензионное
Справочно-правовые системы:		
КонсультантПлюс: справочно-правовая система	отечественное	лицензионное
Средства антивирусной защиты:		
Kaspersky Endpoint Security для бизнеса Стандартный Russian Edition	отечественное	лицензионное
Среды разработки:		

Visual Studio Code	зарубежное	свободно распространяемое
Bash (Unix shell)	зарубежное	свободно распространяемое
Anaconda	зарубежное	свободно распространяемое
Robotic Operating System	зарубежное	свободно распространяемое
CopelliaSim	зарубежное	свободно распространяемое
Google Colaboratory	зарубежное	свободно распространяемое
Пакеты программных средств и библиотек:		
AutoPsy	зарубежное	свободно распространяемое
Interactive Disassembler (IDA)	зарубежное	свободно распространяемое
Системы управления библиографической информацией:		
Zotero	зарубежное	свободно распространяемое
Сервисы и службы:		
Bind	зарубежное	свободно распространяемое
Docker	зарубежное	свободно распространяемое

7. Методические и оценочные материалы

Методические указания для обучающихся по освоению дисциплины (модуля)

В процессе изучения дисциплины (модуля) «Речевые технологии» в рамках текущего контроля успеваемости используются такие виды учебной работы, как лекции, семинары, домашние задания, а также различные виды самостоятельной работы обучающихся по заданию преподавателя, направленные на развитие навыков профессиональной лексики, закрепление практических профессиональных компетенций, поощрение инициатив.

Лекция – систематическое, последовательное, монологическое изложение преподавателем учебного материала, как правило, теоретического характера.

В процессе лекций рекомендуется вести конспект лекций: кратко и схематично фиксировать основные идеи, выводы и обобщения лекции; выделять важные мысли, ключевые слова и термины. Необходимо отметить вопросы или материалы, которые вызывают затруднения, и попытаться найти ответы в рекомендованной литературе. Если разобраться в материале не удастся, следует сформулировать вопрос и задать его преподавателю на консультации или во время семинарского (практического) занятия.

Семинар – это форма учебной деятельности, проводимая в учебном заведении под руководством преподавателя, где студенты активно участвуют в обсуждениях, практических заданиях и других формах взаимодействия.

Для успешной подготовки к семинару рекомендуется заранее ознакомиться с темой занятия и основными материалами, чтобы иметь возможность активно участвовать в обсуждении. Также полезно подготовить вопросы и идеи для обсуждения, что поможет глубже понять материал и продемонстрировать заинтересованность.

Домашнее задание – набор задач по темам недели.

При работе над домашними заданиями важно внимательно ознакомиться с требованиями и сроками выполнения. Рекомендуется разбивать задания на этапы, чтобы избежать перегрузки и лучше усвоить материал. Использовать различные источники информации, включая учебники и онлайн-ресурсы, для более глубокого понимания темы.

Бонусные баллы — это оценки, которые студенты могут получить за выполнение дополнительных заданий.

Формат бонусных баллов позволяет студентам улучшить общую оценку по дисциплине (модулю) и стимулирует углубленное изучение материала.

Самостоятельная работа – работа студентов, направленная на углубленное

изучение отдельных тем и вопросов учебной дисциплины (модуля).

В процессе самостоятельной работы студенты взаимодействуют с рекомендованными материалами при минимальном участии преподавателя. Задачи студента включают работу с конспектами лекций (обработка текста), повторное изучение учебных материалов планов и тезисов ответов, изучение дополнительных тем, выполнение учебно-исследовательских заданий и другое.

Система оценивания результатов обучения по дисциплине (модулю)

Критерии получения уровня и оценивания сформированности компетенций по дисциплине (модулю) «Речевые технологии»

Оценивание уровня учебных достижений обучающихся по дисциплине (модулю) осуществляется в виде текущего контроля успеваемости и промежуточной аттестации.

Промежуточная аттестация по дисциплине (модулю) осуществляется в форме *зачета*, при этом проводится оценка компетенций, сформированных по дисциплине.

Для оценивания текущего контроля успеваемости и промежуточной аттестации используется десятибалльная шкала оценивания, которая соотносится с традиционной пятибалльной шкалой следующим образом:

Десятибалльная оценка	Пятибалльная оценка	Оценка за зачет	Общая характеристика результата обучения по дисциплине (модулю)
10	Отлично	Зачтено	Студент полностью владеет знаниями, изложенными в рабочей программе, и глубоко осмысляет дисциплину (модуль). Он самостоятельно и логически последовательно отвечает на все вопросы, акцентируя внимание на наиболее важном. Умеет анализировать, сравнивать, классифицировать, обобщать, конкретизировать и систематизировать изученный материал, выделяя ключевые моменты и устанавливая причинно-следственные связи. Четко формулирует ответы, уверенно интерпретирует результаты анализов и других исследований, а также решает сложные задачи. Студент хорошо знаком с методами исследования, необходимыми для практической деятельности, и умеет связывать теоретические аспекты дисциплины (модуля) с практическими задачами.
9	Отлично	Зачтено	
8	Отлично	Зачтено	
7	Хорошо	Зачтено	Студент обладает знаниями предмета почти в полном объеме рабочей программы и самостоятельно, логически последовательно и всесторонне отвечает на все вопросы, акцентируя внимание на наиболее значимых моментах. Он умеет анализировать, сравнивать, классифицировать, обобщать,
6	Хорошо	Зачтено	

Десятибалльная оценка	Пятибалльная оценка	Оценка за зачет	Общая характеристика результата обучения по дисциплине (модулю)
			конкретизировать и систематизировать изученный материал, выделяя его ключевые аспекты и устанавливая причинно-следственные связи. Формулирует свои ответы, уверенно интерпретирует результаты анализов и других исследований, а также решает сложные ситуационные задачи. Студент хорошо знаком с методами исследования, необходимыми для практической деятельности, и умеет связывать теоретические аспекты предмета с практическими задачами.
5	Удовлетворительно	Зачтено	Студент обладает базовыми знаниями по дисциплине (модулю), но испытывает трудности при самостоятельных ответах и использует неточные формулировки. В ходе ответов он допускает ошибки, касающиеся сути вопросов. Студент способен решать только самые простые задачи и владеет лишь минимальным набором методов исследования.
4	Удовлетворительно	Зачтено	
3	Не сдан	Не зачтено	Студент не овладел обязательным минимумом знаний по предмету и не может ответить на вопросы, даже если преподаватель задает дополнительные наводящие вопросы.
2	Не сдан	Не зачтено	
1	Не сдан	Не зачтено	

Дисциплина (модуль) «Речевые технологии» оценивается следующим образом:

Активность	Вес	Описание
Домашние задания	80%	Набор задач по темам недели
Зачет	20%	Письменная или устная работа над заданием, направленным на проверку полученных знаний и навыков по дисциплине (модулю)

В рамках изучения дисциплины (модуля) возможно получение бонусных баллов.

Формула расчёта итоговой оценки по дисциплине (модулю) «Речевые технологии»: $\langle 0,8 \times \text{среднее за домашние задания} + 0,2 \times \text{зачет} \rangle$.

Текущий контроль успеваемости обучающихся по дисциплине (модулю)

Примерные домашние задания

Домашнее задание

Реализуйте и сравните систему распознавания речи на основе классического подхода (MFCC + CTC/RNN или HMM-GMM) и предобученной self-supervised модели (например,

Wav2Vec2 или Whisper) на одном и том же подмножестве датасета, рассчитав WER и проанализировав различия в качестве и типах ошибок.

Домашнее задание

Постройте систему голосовой биометрии с извлечением эмбедингов диктора (например, ECAPA-TDNN), реализуйте проверку говорящего по косинусному сходству с расчетом EER и дополнительно выполните диаризацию многоголосовой записи, проанализировав точность сегментации.

Домашнее задание

Реализуйте пайплайн нейросетевого синтеза речи (текст → акустическая модель → нейровокодер), дополнительно исследуйте использование нейрокодека для дискретного представления аудио и сравните качество генерации с классическим подходом по спектрограммам и субъективной оценке.

Домашнее задание

Загрузите речевой аудиофайл, выполните его цифровое представление и базовую обработку (ресемплинг, нормализация, оконное преобразование Фурье), постройте спектрограмму и mel-спектрограмму, извлеките MFCC и проанализируйте, как выбор параметров дискретизации и окна влияет на временно-частотное представление сигнала.

Домашнее задание

Выберите предобученную большую аудио или мультимодальную модель (например, Whisper, AudioLM или CLAP), выполните инференс на примерах аудио и текста, проанализируйте архитектуру и механизм обучения модели, а также продемонстрируйте адаптацию (fine-tuning или prompting) для конкретной прикладной задачи с оценкой качества результата.

Задания для промежуточной аттестации по дисциплине (модулю)

№ п/п	Задание	Ответ	Компетенция
1.	Если Вы регулярно улучшаете модель через повторные циклы обучения и анализа ошибок, как одним словом называется такой подход?	итеративность	УК-6
2.	Если Вы анализируете причины ошибок модели распознавания после эксперимента, как одним словом называется такой процесс?	рефлексия	УК-6
3.	Если Вы определяете порядок изучения архитектур LSTM и Transformer для работы с речью, как одним словом называется такой процесс?	приоритизация	УК-6
4.	Если Вы распределяете время между реализацией ASR и TTS-моделей при подготовке проекта, как одним словом называется такой процесс?	планирование	УК-6
5.	Если Вы оцениваете уровень своего понимания научной статьи по диффузионным вокодерам, как одним словом называется такой анализ?	самооценка	УК-6
6.	Если речевой сигнал представляется в виде дискретных отсчетов через равные интервалы времени, как одним словом называется такой процесс цифровой обработки?	дискретизация	ОПК-2
7.	Если для анализа спектра речевого сигнала применяется кратковременное преобразование Фурье,	STFT	ОПК-2

	как сокращенно называется этот метод?		
8.	Если акустические признаки вычисляются на основе мел-шкалы и логарифма энергии спектра, как одним словом называются такие коэффициенты?	MFCC	ОПК-2
9.	Если модель распознавания речи обучается предсказывать следующий токен по предыдущим, как одним словом называется такой тип моделирования?	авторегрессия	ОПК-2
10.	Если в синтезе речи используется модель, восстанавливающая сигнал из спектрограммы, как одним словом называется такой компонент?	вокодер	ОПК-2
11.	Если доля неправильно распознанных слов в системе ASR рассчитывается по специальной формуле, как называется такая метрика одним словом?	WER	ПК-3
12.	Если при голосовой биометрии определяется вероятность равенства ложных допусков и отказов, как называется такая точка одним словом?	EER	ПК-3
13.	Если качество синтезированной речи оценивается пользователями по шкале субъективного восприятия, как называется такая оценка одним словом?	MOS	ПК-3
14.	Если в многоголосовой записи определяется, кто говорит в каждый момент времени, как одним словом называется такой процесс?	диаризация	ПК-3
15.	Если после внедрения новой модели ASR наблюдается снижение WER, как одним словом называется решение о расширении её использования в продукте?	масштабирование	ПК-3
16.	Если Вы представляете архитектуру нейросетевой модели синтеза речи с помощью блок-схемы, как одним словом называется такой способ представления?	визуализация	ПК-4
17.	Если Вы кратко формулируете основные результаты эксперимента по распознаванию речи в начале выступления, как одним словом называется такая часть доклада?	вывод	ПК-4
18.	Если Вы сопровождаете доклад сравнением моделей по метрикам качества, как одним словом называется такой формат представления?	обзор	ПК-4
19.	Если результаты исследования по синтезу речи оформляются в структурированном документе с графиками и таблицами, как одним словом называется такой документ?	отчет / отчёт	ПК-4
20.	Если при публичной защите проекта Вы отвечаете на вопросы о переобучении модели, как одним словом называется такой этап?	обсуждение	ПК-4