

Министерство науки и высшего образования Российской Федерации
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ (НИ ТГУ)

Институт прикладной математики и компьютерных наук

УТВЕРЖДЕНО:
Директор
А. В. Замятин



Рабочая программа дисциплины

Представление знаний и визуализация данных

по направлению подготовки

02.04.02 Фундаментальная информатика и информационные технологии

Направленность (профиль) подготовки:

Математика беспроводных сетей связи и интернета вещей

Форма обучения

Очная

Квалификация

Магистр

Год приема

2024

СОГЛАСОВАНО:
Руководитель ОП
С.П. Моисеева



Председатель УМК
С.П. Сущенко



Томск – 2024

1. Цель и планируемые результаты освоения дисциплины

Целью освоения дисциплины является формирование следующих компетенций:

ОПК-4 Способен оптимальным образом комбинировать существующие информационно-коммуникационные технологии для решения задач в области профессиональной деятельности с учетом требований информационной безопасности.

ПК-4 Способен управлять получением, хранением, передачей, обработкой больших данных.

Результатами освоения дисциплины являются следующие индикаторы достижения компетенций:

ИОПК-4.1 Анализирует задачи профессиональной деятельности средствами информационных технологий

ИПК-4.1 Осуществляет мониторинг и оценку производительности обработки больших данных

ИПК-4.2 Использует методы и инструменты получения, хранения, передачи, обработки больших данных

2. Задачи освоения дисциплины

- Освоить аппарат решения прикладных профессиональных задач с использованием методов и средств визуализации данных и знаний;
- Научиться анализировать задачи профессиональной деятельности средствами информационных технологий;
- Научиться интерпретировать результаты визуализации и грамотно их описывать;
- Научиться применять понятийный аппарат при использовании библиотек современных скриптовых языков (Python, R) для визуализации данных и знаний.

3. Место дисциплины в структуре образовательной программы

Дисциплина относится к Блоку 1 «Дисциплины (модули)».

Дисциплина относится к части образовательной программы, формируемой участниками образовательных отношений, предлагается обучающимся на выбор. Дисциплина входит в модуль Введение в интеллектуальный анализ больших данных.

4. Семестр(ы) освоения и форма(ы) промежуточной аттестации по дисциплине

Второй семестр, зачет

5. Входные требования для освоения дисциплины

Для успешного освоения дисциплины требуются компетенции, сформированные в ходе освоения образовательных программ предшествующего уровня образования.

6. Язык реализации

Русский

7. Объем дисциплины

Общая трудоемкость дисциплины составляет 3 з.е., 108 часов, из которых:

-лекции: 16 ч.

-лабораторные: 16 ч.

Объем самостоятельной работы студента определен учебным планом.

8. Содержание дисциплины, структурированное по темам

Тема 1. Концептуальные основы визуализации данных.

Концепции и цели визуализации, задачи визуализации данных, классификация по визуализируемым объектам. Выбор правильного типа визуализации.

Тема 2. Способы и инструменты визуализации данных. Библиотеки Python (или R) для решений задач визуализации. Правильное оформление визуализации. Сложная визуализация – несколько графиков на одном полотне. Сравнение полученных визуальных образов. Первичный анализ данных с использованием методов визуализации

Тема 3. Управление компоновкой диаграмм. Интерактивная визуализация. Примеры задач с использованием интерактивной визуализации. Библиотеки Python, поддерживающие интерактивную визуализацию.

Тема 4. Способы визуализации многомерных данных. Визуализация многомерных данных с использованием параллельных координат, лиц Чернова, диаграмм Эндрюса. Способы снижения размерности в визуализации многомерных данных.

Тема 5. Научная визуализация. Визуализация в научных исследованиях. Применение визуализации в научных исследованиях, примеры, разбор кейсов.

9. Текущий контроль по дисциплине

Текущий контроль по дисциплине проводится путем контроля посещаемости, проведения контрольных работ, тестов по лекционному материалу, выполнения лабораторных работ и фиксируется в форме контрольной точки не менее одного раза в семестр.

Оценочные материалы текущего контроля размещены на сайте ТГУ в разделе «Информация об образовательной программе» – <https://www.tsu.ru/sveden/education/eduop/>.

10. Порядок проведения и критерии оценивания промежуточной аттестации

Зачет во втором семестре проводится на основе оценки выполнения практического задания. Продолжительность зачета 1 час.

11. Учебно-методическое обеспечение

а) Электронный учебный курс по дисциплине в электронном университете «LMS IDO»

б) Оценочные материалы текущего контроля и промежуточной аттестации по дисциплине.

в) Методические указания по проведению лабораторных работ.

Выполнение лабораторной работы предусмотрено по каждой изучаемой теме дисциплины. Методические пособия размещены в локальной сети ИПМКН в каталоге X:\Workspace\Магистратура\ Представление знаний и визуализация данных.

г) Методические указания по организации самостоятельной работы студентов.

Для укрепления изученного материала предусмотрено выполнение индивидуального проекта в рамках часов самостоятельной работы. Проект может быть выполнен как индивидуально, так и в мини-группе (2-3 чел.), при условии, что объем работы также будет увеличен. В конце семестра по каждому проекту представляется мини-презентация о результатах работы.

12. Перечень учебной литературы и ресурсов сети Интернет

а) основная литература:

- Эрик Мэтиз. Изучаем Python: программирование игр, визуализация данных, веб-приложения. – СПб.: Питер, 2021. – 587 с.
- Маккинни, У. Python и анализ данных. – М.: ДМК Пресс, 2020. – 540 с.
- Мастицкий, С.Э., Шитиков В.К. Статистический анализ и визуализация данных с помощью R. – ДМК Пресс, 2015. – 496 с.

б) дополнительная литература

- Мастицкий, С.Э. Визуализация данных с помощью ggplot2. – М.: ДМК Пресс, 2017. – 222 с.

в) ресурсы сети Интернет:

- Общероссийская Сеть КонсультантПлюс Справочная правовая система.
<http://www.consultant.ru>

13. Перечень информационных технологий

а) лицензионное и свободно распространяемое программное обеспечение:

- Microsoft Office Standart 2013 Russian: пакет программ. Включает приложения: MS Office Word, MS Office Excel, MS Office PowerPoint, MS Office On-eNote, MS Office Publisher, MS Outlook, MS Office Web Apps (Word Excel MS PowerPoint Outlook);
- публично доступные облачные технологии (Google Docs, Яндекс диск и т.п.).
- средства программирования на Python3 и R.

б) информационные справочные системы:

- Представление знаний и визуализация данных
<https://moodle.ido.tsu.ru/course/view.php?id=1403>
- 7 самых популярных Python-библиотек для визуализации данных в Data Science
<https://medium.com/@bigdataschool/7-самых-популярных-python-библиотек-для-визуализации-данных-в-data-science-c0a4875df93b>
- Обзор пакетов для визуализации данных на Python
<https://medium.com/@alibekashirali/обзор-пакетов-для-визуализации-данных-на-python-a1525b2a5cc7>
- Каталог Визуализации данных <https://datavizcatalogue.com/>
- Зачем и как использовать визуализацию данных?
<https://infogra.ru/infographics/zachem-i-kak-ispolzovat-vizualizatsiyu-dannyh>
- Открытый курс машинного обучения
<https://habr.com/ru/company/ods/blog/323210/>
- Обзор методов визуализации многомерных данных
<https://cyberleninka.ru/article/n/obzor-metodov-vizualizatsii-mnogomernyh-dannyh>

в) профессиональные базы данных (*при наличии*):

- Искусственный интеллект и сферы его применения. Новости разработки квантовых компьютеров. Исследования искусственных нейронных сетей.
<https://ai-news.ru> Свободный доступ
- Портал мировых соревнований команд по DM <https://www.kaggle.com/>
Свободный доступ
- Репозиторий по машинному обучению <https://archive.ics.uci.edu/ml/index.php>
Свободный доступ

- Искусственный интеллект и сферы его применения. Новости разработки квантовых компьютеров. Исследования искусственных нейронных сетей. <https://ai-news.ru> Свободный доступ
- Портал мировых соревнований команд по DM <https://www.kaggle.com/> Свободный доступ
- Репозиторий по машинному обучению <https://archive.ics.uci.edu/ml/index.php> Свободный доступ

14. Материально-техническое обеспечение

Аудитории для проведения занятий лекционного типа.

Аудитории для проведения лабораторных занятий, индивидуальных и групповых консультаций, текущего контроля и промежуточной аттестации.

Помещения для самостоятельной работы, оснащенные компьютерной техникой и доступом к сети Интернет, в электронную информационно-образовательную среду и к информационным справочным системам.

Для материально-технического обеспечения дисциплины требуется наличие компьютерной техники с установленным соответствующим программным обеспечением и другого оборудования, поддерживающего проведение презентаций и выходом в сеть Интернет.

15. Информация о разработчиках

Марухина Ольга Владимировна, канд. техн. наук, доцент, кафедра теоретических основ информатики ТГУ, доцент