

Методические рекомендации к курсу «Python для анализа данных» для направления 49.04.01 «Физическая культура»

Краева Яна Александровна

2021

Цель дисциплины: Целями освоения дисциплины является развитие навыков программирования на языке Python, представления о сборе, обработке и анализе данных в интерактивной среде Jupyter, введение в автоматизированные методы работы с данными, которые будут использоваться для обучения моделей машинного обучения, а также основные принципы создания систем искусственного интеллекта.

Задачи дисциплины: изучить основы работы в Python для анализа данных; изучить специальные библиотеки языка Python для анализа данных: Numpy; Scipy; Matplotlib; Pandas. Приобрести навыки работы с текстовыми данными. Изучить методики сбора данных из открытых источников и их предварительная обработка. Изучить навыки текстового анализа данных. Изучить навыки разработки систем искусственного интеллекта.

Основная литература по курсу:

1. Маккинни, У. Python и анализ данных / У. Маккинни; перевод с английского А. А. Слинкина. — 2-ое изд., испр. и доп. — Москва: ДМК Пресс, 2020. — 540 с. — ISBN 978-5-97060-590-5. — Текст: электронный // Лань: электронно-библиотечная система. — Режим доступа: для авториз. пользователей. <https://e.lanbook.com/book/131721>
2. Полупанов, Д. В. Программирование в Python 3: учебное пособие / Д. В. Полупанов, С. Р. Абдюшева, А. М. Ефимов. — Уфа: БашГУ, 2020. — 164 с. — ISBN 978-5-7477-5230-6. — Текст: электронный // Лань: электронно-библиотечная система. — Режим доступа: для авториз. пользователей. <https://e.lanbook.com/book/179915>

Дополнительная литература по курсу:

1. Борзунов, С. В. Алгебра и геометрия с примерами на Python: учебное пособие для вузов / С. В. Борзунов, С. Д. Кургалин. — 2-е изд., стер. — Санкт-Петербург: Лань, 2021. — 444 с. — ISBN 978-5-8114-7961-0. — Текст: электронный // Лань: электронно-библиотечная система. — Режим доступа: для авториз. пользователей. <https://e.lanbook.com/book/169808>

Объем и виды учебной работы:

- Семестр: 2.
- Общая трудоёмкость дисциплины: 72 часа.
- Лекции: 16 часов.
- Практические занятия: 32 часа.

Компетенции:

1. УК-91 Способен понимать фундаментальные принципы работы современных систем искусственного интеллекта, разрабатывать правила и стандарты взаимодействия человека и искусственного интеллекта и использовать их в социальной и профессиональной деятельности:

Умеет: применять современные методы и инструменты для представления результатов научно-исследовательской деятельности.

Рекомендации к выполнению практических работ:

Рекомендации к выполнению **Практической работы 1** «Базовые конструкции языка Python для анализа данных»:

Вопросы для проверки усвоения материала практической работы № 1:

1. Перечислите базовые типы данных.
2. Дайте определение и перечислите основные особенности списков.
3. Дайте определение и перечислите основные особенности кортежей.
4. Дайте определение и перечислите основные особенности словарей.
5. Дайте определение и перечислите основные особенности множеств.
6. Дайте определение функции.
7. Дайте определение классу и объекту.
8. Опишите синтаксис подключения библиотек.
9. Опишите функции ввода-вывода данных.

Рекомендации к выполнению **Практической работы 2** «Основные библиотеки (Math, SymPy) и функции для работы с математическим аппаратом в языке Python»:

Вопросы для проверки усвоения материала практической работы № 2:

1. Какая функция округляет до ближайшего большего числа (Math).
2. Какая функция округляет вниз (Math).
3. Какая функция используется для взятия модуля числа (Math).
4. Какая функция извлекает квадратный корень из X (Math).
5. Для чего применяется библиотека SymPy.
6. Особенности символьных вычислений в SymPy.
7. Перечислите типы данных с которыми работает SymPy.
8. Перечислите функции упрощения и их назначение (SymPy).

Рекомендации к выполнению **Практической работы 3** «Основные функции библиотеки NumPy в задаче подготовки числовых наборов данных (random, linalg и др.)»:

Вопросы для проверки усвоения материала практической работы

№ 3:

1. Особенности многомерных массивов.
2. Особенности векторных вычислений.
3. Правила индексирования и вырезания.
4. перечислите универсальные функции: быстрые поэлементные операции над массивами.
5. Перечислите функции для сортировки.
6. Устранение дубликатов и другие теоретико-множественные операции.
7. Файловый ввод-вывод массивов.
8. Линейная алгебра.
9. Генерация случайных чисел.

Рекомендации к выполнению **Практической работы 4** «Основные функции библиотеки Scipy в задаче подготовки числовых наборов данных (ndimage, cluster и др.)»:

Вопросы для проверки усвоения материала практической работы

№ 4:

1. Для чего предназначена библиотека Scipy.
2. Как правильно импортировать библиотеку Scipy.
3. Какие функции присутствуют в библиотеке для решения задач математической физики.
4. Какие функции существуют в библиотеке для решения задач линейной алгебры.
5. Можно ли использовать функции NumPy из библиотеки Scipy.

Рекомендации к выполнению **Практической работы 5** «Основные методы визуализации подготовленных данных с использованием библиотеки Matplotlib. Обработка изображений в различных форматах (JPEG, PDF, PNG и др.)»:

Вопросы для проверки усвоения материала практической работы

№ 5:

1. Для чего предназначена библиотека Matplotlib.
2. Какой командой подключается библиотека Matplotlib.
3. Приведите код для построения простого линейного графика.
4. Приведите код для построения простого диаграммы рассеивания.
5. Приведите код для построения простого гистограммы.
6. Какой командой можно добавить надпись на графике.
7. Какой командой можно поменять цвет графика.
8. Приведите код для построения простой круговой диаграммы.
9. Приведите код для построения простой тепловой карты.

Рекомендации к выполнению **Практической работы 6** «Основные методы подготовки, преобразования и обработки наборов данных с использованием библиотеки Pandas для дальнейшего их использования в обучении систем искусственного интеллекта. Объекты Series и DataFrame.»:

Вопросы для проверки усвоения материала практической работы

№ 6:

1. Для чего предназначена библиотека Pandas.
2. Какой командой импортировать библиотеку Pandas.
3. Дайте описание объекту Series.
4. Дайте описание объекту DataFrame.
5. Дайте описание объекту Index.
6. Расскажите анатомию Python Pandas DataFrame — Column, Index, Data.
7. Что означает выбор подмножества.
8. Напишите код для извлечения отдельных компонентов DataFrame.
9. Какие типы данных могут быть у компонентов.
10. Напишите код для выбора нескольких столбцов с помощью операции индексации.

11. Напишите код для выборки данных из Pandas DataFrame с помощью .loc.

12. Напишите код для выбора строк и столбцов одновременно с .loc.

Рекомендации к выполнению **Практической работы 7** «Основные методы для создания текстовых наборов данных.»:

Вопросы для проверки усвоения материала практической работы № 7:

1. Инструменты ввода-вывода (текст, CSV, HDF5 и др.).
2. Чем тип StringDtype отличается от типа object.
3. Перечислите основные методы для работы со строками.
4. Опишите метод split и приведите код для примера.
5. Опишите метод replace и приведите код для примера.
6. Как провести индексирование с помощью .str, приведите код для примера.

Рекомендации к выполнению **Практической работы 8** «Основные этапы разработки и внедрения систем искусственного интеллекта в профессиональную область»:

Вопросы для проверки усвоения материала практической работы № 8:

1. Перечислите основные этапы разработки систем искусственного интеллекта для работы с изображениями.
2. Перечислите основные этапы разработки систем искусственного интеллекта для работы со статистическими данными.
3. Перечислите основные этапы разработки систем искусственного интеллекта для работы с видеоданными.
4. Перечислите основные принципы внедрения систем искусственного интеллекта в системы судейства.
5. Перечислите основные принципы внедрения систем искусственного интеллекта в системы для мониторинга состояния спортсмена.

6. Перечислите основные принципы внедрения систем искусственного интеллекта в системы для прогнозирования состояния спортсмена.

Вопросы к экзамену:

1. Перечислите базовые типы данных.
2. Дайте определение и перечислите основные особенности списков.
3. Дайте определение и перечислите основные особенности кортежей.
4. Дайте определение и перечислите основные особенности словарей.
5. Дайте определение и перечислите основные особенности множеств.
6. Дайте определение функции.
7. Дайте определение классу и объекту.
8. Опишите синтаксис подключения библиотек.
9. Опишите функции ввода-вывода данных.
10. Какая функция округляет до ближайшего большего числа (Math).
11. Какая функция округляет вниз (Math).
12. Какая функция используется для взятия модуля числа (Math).
13. Какая функция извлекает квадратный корень из X (Math).
14. Для чего применяется библиотека SymPy.
15. Особенности символьных вычислений в SymPy.
16. Перечислите типы данных с которыми работает SymPy.
17. Перечислите функции упрощения и их назначение (SymPy).
18. Особенности многомерных массивов.
19. Особенности векторных вычислений.
20. Правила индексирования и вырезания.
21. перечислите универсальные функции: быстрые поэлементные операции над массивами.
22. Перечислите функции для сортировки.
23. Устранение дубликатов и другие теоретико-множественные операции.
24. Файловый ввод-вывод массивов.
25. Линейная алгебра.
26. Генерация случайных чисел.
27. Для чего предназначена библиотека SciPy.
28. Как правильно импортировать библиотеку SciPy.

29. Какие функции присутствуют в библиотеке для решения задач математической физики.
30. Какие функции существуют в библиотеке для решения задач линейной алгебры.
31. Можно ли использовать функции NumPy из библиотеки Scipy.
32. Для чего предназначена библиотека Matplotlib.
33. Какой командой подключается библиотека Matplotlib.
34. Приведите код для построения простого линейного графика.
35. Приведите код для построения простого диаграммы рассеивания.
36. Приведите код для построения простого гистограммы.
37. Какой командой можно добавить надпись на графике.
38. Какой командой можно поменять цвет графика.
39. Приведите код для построения простой круговой диаграммы.
40. Приведите код для построения простой тепловой карты.
41. Для чего предназначена библиотека Pandas.
42. Какой командой импортировать библиотеку Pandas.
43. Дайте описание объекту Series.
44. Дайте описание объекту DataFrame.
45. Дайте описание объекту Index.
46. Расскажите анатомию Python Pandas DataFrame — Column, Index, Data.
47. Что означает выбор подмножества.
48. Напишите код для извлечения отдельных компонентов DataFrame.
49. Какие типы данных могут быть у компонентов.
50. Напишите код для выбора нескольких столбцов с помощью операции индексации.
51. Напишите код для выборки данных из Pandas DataFrame с помощью .loc.
52. Напишите код для выбора строк и столбцов одновременно с .loc.
53. Инструменты ввода-вывода (текст, CSV, HDF5 и др.).
54. Чем тип StringDtype отличается от типа object.

- 55.Перечислите основные методы для работы со строками.
- 56.Опишите метод `split` и приведите код для примера.
- 57.Опишите метод `replace` и приведите код для примера.
- 58.Как провести индексирование с помощью `.str`, приведите код для примера.
- 59.Перечислите основные этапы разработки систем искусственного интеллекта для работы с изображениями.
- 60.Перечислите основные этапы разработки систем искусственного интеллекта для работы со статистическими данными.
- 61.Перечислите основные этапы разработки систем искусственного интеллекта для работы с видеоданными.
- 62.Перечислите основные принципы внедрения систем искусственного интеллекта в системы судейства.
- 63.Перечислите основные принципы внедрения систем искусственного интеллекта в системы для мониторинга состояния спортсмена.
- 64.Перечислите основные принципы внедрения систем искусственного интеллекта в системы для прогнозирования состояния спортсмена.