

АННОТАЦИЯ рабочей программы дисциплины

Б1.В.09 Технологии обработки языка, звуковых данных, включая распознавание и синтез речи

Объем трудоемкости: 4 з.е.

Цели дисциплины:

Формирование у студентов системных знаний, практических навыков и компетенций в области обработки естественного языка и звуковых данных, включая распознавание и синтез речи – ключевых направлений современного искусственного интеллекта. Освоение дисциплины направлено на подготовку специалистов, способных разрабатывать, обучать и интегрировать модели обработки речевых и текстовых сигналов в прикладные интеллектуальные системы, что необходимо для профессиональной деятельности аналитика данных, AI-инженера, разработчика машинного обучения и инженера по речевым технологиям.

Задачи дисциплины: в соответствии с поставленной целью состоят в следующем:

- Ознакомить студентов с архитектурой и принципами работы современных систем обработки естественного языка и звуковых данных, включая модели автоматического распознавания речи (ASR), синтеза речи (TTS), и языковые модели (LLM).
- Изучить методы представления, анализа и предобработки звуковых сигналов и текстовых данных, включая извлечение признаков (MFCC, spectrogram, embeddings) и нормализацию текстов.
- Научить использовать современные инструменты и библиотеки (SpeechRecognition, torchaudio, librosa, Hugging Face Transformers, FastSpeech и др.) для решения прикладных задач распознавания и генерации речи.
- Развить навыки построения и интеграции пайплайнов для обработки речевых и текстовых данных – от сбора и подготовки данных до обучения, оценки качества и внедрения моделей.
- Подготовить студентов к проектированию и анализу систем речевого взаимодействия (чат-боты, голосовые ассистенты, системы субтитрования и озвучивания).
- Показать профессиональные кейсы применения технологий распознавания и синтеза речи в индустрии (автоматизация колл-центров, доступность контента, медиасервисы, аналитика звонков и др.).

Место дисциплины в структуре ООП ВО

Дисциплина «Технологии обработки языка, звуковых данных, включая распознавание и синтез речи» относится к части, формируемой участниками образовательных отношений Блока 1 «Дисциплины (модули)» учебного плана. Для успешного освоения данной дисциплины необходимы знания следующих дисциплин: Математический анализ, Векторная алгебра, Основы программирования, Обработка данных на Python, Математические модели нейронных сетей, Нейросетевые технологии, Аналитика данных, A/B-тестирование и Uplift-моделирование.

Требования к уровню освоения дисциплины

Изучение данной учебной дисциплины направлено на формирование у обучающихся следующих компетенций:

Код, уровень и формулировка компетенции	Индикаторы	Уровни освоения индикаторов компетенции
DL-4 П Способен применять и (или) разрабатывать алгоритмы, методы и технологии	DL-4.1 Применяет (проводя выбор и эксперименты) известные алгоритмы и библиотеки для	Выбирает и обосновывает использование современных моделей обработки текста (BERT, RoBERTa, GPT и др.) под конкретную задачу; проводит

обработки естественного языка	обработки естественного языка, предобученные глубокие нейросетевые модели для прикладных задач анализа текстов, при необходимости дообучая и валидируя на собственных наборах данных	дообучение и оценку качества; анализирует результаты с точки зрения архитектуры, набора данных и применимости модели в реальных условиях.
FC-2 Б Способен проводить фронтальные исследования в области фундаментальных и генеративных моделей	FC-2.1 Исследует и разрабатывает большие языковые модели (LLM) и другие модели для символьных данных	Проводит эксперименты с моделями, выдвигает гипотезы
LLM-1 Б Способен применять и (или) разрабатывать генеративные модели и БЯМ	LLM-1.1 Знает архитектуры генеративных моделей	Описывает принципы и отличия архитектур современных генеративных моделей (Transformer, seq2seq, Diffusion, GAN, VAE, Tacotron, Whisper и др.). Сопоставляет их применимость для задач синтеза и распознавания речи, генерации текста, мультимодальных сценариев; обосновывает выбор архитектуры под конкретную задачу.

Содержание дисциплины:

Распределение видов учебной работы и их трудоемкости по разделам дисциплины.

№	Наименование разделов (тем)	Количество часов				
		Всего	Аудиторная работа			Внеаудиторная работа
			Л	ПЗ	ЛР	
1	2	3	4	5	6	7
1.	Введение в технологии обработки языка и звуковых данных	8	2		2	4
2.	Основы цифровой обработки звуковых сигналов	12	4		4	4
3.	Предобработка и анализ речевых и текстовых данных	12	4		4	4
4.	Архитектуры и принципы работы моделей распознавания речи (ASR)	12	4		4	4
5.	Модели синтеза речи (TTS) и генерация звуковых сигналов	12	4		4	4
6.	Языковые модели и интеграция с речевыми системами	12	4		4	4
7.	Методы оценки качества распознавания и синтеза речи	12	4		4	4
8.	Интеграция речевых технологий в прикладные системы	12	4		4	4
9.	Этические и социальные аспекты применения речевых технологий	12	4		4	4
ИТОГО по разделам дисциплины		104	34		34	36
Контроль самостоятельной работы (КСР)		4				
Промежуточная аттестация (ИКР)		0,3				
Подготовка к экзамену		35,7				
Общая трудоемкость по дисциплине		144				

Курсовые работы: не предусмотрены

Форма проведения аттестации по дисциплине: экзамен.

Автор: Романюта Д.Ю., ст. преподаватель.