



CS634

ИНСТРУМЕНТИ ЗА ВИЗУАЛИЗИРАНЕ И АНАЛИЗ НА ДАННИ

Приета: прот. №29 от 29.06.2023 г.

Преподавател: ас. д-р Александър Иванов

АНОТАЦИЯ

В курса се представят програмни средства за обработка, визуализация и анализ на данни. Акцентът е върху езика Python.

ОСНОВНИ ЦЕЛИ

Цел на настоящия курс е да представи на студентите основните средства на езика Python за математически изчисления, визуализиране на числови данни, обработка на изображения, техники за анализ на данни. По-конкретно, студентите трябва:

- да разберат същността на основните синтактични правила на езика Python;
- да усвоят основните библиотеки за математически изчисления
- да придобият практически умения за визуализиране и обработка на различни типове данни

ПРЕДПОСТАВКИ

За успешно усвояване на учебния курс са необходими знания по: Програмиране, Вероятности и статистика, Приложно статистическо моделиране .

СТАТУТ И СТРУКТУРА

специалност	статут	Кредити	редовно обучение				задочно обучение			
			л	с	у	общ	л	с	у	общ
АД	Задължителна	6	30	30		60	15	15		30

СЪДЪРЖАНИЕ НА КУРСА

А. Лекции

Тема 1. Езикът Python – общи характеристики. Структура и изпълнение на проста програма на Python. Среди за разработка. Примитивни типове от данни: константи, променливи и операции. Преобразуване на стойностите между примитивни типове. Базов вход и изход.

Тема 2. Оператори. Условни оператори — if, else. Вложени проверки. Логически изрази. Оператори за цикли — for, while. Вложени цикли.

Тема 3. Линейни структури от данни — структура List (листа, списък). Основни операции с листи. Структура речник. Структура множество.

Тема 4. Функции. Математическа библиотека math. Работа със случайни числа чрез библиотеката random. Декоратори. Ламбда-функции

Тема 5. Пакети за визуализация на данни — matplotlib. Изчертаване на различни видове графики. Изчертаване на графики на математически функции

Тема 6. Библиотека NumPy — вектори, матрици, операции с вектори и матрици. Манипулации на масиви от данни. Генериране на данни. Символна математика с библиотека SymPy

Тема 7. Изключения (Exception). Обработка на изключения. Работа с файлове. CSV файлов формат. Пакет Pandas. Структура dataframe. Мащабирне на числови данни. Етикиране на данни

Тема 8. Работа с изображения — библиотека PIL Методи за обработка на изображения: ротация, преоразмеряване, изрязване, цветови манипулации и др.

Тема 9. Езикът R - Основни положения. Визуализация на данни. Статистически инструменти

Тема 10. Корелационен анализ — библиотека Scikit-learn. Анализ на главни компоненти. Инструменти за намаляване на размерност

Тема 11. Клъстърен анализ — алгоритми за клъстеризация. Мярка «силует». Инструменти в пакета Scikit-learn. Диаграми на Вороной.

Тема 12. Генератори и итератори.

В. Семинарни упражнения

Тема 1. Вход и изход. Основни примитивни типове данни, аритметични, логически изрази. Преобразуване на стойностите между примитивни типове.

Тема 2. Създаване на прости програми с проверки. Решаване на задачи чрез цикли.

Тема 3. Упражнение върху линейни структури от данни

Тема 4. Упражнение върху функции. Изчертаване на графики на математически функции

Тема 5. Четене на стойности от файлове. Корелационен анализ на данни

Тема 6. Алгебра с програмни средства

Тема 7. Манипулации на изображения

Тема 8. Статистически анализи на данни чрез R

Тема 9. Обработка на данни с алгоритми за машинно обучение — клъстеризация, намаляване на размерност и др.

Тема 10. Работа по курсова задача

С. Курсова задача

Изпълнението на курсовата работа започва с получаване от всеки студент на задание за курсово проектиране. Темата се уточнява с преподавателя. Студентът трябва да реализира приложение, реализиращо съответната функционалност от заданието.

ПЛАНИРАНИ УЧЕБНИ ДЕЙНОСТИ И МЕТОДИ ЗА ОБУЧЕНИЕ

Планираните учебни дейности в този курс са съсредоточени основно към трайното усвояване на учебния материал. За тази цел са предвидени:

- На първата лекция студентите получават пълна информация за съдържанието на програмата на дисциплината, изискванията за подготовката на дисциплината и конспекта;
- Лекциите са по въпроси от конспекта, обединени по теми, които завършват с обобщени въпроси за самоподготовка;
- Работа със студентите в малки групи (на семинарни упражнения) или единично с цел по-трайно усвояване на материала;
- Провеждане на междинни тестове, които показват нивото на студентите;
- В края на учебния материал се въвеждат курсови задачи с практически примери, като за всеки от студентите се изпълнява различно задание;
- Електронни материали подпомагат учебния процес на студентите и разширява възможностите на преподавателя при осъществяването на текущ контрол.

МЕТОДИ И КРИТЕРИИ ЗА ОЦЕНЯВАНЕ

Критерият за оценка е комплексен и се базира на следните фактори:

- Текущи тестове: 2 теста по 10 точки. Общо 20 точки
- Текущи задания: 2 броя практически задания по 10 точки. Общо 20 точки
- Курсова задача: 60 точки

СКАЛА ЗА ОЦЕНЯВАНЕ

<50 т. : Слаб 2

50-62т. : Среден 3

63-75т.: Добър 4

76-88т.: Мн. добър 5

89-100т.: Отличен 6

ПРЕПОРЪЧИТЕЛНА ЛИТЕРАТУРА

1. <https://www.learnpython.org>
2. <https://www.python.org/>
3. <https://www.programiz.com/python-programming>
4. <https://python-book.softuni.bg>
5. <https://github.com/SoftUni/Programming-Basics-Book-Python-BG>
6. <https://www.tutorialspoint.com/python/index.htm>
7. <https://greenteapress.com/wp/think-python/>
8. <https://matplotlib.org/>
9. <https://pandas.pydata.org/>
10. <https://scikit-learn.org/stable/tutorial/index.html>
11. <https://www.w3schools.com/r/>



CS634

ИНСТРУМЕНТИ ЗА ВИЗУАЛИЗИРАНЕ И АНАЛИЗ НА ДАННИ

Приета: прот. №29 от 29.06.2023 г.

Преподавател: ас. д-р Александър Иванов

Тема 1. Езикът Python – общи характеристики. Структура и изпълнение на проста програма на Python. Среди за разработка. Примитивни типове от данни: константи, променливи и операции. Преобразуване на стойностите между примитивни типове. Базов вход и изход.

Тема 2. Оператори. Условни оператори — if, else. Вложени проверки. Логически изрази. Оператори за цикли — for, while. Вложени цикли.

Тема 3. Линейни структури от данни — структура List (листа, списък). Основни операции с листи. Структура речник. Структура множество.

Тема 4. Функции. Математическа библиотека math. Работа със случайни числа чрез библиотеката random. Декоратори. Ламбда-функции

Тема 5. Пакети за визуализация на данни — matplotlib. Изчертаване на различни видове графики. Изчертаване на графики на математически функции

Тема 6. Библиотека NumPy — вектори, матрици, операции с вектори и матрици. Манипулации на масиви от данни. Генериране на данни. Символна математика с библиотека SymPy

Тема 7. Изключения (Exception). Обработка на изключения. Работа с файлове. CSV файлов формат. Пакет Pandas. Структура dataframe. Мащабирне на числови данни. Етикиране на данни

Тема 8. Работа с изображения — библиотека PIL Методи за обработка на изображения: ротация, преоразмеряване, изрязване, цветови манипулации и др.

Тема 9. Езикът R - Основни положения. Визуализация на данни. Статистически инструменти

Тема 10. Корелационен анализ — библиотека Scikit-learn. Анализ на главни компоненти. Инструменти за намаляване на размерност

Тема 11. Клъстърен анализ — алгоритми за клъстеризация. Мярка «силует». Инструменти в пакета Scikit-learn. Диаграми на Вороной.

Тема 12. Генератори и итератори.

ПРЕПОРЪЧИТЕЛНА ЛИТЕРАТУРА

1. <https://www.learnpython.org>
2. <https://www.python.org/>
3. <https://www.programiz.com/python-programming>
4. <https://python-book.softuni.bg>
5. <https://github.com/SoftUni/Programming-Basics-Book-Python-BG>
6. <https://www.tutorialspoint.com/python/index.htm>
7. <https://greenteapress.com/wp/think-python/>
8. <https://matplotlib.org/>
9. <https://pandas.pydata.org/>
10. <https://scikit-learn.org/stable/tutorial/index.html>
11. <https://www.w3schools.com/r/>