

Документ подписан простой электронной подписью
Информация о владельце:

ФИО: Исаев Игорь Магомедович

Должность: Проректор по учебной и научной работе

Дата подписания: 29.08.2023 14:30:16

Уникальный идентификатор документа:

d7a26b9e8ca85e98ec3de2eb454b4659d061f249

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное автономное образовательное учреждение
высшего образования

«Национальный исследовательский технологический университет «МИСИС»

Рабочая программа дисциплины (модуля)

Практика машинного обучения. Дизайн новых материалов

Закреплена за подразделением Кафедра магистерская школа информационных бизнес систем

Направление подготовки 09.04.02 ИНФОРМАЦИОННЫЕ СИСТЕМЫ И ТЕХНОЛОГИИ

Профиль Экосистема больших данных для цифровой трансформации

Квалификация **Магистр**

Форма обучения **очная**

Общая трудоемкость **5 ЗЕТ**

Часов по учебному плану 180

в том числе:

аудиторные занятия 40

самостоятельная работа 86

часов на контроль 54

Формы контроля в семестрах:
экзамен 3

Распределение часов дисциплины по семестрам

Семестр (<Курс>.<Семестр на курсе>)	3 (2.1)		Итого	
	УП	РП	УП	РП
Неделя	19			
Вид занятий	УП	РП	УП	РП
Лекции	4	4	4	4
Практические	36	36	36	36
В том числе инт.	36		36	
Итого ауд.	40	40	40	40
Контактная работа	40	40	40	40
Сам. работа	86	86	86	86
Часы на контроль	54	54	54	54
Итого	180	180	180	180

Программу составил(и):
ассистент, Сиднов К.

Рабочая программа

Практика машинного обучения. Дизайн новых материалов

Разработана в соответствии с ОС ВО:

Самостоятельно устанавливаемый образовательный стандарт высшего образования - магистратура Федеральное государственное автономное образовательное учреждение высшего образования «Национальный исследовательский технологический университет «МИСиС» по направлению подготовки 09.04.02 ИНФОРМАЦИОННЫЕ СИСТЕМЫ И ТЕХНОЛОГИИ (приказ от 05.03.2022 г. № 95 о.в.)

Составлена на основании учебного плана:

09.04.02 ИНФОРМАЦИОННЫЕ СИСТЕМЫ И ТЕХНОЛОГИИ, 09.04.02-МИСТ-23-2.plx Экосистема больших данных для цифровой трансформации, утвержденного Ученым советом НИТУ МИСИС в составе соответствующей ОПОП ВО 22.06.2023, протокол № 5-23

Утверждена в составе ОПОП ВО:

09.04.02 ИНФОРМАЦИОННЫЕ СИСТЕМЫ И ТЕХНОЛОГИИ, Экосистема больших данных для цифровой трансформации, утвержденной Ученым советом НИТУ МИСИС 22.06.2023, протокол № 5-23

Рабочая программа одобрена на заседании

Кафедра магистерская школа информационных бизнес систем

Протокол от 24.06.2022 г., №10

Руководитель подразделения М.И. Нежурина

1. ЦЕЛИ ОСВОЕНИЯ

1.1	Цель освоения дисциплины обучающимся заключается в ознакомлении с принципами построения моделей машинного обучения на основе результатов моделирования материалов из первых принципов. Формирование профессиональной компетенции в прикладной области подразумевает ознакомление обучающихся с основами работы с базами данных результатов моделирования материалов в рамках теории функционала электронной плотности для построения моделей машинного обучения, а также обработкой и анализом результатов применения моделей машинного обучения для дизайна новых материалов.
-----	--

2. МЕСТО В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

Блок ОП:		Б1.В.ДВ.03
2.1	Требования к предварительной подготовке обучающегося:	
2.1.1	Интеллектуальный анализ данных	
2.1.2	Миграция данных и расширенный SQL	
2.1.3	Научно-исследовательская работа. Учебный проект	
2.1.4	Программные платформы и технологии больших данных	
2.1.5	СУБД.Продвинутый уровень SQL	
2.1.6	Языки программирования для работы с большими данными	
2.1.7	Информационные технологии в офисной деятельности	
2.1.8	Новые направления и технологии современных СУБД	
2.1.9	Специальные главы математики. Часть 1	
2.2	Дисциплины (модули) и практики, для которых освоение данной дисциплины (модуля) необходимо как предшествующее:	
2.2.1	Научно-исследовательская работа	
2.2.2	Подготовка к процедуре защиты и защита выпускной квалификационной работы	
2.2.3	Производственная проектная практика	

3. РЕЗУЛЬТАТЫ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ, СООТНЕСЕННЫЕ С ФОРМИРУЕМЫМИ КОМПЕТЕНЦИЯМИ

ПК-4: Способен выполнять работы по сбору, обработке и анализу больших данных в междисциплинарных областях	
Знать:	
ПК-4-31	Способы формализации признакового описания объектов в предметной области
ОПК-2: Способен проектировать и разрабатывать оригинальные алгоритмы и программные средства, в том числе с использованием современных интеллектуальных технологий, для решения профессиональных задач в условиях неопределенности и альтернативных решений в рамках междисциплинарных областей	
Знать:	
ОПК-2-32	Основные типы алгоритмов машинного обучения и особенности их применения в предметной области
ОПК-2-31	Основные постановки современных задач машинного обучения для дизайна новых материалов
ПК-4: Способен выполнять работы по сбору, обработке и анализу больших данных в междисциплинарных областях	
Уметь:	
ПК-4-У1	Решать задачи кластеризации, регрессии, прогнозирования, снижения размерности и ранжирования данных
ОПК-2: Способен проектировать и разрабатывать оригинальные алгоритмы и программные средства, в том числе с использованием современных интеллектуальных технологий, для решения профессиональных задач в условиях неопределенности и альтернативных решений в рамках междисциплинарных областей	
Уметь:	
ОПК-2-У1	Использовать современные свободно распространяемые программные средства для визуализации, интерпретации и анализа баз данных в предметной области

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. Практика машинного обучения для дизайна новых материалов							
1.1	Признаковое описание объектов в предметной области; работа с базами данных материалов; визуализация и интерпретация больших данных; основные типы алгоритмов машинного обучения, их особенности и принципы работы /Лек/	3	2	ОПК-2-31	Л1.1Л3.1 Э1 Э2	Доп. материал: конспект лекций		
1.2	Выбор ключевых признаков для прогнозирования упругих свойств бинарных и тройных сплавов на основе алюминия; визуализация и интерпретация результатов моделирования материалов из первых принципов на основе базы данных "Materials Project" /Пр/	3	18	ОПК-2-31 ОПК-2-У1	Л2.1 Л2.2 Э1 Э2	Доп.материал: Сборник методов в формате ".ipynb". совместимый с открытой веб-платформой "Google Colab"		
1.3	Выбор термодинамических параметров для использования в качестве ключевых признаков при построении моделей машинного обучения для прогнозирования упругих свойств, выполнение домашнего задания №1 /Ср/	3	25	ОПК-2-У1	Л3.1 Э1			
1.4	Анализ методов машинного обучения для прогнозирования упругих свойств бинарных и тройных соединений /Пр/	3	9	ОПК-2-У1	Э2 Э3	Доп.материал: Сборник методов в формате ".ipynb". совместимый с открытой веб-платформой "Google Colab"		
1.5	Выбор оптимальной модели машинного обучения для прогнозирования упругих свойств бинарных и тройных соединений, выполнение домашнего задания №2, выполнение итогового проекта /Ср/	3	25	ОПК-2-У1 ПК -4-У1	Э3 Э4			
1.6	Создание моделей машинного обучения, анализ и интерпретация результатов машинного обучения для дизайна новых материалов /Лек/	3	2	ОПК-2-32	Л1.2 Э3 Э4	Доп. материал: конспект лекций		

1.7	Создание модели машинного обучения для прогнозирования упругих свойств бинарных и тройных соединений /Пр/	3	9	ПК-4-У1	Э2	Доп.материал: Сборник методов в формате ".ipynb". совместимый с открытой веб-платформой "Google Colab"		
1.8	Классификация и анализ результатов машинного обучения для дизайна новых материалов, выполнение домашнего задания №3, выполнение итогового проекта /Ср/	3	36	ОПК-2-У1 ПК-4-У1	Л1.3 Э1 Э2 Э3 Э4			

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	Защита итогового проекта	ОПК-2-31;ОПК-2-У1;ПК-4-31;ПК-4-У1	<p>1. Выбор параметров и результатов первопринципных расчетов для построения моделей машинного обучения с целью предсказания свойств материалов</p> <p>2. Определение целевых свойств материалов, полученных в результате первопринципных расчетов для прогнозирования методами машинного обучения</p> <p>3. Особенности обработки, анализа и визуализации содержания баз данных результатов первопринципных расчётов</p> <p>4. Применение существующих систем прогнозирования свойств материалов, построенных с использованием машинного обучения</p> <p>5. Основы применения "классических" дескрипторов свойств материалов для построения моделей машинного обучения на основе баз данных результатов первопринципных расчетов</p> <p>6. Основы разработки дескрипторов свойств материалов для построения моделей машинного обучения на основе баз данных результатов первопринципных расчетов</p> <p>7. Основные виды и особенности применения моделей машинного обучения для разработки и моделирования материалов с применением результатов расчётов в рамках теории функционала электронной плотности и классической молекулярной динамики</p>

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
P1	Работа №1	ОПК-2-31;ПК-4-31	Признаковое описание объектов в предметной области
P2	Работа №2	ОПК-2-31;ОПК-2-У1	Работа с базами данных материалов: визуализация и интерпретация больших данных
P3	Работа №3	ПК-4-У1;ОПК-2-32	Оценка точности моделей для предсказания свойств сплавов

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

Промежуточная аттестация по дисциплине проводится в форме защиты группового проекта.

Задания для групповых проектов включают теоретическую и практическую составляющие из списка вопросов для подготовки к промежуточной аттестации.

Каждая тематика итогового проекта отводится на выполнение подгруппой, состоящей из не более, чем пяти человек. Темы итоговых проектов выдаются не менее, чем за две недели до планируемой даты представления результатов их выполнения.

Пример задания для итогового проекта:

На основании пройденного материала и рассмотренных практических примеров, представить реализацию алгоритма обработки и анализа базы данных, содержащей результаты молекулярно-динамического моделирования заданной системы. Основной задачей необходимого алгоритма является интеллектуальная обработка описания состояний заданной системы в изменяющихся внешних условиях.

Функционал имплементированного алгоритма должен включать визуализирующие составляющие, отражающие работу его аналитической части:

- Возможность визуализации структур, полученных усреднением координат определенного пользователем диапазона шагов молекулярно-динамического моделирования;
- Возможность визуализации заданных проекций траекторий движения атомов на любую плоскость, образованную парой векторов трансляции;

После подготовки реализации алгоритма, варьируя исследуемые участки процесса, и используя файл с известной температурой процесса ($T = 1500 \text{ K}$), попробуйте предположить при какой температуре происходил процесс, в заданном по заданию файле для анализа. Данный ответ необходимо объяснить (используйте для этого также справочную информацию о заданной системе).

При помощи имплементированного алгоритма и на основании найденной информации, полученного вывода, визуализаций начальной и конечной структур определите, какой процесс протекал при моделировании:

- стационарный процесс;
- фазовый переход (между кристаллическими фазами);
- плавление;

По итогу работы необходимо подготовить презентацию демонстрирующую решение указанных задач.

Презентация должна включать следующие основные тезисы:

- 1) Принцип работы предложенного алгоритма обработки траекторий, с учетом случаев пересечения периодических границ описанной ячейки;
- 2) Визуализацию начальной, усредненной (по всем шагам) и конечной структуры (по последним 500-1000 шагам);
- 3) Графики проекций траекторий на одну из плоскостей (кривые), образованных векторами трансляции, а также проекции координат атомов нулевого шага (точки) и последнего шага молекулярной динамики (точки);
- 4) Найденную температуру (диапазон температур) процесса и обоснование;
- 5) Определенный тип протекающего процесса и обоснование.

Доклад планируется участниками подгруппы таким образом, чтобы каждый участник осветил какую-либо часть выполненного задания. После ответа каждого участника подгруппы, ему будет задан ряд дополнительных вопросов, относящихся к освещаемой им теме.

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

Промежуточная аттестация по дисциплине предусмотрена в форме защиты итогового проекта.

Для прохождения аттестации необходимо выполнение следующих условий:

1. Выполнение всех предусмотренных по дисциплине домашних заданий (балльная система оценивания).
2. Защита итогового группового проекта (балльная система оценивания).

Система оценивания видов учебной деятельности:

1. Методика оценки домашнего задания №1: "Признаковое описание объектов в предметной области"

Максимальное число баллов – 15 баллов.

2. Методика оценки домашнего задания №2: на тему "Работа с базами данных материалов: визуализация и интерпретация больших данных"

Максимальное число баллов – 15 баллов.

3. Методика оценки домашнего задания №3: на тему "Оценка точности моделей для предсказания свойств сплавов"

Максимальное число баллов – 20 баллов.

4. Методика оценки защиты группового проекта:

Максимальная итоговая индивидуальная оценка - 50 баллов.

Доклад по итогам выполнения группового проекта планируется таким образом, чтобы каждый участник осветил какой-либо из вопросов, предложенных в тексте задания. После ответа каждого участника подгруппы, ему будет задан ряд дополнительных вопросов, относящихся к освещаемой им теме. В случае, если равномерное распределение вопросов, среди участников подгруппы по каким-либо причинам невозможно, то оставшимся участникам будут заданы дополнительные вопросы эквивалентной сложности. Если, по какой-либо причине, участник подгруппы не смог ответить на заданный основной или дополнительный вопрос, остальные участники могут ответить за него, повысив тем самым собственный вклад в "общую оценку" групповой работы (см. ниже), но согласно методике расчета итоговых индивидуальных оценок (ниже), это снизит итоговые индивидуальные оценки каждого участника группы.

По результатам защиты итогового проекта, каждый участник подгруппы получит собственную "промежуточную индивидуальную оценку" с максимальным значением, вычисляемым следующим образом:

$$\text{ПРОМЕЖУТОЧНАЯ ИНДИВИДУАЛЬНАЯ ОЦЕНКА} = 100 / [\text{КОЛИЧЕСТВО УЧАСТНИКОВ ПОДГРУППЫ}].$$

Сумма всех промежуточных оценок составит "ОБЩУЮ ОЦЕНКУ" групповой работы.

Итоговая индивидуальная оценка определяется следующим образом:

$$\text{ИНДИВИДУАЛЬНАЯ ОЦЕНКА} = (50 * [\text{ОБЩАЯ ОЦЕНКА}] / 100) * 0.5 + (50 * [\text{ПРОМЕЖУТОЧНАЯ ИНДИВИДУАЛЬНАЯ ОЦЕНКА}] / (100 / [\text{КОЛИЧЕСТВО УЧАСТНИКОВ ПОДГРУППЫ}])) * 0.5$$

Если кем-либо из участников подгруппы не был внесен какой-либо индивидуальный вклад в групповую работу без веской причины, то оцениваемый вклад данного участника может быть исключен из расчета общей оценки, а исключенный участник подгруппы получит 0 баллов за защиту итогового проекта.

Оценкой по результатам итогового контроля считается "индивидуальная оценка".

По итогам контроля знаний по сумме набранных баллов студенту выставляется оценка:

- 2 - «неудовлетворительно» – от 0 до 60 баллов;
- 3 - «удовлетворительно» – от 61 до 69 баллов;
- 4 - «хорошо» – от 70 до 84 баллов;
- 5 - «отлично» – от 85 до 100 баллов.

Максимальное число баллов по дисциплине – 100 баллов.

6. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ

6.1. Рекомендуемая литература

6.1.1. Основная литература

	Авторы, составители	Заглавие	Библиотека	Издательство, год
Л1.1	Киттель Ч.	Введение в физику твердого тела	Электронная библиотека	Москва: Наука, 1978
Л1.2	Алпайдин Э.	Машинное обучение: новый искусственный интеллект: пер. с англ.	Библиотека МИСиС	М.: Альпина Паблицер, 2017
Л1.3	Ландау Л. Д., Лифшиц Е. М.	Т. 7: Теория упругости	Библиотека МИСиС	, 1965

6.1.2. Дополнительная литература

	Авторы, составители	Заглавие	Библиотека	Издательство, год
--	---------------------	----------	------------	-------------------

	Авторы, составители	Заглавие	Библиотека	Издательство, год
Л2.1	Шелудько В. М.	Основы программирования на языке высокого уровня Python: учебное пособие	Электронная библиотека	Ростов-на-Дону, Таганрог: Южный федеральный университет, 2017
Л2.2	Шелудько В. М.	Язык программирования высокого уровня Python: функции, структуры данных, дополнительные модули: учебное пособие	Электронная библиотека	Ростов-на-Дону, Таганрог: Южный федеральный университет, 2017

6.1.3. Методические разработки

	Авторы, составители	Заглавие	Библиотека	Издательство, год
Л3.1	Трахтенброт Б. А.	Алгоритмы и машинное решение задач: научно-популярное издание	Электронная библиотека	Москва: Государственное издательство технико-теоретической литературы, 1957

6.2. Перечень ресурсов информационно-телекоммуникационной сети «Интернет»

Э1	База данных результатов моделирования в рамках теории электронной плотности	https://materialsproject.org/
Э2	Google Colab	https://colab.research.google.com/
Э3	Обзорная статья о применении машинного обучения для дизайна новых материалов (Nature, 2019)	https://www.nature.com/articles/s41524-019-0221-0#:~:text=Machine%20learning%20in%20materials%20science%20is%20mostly%20concerned%20with%20supervised,major%20challenges%20in%20material%20informatics.&text=This%20allows%20the%20application%20of%20the%20data%20for%20new%20purposes.
Э4	Общая информация о применении машинного обучения в моделировании и разработке новых материалов (InfoMat, 2019)	https://www.researchgate.net/publication/335717921_Machine_learning_in_materials_science

6.3 Перечень программного обеспечения

П.1	ОС Linux (Ubuntu) / Windows
П.2	Microsoft Office
П.3	LMS Canvas

6.4. Перечень информационных справочных систем и профессиональных баз данных

И.1	https://materialsproject.org/
-----	---

7. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ

Ауд.	Назначение	Оснащение
М-105	Компьютерный класс	рабочее место преподавателя, оборудованное персональным компьютером, пакет лицензионных программ MS Office; проектор; экран; маркерная доска; комплект учебной мебели
М-105	Компьютерный класс	рабочее место преподавателя, оборудованное персональным компьютером, пакет лицензионных программ MS Office; проектор; экран; маркерная доска; комплект учебной мебели
М-102	Учебная аудитория	Комплект учебной мебели на 15 рабочих мест, ноутбуки с подключением к сети «Интернет» и доступом в электронную информационно-образовательную среду университета

8. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ

Большая часть работы по освоению программы дисциплины отводится для самостоятельной работы. Качественное освоение дисциплины возможно только при систематической самостоятельной работе.

Работа над конспектом лекции.

Основу теоретического обучения студентов составляют лекции. Они дают систематизированные знания студентам о наиболее сложных и актуальных проблемах изучаемой дисциплины. На лекциях особое внимание уделяется усвоению студентами изучаемых проблем, развитию их профессионально-значимых свойств и качеств. Лекции по учебной дисциплине проводятся в форме диалога, с использованием подготовленного материала – презентации. Работу над конспектом следует начинать с его доработки, желательно в тот же день, пока материал еще легко воспроизводим в памяти. С целью доработки необходимо прочитать записи, восстановить текст в памяти, а также исправить описки, расшифровать неприятные ранее сокращения, заполнить пропущенные места, понять текст, вникнуть в его смысл. Далее прочитать

материал по рекомендуемой литературе, разрешая в ходе чтения возникшие ранее затруднения, вопросы, а также дополняя и исправляя свои записи. Записи должны быть наглядными, для чего следует применять различные способы выделений. В ходе доработки конспекта углубляются, расширяются и закрепляются знания, а также дополняется, исправляется и совершенствуется конспект. Подготовленный конспект и рекомендуемая литература используются при подготовке к практическим занятиям. Подготовка сводится к внимательному прочтению учебного материала, к выводу с карандашом в руках всех утверждений и формул, к решению примеров, задач, к ответам на вопросы. Примеры, задачи, вопросы по теме являются средством самоконтроля.

Подготовка к практическому занятию.

Для успешного освоения материала студентам рекомендуется сначала ознакомиться с учебным материалом, изложенным в лекциях и основной литературе, затем выполнить самостоятельные задания, при необходимости обращаясь к дополнительной литературе. При подготовке к практическим занятиям можно выделить два этапа:

- организационный, на котором студент планирует свою самостоятельную работу,
- рабочий, на котором осуществляется непосредственная подготовка студента к занятию.

Самостоятельная работа с рекомендованной литературой.

При работе с основной и дополнительной литературой целесообразно придерживаться такой последовательности. Сначала ознакомиться с содержанием всего заданного материала, чтобы составить представление об изучаемых разделах научного знания. После чего, используя конспект лекций выделить разделы, представляющие собой развернутое изложение темы лекции или методов, используемых на практических занятиях. Затем прочитать выделенные разделы, чтобы в ходе чтения понять и запомнить смысл каждого положения и вопроса в целом. Продуктивно сопровождать чтение записями (план прочитанного текста, тезисы, выписки, конспектирование и др.). Если содержание материала несложное, легко усваиваемое, можно ограничиться составлением плана. Если материал содержит новую и трудно усваиваемую информацию, целесообразно его законспектировать.

Для усвоения курса, в качестве вспомогательной литературы рекомендуются к использованию:

- 1 R. M. Martin, *Electronic Structure: Basic Theory and Methods*, Cambridge University Press, Cambridge, U.K., 2004
- 2 M. C. Payne, M. P. Teter, D. C. Allan, T. A. Arias, and J. D. Joannopoulos, *Iterative minimization techniques for ab initio total-energy calculations: molecular dynamics and conjugate gradients*, *Rev. Mod. Phys.* 64:1045–1097, 1992
- 3 Бринк Х., Ричардс Дж., Феверолф М. *Машинное обучение*. СПб: Питер, 2017.