

Документ подписан простой электронной подписью
Информация о владельце:

ФИО: Исаев Игорь Магомедович

Должность: Проректор по безопасности и общим вопросам

Дата подписания: 25.09.2023 17:31:57

Уникальный программный ключ:

d7a26b9e8ca85e98ac3de2ab454b4659d961f749

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное автономное образовательное учреждение
высшего образования

«Национальный исследовательский технологический университет «МИСиС»

Рабочая программа дисциплины (модуля)

Обработка естественного языка

Закреплена за подразделением

Кафедра инженерной кибернетики

Направление подготовки

01.03.04 ПРИКЛАДНАЯ МАТЕМАТИКА

Профиль

Алгоритмы и методы наукоемкого программного обеспечения

Квалификация **Бакалавр**

Форма обучения **очная**

Общая трудоемкость **4 ЗЕТ**

Часов по учебному плану 144

Формы контроля в семестрах:
экзамен 7

в том числе:

аудиторные занятия 34

самостоятельная работа 74

часов на контроль 36

Распределение часов дисциплины по семестрам

Семестр (<Курс>.<Семестр на курсе>)	7 (4.1)		Итого	
	18			
Неделя	УП	РП	УП	РП
Лекции	17	17	17	17
Практические	17	17	17	17
Итого ауд.	34	34	34	34
Контактная работа	34	34	34	34
Сам. работа	74	74	74	74
Часы на контроль	36	36	36	36
Итого	144	144	144	144

Программу составил(и):

к.т.н., доц., Тригуб Н.А.

Рабочая программа

Обработка естественного языка

Разработана в соответствии с ОС ВО:

Самостоятельно устанавливаемый образовательный стандарт высшего образования - бакалавриат Федеральное государственное автономное образовательное учреждение высшего образования «Национальный исследовательский технологический университет «МИСиС» по направлению подготовки 01.03.04 ПРИКЛАДНАЯ МАТЕМАТИКА (приказ от 02.04.2021 г. № 119 о.в.)

Составлена на основании учебного плана:

01.03.04 ПРИКЛАДНАЯ МАТЕМАТИКА, 01.03.04-БПМ-22.plx Алгоритмы и методы наукоемкого программного обеспечения, утвержденного Ученым советом ФГАОУ ВО НИТУ "МИСиС" в составе соответствующей ОПОП ВО 22.09.2022, протокол № 8-22

Утверждена в составе ОПОП ВО:

01.03.04 ПРИКЛАДНАЯ МАТЕМАТИКА, Алгоритмы и методы наукоемкого программного обеспечения, утвержденной Ученым советом ФГАОУ ВО НИТУ "МИСиС" 22.09.2022, протокол № 8-22

Рабочая программа одобрена на заседании

Кафедра инженерной кибернетики

Протокол от 23.06.2021 г., №11

Руководитель подразделения Ефимов Альберт Рувимович, к.филос.н.

1. ЦЕЛИ ОСВОЕНИЯ

1.1	Изучить методы и подходы из области искусственного интеллекта к анализу и синтезу речевых конструкций на естественном языке с формированием критериев оценивания возможности применения для решения практических задач в том числе после знакомства с современными сервисами и платформами интеллектуальной обработки естественного языка.
-----	--

2. МЕСТО В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

Блок ОП:		Б1.В.ДВ.06
2.1	Требования к предварительной подготовке обучающегося:	
2.1.1	Дискретные и нелинейные системы автоматического управления	
2.1.2	Имитационное моделирование	
2.1.3	Машинное обучение II	
2.1.4	Методы и средства обработки изображений	
2.1.5	Методы оптимизации	
2.1.6	Прикладной статистический анализ	
2.1.7	Программирование роботов I	
2.1.8	Фрактальный анализ	
2.1.9	Математическое моделирование	
2.1.10	Основы теории информации и автоматов	
2.1.11	Основы электротехники и электроники	
2.1.12	Современные технологии разработки мобильных приложений	
2.1.13	Теория систем автоматического управления	
2.1.14	Теория случайных процессов	
2.1.15	Функциональный анализ	
2.1.16	Численные методы	
2.1.17	Алгоритмы дискретной математики	
2.1.18	Математика	
2.1.19	Операционные системы и среды	
2.1.20	Основы теории информации и автоматов	
2.1.21	Разработка клиент-серверных приложений	
2.1.22	Сетевые технологии	
2.1.23	Учебная практика по ознакомлению с технологиями разработки наукоемкого ПО	
2.1.24	Учебная практика по ознакомлению с технологиями разработки робототехнических и киберфизических систем	
2.1.25	Базы данных	
2.1.26	Комбинаторика и теория графов	
2.1.27	Технологии программирования	
2.1.28	Физика	
2.1.29	Инженерная компьютерная графика	
2.1.30	Объектно-ориентированное программирование	
2.1.31	Основы дискретной математики	
2.1.32	Вычислительные машины, сети и системы	
2.1.33	Программирование и алгоритмизация	
2.2	Дисциплины (модули) и практики, для которых освоение данной дисциплины (модуля) необходимо как предшествующее:	
2.2.1	Глубокое обучение	
2.2.2	Искусственный интеллект и мультиагентные системы	
2.2.3	Киберфизические сети	
2.2.4	Параллельные вычисления	
2.2.5	Подготовка к процедуре защиты и защита выпускной квалификационной работы	
2.2.6	Подготовка к процедуре защиты и защита выпускной квалификационной работы	
2.2.7	Преддипломная практика для апробации темы выпускной квалификационной работы	
2.2.8	Преддипломная практика для апробации темы выпускной квалификационной работы	
2.2.9	Системы обеспечения информационной безопасности и блокчейн	

2.2.10	Современные инструменты DevOps
2.2.11	Специальные главы баз данных

3. РЕЗУЛЬТАТЫ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ, СООТНЕСЕННЫЕ С ФОРМИРУЕМЫМИ КОМПЕТЕНЦИЯМИ

ОПК-4: Способен понимать принципы работы современных информационных технологий и использовать их для решения задач профессиональной деятельности, разрабатывать алгоритмы и компьютерные программы, пригодные для практического применения, выбирать и применять методики проектирования и актуальные инструментальные средства разработки

Знать:

ОПК-4-31 Методы ИИ для обработки ЕЯ

ПК-5: Способен использовать и развивать методы искусственного интеллекта для решения трудно-формализуемых задач

Знать:

ПК-5-31 - Современные направления развития Natural Language Processing.

ОПК-1: Способен применять знание фундаментальной математики и естественно-научных дисциплин при решении задач в области естественных наук и инженерной практике

Знать:

ОПК-1-31 Архитектуры нейронных сетей применяемых для обработки ЕЯ

ПК-5: Способен использовать и развивать методы искусственного интеллекта для решения трудно-формализуемых задач

Уметь:

ПК-5-У1 - Определять области применения современных сервисов и библиотек ИИ для решения поставленной практической задачи обработки естественного языка.

ОПК-4: Способен понимать принципы работы современных информационных технологий и использовать их для решения задач профессиональной деятельности, разрабатывать алгоритмы и компьютерные программы, пригодные для практического применения, выбирать и применять методики проектирования и актуальные инструментальные средства разработки

Уметь:

ОПК-4-У1 Формировать новые алгоритмы для интеллектуальной обработки ЕЯ

ОПК-1: Способен применять знание фундаментальной математики и естественно-научных дисциплин при решении задач в области естественных наук и инженерной практике

Уметь:

ОПК-1-У1 Проводить обучений нейронной сети для обработки ЕЯ

Владеть:

ОПК-1-В1 Навыками подготовки датасета для обучения нейронной сети на обработки ЕЯ текстовой информации

ОПК-4: Способен понимать принципы работы современных информационных технологий и использовать их для решения задач профессиональной деятельности, разрабатывать алгоритмы и компьютерные программы, пригодные для практического применения, выбирать и применять методики проектирования и актуальные инструментальные средства разработки

Владеть:

ОПК-4-В1 Математическим аппаратом для оценки применимости в решении практической задачи обработки ЕЯ набора методов ИИ

4. СТРУКТУРА И СОДЕРЖАНИЕ

Код занятия	Наименование разделов и тем /вид занятия/	Семестр / Курс	Часов	Формируемые индикаторы компетенций	Литература и эл. ресурсы	Примечание	КМ	Выполняемые работы
	Раздел 1. Современные тенденции развития и практические разработки в NLP							

1.1	Анализ рынка современных разработок для анализа\синтеза речи на русском языке (групповая работа1) /Ср/	7	18	ОПК-1-В1 ОПК-4-У1 ОПК-4-В1	Л1.1Л2.3 Э3 Э4		КМ3	Р1
1.2	Компании разрабатывающие в внедряющие решения связанные с интеллектуальной разработкой ЕЯ /Лек/	7	2	ОПК-4-31 ОПК-1-31 ПК-5-31	Л1.1Л2.4 Э1 Э2		КМ3	Р1
Раздел 2. Введение в NLP								
2.1	Основные термины и определения NLP /Лек/	7	4	ОПК-4-31 ОПК-4-В1 ОПК-1-У1	Л1.1Л2.5			
Раздел 3. Математический аппарат как базис для методов NLP								
3.1	Архитектуры нейросетей и особенности обучения для NLP /Лек/	7	4	ОПК-1-31	Л2.4Л1.1			
3.2	Обработка звука и вероятностные модели обработки ЕЯ /Лек/	7	3	ОПК-4-31 ОПК-1-31 ПК-5-31	Л2.3Л1.1			
3.3	Решения задачи поиска фрагментов текста по поисковым запросам на русском языке /Пр/	7	6	ПК-5-У1 ОПК-1-В1 ОПК-1-У1 ОПК-4-У1	Л2.5Л2.4			
3.4	Моделирование структуры рекурсивной нейронной сети. (Лабраб1) /Ср/	7	12	ОПК-4-У1 ОПК-4-В1 ОПК-1-31 ОПК-1-У1 ОПК-1-В1	Л1.1Л2.3		КМ2	Р3
3.5	Разработка программного кода для реализации самоорганизующиеся карты Кохонена. (Лабраб2) /Ср/	7	16	ОПК-4-У1	Л2.3Л1.1		КМ1	Р4
3.6	Подготовка корпуса текста для обучение на свободном корпусе текстов. (Лабраб3) /Ср/	7	22	ОПК-1-У1 ОПК-1-В1	Л2.5Л2.4		КМ2	Р5
Раздел 4. Методы и подходы в NLP								
4.1	Байесовские методы машинного обучения. Вероятностные модели /Лек/	7	4	ОПК-4-31	Л2.3Л2.4			
4.2	Решение задачи построения объектной структуры с семантическими связями их формализованного текста /Пр/	7	2	ПК-5-У1 ПК-5-31 ОПК-1-В1	Л2.3Л2.5			
4.3	Сравнительный анализ работы Байесовских и нейронных рекурсивных сетей на эталонном корпусе текстов в задачи категоризации текста. (групповая работа 2) /Ср/	7	6	ОПК-4-В1	Л2.4Л2.3		КМ1	Р2
Раздел 5. Современные сервисы и платформы для ИИ обработки ЕЯ								
5.1	Возможности ML Spase /Пр/	7	2	ПК-5-У1 ПК-5-31 ОПК-4-У1	Л2.5Л1.1		КМ1	Р1

5.2	Категоризация свободного текста на русском языке /Пр/	7	7	ОПК-4-В1 ОПК-1-У1	Л1.Л2.3			
-----	---	---	---	----------------------	---------	--	--	--

5. ФОНД ОЦЕНОЧНЫХ МАТЕРИАЛОВ

5.1. Контрольные мероприятия (контрольная работа, тест, коллоквиум, экзамен и т.п), вопросы для самостоятельной подготовки

Код КМ	Контрольное мероприятие	Проверяемые индикаторы компетенций	Вопросы для подготовки
КМ1	Контрольная работа2. Нейронные сети в интеллектуальной обработке текстов на ЕЯ	ОПК-1-У1;ОПК-1-В1;ОПК-1-З1;ОПК-4-В1	1. Перечислите архитектуры нейронных сетей. 2. Что такое искусственная нейронная сеть 3. Что такое персептрон 4. Покажите архитектуру многослойного искусственного персептрона 5. Покажите архитектуру сверточной нейронной сети 6. Покажите архитектуру рекурсивной нейронной сети 7. Покажите архитектуру самообучающейся нейронной сети
КМ2	Контрольная работа1. Байесовские сети в интеллектуальной обработке на ЕЯ	ОПК-4-З1;ОПК-4-У1;ПК-5-З1	1. Докажите теорема Байеса 2. Вероятностная модель в теории Байеса 3. Покажите на примере применений теоремы Байеса 4. покажите архитектуру Байесовской сети 5. Формализуйте задачу классификации 6. Покажите сходимость байесовского метода классификации 7. Пересилите методы байесовской классификации доступные в модулях Python
КМ3	Тест. ИИ в обработке текстовой информации на ЕЯ	ПК-5-У1;ПК-5-З1	1. Искусственный интеллект это 2. Кто создал основополагающие работы в области искусственного интеллекта -кибернетике 3. Назовите главное "мыслящее" устройство направления исследования в области искусственного интеллекта 4. Какие подходы к определению понятия «искусственный интеллект» существуют 5. Какие системы искусственного интеллекта (СИИ) входят в состав систем, основанных на языках 6. Какие системы искусственного интеллекта (СИИ) входят в состав систем эвристического поиска 7. Какие системы искусственного интеллекта (СИИ) входят в состав систем, основанных на языках 8. Какими характерными особенностями обладают системы искусственного интеллекта 9. Научное направление, связанное с попытками формализовать мышление человека называется ... 10. С каким объектом изучения тесно связаны термины "интеллект" и "информатика"

5.2. Перечень работ, выполняемых по дисциплине (Курсовая работа, Курсовой проект, РГР, Реферат, ЛР, ПР и т.п.)

Код работы	Название работы	Проверяемые индикаторы компетенций	Содержание работы
Р1	Групповая работа 1. Анализ рынка современных разработок для анализа и синтеза реси на ЕЯ	ОПК-4-В1;ПК-5-У1;ПК-5-З1	1. Анализ современного рынка компаний, занимающихся разработкой ИС для обработки информации на русском языке 2. Формирование системы критериев оценки аналогичных разработок в области анализа информации на ЕЯ 3. Сравнительный анализ 3х произвольных разработок на эталонном корпусе тестовой информации на ЕЯ

P2	Групповая работа 2. Сравнительный анализ работы Байесовских и нейронных рекурсивных сетей на эталонном корпусе текстов в задачи категоризации текста.	ОПК-4-31;ОПК-4-В1;ОПК-4-У1;ОПК-1-31;ОПК-1-У1;ОПК-1-В1	1. Анализ особенностей подходов для анализа информации на ЕЯ: на нейронных сетях и байесовских методах 2. Формирование системы критериев оценки аналогичных разработок в области анализа информации на ЕЯ основанных на нейро и Байесовских сетях 3. Сравнительный анализ разработок в области анализа информации на ЕЯ основанных на нейро и Байесовских сетях на эталонном корпусе текстовой информации на ЕЯ
P3	Лабораторная работа 1. Моделирование структуры рекурсивной нейронной сети.	ОПК-1-31;ОПК-1-В1;ОПК-1-У1	1. Исследование архитектур нейронных сетей 2. Анализ особенностей рекурсивной нейронной сети 3. Моделирование много слойной рекурсивной нейронной сети
P4	Лабораторная работа 2. Разработка программного кода для реализации самоорганизующиеся карты Кохонена	ОПК-4-31;ОПК-4-У1;ОПК-4-В1	1. Исследование архитектур самоорганизующийся карты Кохонена 2. Анализ особенностей самоорганизующийся карты Кохонена 3. Моделирование самоорганизующийся карты Кохонена
P5	Лабораторная работа 3. Подготовка корпуса текста для обучение на свободном корпусе текстов	ОПК-1-В1;ОПК-1-У1;ПК-5-У1	1. Постановка задачи классификации текстовой информации на русском языке 2. Сбор корпуса текстов на русском языке 3. Разметка корпуса текстов на русском языке 4. Формирование обучающей, тестовой и контрольной выборок из корпуса текстов на русском языке

5.3. Оценочные материалы, используемые для экзамена (описание билетов, тестов и т.п.)

1. Как называется область информационной технологии, изучающая методы превращения знаний в объект обработки на компьютере
2. В чем состоит главное назначение инженерии знаний ...
3. Как называются знания о конкретной ситуации в форме числовых, текстовых данных или простых утверждений ...
4. Как называются программы для ЭВМ, обладающие компетентностью, символическими рассуждениями, глубиной и самосознанием ...
5. Как называется искусственная система, имитирующая решение человеком сложных задач в процессе его жизнедеятельности ...
6. Укажите разрядность нейропроцессора
7. Укажите основные концепции развития СИИ
8. Что такое нечеткое множество
9. Состав базы знаний
10. Какие операции можно проводить с нечеткими знаниями
11. Укажите основные направления исследований СИИ
12. Безымянная функция для связывания формальных и фактических параметров на время
13. Сколько и какие точки зрения существуют на развитие СИИ
14. Укажите основные элементы нейропроцессора
15. Среди многих точек зрения, которые описывает научную область исследования СИИ перечислите доминирующие три...
16. Целью построения систем "Разработка интеллектуальных информационных систем или систем, основанных на знаниях" является...
17. Целью построения систем "Генерации и распознавание речи" является...
18. Целью построения систем "Разработка естественно-языковых интерфейсов и машинный перевод" является...
19. Целью построения систем "Обработка визуальной информации" является...
20. Целью построения систем "Обучение и самообучение" является...
21. Целью построения систем "Распознавание образов" является...
22. Целью построения систем "Игры и машинное творчество" является...

5.4. Методика оценки освоения дисциплины (модуля, практики. НИР)

Шкала оценивания дисциплины:

70 -80 баллов - удовлетворительно

80-90 баллов - хорошо

90- 100 баллов - отлично

Максимальные баллы по выполняемым мероприятиям.

1. Групповые работы:

Первая - 10 баллов

Вторая - 10 баллов

2. Лабораторные работы:

Первая - 10 баллов

Вторая - 10 баллов

Третья - 10 баллов

3. Контрольная работа:

Первая - 15 баллов

Вторая - 15 баллов

4. Общее тестирование: итого 10 баллов

5. Итоговый контроль: 10 баллов

6. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ**6.1. Рекомендуемая литература****6.1.2. Дополнительная литература**

	Авторы, составители	Заглавие	Библиотека	Издательство, год
Л2.1	Петровичев Е. И.	Введение в искусственные нейронные сети: учеб. пособие по дисц. "Нейротехнологии в управлении"	Библиотека МИСиС	М.: Изд-во МГТУ, 2008
Л2.2	Ясницкий Л. Н.	Введение в искусственный интеллект: учеб. пособие для студ. вузов, обуч. по спец. 010100 "Математика"	Библиотека МИСиС	М.: ACADEMIA, 2005
Л2.3	Алпайдин Э.	Машинное обучение: новый искусственный интеллект: пер. с англ.	Библиотека МИСиС	М.: Альпина Паблицер, 2017
Л2.4	Галушкин А. И.	Нейронные сети: основы теории: монография	Библиотека МИСиС	М.: Горячая линия - Телеком, 2010
Л2.5	Смолин Д. В.	Введение в искусственный интеллект: конспект лекций: курс лекций	Электронная библиотека	Москва: Физматлит, 2007

6.2. Перечень ресурсов информационно-телекоммуникационной сети «Интернет»

Э1	В продуктах компании используются технологии искусственного интеллекта для решения бизнес-задач.	https://www.abbyy.com/blog/libor-transition-why-ai-matters/
Э2	«ЦРТ-инновации» — научно-исследовательская компания, разработчик голосовых и бимодальных биометрических систем	https://www.speechpro.ru/
Э3	Алиса от Яндекс	https://yandex.ru/alice
Э4	Яндекс. Переводчик	https://translate.yandex.ru/collections

6.3 Перечень программного обеспечения

П.1	ТЕХНОЛОГИИ УПРАВЛЕНИЯ ПЕРЕВОДОМ ООО
П.2	Microsoft Visio 2016
П.3	Microsoft Visual Studio 2015
П.4	Microsoft SQL server 2016
П.5	Microsoft Office
П.6	LMS Canvas
П.7	MS Teams
П.8	R Studio
П.9	Python

П.10	PgAdmin III
П.11	Statistica Neural Networks
6.4. Перечень информационных справочных систем и профессиональных баз данных	
И.1	ML Space

7. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ

Ауд.	Назначение	Оснащение
Б-904а	Учебная аудитория:	20 стационарных компьютеров (core i5-3470 8gb RAM), пакет лицензионных программ MS Office, демонстрационное оборудование: доска, проектор мультимедийный, экран, колонки, комплект учебной мебели
Б-907	Учебная аудитория:	1 стационарный компьютер, пакет лицензионных программ MS Office, комплект учебной мебели на 42 посадочных места, демонстрационное оборудование: доска, проектор мультимедийный x 2, экран x 2, колонки
Б-934	Лекционная аудитория: мультимедийная	4 кабины для синхронного перевода, мультимедийные экраны и проектор, ноутбук, пакет лицензионных программ MS Office, комплект учебной мебели на 130 посадочных мест
Б-902	Учебная аудитория:	12 стационарных компьютеров (2 x core i5-3470 8gb RAM, 10 x ryzen5 2400g 32gb RAM), пакет лицензионных программ MS Office, демонстрационное оборудование: доска, проектор мультимедийный, комплект учебной мебели

8. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ

В дисциплине практические занятия проводятся на сервисах машинного обучения сторонних разработчиков. Необходима предварительная регистрация индивидуального аккаунта каждого ученика. А также уделить время перед началом практических занятий освоения графического интерфейса соответствующего сервиса указанного преподавателем.

Все лабораторные работы проводятся в самостоятельно м режиме с выбранной и согласованной с преподавателем задачей и на указанных характеристиках текстовой информации с выбранным естественным языкам