

АННОТАЦИЯ рабочей программы дисциплины

К.М.01.08 «Технологии обработки языка, звуковых данных, включая распознавание и синтез речи»

Объем трудоемкости: 2 з.е.

Цели изучения дисциплины определены государственным образовательным стандартом высшего образования и соотнесены с общими целями ООП ВО по направлению подготовки 09.03.03 Прикладная информатика, в рамках которой преподается дисциплина.

Цели дисциплины Технологии обработки языка, звуковых данных, включая распознавание и синтез речи:

- познакомить студентов с основами обработки естественного языка (NLP);
- познакомить студентов с основами обработки звуковых данных;
- рассмотреть задачу распознавания и синтеза речи и способы её решения.

Задачи дисциплины в соответствии с поставленной целью состоят в следующем:

- изучить основные архитектуры для решения задач NLP и работы с звуковыми данными;
- изучить большие языковые модели;
- научиться работать с библиотеками Python для обработки естественного языка и звуковых данных: gensim, NLTK, PyMorphu, а также с фреймворком PyTorch.

Место дисциплины в структуре ООП ВО

Дисциплина «Технологии обработки языка, звуковых данных, включая распознавание и синтез речи» относится к вариативной части Блока 1 «Дисциплины (модули)» учебного плана. Для успешного освоения данной дисциплины необходимы знания следующих дисциплин: Б1.О.05 Математический анализ, Б1.О.06 Векторная алгебра, Б1.О.08 Курс теории вероятностей, К.М.01.01 Математические модели нейронных сетей и К.М.01.03 Нейросетевые технологии.

Требования к уровню освоения дисциплины

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

Код и наименование индикатора*	Результаты обучения по дисциплине
ПК-4 Способность использовать знание основных методов искусственного интеллекта в последующей профессиональной деятельности в качестве научных сотрудников, преподавателей образовательных организаций высшего образования, инженеров, технологов	
ИПК-4.3 (40.011 А/02.5 Зн.2) Отечественный и международный опыт в области искусственного интеллекта в прикладных областях	Знает: основные достижения в области обработки естественного языка
	Умеет: применять отечественный и международный опыт по разработке систем искусственного интеллекта
	Владет: необходимыми знаниями отечественного и международного опыта по анализу данных
ИПК-4.10 (40.011 А/02.5 Тд.2) Проведение наблюдений и измерений, составление их описаний и формулировка выводов при проведении исследований в области ИИ в прикладных областях	Знает: как проводить наблюдения и измерения в области обработки естественного языка
	Умеет: формулировать выводы по разработанным на основе ИИ моделям
	Владет: знаниями о методах проведения измерений качества разработанных моделей
ПК-5 Способен выявить естественнонаучную сущность проблем, возникающих в ходе профессиональной деятельности в области моделирования и анализа сложных естественных и искусственных систем	
ИПК-5.3 (40.011 А/02.5 Зн.2) Отечественный и международный опыт в области моделирования и анализа сложных естественных и искусственных систем в прикладных областях	Знает: основные задачи в сфере обработки естественного языка
	Умеет: использовать знания из области обработки естественного языка в практической деятельности
	Владет: необходимыми знаниями библиотек (gensim, NLTK, PyMorphu и т.д.), с помощью которых можно решать задачи из области обработки естественного языка и звуковых данных

Код и наименование индикатора*	Результаты обучения по дисциплине
ИПК-5.10 (40.011 А/02.5 Тд.2) Проведение наблюдений и измерений, составление их описаний и формулировка выводов при проведении исследований в области моделирования и анализа сложных естественных и искусственных систем в прикладных областях	Знает: классические и нейросетевые подходы к решению задач обработки языка и звуковых данных
	Умеет: формулировать выводы по разработанным на основе ИИ моделям
	Владеет: знаниями о методах проведения измерений качества разработанных моделей

Основные разделы дисциплины:

№	Наименование разделов (тем)	Всего	Количество часов			
			Аудиторная работа			Внеаудиторная работа
			Л	ПЗ	ЛР	СРС
1	2	3	4	5	6	7
1.	Обработка естественного языка с помощью классических методов машинного обучения	16	4		10	6
2.	Архитектуры нейронных сетей для работы с речью и звуком	14	4		8	2
3.	Большие языковые модели	24	6		14	6
4.	Синтез речи нейросетевыми методами	11,8	2			5,8
ИТОГО по разделам дисциплины		69,8	16		34	19,8
Контроль самостоятельной работы (КСР)		2				
Промежуточная аттестация (ИКР)		0,3				
Подготовка к экзамену		-				
Общая трудоемкость по дисциплине		72				

Курсовые работы: не предусмотрены.

Форма проведения аттестации по дисциплине: зачет.

Автор: Казаковцева Е.В. –к.ф.-м.н., ст.преподаватель КАДИИ