

Документ подписан простой электронной подписью Информация о владельце: ФИО: Таскаев Сергей Валерьевич Должность: Ректор Дата подписания: 18.03.2025 14:53:16 Уникальный программный ключ: 04c19ed8bfb98f3b6cb77a486b9a8788b87272737	МИНИСТЕРСТВО НАУКИ РОССИИ Федеральное государственное бюджетное образовательное учреждение высшего образования «Челябинский государственный университет» (ФГБОУ ВО «ЧелГУ»)	Рабочая программа дисциплины "Программирование для анализа данных" по направлению подготовки (специальности) 01.03.02 "Прикладная математика и информатика" направленности (профилю) Прикладная математика и искусственный интеллект ФГБОУ ВО «ЧелГУ»	стр. 1
-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--------

Рабочая программа дисциплины (модуля)*
Программирование для анализа данных

 Направление подготовки (специальность)
01.03.02 Прикладная математика и информатика

 Направленность (профиль)
Прикладная математика и искусственный интеллект

 Присваиваемая квалификация (степень)
бакалавр

 Форма обучения
очная

 Год набора 2024

*Рабочая программа дисциплины (модуля) адаптирована для инклюзивного обучения инвалидов и лиц с ограниченными возможностями здоровья

Челябинск 2024 г.



Содержание

1. Цели освоения дисциплины
2. Место дисциплины в структуре ОПОП
3. Компетенции обучающегося, формируемые в результате освоения дисциплины (модуля)
4. Объем дисциплины (модуля)
5. Структура и содержание дисциплины (модуля)
6. Фонд оценочных средств
 - 6.1. Перечень видов оценочных средств
 - 6.2. Типовые контрольные задания и иные материалы для текущей аттестации
 - 6.3. Типовые контрольные вопросы и задания для промежуточной аттестации
 - 6.4. Критерии оценивания
7. Учебно-методическое и информационное обеспечение дисциплины (модуля)
 - 7.1. Рекомендуемая литература
 - 7.2. Перечень ресурсов информационно-телекоммуникационной сети "Интернет"
 - 7.3. Перечень информационных технологий
8. Материально-техническое обеспечение дисциплины (модуля)
9. Методические указания для обучающихся по освоению дисциплины (модуля)
10. Специальные условия освоения дисциплины обучающимися с инвалидностью и ограниченными возможностями здоровья



1. ЦЕЛИ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Целью дисциплины является формирование базовых представлений, знаний и умений в области анализа данных. Основные задачи дисциплины: ознакомить студента с основными методами и подходами сбора и подготовки данных для анализа, дать описание основных характеристик наборов данных, показать способы предварительной обработки данных.
Результаты обучения по дисциплине направлены на достижение следующих индикаторов компетенций ПК-4 (ПК-2 модели), ПК-8 (ПК-5 модели):
ПК-4.2. Разрабатывает приложения систем искусственного интеллекта
ПК-8.1. Осуществляет оценку и выбор инструментальных средств для решения поставленной задачи

2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОПОП

Цикл (раздел) ОПОП: Б1.В.1.01

2.1 Требования к предварительной подготовке обучающегося:

2.2 Дисциплины и практики, для которых освоение данной дисциплины (модуля) необходимо как предшествующее:

Функциональное и логическое программирование
Производственная практика (преддипломная) (8 семестр)
Алгоритмы машинного обучения
Современные технологии разработки программных систем искусственного интеллекта
Высокопроизводительные параллельные вычисления на кластерных системах
Современные архитектуры глубоких искусственных нейронных сетей
Анализ требований и проектирование систем искусственного интеллекта
Практика по получению профессиональных умений и опыта профессиональной деятельности

3. КОМПЕТЕНЦИИ ОБУЧАЮЩЕГОСЯ, ФОРМИРУЕМЫЕ В РЕЗУЛЬТАТЕ ОСВОЕНИЯ ДИСЦИПЛИНЫ (МОДУЛЯ)

ПК-4: Способен разрабатывать и тестировать программные компоненты решения задач в системах искусственного интеллекта

Знать:

-

Уметь:

Для достижения ПК-4.2. [ПК-2.2. У-1.]: умеет разрабатывать программные приложения систем искусственного интеллекта, с использованием современных языков программирования, библиотек и программных платформ функционального, логического, объектно- ориентированного программирования (Python, R, C++, C#)

Владеть:

Для достижения ПК-4.2.: имеет практический опыт участия в разработке программных приложений систем искусственного интеллекта, с использованием современных языков программирования.

ПК-8: Способен использовать инструментальные средства для решения задач машинного обучения

Знать:

Для достижения ПК-8.1. [ПК-5.1. З-1.]: знает возможности современных инструментальных средств и систем программирования для решения задач анализа данных и машинного обучения.

Уметь:

-

Владеть:

-

В результате освоения дисциплины обучающийся должен

3.1 Знать:



Рабочая программа дисциплины "Программирование для анализа данных" по направлению подготовки (специальности) 01.03.02 "Прикладная математика и информатика" направленности (профилю) Прикладная математика и искусственный интеллект ФГБОУ ВО «ЧелГУ»

стр. 4

3.1.1 возможности современных инструментальных средств и систем программирования для решения задач анализа данных и машинного обучения.

3.2 Уметь:

3.2.1 разрабатывать программные приложения систем искусственного интеллекта, с использованием современных языков программирования (Python, R)

3.3 Владеть:

3.3.1 участия в разработке программных приложений систем искусственного интеллекта, с использованием современных языков программирования.

4. ОБЪЕМ ДИСЦИПЛИНЫ (МОДУЛЯ)

Общая трудоемкость	4 ЗЕТ
Часов по учебному плану : 144 в том числе : аудиторные занятия : 64 самостоятельная работа : 71,5 : контактная работа: 72,5 ИКР: 8,5	Виды контроля в семестрах: зачеты с оценкой 3

5. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

Код занятия	Наименование разделов и тем /вид занятия/	Семестр / Курс	Часов	Литература
Раздел 1. Введение в анализ данных				
1.1	Введение в интеллектуальный анализ данных: основные понятия, области применения современных технологий обработки и интеллектуального анализа больших данных /Лек/	3	4	Л1.3Л2.1 Э1 Э2 Э3
1.2	Организация сбора и хранения больших наборов данных. /Лаб/	3	2	Л1.3 Э1 Э2 Э3
1.3	Основные компоненты статистической среды R /Лаб/	3	2	Л1.1 Э1 Э2 Э3
Раздел 2. Очистка, интеграция и преобразование данных				
2.1	Этапы анализа данных. Структурированные и неструктурированные данные. Сбор и подготовка данных. Библиотека pandas. Структуры данных в pandas, работа со структурами данных. Операции над данными. Комбинирование данных из разных источников. Обработка пропущенных значений. /Лек/	3	2	Л1.3 Э1 Э2 Э3
2.2	Виды данных и их загрузка.. Структурированные и неструктурированные типы данных. Основные форматы файлов - txt, csv, xls, sav. Структура файлов. Виды разделителей, символы окончания строки. Проблема кодировок и различия операционных систем. SQL-базы данных. Удаленные базы данных (API). Неструктурированные данные - json, xml. Сохранение или запись файлов, представление в внешних веб-приложения /Лек/	3	4	Л1.3 Э1 Э2 Э3
2.3	Работа с данными в pandas /Лаб/	3	2	Л1.2 Л1.3 Э1 Э2 Э3
2.4	Импорт и преобразование данных /Лаб/	3	2	Л1.2 Л1.3 Э1 Э2 Э3
Раздел 3. Программные модули и пакеты для работы с данными				
3.1	Введение в визуализацию данных. Нюансы визуализации данных и принципы человеческого восприятия. Правила создания хороших визуализаций. /Лек/	3	2	Л1.3 Э1 Э2 Э3



3.2	Обзор различных видов графиков (гистограмма, бар-чарт, секторная диаграмма, линейные графики, график рассеяния, тепловая карта и т.д.). Особенности разных видов графиков и их использования. /Лек/	3	2	Л1.3 Э1 Э2 Э3
3.3	Задачи визуализации данных. Статичные графики, интерактивные визуализации, инфографика. Виды графиков. Принципы визуальной презентации данных. Ошибки в использовании линейных графиков, гистограмм, круговых и объемных диаграмм. Палитры для графиков. /Лек/	3	2	Л1.3 Э1 Э2 Э3
3.4	Библиотеки NumPy, SciPy: основные функции. /Лаб/	3	2	Л1.2 Л1.3 Э1 Э2 Э3
3.5	Визуализация данных с Matplotlib и Pandas /Лаб/	3	2	Л1.2 Л1.3 Э1 Э2 Э3
3.6	Визуализация данных. Базовые графики в R. /Лаб/	3	2	Л1.1 Э1 Э2 Э3
3.7	Динамические визуализации в R. Создание интерактивных визуализаций и отчетов с помощью инструмента Plotly. /Лаб/	3	4	Л1.1 Э1 Э2 Э3
Раздел 4. Разведочный анализ и создание отчетов				
4.1	Статистика: описание одиночного набора данных, показатели центра распределения, показатели вариации, корреляция и причинная зависимость. Проверка статистических гипотез. /Лек/	3	4	Л1.1Л2.1 Э1 Э2 Э3
4.2	Оценка выборочных параметров с использованием специальных функций. Законы распределения вероятностей. Проверка на нормальность распределения. /Лек/	3	4	Л1.1Л2.1 Э1 Э2 Э3
4.3	Обзор языка разметки markdown. Экспорт данных и отчетов в Python. /Лаб/	3	2	Л1.2 Л1.3 Э1 Э2 Э3
4.4	Создание отчетов. R Markdown и Quarto. /Лаб/	3	4	Л1.1 Э1 Э2 Э3
Раздел 5. Многомерные методы и основы машинного (статистического) обучения				
5.1	Классические методы и критерии статистики. Гипотеза о равенстве средних двух генеральных совокупностей. Рандомизация, бутстрап и оценка статистической мощности. Введение в дисперсионный анализ. Оценка корреляций двух случайных величин. Критерий хи- квадрат. /Лек/	3	2	Л1.1Л2.1 Э1 Э2 Э3
5.2	Регрессионные модели зависимостей между количественными переменными. Модель простой линейной регрессии. Модели регрессии при различных видах функции потерь. Полиномиальные и нелинейные модели регрессии. Сравнение эффективности различных моделей при прогнозировании /Лек/	3	2	Л1.2Л2.1 Э1 Э2 Э3
5.3	Модели для оценки временных рядов. Определение стационарности временных рядов. Определение ложных регрессий. Работа с трендами. Гармонический и спектральный анализ. Построение и прогнозирование авторегрессионных моделей ARIMA. Модели ADL, ARCH, GARCH. /Лек/	3	2	Л1.3Л2.1 Э1 Э2 Э3
5.4	Панельное исследование. Линейная панель. Структура данных. Расчетный подход к оценке. Трансформация данных. Динамические формулы. Оценка простейших моделей. /Лек/	3	2	Л1.3Л2.1 Э1 Э2 Э3
5.5	Классические методы и критерии статистики. /Лаб/	3	2	Л1.1Л2.1 Э1 Э2 Э3



5.6	Регрессионные модели зависимостей между количественными переменными. /Лаб/	3	2	Л1.2Л2.1 Э1 Э2 Э3
5.7	Модели для оценки временных рядов /Лаб/	3	2	Л1.3Л2.1 Э1 Э2 Э3
5.8	Многомерные методы анализа данных /Лаб/	3	2	Л1.3Л2.1 Э1 Э2 Э3
Раздел 6. Иная контактная работа				
6.1	Индивидуальные консультации, текущий контроль /ИКР/	3	8,5	Л1.1 Л1.2 Л1.3Л2.1 Э1 Э2 Э3
6.2	Подготовка отчетов по лабораторным работам /Ср/	3	25	Л1.1 Л1.2 Л1.3Л2.1 Э1 Э2 Э3
6.3	Проработка лекционного материала. Подготовка к выполнению лабораторных работ /Ср/	3	46,5	Л1.1 Л1.2 Л1.3Л2.1 Э1 Э2 Э3

6. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

6.1. Перечень видов оценочных средств

Лабораторные работы 1-16
Вопросы к зачету

6.2. Типовые контрольные задания и иные материалы для текущей аттестации

Лабораторная работа № 1. Организация сбора и хранения больших наборов данных.
Лабораторная работа № 2. Основные компоненты статистической среды R.
Лабораторная работа № 3. Работа с данными в Pandas.
Лабораторная работа № 4. Импорт и преобразование данных.
Лабораторная работа № 5. Библиотеки NumPy, SciPy: основные функции.
Лабораторная работа № 6. Визуализация данных с Matplotlib и Pandas.
Лабораторная работа № 7. Визуализация данных. Базовые графики в R.
Лабораторная работа № 8-9. Динамические визуализации в R. Создание интерактивных визуализаций и отчетов с помощью инструмента Plotly.
Лабораторная работа № 10. Обзор языка разметки Markdown. Экспорт данных и отчетов в Python.
Лабораторная работа № 11-12. Создание отчетов. R Markdown и Quarto.
Лабораторная работа № 13. Классические методы и критерии статистики.
Лабораторная работа № 14. Регрессионные модели зависимостей между количественными переменными.
Лабораторная работа № 15. Модели для оценки временных рядов.
Лабораторная работа № 16. Многомерные методы анализа данных.

Образец лабораторной работы приведен в приложении.

6.3. Типовые контрольные вопросы и задания для промежуточной аттестации

Вопросы к зачету

1. Введение в интеллектуальный анализ данных: основные понятия, области применения современных технологий обработки и интеллектуального анализа больших данных.
2. Этапы анализа данных.
3. Структурированные и неструктурированные данные. Сбор и подготовка данных.
4. Библиотека pandas.
5. Структуры данных в pandas, работа со структурами данных.
6. Операции над данными.
7. Комбинирование данных из разных источников.
8. Обработка пропущенных значений.
9. Виды данных и их загрузка.
10. Структурированные и неструктурированные типы данных. Основные форматы файлов - txt, csv, xls, sav.
11. Структура файлов.
12. Виды разделителей, символы окончания строки.
13. Проблема кодировок и различия операционных систем. SQL-базы данных. Удаленные базы данных (API).



14. Неструктурированные данные - json, xml.
15. Сохранение или запись файлов, представление в внешних веб-приложения
16. Введение в визуализацию данных.
17. Нюансы визуализации данных и принципы человеческого восприятия.
18. Правила создания хороших визуализаций.
19. Обзор различных видов графиков (гистограмма, бар-чарт, секторная диаграмма, линейные графики, график рассеяния, тепловая карта и т.д.).
20. Особенности разных видов графиков и их использования.
21. Задачи визуализации данных.
22. Статичные графики, интерактивные визуализации, инфографика.
23. Виды графиков.
24. Принципы визуальной презентации данных.
25. Ошибки в использовании линейных графиков, гистограмм, круговых и объемных диаграмм. Палитры для графиков.
26. Статистика: описание одиночного набора данных, показатели центра распределения, показатели вариации, корреляция и причинная зависимость.
27. Проверка статистических гипотез.
28. Оценка выборочных параметров с использованием специальных функций.
29. Законы распределения вероятностей в R.
30. Проверка на нормальность распределения.
31. Классические методы и критерии статистики.
32. Гипотеза о равенстве средних двух генеральных совокупностей.
33. Рандомизация, бутстрап и оценка статистической мощности. Введение в дисперсионный анализ.
34. Оценка корреляций двух случайных величин. Критерий хи-квадрат.
35. Регрессионные модели зависимостей между количественными переменными.
36. Модель простой линейной регрессии.
37. Модели регрессии при различных видах функции потерь.
38. Полиномиальные и нелинейные модели регрессии.
39. Сравнение эффективности различных моделей при прогнозировании.
40. Модели для оценки временных рядов.
41. Определение стационарности временных рядов.
42. Определение ложных регрессий.
43. Работа с трендами.
44. Гармонический и спектральный анализ.
45. Построение и прогнозирование авторегрессионных моделей ARIMA.
46. Модели ADL, ARCH, GARCH.
47. Многомерные методы анализа данных

6.4. Критерии оценивания

Учебным планом по данной дисциплине предусмотрен зачет с оценкой.

В течение учебного семестра студенты за каждый вид работы получают баллы. Кроме этого на зачете максимально можно получить 20 баллов. Итоговая оценка складывается из суммы баллов, полученных за работу в семестре и за ответ на зачете. Затем полученная сумма баллов переводится в оценку. При этом допускается получение студентом автоматической оценки только по результатам работы в семестре.

Набранные баллы Оценка

25 – 49	неудовлетворительно
50 – 69	удовлетворительно
70 – 90	хорошо
91 – 100	отлично

Начисляемые баллы за выполнение плановых заданий

Выполнение лабораторной работы №1	- 5 баллов
Выполнение лабораторной работы №2	- 5 баллов
Выполнение лабораторной работы №3	- 5 баллов
Выполнение лабораторной работы №4	- 5 баллов
Выполнение лабораторной работы №5	- 5 баллов
Выполнение лабораторной работы №6	- 5 баллов
Выполнение лабораторной работы №7	- 5 баллов



Выполнение лабораторной работы №8-9 - 10 баллов
Выполнение лабораторной работы №10 - 5 баллов
Выполнение лабораторной работы №11-12 - 10 баллов
Выполнение лабораторной работы №13 - 5 баллов
Выполнение лабораторной работы №14 - 5 баллов
Выполнение лабораторной работы №15 - 5 баллов
Выполнение лабораторной работы №16 - 5 баллов
Выполнение заданий на зачете - 20

Порядок начисления баллов за лабораторные работы №1-7, 10, 13-16.

Лабораторная работа проводится в течение одной пары. По окончании пары студент предоставляет отчет. Отчет по заданию высылается в виде документа формата PDF либо ZIP архива со всеми необходимыми файлами.

Максимальная оценка за задание 5 баллов. Баллы выставляются следующим образом:

1) Своевременность предоставления ответа. В зачет идет время последней попытки в случае повторной отправки ответа.

1 балл - отчет предоставлен в течении двух недель начиная с даты проведения лабораторной работы;

0 баллов - за предоставление ответа по прошествии двух недель после даты проведения лабораторной работы.

2) Правильность выполнения лабораторной работы.

2 балла - все пункты лабораторной работы выполнены верно;

1 балл - все пункты лабораторной работы выполнены верно, но имеются недочеты в реализации;

0 баллов - не все пункты лабораторной работы выполнены или имеются существенные недостатки в реализации.

3) Индивидуальность выполнения задания.

1 балл - работа выполнена индивидуально;

0 баллов - работа выполнена в группе.

4) Полнота предоставления отчета по работе.

1 балл - все разделы отчета присутствуют или все необходимые файлы присутствуют в случае отправки ZIP архива;

0 баллов - не все разделы отчета присутствуют или недостает некоторых требуемых файлов в ZIP архиве.

Порядок начисления баллов за лабораторные работы №8-9, 11-12.

Лабораторная работа проводится в течение двух пар. По окончании второй пары студент предоставляет отчет. Отчет по заданию высылается в виде документа формата PDF либо ZIP архива со всеми необходимыми файлами.

Максимальная оценка за задание 10 баллов. Баллы выставляются следующим образом:

1) Своевременность предоставления ответа. В зачет идет время последней попытки в случае повторной отправки ответа.

2 балл - отчет предоставлен в течении двух недель начиная с даты проведения второй пары лабораторной работы;

0 баллов - за предоставление ответа по прошествии двух недель после даты проведения второй пары лабораторной работы.

2) Правильность выполнения лабораторной работы.

4 балла - все пункты лабораторной работы выполнены верно;

2 балл - все пункты лабораторной работы выполнены верно, но имеются недочеты в реализации;

0 баллов - не все пункты лабораторной работы выполнены или имеются существенные недостатки в реализации.

3) Индивидуальность выполнения задания.

2 балл - работа выполнена индивидуально;

0 баллов - работа выполнена в группе.

4) Полнота предоставления отчета по работе.

2 балл - все разделы отчета присутствуют или все необходимые файлы присутствуют в случае отправки ZIP архива;

0 баллов - не все разделы отчета присутствуют или недостает некоторых требуемых файлов в ZIP архиве.

На дифференцированном зачете происходит оценивание учебной деятельности обучающихся по дисциплине на основе полученных оценок за контрольно-рейтинговые мероприятия текущего контроля. Оценка за дифференцированный зачет выставляется по текущему контролю при соблюдении следующих условий:

1) выполнены и сданы все лабораторные работы;

2) количества баллов, полученных за контрольные мероприятия текущего контроля, достаточно для выставления положительной оценки.

Если студенту не хватает баллов для выставления оценки по текущему контролю, то при условии выполнения всех лабораторных работ студенту предлагается пройти опрос. Студенту задается 5 вопросов из разных тем курса. На подготовку ответов дается 45 минут. Использование литературы для подготовки ответа не допускается. По



окончании отведенного времени студент озвучивает ответы. Правильный ответ на вопрос соответствует 4 баллам. Неправильный ответ на вопрос соответствует 0 баллов.

7. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

7.1. Рекомендуемая литература

7.1.1. Основная литература

	Авторы, составители	Заглавие	Издательство, год	Ресурс
Л1.1	Буховец А. Г., Москалев П. В.	Алгоритмы вычислительной статистики в системе R: учебное пособие для вузов	Санкт-Петербург [и др.]: Лань, 2015	
Л1.2	Рашка С.	Python и машинное обучение: крайне необходимое пособие по новейшей предсказательной аналитике, обязательное для более глубокого понимания методологии машинного обучения (https://e.lanbook.com/book/100905)	Москва : ДМК Пресс, 2017	ЭБС
Л1.3	Маккинни У.	Python и анализ данных (https://e.lanbook.com/book/131721)	Москва : ДМК Пресс, 2020	ЭБС

7.1.2. Дополнительная литература

	Авторы, составители	Заглавие	Издательство, год	Ресурс
Л2.1	Миркин Б. Г.	Введение в анализ данных: учебник и практикум (https://urait.ru/bcode/511121)	Москва : Юрайт, 2023	ЭБС

7.2. Перечень ресурсов информационно-телекоммуникационной сети "Интернет"

Э1	Лань [Электронный ресурс] : электронно-библиотечная система (ЭБС) / издательство Лань. – URL: http://e.lanbook.com/ http://e.lanbook.com/
Э2	Университетская библиотека онлайн [Электронный ресурс] : электронно-библиотечная система (ЭБС) / ООО Директмедиа Паблишинг. – URL: http://biblioclub.ru/ http://biblioclub.ru
Э3	Znanium.com [Электронный ресурс] : электронно-библиотечная система (ЭБС) / Научно-издательский центр ИНФРА-М. – URL: http://znanium.com/ http://znanium.com/

7.3 Перечень информационных технологий

7.3.1 Программное обеспечение

LMS Moodle

LibreOffice

Python

R

7.3.2 Профессиональные базы данных и информационно-справочные системы

1. Научная электронная библиотека eLIBRARY.RU (<https://elibrary.ru/defaultx.asp?>) eLIBRARY.RU : научная электронная библиотека : сайт. – Москва, 2000 – . – URL: <https://elibrary.ru.> – Режим доступа: для зарегистрир. пользователей. – Текст : электронный.
2. Реферативная база по математике MathSciNet (<https://mathscinet.ams.org/mathscinet/>) Mathematical Reviews (MR) : реферативная база данных / American Mathematical Society. – URL: <http://www.ams.org/mathscinet/>. – Яз. рус., англ. – Режим доступа: для зарегистрир. пользователей ЧелГУ. – Текст : электронный

8. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

Для реализации дисциплины используются учебные аудитории для проведения занятий лекционного типа, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, компьютерные классы для проведения лабораторных работ, а также помещения для самостоятельной работы.

Учебные аудитории укомплектованы специализированной мебелью (подразумевается наличие стандартных рабочих (посадочных) мест) и техническими средствами обучения (переносное и / или стационарное мультимедийное оборудование: экран, ноутбук, проектор).

Для проведения занятий лекционного типа предлагаются наборы демонстрационного оборудования и учебно-наглядных пособий (мультимедийные презентации по отдельным темам, рисунки, таблицы, схемы и т.д.).



Помещения для самостоятельной работы обучающихся оснащены компьютерной техникой с подключением к сети "Интернет" и обеспечением доступа в электронную информационно-образовательную среду университета

9. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ (МОДУЛЯ)

Учебным планом предусмотрена самостоятельная работа студентов. Самостоятельная работа проводится с целью углубления знаний по дисциплине и предусматривает:

- проработку теоретического материала по учебникам или конспекту лекций с обязательным разбором приведенных примеров;
- подготовку к лабораторным работам;
- подготовку к сдаче зачета.

При планировании времени на самостоятельную работу студентам необходимо предусмотреть регулярное повторение пройденного материала.

В случае применения при изучении дисциплины электронного обучения, дистанционных образовательных технологий общение обучающихся и преподавателя осуществляется в режиме реального или отложенного времени, при этом используются возможности системы дистанционного обучения Moodle и электронная почта.

Большую часть времени обучающиеся самостоятельно работают с учебно-методическими материалами. Студенты имеют возможность консультироваться с преподавателем по всем вопросам, возникающим в ходе самостоятельной работы, посредством электронной почты, сообщений системы дистанционного обучения Moodle.

Доступ обучающегося к учебным ресурсам в режиме отложенного времени, самостоятельной работы осуществляется через сеть Интернет в удобном для него месте, времени и темпе.

При обучении лиц с ограниченными возможностями здоровья электронное обучение, дистанционные образовательные технологии предусматривают возможность приема-передачи информации в доступных для них формах.

Реализация дисциплины с применением электронного обучения, дистанционных образовательных технологий (далее – ЭО, ДОТ) осуществляется на основании «Положения о реализации основных и дополнительных образовательных программ с применением электронного обучения и дистанционных образовательных технологий в федеральном государственном бюджетном образовательном учреждении высшего образования «Челябинский государственный университет», «Положения о порядке зачета обучающимся по основным профессиональным образовательным программам высшего образования в ФГБОУ ВО «ЧелГУ» результатов освоения в организациях, осуществляющих образовательную деятельность, учебных предметов, курсов, дисциплин (модулей), практик, дополнительных образовательных программ» посредством электронной информационно-образовательной среды ФГБОУ ВО «ЧелГУ». В исключительных случаях (форс-мажор и т.п.) при реализации образовательной деятельности с применением ЭО, ДОТ могут применять компоненты, не входящие в перечень электронной информационно-образовательной среды.

10. СПЕЦИАЛЬНЫЕ УСЛОВИЯ ОСВОЕНИЯ ДИСЦИПЛИНЫ ОБУЧАЮЩИМИСЯ С ИНВАЛИДНОСТЬЮ И ОГРАНИЧЕННЫМИ ВОЗМОЖНОСТЯМИ ЗДОРОВЬЯ

Освоение дисциплины инвалидами и лицами с ограниченными возможностями здоровья осуществляется с использованием специальных технических средств и голо информационных технологий, предоставляемых Ресурсным учебно-методическим центром по обучению инвалидов и лиц с ограниченными возможностями здоровья ЧелГУ по запросу обучающегося.

1. Мобильные специальные технические средства для лиц с нарушениями зрения: портативный компьютер с вводом/выводом шрифтом Брайля с синтезатором речи «EIBraile-W14J G2»; ноутбуки с программой экранного доступа NVDA; электронные увеличители для удаленного просмотра; видеоувеличители портативные; тифлоплеер; цифровые диктофоны.

2. Мобильные специальные технические средства для лиц с нарушениями слуха: система свободного звукового поля со встроенной совместимостью с FM-устройствами; радиоклассы «Сонет-PCM» с передатчиком, заушным индуктором и индукционной петлей; система информационная для слабослышащих переносная «Исток» А2 со встроенным плеером – звуковым информатором; документ-камера; программируемые слуховые аппараты индивидуального пользования.

3. Ассистивные информационные технологии: программное обеспечение экранного доступа с синтезом речи NVDA; программы экранного увеличения; программы речевого синтеза для компьютеров и ноутбуков; программы речевого синтеза для мобильных устройств; экранная клавиатура; экранная лупа.

При необходимости для обучающихся с нарушениями зрения на рабочих местах для проведения практических или лабораторных занятий устанавливается специальное программное обеспечение (программа речевой навигации NVDA, речевые синтезаторы, экранные лупы).



В учебные аудитории обеспечивается беспрепятственный доступ для обучающихся инвалидов и обучающихся с ограниченными возможностями здоровья. В каждой аудитории, где обучаются инвалиды и лица с ограниченными возможностями здоровья, предусматривается соответствующее количество мест для обучающихся с учетом нарушений их здоровья.

Для освоения дисциплины инвалидам и лицам с ограниченными возможностями здоровья предоставляется доступ к печатным источникам, имеющимся в научной библиотеке ЧелГУ, с помощью специальных технических средств; доступ к электронным источникам, представленным в форме электронного документа в фонде научной библиотеки ЧелГУ или электронно-библиотечных системах, с помощью специальных технических и программных средств (рабочее место для незрячего пользователя с программным обеспечением экранного доступа с синтезом речи NVDA, рабочее место с компьютерным роллером и клавиатурой Clevy с большими кнопками и с разделяющей клавиши накладкой).

Учебно-методические материалы для обучающихся из числа инвалидов и лиц с ограниченными возможностями здоровья предоставляются в формах, адаптированных к ограничениям их здоровья и восприятия информации:

Для лиц с нарушениями зрения:

- в печатной форме увеличенным шрифтом,
- в форме электронного документа,
- в форме аудиофайла,
- в печатной форме шрифтом Брайля.

Для лиц с нарушениями слуха:

- в печатной форме,
- в форме электронного документа.

Для лиц с нарушениями опорно-двигательного аппарата:

- в печатной форме,
- в форме электронного документа,
- в форме аудиофайла.

Данный перечень может быть конкретизирован в зависимости от контингента обучающихся.

Для инвалидов и лиц с ограниченными возможностями здоровья освоение дисциплины может быть частично или полностью осуществлено с использованием дистанционных образовательных технологий (Moodle, Adobe Connect Pro и пр.).

В освоении дисциплины инвалидами и лицами с ограниченными возможностями здоровья используется индивидуальная работа. Под индивидуальной работой подразумевается две формы взаимодействия с преподавателем: индивидуальная учебная работа (консультации), т.е. дополнительное разъяснение учебного материала и углубленное изучение материала с теми обучающимися, которые в этом заинтересованы, и индивидуальная воспитательная работа. Индивидуальные консультации направлены на индивидуализацию обучения и установлению воспитательного контакта между преподавателем и обучающимся инвалидом или обучающимся с ограниченными возможностями здоровья.

При проведении процедуры оценивания результатов обучения инвалидов и лиц с ограниченными возможностями здоровья по дисциплине обеспечивается выполнение следующих дополнительных требований в зависимости от индивидуальных особенностей, обучающихся:

а) инструкция по порядку проведения процедуры оценивания предоставляется в доступной форме (устно, в письменной форме, в письменной форме шрифтом Брайля, устно с использованием услуг сурдопереводчика);

б) доступная форма предоставления заданий оценочных средств (в печатной форме, в печатной форме увеличенным шрифтом, в печатной форме шрифтом Брайля, в форме электронного документа, задания зачитываются ассистентом, задания предоставляются с использованием сурдоперевода);

в) доступная форма предоставления ответов на задания (письменно на бумаге, набор ответов на компьютере, письменно шрифтом Брайля, с использованием услуг ассистента, устно).

При проведении процедуры оценивания результатов обучения инвалидов и лиц с ограниченными возможностями здоровья предусматривается использование технических средств, необходимых им в связи с их индивидуальными особенностями. Эти средства могут быть предоставлены ЧелГУ или могут использоваться собственные технические средства. При необходимости инвалидам и лицам с ограниченными возможностями здоровья предоставляется дополнительное время для подготовки ответа на задания, процедура оценивания результатов обучения по дисциплине может проводиться в несколько этапов.

Проведение процедуры оценивания результатов обучения инвалидов и лиц с ограниченными возможностями здоровья допускается с использованием дистанционных образовательных технологий.

Лабораторная работа № 1.

Организация сбора и хранения больших наборов данных.

1. Прочтите данные из файла `data.csv`

Функции, которые могут пригодиться при решении: `pd.read_csv(..., delimiter=',')`

2. Выведите описание прочтенных данных.

Функции, которые могут пригодиться при решении: `.describe()`

3. Отобразите несколько первых и несколько последних записей.

Функции, которые могут пригодиться при решении: `.head()`, `.tail()`

Какие параметры можно передать этим функциям?

4. Прочтите в файле `DataDictionary-ru.txt`, что означают столбцы матрицы. Какому типу принадлежит каждый столбец (вещественный, целый, категориальный)?

Заметьте, что столбец `DebtRatio` содержит неправдоподобные данные. Только значения, соответствующие известному месячному доходу, являются отношениями. Остальные – абсолютные значения месячных выплат процентов.

5. Исправьте данные, сделав все значения столбца `DebtRatio` абсолютными (умножьте их на `MonthlyIncome`). Чтобы ваша программа быстро работала на полных данных, постарайтесь не использовать цикл.

Лабораторная работа № 2

. Основные компоненты статистической среды R

Цель работы: научиться загружать наборы данных, изменять тип данных, делать срезы данных по условию.

Задание

1. Используя ресурс kaggle: <https://www.kaggle.com/>, выбрать один из наборов данных. Загрузить этот набор в рабочую директорию. Считать данные, определить тип данных.

2. Описать данные набора: какие переменные в нем присутствуют, какой тип данных у этих переменных.

Контрольные вопросы и задания

1. Как выполняется считывание данных?
2. Какие проблемы при импорте данных могут возникнуть, какие способы их решения могут быть использованы?
3. Есть ли в наборе данных количественная дискретная переменная?
4. Есть ли в данных категориальные переменные? Как преобразовать их в факторы?
5. Есть ли в данных количественные переменные? Какой у них тип?
6. Как построить срез данных по условию? Как провести сортировку набора данных по одному из столбцов?

Лабораторная работа № 3

Работа с данными в Pandas

1. Прочтите данные из файла `data.csv`
2. Выведите описание прочтенных данных.
Функции, которые могут пригодиться при решении: `.describe()`
3. Отобразите несколько первых и несколько последних записей.
4. Прочтите в файле `DataDictionary-ru.txt`, что означают столбцы матрицы.
5. Исправьте данные, сделав все значения столбца `DebtRatio` абсолютными (умножьте их на `MonthlyIncome`).
6. Вычислите средний ежемесячный доход и присвойте всем клиентам с неизвестным доходом полученное число.
7. Используя метод `groupby`, оцените вероятности невозврата кредита (`SeriousDlqin2yrs=1`) для различных значений количества иждивенцев (`NumberOfDependents`).
8. Прodelайте аналогичную процедуру для различных значений столбца `NumberRealEstateLoansOrLines`

Лабораторная работа № 4

Импорт и преобразование данных

Задание:

1. Скачать файл по ссылке: <https://www.kaggle.com/mirichoi0218/insurance>
2. Загрузить в рабочее пространство набор данных из файла `insurance.csv` предварительно размещенного в рабочей директории.
3. Построить срезы на основе отбора строк и столбцов и на основе фильтров на загруженном наборе данных.
4. Упорядочить данные по указанному столбцу в соответствии с вариантом.
5. Вывести первые / последние 10 строк набора данных.
6. Выполнить индивидуальное задание
7. Оформить отчет

Индивидуальное задание

1. Скачать файл по ссылке:
https://raw.githubusercontent.com/Pozdniakov/tidy_stats/master/data/heroes_information.csv
2. Выполнить импорт данных и проверить их (в зависимости от варианта: импорт простой; импорт таблиц в бинарном формате: таблицы Excel, SPSS; импорт данных из Google Sheets; быстрый импорт данных).

Лабораторная работа № 5

Библиотеки NumPy, SciPy: основные функции

Задание:

а) Решить СЛАУ при помощи NumPy.

б) Построить график функции и её фурье-образа (модуль, фаза, Re/Im-части).

в) Разложить дробно-рациональную функцию на множители, на слагаемые (в символьном виде).

Порядок выполнения работы:

а) Решить СЛАУ при помощи NumPy.

Каждый обучаемый получает свой вариант задания открытого типа с системой линейных алгебраических уравнений (5–10 порядка).

Необходимо написать сценарий на языке Python, в котором коэффициенты СЛАУ вводятся в виде матрицы NumPy и выполняется вызов `numpy.linalg.solve`. Полученный результат необходимо проверить (перемножить матрицы).

Результатом работы по данному пункту является листинг сценария и численный ответ – решение предложенной СЛАУ.

б) Построить график функции и её Фурье-образа (модуль, фаза, Re/Im-части). Каждый обучаемый получает свой вариант задания открытого типа с функцией, заданной в аналитическом виде.

Необходимо написать сценарий на языке Python, в котором для заданной функции генерируется массив отсчётов. Над этим массивом отсчётов выполняется БПФ (`numpy.fft`).

Полученные массивы используются для построения графиков при помощи библиотеки Matplotlib.

Результатом работы по данному пункту является листинг сценария и графики: исходной функции, модуля Фурье-образа, аргумента Фурье-образа, вещественной части Фурье-образа, мнимой части Фурье-образа.

в) Разложить дробно-рациональную функцию на множители, на слагаемые (в символьном виде).

Каждый обучаемый получает свой вариант задания открытого типа с дробно-рациональной функцией, заданной в аналитическом виде.

Необходимо написать сценарий на языке Python, в котором заданная функция вводится в виде полиномов (`numpy.poly1d`). Путём нахождения корней полиномов числителя и знаменателя (`numpy.roots`) функция записывается в виде произведения дробно-рациональных функций меньшего порядка (это полезно при моделировании БИХ-фильтров, представляя их в виде каскадной схемы). Путём разложения дробно-рациональной функции на простые дроби (`scipy.signal.residue`) исходная функция записывается в виде суммы дробно-рациональных функций меньшего порядка (это полезно при моделировании БИХ-фильтров, представляя их в виде параллельной схемы).

Результатом работы по данному пункту является листинг сценария и аналитическая запись разложений заданной функции.

Варианты заданий к пункту А

<p>1) $\begin{cases} 3x_1 + x_2 = 1 \\ x_1 + 2x_2 + x_3 = 4 \\ x_2 + 3x_3 + x_4 = 10 \\ x_3 + 2x_4 + x_5 = 10 \\ x_4 + 3x_5 + x_6 = 10 \\ x_5 + 2x_6 + x_7 = 4 \\ x_6 + 3x_7 = 1 \end{cases}$</p>	<p>6) $\begin{cases} 4x_1 + x_2 = 6 \\ x_1 + 4x_2 + x_3 = 10 \\ x_2 + 4x_3 + x_4 = 8 \\ x_3 + 4x_4 + x_5 = 10 \\ x_4 + 4x_5 + x_6 = 8 \\ x_5 + 4x_6 + x_7 = 6 \\ x_6 + 4x_7 = 5 \end{cases}$</p>
<p>2) $\begin{cases} 12x_1 + x_2 + x_3 + x_4 + x_5 + x_6 = 16 \\ x_1 + 12x_2 + x_3 + x_4 + x_5 + x_6 = 16 \\ x_1 + x_2 + 12x_3 + x_4 + x_5 + x_6 = 16 \\ x_1 + x_2 + x_3 + 12x_4 + x_5 + x_6 = 16 \\ x_1 + x_2 + x_3 + x_4 + 12x_5 + x_6 = 16 \\ x_1 + x_2 + x_3 + x_4 + x_5 + 12x_6 = 5 \end{cases}$</p>	<p>7) $\begin{cases} 3x_1 + x_2 + x_3 + x_4 + x_5 + x_6 = 2 \\ x_1 + 3x_2 + x_3 + x_4 + x_5 + x_6 = 4 \\ x_1 + x_2 + 3x_3 + x_4 + x_5 + x_6 = 6 \\ x_1 + x_2 + x_3 + 3x_4 + x_5 + x_6 = -6 \\ x_1 + x_2 + x_3 + x_4 + 3x_5 + x_6 = -4 \\ x_1 + x_2 + x_3 + x_4 + x_5 + 3x_6 = -2 \end{cases}$</p>
<p>3) $\begin{cases} 5x_1 + 0.1x_2 + 0.1x_3 + 0.1x_4 + 0.1x_5 + 0.1x_6 = 7 \\ 0.1x_1 + 5x_2 + 0.1x_3 + 0.1x_4 + 0.1x_5 + 0.1x_6 = 11.9 \\ 0.1x_2 + 5x_3 + 0.1x_4 + 0.1x_5 + x_6 = 16.7 \\ 0.1x_3 + 5x_4 + 0.1x_5 + 0.1x_6 = 21.4 \\ 0.1x_4 + 5x_5 + 0.1x_6 = 26 \\ 0.1x_5 + 5x_6 = 30.5 \end{cases}$</p>	<p>8) $\begin{cases} 3x_1 - x_2 = -1 \\ -x_1 + 3x_2 - x_3 = 2 \\ -x_2 + 3x_3 - x_4 = 2 \\ -x_3 + 3x_4 - x_5 = -2 \\ -x_4 + 3x_5 - x_6 = 2 \\ -x_5 + 3x_6 = 2 \end{cases}$</p>
<p>4) $\begin{cases} 2x_1 - x_2 = 0 \\ -x_1 + 2x_2 - x_3 = 2 \\ -x_2 + 2x_3 - x_4 = -2 \\ -x_3 + 2x_4 - x_5 = 2 \\ -x_4 + 2x_5 - x_6 = -2 \\ -x_5 + 2x_6 = 3 \end{cases}$</p>	<p>9) $\begin{cases} x_1 + 0.1x_2 + 0.1x_3 + 0.1x_4 + 0.1x_5 + 0.1x_6 = 0.3 \\ 0.2x_1 + x_2 + 0.1x_3 + 0.1x_4 + 0.1x_5 + 0.1x_6 = 0.4 \\ 0.2x_2 + x_3 + 0.1x_4 + 0.1x_5 + x_6 = 0.49 \\ 0.2x_3 + x_4 + 0.1x_5 + 0.1x_6 = 0.57 \\ 0.2x_4 + x_5 + 0.1x_6 = 0.64 \\ 0.2x_5 + x_6 = 0.7 \end{cases}$</p>
<p>5) $\begin{cases} 3x_2 + 4x_3 + 5x_4 + 6x_5 + 7x_6 = 25 \\ 3x_1 + 5x_3 + 6x_4 + 7x_5 + 8x_6 = 29 \\ 4x_1 + 5x_2 + 7x_4 + 8x_5 + 9x_6 = 33 \\ 5x_1 + 6x_2 + 7x_3 + 9x_5 + 10x_6 = 37 \\ 6x_1 + 7x_2 + 8x_3 + 9x_4 + 11x_6 = 41 \\ 7x_1 + 8x_2 + 9x_3 + 10x_4 + 11x_5 = 45 \end{cases}$</p>	<p>10) $\begin{cases} x_2 + 2x_3 + 3x_4 + 4x_5 + 5x_6 = 70 \\ x_1 + 2x_3 + 3x_4 + 4x_5 + 5x_6 = 69 \\ x_1 + 2x_2 + 3x_4 + 4x_5 + 5x_6 = 67 \\ x_1 + 2x_2 + 3x_3 + 4x_5 + 5x_6 = 64 \\ x_1 + 2x_2 + 3x_3 + 4x_4 + 5x_6 = 60 \\ x_1 + 2x_2 + 3x_3 + 4x_4 + 5x_5 = 55 \end{cases}$</p>

Варианты к пункту Б

1)	$f(x) = 0.5e^{-(x-1)^2} + 0.7e^{-(x+1)^2}$	6)	$f(x) = \frac{0.8}{\sqrt{0.1 + 0.5(x-1)^2}} + \frac{0.6}{\sqrt{0.1 + 0.4(x+1)^2}}$
2)	$f(x) = \frac{1.2x^2 - 0.4x + 2.4}{x^4 + 4}$	7)	$f(x) = 0.7e^{-2(x-1)^2} + 0.3e^{-3(x+1)^2}$
3)	$f(x) = \frac{0.5}{\sqrt{0.1 + (x-1)^2}} + \frac{0.7}{\sqrt{0.1 + (x+1)^2}}$	8)	$f(x) = \frac{2.7x^2 + 3x + 3.7}{6x^4 - 7x^2 + 2x + 12}$
4)	$f(x) = 0.8e^{-0.5(x-1)^2} + 0.6e^{-0.4(x+1)^2}$	9)	$f(x) = \frac{0.7}{\sqrt{0.1 + 2(x-1)^2}} + \frac{0.3}{\sqrt{0.1 + 3(x+1)^2}}$
5)	$f(x) = \frac{0.62x^2 + 0.04x + 2.02}{0.2x^4 + 0.5x^2 - 0.2x + 2.1}$	10)	$f(x) = \frac{5x^2 + 4x + 6}{21x^4 - 32x^2 + 8x + 32}$

Варианты к пункту В

1)	$H(z) = \frac{z^3 + 3z^2 + 3z + 1}{z^3 + 0.351z^2 + 0.195z + 0.013}$	6)	$H(z) = \frac{z^3 + 1.787z^2 + 1.787z + 1}{z^3 + 0.768z^2 + 0.974z + 0.131}$
2)	$H(z) = \frac{z^3 + 1.865z^2 + 1.865z + 1}{z^3 - 0.00085z^2 + 0.421z + 0.0343}$	7)	$H(z) = \frac{z^3 + 3z^2 + 3z + 1}{z^3 + 2.616z^2 + 2.290z + 0.670}$
3)	$H(z) = \frac{z^3 + 0.695z^2 + 0.695z + 1}{z^3 - 1.251z^2 + 1.281z - 0.532}$	8)	$H(z) = \frac{z^3 + 0.652z^2 + 0.652z + 1}{z^3 - 1.145z^2 + 0.727z - 0.1205}$
4)	$H(z) = \frac{z^3 + 3z^2 + 3z + 1}{z^3 + 2.263z^2 + 1.736z + 0.449}$	9)	$H(z) = \frac{z^3 - 0.614z^2 - 0.614z + 1}{z^3 - 2.281z^2 + 2.070z - 0.662}$
5)	$H(z) = \frac{z^3 + 2.706z^2 + 2.706z + 1}{z^3 + 1.415z^2 + 0.959z + 0.236}$	10)	$H(z) = \frac{z^3 + 3z^2 + 3z + 1}{z^3 - 1.499z^2 + 1.074z - 0.296}$

Лабораторная работа № 6

Визуализация данных с Matplotlib и Pandas

Задание:

1. Подберите набор данных на ресурсах [3, 4] и согласуйте свой выбор с преподавателем. Студент может предложить синтезированный набор данных.
2. Проведите первичный анализ данных. Особое внимание следует уделить графическому представлению распределений признаков, визуализации взаимосвязей, позволяющие судить о наборе данных. Построение графиков желательно произвести по нескольким проекциям. При анализе данных использовать как можно более разнообразные типы графиков.

Примечания

3.2 Важные замечания

1. Статья о типах графиков при первичном анализе данных:
<https://medium.com/open-machine-learning-course/open-machine-learning-course-topic-2-visual-data-analysis-in-python-846b989675cd>
2. При выборе набора данных (data set) на ресурсах [3, 4] необходимо согласовать свой выбор с другими студентами группы и преподавателем, так как работа над одинаковыми наборами данных недопустима.
3. В рамках данной лабораторной работы рекомендуется использовать инструментарий Python (библиотеки, среду разработки) для решения поставленных задач.

Лабораторная работа № 7

Визуализация данных. Базовые графики в R.

Задание:

1. Подберите набор данных на ресурсах [3, 4] и согласуйте свой выбор с преподавателем. Студент может предложить синтезированный набор данных.
2. Проведите первичный анализ данных. Особое внимание следует уделить графическому представлению распределений признаков, визуализации взаимосвязей, позволяющие судить о наборе данных. Построение графиков желательно произвести по нескольким проекциям. При анализе данных использовать рассмотренные на лекции функции.

Примечания

1. Статья о типах графиков при первичном анализе данных:
<https://medium.com/open-machine-learning-course/open-machine-learning-course-topic-2-visual-data-analysis-in-python-846b989675cd>
2. При выборе набора данных (data set) на ресурсах [3, 4] необходимо согласовать свой выбор с другими студентами группы и преподавателем, так как работа над одинаковыми наборами данных недопустима.
3. В рамках данной лабораторной работы рекомендуется использовать инструментарий Python (библиотеки, среду разработки) для решения поставленных задач.

Лабораторная работа № 8-9

Динамические визуализации в R. Создание интерактивных визуализаций и отчетов с помощью инструмента Plotly.

Задание:

1. Скачать набор данных:

https://raw.githubusercontent.com/Pozdniakov/tidy_stats/master/data/heroes_information.csv

2. Загрузить этот набор в рабочую директорию. Считать данные, определить тип данных.

3. Выполнить динамическую визуализацию с помощью пакета `plotly`.

Примечания

1. Студент может предложить свой набор данных.

2. При выборе набора данных (data set) на других ресурсах необходимо согласовать свой выбор с другими студентами группы и преподавателем, так как работа над одинаковыми наборами данных недопустима.

3. В рамках данной лабораторной работы рекомендуется использовать инструментарий Python (библиотеки, среду разработки) для решения поставленных задач.

Лабораторная работа № 10

Обзор языка разметки Markdown. Экспорт данных и отчетов в Python.

Задание:

1. Сделать отчет по предыдущей лабораторной работе (Python) в формате Markdown.
2. В качестве отчета по ЛР10 предоставить отчет по предыдущей лабораторной работе в трех форматах: docx, pdf и md (в zip-архиве). Отчет должен содержать скриншоты, makefile и др.
3. Ответить на вопросы преподавателя.

Содержание отчета

Сформированный отчет должен содержать:

- 1) титульный лист с указанием номера лабораторной работы и наименования, ФИО студента и ФИО преподавателя;
- 2) формулировку задания;
- 3) описание результатов выполнения работы:
 - скриншоты, фиксирующие выполнение лабораторной работы;
 - ответы на вопросы;
- 4) выводы, которые должны быть согласованы с заданием.

Лабораторная работа № 11-12

Создание отчетов. R Markdown и Quarto

Задание:

1. Сделать отчет по предыдущей лабораторной работе (R) в формате Markdown.

2. В качестве отчета по ЛР10 предоставить отчет по предыдущей лабораторной работе в трех форматах: .docx, .pdf и .rmd (в zip-архиве). Отчет должен содержать скриншоты, makefile и др.

3. В отчеты должны быть использованы заголовки разных уровней (где уместно), списки, цитаты, таблица (небольшая из вашего датасета). Можно использовать графики и динамические таблицы.

4.

5. Ответить на вопросы преподавателя.

Содержание отчета

Сформированный отчет должен содержать:

- 1) титульный лист с указанием номера лабораторной работы и наименования, ФИО студента и ФИО преподавателя;
- 2) формулировку задания;
- 3) описание результатов выполнения работы:
 - скриншоты, фиксирующие выполнение лабораторной работы;
 - ответы на вопросы;
- 4) выводы, которые должны быть согласованы с заданием.

Замечания:

1. Если вы выбрали в качестве формата PDF (презентацию или документ), то вам понадобится еще установить LaTeX на компьютер.

Это тоже можно сделать с помощью специального пакета:

```
install.packages('tinytex')
tinytex::install_tinytex() # install TinyTeX
```

2. В RStudio есть подсказка по синтаксису Markdown, для ее вызова нужно нажать Help - Markdown Quick Reference.

3.

Лабораторная работа № 13

Классические методы и критерии статистики

Задание:

1. Построить гистограмму плотности случайной величины и выдвинуть три разные гипотезы о законе её распределения.
2. Проверить выдвинутые гипотезы с помощью критерия хи-квадрат.
3. Самостоятельно найти функцию для проверки гипотезы о законе распределения случайной величины с помощью критерия Колмогорова-Смирнова в библиотеке `scipy.stats`. Применить её для рассматриваемых данных и сравнить результаты.

Лабораторная работа №14

Регрессионные модели зависимостей между количественными переменными

Задание:

1. Подберите набор данных на ресурсах [5-7] и согласуйте свой выбор с преподавателем. Студент может предложить синтезированный набор данных.
2. Постройте модель регрессии на основе универсального пайплайна.
3. Постройте модель многомерной регрессии с использованием стратегии `backwardelimination`.

Лабораторная работа № 15

Модели для оценки временных рядов

Цель работы: научиться использовать готовые решения по прогнозированию временных рядов на языке R.

Задание

- 1) загрузить данные из csv-файла;
- 2) преобразовать данные во временной ряд;
- 3) провести прогноз временного ряда;
- 4) вывести график с прогнозом;
- 5) рассчитать оценки качества прогноза.

Лабораторная работа № 16

Многомерные методы анализа данных

Задание

1) подготовить данные, рассчитать средние и стандартные отклонения, рассчитать матрицу евклидовых расстояний, выполнить кластеризацию по методу ближайшего соседа, построить дендрограмму.

2) рассчитать суммы внутригрупповых расстояний, показать на графике;

3) написать функцию, которая строит модель для заданного числа кластеров, проводит анализ этой модели и строит некоторые графики;

4) оценить качество разбиения на группы по методу главных компонент.

