

Документ подписан простой электронной подписью
Информация о владельце:

ФИО: Исаев Игорь Магомедович

Должность: Проректор по учебной и научной работе

Дата подписания: 10.10.2023 14:47:43

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)

Python для анализа данных

Закреплена за подразделением

Кафедра АСУ

Направление подготовки

22.04.02 МЕТАЛЛУРГИЯ

Профиль

Металлы высоких технологий

Квалификация

Магистр

Форма обучения

очная

Общая трудоемкость

2 ЗЕТ

Часов по учебному плану

72

Формы контроля в семестрах:

в том числе:

зачет 2

аудиторные занятия

36

самостоятельная работа

36

Распределение часов дисциплины по семестрам

Семестр (<Курс>.<Семестр на курсе>)	2 (1.2)		Итого	
	Неделя			
Неделя	18			
Вид занятий	УП	РП	УП	РП
Практические	36	36	36	36
Итого ауд.	36	36	36	36
Контактная работа	36	36	36	36
Сам. работа	36	36	36	36
Итого	72	72	72	72

Программу составил(и):
-, *ст.преп., Агабубаев А.Т.*

Рабочая программа
Python для анализа данных

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 22.04.02 МЕТАЛЛУРГИЯ (приказ от 05.03.2020 г. № 95 о.в.)

Составлена на основании учебного плана:

22.04.02 Металлургия, 22.04.02-ММТ-23-18.plx Металлы высоких технологий, утвержденного Ученым советом НИТУ МИСИС в составе соответствующей ОПОП ВО 22.06.2023, протокол № 5-23

Утверждена в составе ОПОП ВО:

22.04.02 Металлургия, Металлы высоких технологий, утвержденной Ученым советом НИТУ МИСИС 22.06.2023, протокол № 5-23

Рабочая программа одобрена на заседании
Кафедра АСУ

Протокол от 10.04.2023 г., №5

Руководитель подразделения Темкин Игорь Олегович

1. ЦЕЛИ ОСВОЕНИЯ

1.1	Изучение методологических основ сбора, обработки и анализа данных с использованием специализированных библиотек Python;
1.2	Получение навыков в решении прикладных задач предобработки и визуализации данных;
1.3	Получение навыков применения специализированных библиотек Python в задачах машинного обучения.

2. МЕСТО В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

Блок ОП:		ФТД
2.1	Требования к предварительной подготовке обучающегося:	
2.2	Дисциплины (модули) и практики, для которых освоение данной дисциплины (модуля) необходимо как предшествующее:	
2.2.1	Автогенные процессы цветной металлургии	
2.2.2	Защита интеллектуальной собственности и патентоведение	
2.2.3	Моделирование и оптимизация металлургических процессов	
2.2.4	Новые направления экстрактивной металлургии	
2.2.5	Получение особо чистых веществ	
2.2.6	Потребительские свойства металлургической продукции	
2.2.7	Ресурсо- и энергосбережение в производстве легких редких металлов, ч.2	
2.2.8	Ресурсо- и энергосбережение в производстве редкоземельных металлов, ч.2	
2.2.9	Ресурсо- и энергосбережение в производстве тугоплавких редких металлов, ч.2	
2.2.10	Ресурсо- и энергосбережение в производстве тяжелых цветных металлов и сопутствующих элементов, ч.2	
2.2.11	Управление проектами	
2.2.12	Цифровизация производства	
2.2.13	Экономика и организация производства	
2.2.14	Подготовка к процедуре защиты и защита выпускной квалификационной работы	
2.2.15	Преддипломная практика	
2.2.16	Экоаудит металлургических технологий	

3. РЕЗУЛЬТАТЫ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ, СООТНЕСЕННЫЕ С ФОРМИРУЕМЫМИ КОМПЕТЕНЦИЯМИ

ОПК-4: Способен находить и перерабатывать информацию, требуемую для принятия решений в научных исследованиях и в практической технической деятельности, осуществлять моделирование, анализ и эксперименты в целях проведения детального исследования для решения сложных задач в профессиональной области

Знать:

ОПК-4-31 Основы теории систем и системного анализа
Особенности моделирования в различных задачах системного анализа
Основы математического анализа и линейной алгебры
Основы теории вероятности и прикладной статистики

Уметь:

ОПК-4-У1 Применять методы и подходы системного анализа в прикладных задачах обработки и анализа данных
Применять методы математического анализа в задачах обработки и анализа данных

Владеть:

ОПК-4-В1 Навыками применения специализированных библиотек Python в прикладных задачах разработки и оптимизации аналитических алгоритмов
Навыками применение программных инструментов формализации бизнес/производственных - процессов

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. Введение в анализ данных							
1.1	Python для на анализа данных /Пр/	2	1	ОПК-4-31 ОПК-4-У1 ОПК-4-В1	Л1.2 Л1.4 Л1.6 Л1.8			

1.2	Применение методов и алгоритмов машинного обучения в бизнес-задачах /Пр/	2	2	ОПК-4-31 ОПК-4-У1 ОПК-4-В1				
1.3	Основы Pandas /Пр/	2	3	ОПК-4-31 ОПК-4-У1 ОПК-4-В1				
1.4	Первичный анализ и визуализация данных /Пр/	2	2	ОПК-4-31 ОПК-4-У1 ОПК-4-В1				
1.5	Анализ данных. Соревнования на kaggle.com /Ср/	2	18	ОПК-4-31 ОПК-4-У1 ОПК-4-В1			КМ1	
	Раздел 2. Методы, модели и алгоритмы в задачах анализа данных							
2.1	Классификация, деревья решений и метод ближайших соседей /Пр/	2	2	ОПК-4-31 ОПК-4-У1 ОПК-4-В1	Л1.1 Л1.3 Л1.5 Л1.7			
2.2	Линейные модели классификации и регрессии /Пр/	2	2	ОПК-4-31 ОПК-4-У1 ОПК-4-В1				
2.3	Композиции: бэггинг, случайный лес /Пр/	2	2	ОПК-4-31 ОПК-4-У1 ОПК-4-В1				
2.4	Построение и отбор признаков /Пр/	2	2	ОПК-4-31 ОПК-4-У1 ОПК-4-В1				
2.5	Обработка и анализ данных. Проектная работа /Ср/	2	18	ОПК-4-31 ОПК-4-У1 ОПК-4-В1			КМ2	
	Раздел 3. Практикум							
3.1	Установка Jupiter Notebook и настройка конфигурации /Пр/	2	2	ОПК-4-31 ОПК-4-У1 ОПК-4-В1	Л1.2 Л1.4 Л1.6 Л1.8 Л1.9 Л1.10 Э1			Р1
3.2	Основы программирования на Python (Типы данных, циклы, функции и т.д.) /Пр/	2	2	ОПК-4-31 ОПК-4-У1 ОПК-4-В1				Р2
3.3	Библиотеки Python (NumPy, SciPy, matplotlib, Pandas, scikit-learn, statsmodels) /Пр/	2	2	ОПК-4-31 ОПК-4-У1 ОПК-4-В1				Р3
3.4	Структуры данных, функции и файлы /Пр/	2	2	ОПК-4-31 ОПК-4-У1 ОПК-4-В1				Р4
3.5	Основы NumPy: массивы и векторные вычисления /Пр/	2	2	ОПК-4-31 ОПК-4-У1 ОПК-4-В1				Р5
3.6	Введение в структуры данных Pandas /Пр/	2	2	ОПК-4-31 ОПК-4-У1 ОПК-4-В1				Р6
3.7	Pandas. Базовая функциональность /Пр/	2	2	ОПК-4-31 ОПК-4-У1 ОПК-4-В1				Р7
3.8	Чтение и запись данных, форматы файлов /Пр/	2	2	ОПК-4-31 ОПК-4-У1 ОПК-4-В1				Р8
3.9	Очистка и подготовка данных /Пр/	2	2	ОПК-4-31 ОПК-4-У1 ОПК-4-В1				Р9
3.10	Визуализация данных /Пр/	2	2	ОПК-4-31 ОПК-4-У1 ОПК-4-В1				Р10

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 работа № 1		https://www.kaggle.com/competitions/click-prediction-2/overview
КМ2	Коллоквиум	ОПК-4-31;ОПК-4-У1	<ol style="list-style-type: none"> 1. Python - интерпретируемый язык или компилируемый? 2. Какие есть меняющиеся и постоянные типы данных? 3. Что такое область видимости переменных? 4. Что такое introspection? 5. Разница между is и ==? 6. Разница между __init () и __new__ ()? 7. В чем разница между потоками и процессами? 8. Какие есть виды импорта? 9. Что такое класс, итератор, генератор? 10. Что такое метакласс, переменная цикла? 11. В чем разница между итераторами и генераторами? 12. В чем разница между staticmethod и classmethod? 13. Как работают декораторы, контекстные менеджеры? 14. Как работают dict comprehension, list comprehension и set comprehension? 15. Можно ли использовать несколько декораторов для одной функции? 16. Можно ли создать декоратор из класса? 17. Какие есть основные популярные пакеты (requests, pytest, etc)? 18. Что такое lambda-функции?

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	Лабораторная работа № 1. Установка Jupiter Notebook и настройка конфигурации	ОПК-4-31;ОПК-4-У1;ОПК-4-В1	Установка на ПК Python 3.9., Jupiter Npotebook. Настройка конфигурации блокнота и написание первой программы
P2	Лабораторная работа № 2. Основы программирования на Python	ОПК-4-31;ОПК-4-У1;ОПК-4-В1	Семантика языка. типы данных. Арифметические, логические, условные операторы. Циклы. Работа со списками.
P3	Лабораторная работа № 3. Библиотеки Python	ОПК-4-31;ОПК-4-У1;ОПК-4-В1	Изучение на примерах функциональные возможности библиотек: NumPy, SciPy, matplotlib, Pandas, scikit-learn, statsmodels
P4	Лабораторная работа № 4. Структуры данных, функции и файлы	ОПК-4-31;ОПК-4-У1;ОПК-4-В1	Структуры данных и последовательности. Функции. Файлы и операционная система
P5	Лабораторная работа № 5. Основы NumPy: массивы и векторные вычисления	ОПК-4-31;ОПК-4-У1;ОПК-4-В1	NumPy ndarray: объект многомерного массива. Универсальные функции: быстрые поэлементные операции. Программирование с применением массивов. Файловый ввод-вывод массивов. Генерация псевдослучайных чисел. Пример: случайное блуждание
P6	Лабораторная работа № 6. Введение в структуры данных Pandas	ОПК-4-31;ОПК-4-У1;ОПК-4-В1	Объект Series. Объект DataFrame. Индексные объекты

P7	Лабораторная работа № 7. Pandas. Базовая функциональность	ОПК-4-31;ОПК-4-У1;ОПК-4-В1	Переиндексация. Удаление элементов из оси. Доступ по индексу, выборка и фильтрация. Целочисленные индексы. Арифметические операции и выравнивание данных. Применение функций и отображение. Сортировка и ранжирование. Индексы по осям с повторяющимися значениями.
P8	Лабораторная работа № 8. Чтение и запись данных, форматы файлов	ОПК-4-31;ОПК-4-У1;ОПК-4-В1	Чтение и запись данных в текстовом формате. Двоичные форматы данных. Взаимодействие с HTML и Web API. Взаимодействие с базами данных
P9	Лабораторная работа № 9. Очистка и подготовка данных	ОПК-4-31;ОПК-4-У1;ОПК-4-В1	Обработка отсутствующих данных. Преобразование данных. Манипуляции со строками
P10	Лабораторная работа № 10. Визуализация данных	ОПК-4-31;ОПК-4-У1;ОПК-4-В1	Краткое введение в API библиотеки matplotlib. Построение графиков с помощью pandas и seaborn. Другие средства визуализации для Python.

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

В рамках дисциплины "Python для анализа данных", экзамен не предусмотрен

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

По дисциплине предусмотрен зачет с оценкой

Проектные задачи (варианты индивидуальных заданий хранятся на кафедре);

Соревнование в сервисе kaggle.com (Результаты и учебные материалы хранятся на кафедре);

Отчеты, исходный код хранится в ЭИОС "Canvas".

За текущую учебную деятельность обучающегося при выполнении каждой практической работы (выполнение, защита и предоставление отчета с программным файлом в ЭИОС «Canvas»), самостоятельных заданий (защита и предоставление отчета с программным файлом ЭИОС «Canvas») выставляются оценки по 5-балльной (государственной) шкале.

Итоговая оценка определяется на основе процентного отношения в ЭИОС «Canvas» правильно выполненных обучающимся заданий:

90-100% – «5»,
80-89% – «4»,
60%-79% – «3»,
менее 60% – «2».

Промежуточная аттестация в форме коллоквиума позволяет оценить уровень сформированности компетенций в целом по дисциплине и может осуществляться, как в письменной так и в устной форме.

По окончании изучения дисциплины в системе оценки знаний и умений используются следующие критерии:

«Отлично» – за глубокое и полное овладение содержанием учебного материала, в котором студент легко ориентируется, владение понятийным аппаратом за умение связывать теорию с практикой, решать практические задачи, высказывать и обосновывать свои суждения. Отличная отметка предполагает грамотное, логичное изложение ответа (как в устной, так и в письменной форме), качественное внешнее оформление;

«Хорошо» – если студент полно освоил учебный материал, владеет понятийным аппаратом, ориентируется в изученном материале, осознанно применяет знания для решения практических задач, грамотно излагает ответ, но содержание и форма ответа имеют некоторые неточности;

«Удовлетворительно» – если студент обнаруживает знание и понимание основных положений учебного материала, но излагает его неполно, непоследовательно, допускает неточности в определении понятий, в применении знаний для решения практических задач, не умеет доказательно обосновать свои суждения;

«Неудовлетворительно» – если студент имеет разрозненные, бессистемные знания, не умеет выделять главное и второстепенное, допускает ошибки в определении понятий, искажает их смысл, беспорядочно и неуверенно излагает материал, не может применять знания для решения практических задач; за полное незнание и непонимание учебного материала или отказ отвечать.

6. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ

6.1. Рекомендуемая литература

6.1.1. Основная литература

	Авторы, составители	Заглавие	Библиотека	Издательство, год
Л1.1	Чубукова И. А.	Data Mining: учебное пособие	Электронная библиотека	Москва: Интернет-Университет Информационных Технологий (ИНТУИТ) Бином. Лаборатория знаний, 2008

	Авторы, составители	Заглавие	Библиотека	Издательство, год
Л1.2	Сузи Р. А.	Язык программирования Python: учебное пособие	Электронная библиотека	Москва: Интернет-Университет Информационных Технологий (ИНТУИТ) Бином. Лаборатория знаний, 2007
Л1.3	Крутиков В. Н., Мешечкин В. В.	Анализ данных: учебное пособие	Электронная библиотека	Кемерово: Кемеровский государственный университет, 2014
Л1.4	Балджи А. С., Хрипунова М. Б., Александрова И. А.	Математика на Python: учебно-методическое пособие	Электронная библиотека	Москва: Прометей, 2018
Л1.5	Мыльников Л. А., Краузе Б., Кютц М., Баде К., Шмидт И. А.	Интеллектуальный анализ данных в управлении производственными системами (подходы и методы): монография	Электронная библиотека	Москва: Библио-Глобус, 2017
Л1.6	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
Л1.7	Алпайдин Э.	Машинное обучение: новый искусственный интеллект: пер. с англ.	Библиотека МИСиС	М.: Альпина Паблицер, 2017
Л1.8	Волкова В. М., Семенова М. А., Четвертакова Е. С., Вожов С. С.	Программные системы статистического анализа: обнаружение закономерностей в данных с использованием системы R и языка Python: учебное пособие	Электронная библиотека	Новосибирск: Новосибирский государственный технический университет, 2017
Л1.9	Хахаев И. А.	Практикум по алгоритмизации и программированию на Python: курс: учебное пособие	Электронная библиотека	Москва: Национальный Открытый Университет «ИНТУИТ», 2016
Л1.10	Ушаков Р. М.	Технология Big Data в деятельности органов исполнительной власти: проблемы и перспективы административно-правового регулирования: студенческая научная работа	Электронная библиотека	Саратов: б.и., 2021

6.2. Перечень ресурсов информационно-телекоммуникационной сети «Интернет»

Э1	Курс Data Science	https://habr.com/ru/company/ods/blog/322626/
----	-------------------	---

6.3 Перечень программного обеспечения

П.1	Microsoft Office
П.2	LMS Canvas
П.3	MS Teams
П.4	Python
П.5	R Studio

6.4. Перечень информационных справочных систем и профессиональных баз данных

И.1	https://www.scopus.com
И.2	https://www.kaggle.com
И.3	https://habr.com

7. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ

Ауд.	Назначение	Оснащение
------	------------	-----------

Л-830	Лекционная аудитория	комплект учебной мебели на 60 рабочих мест, рабочее место преподавателя с персональным компьютером, без доступа к ИТС «Интернет»
Л-830	Лекционная аудитория	комплект учебной мебели на 60 рабочих мест, рабочее место преподавателя с персональным компьютером, без доступа к ИТС «Интернет»
Л-830	Лекционная аудитория	комплект учебной мебели на 60 рабочих мест, рабочее место преподавателя с персональным компьютером, без доступа к ИТС «Интернет»
Л-831	Учебная аудитория	комплект учебной мебели на 15 рабочих мест, оборудованных компьютерами, рабочее место преподавателя с персональным компьютером, без доступа к ИТС «Интернет»

8. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ

Дисциплина требует значительного объема самостоятельной работы. Отдельные учебные вопросы выносятся на самостоятельную проработку и контролируются посредством текущей аттестации. При этом организуются групповые и индивидуальные консультации. Качественное освоение дисциплины возможно только при систематической самостоятельной работе, что поддерживается системой текущей и рубежной аттестации. Практикум проводится с широким использованием компьютерных программ, как для выполнения, так и для оформления работы.

В процессе выполнения практических работ необходимо показать умелое применение полученных в процессе обучения знаний и навыков при решении задач. С целью формирования практических навыков, знаний, полученных при изучении дисциплин Основы дискретной математики, Комбинаторика и теория графов, Алгоритмы дискретной математики и Высшая математика.

При выполнении самостоятельных работ акцент делается на формирование навыков работы студентов с научно-технической литературой; работы с документацией библиотек Python; на систематизацию материала для решения поставленных задач; на формирование навыков оформления результатов выполненных работ (пояснительной записки, ссылок на литературные источники, выводов по работе). Индивидуальные задания на самостоятельную работу студент получает у преподавателя в соответствии с прилагаемым перечнем их тематик. Рекомендуемая форма их оформления – отчеты с приложением программного файла ЭИОР «Canvas». Защита работы проводится индивидуально каждым студентом. Студенты делают сообщение и отвечают на вопросы преподавателя.

При подготовке к зачету необходимо опираться на вопросы выходного контроля знаний, основную и дополнительную литературу, другие источники информации.