

МИНОБРНАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

Институт прикладной математики и компьютерных наук

УТВЕРЖДАЮ

Директор института прикладной
математики и компьютерных наук

А.В. Замятин

« 14 » ноября 2021 г.



Визуализация многомерных данных

рабочая программа дисциплины

Закреплена за кафедрой	<i>Теоретических основ информатики</i>
Учебный план	<i>02.03.02 Фундаментальная информатика и информационные технологии, профиль «Искусственный интеллект и разработка программных продуктов»</i>
Форма обучения	<i>очная</i>
Общая трудоёмкость	<i>2 з.е.</i>
Часов по учебному плану	<i>72</i>
в том числе:	
аудиторная контактная работа	<i>33,85</i>
самостоятельная работа	<i>38,15</i>
Вид(ы) контроля в семестрах	<i>6 семестр – зачет</i>

Программу составил:

канд. техн. наук,

доцент кафедры теоретических основ информатики



О.В. Марухина

Рецензент:

канд. техн. наук,

доцент кафедры теоретических основ информатики



С.В. Аксёнов

Рабочая программа дисциплины «Визуализация многомерных данных» разработана в соответствии с образовательным стандартом высшего образования – бакалавриат, самостоятельно устанавливаемым федеральным государственным автономным образовательным учреждением высшего образования «Национальный исследовательский Томский государственный университет» по направлению подготовки 02.03.02 Фундаментальная информатика и информационные технологии (Утвержден Ученым советом НИ ТГУ, протокол от 27.10.2021 г. № 08).

Рабочая программа одобрена на заседании кафедры теоретических основ информатики

Протокол от 04 июня 2021 г. № 05

Заведующий кафедрой теоретических основ информатики,
д-р техн. наук, профессор



А.В. Замятин

Рабочая программа одобрена на заседании учебно-методической комиссии института прикладной математики и компьютерных наук (УМК ИПМКН)

Протокол от 17 июня 2021 г. № 05

Председатель УМК ИПМКН,
д-р техн. наук, профессор



С.П. Сущенко

Цель освоения дисциплины

Цель – формирование у студентов теоретических знаний и практических навыков в области методов, средств, подходов и принципов визуального представления результатов научно-исследовательской деятельности, основанных на основных положениях интеллектуального анализа данных, машинного обучения и реализующихся в выборе инструментов и технологий, к которым можно отнести современные скриптовые языки Python и R.

1. Место дисциплины в структуре ОПОП

Дисциплина «Визуализация многомерных данных» относится к части, формируемой участниками образовательных отношений Блока 1 «Дисциплины», входит модуль «Искусственный интеллект».

Пререквизиты: «Введение в интеллектуальный анализ данных».

Постреквизиты: «Прикладные аспекты машинного обучения», «Технологическая (проектно-технологическая) практика (стационарная)».

2. Компетенции и результаты обучения, формируемые в результате освоения дисциплины

Таблица 1.

Компетенция	Индикатор компетенции	Код и наименование результатов обучения
ОПК-2. Способен применять компьютерные/супер компьютерные методы, современное программное обеспечение, в том числе отечественного происхождения, для решения задач профессиональной деятельности.	ИОПК-2.2. Использует методы высокопроизводительных вычислительных технологий, современного программного обеспечения, в том числе отечественного происхождения.	В результате освоения дисциплины студент будет Знать: – место и роль средств и методов визуализации в научных и прикладных исследованиях (ОР-2.2.1); Уметь: – проводить анализ результатов научного исследования с целью грамотного выбора методов и средств визуализации полученных данных и знаний (ОР-2.2.2); Владеть: – навыками использования библиотек современных скриптовых языков (Python, R) для визуализации данных и знаний (ОР-2.2.3);
	ИОПК-2.3. Использует инструментальные средства высокопроизводительных вычислений в научной и практической деятельности	В результате освоения дисциплины студент будет Знать: – связь методов визуализации и соответствующих математических моделей (ОР-2.3.1) . Уметь: – интерпретировать результаты визуализации, грамотно их описывать (ОР-2.3.2). Владеть: – навыками решения прикладных профессиональных задач с использованием методов и средств визуализации данных и знаний (ОР-2.3.3).

3. Структура и содержание дисциплины

3.1. Структура и трудоемкость видов учебной работы по дисциплине

Общая трудоемкость дисциплины составляет 2 зачетные единицы, 72 часов.

Таблица 2.

Вид учебной работы	Трудоемкость в академических часах	
Общая трудоемкость	72	72
Контактная работа:	33,85	33,85
Лекции (Л):	16	16
Практики (ПЗ)		
Лабораторные работы (ЛР)	16	16
Семинары (СЗ)		
Групповые консультации	1,6	1,6
Индивидуальные консультации		
Промежуточная аттестация	0,25	0,25
Самостоятельная работа обучающегося:	38,15	38,15
- выполнение проекта	14	27
- подготовка доклада, сообщения	6	10
- изучение учебного материала, публикаций	4	15
- подготовка к лабораторным	10	15
- подготовка к рубежному контролю по теме/разделу	4,15	7,15
Вид промежуточной аттестации	зачет	зачет

3.2. Содержание и трудоемкость разделов дисциплины

Таблица 3.

Код занятия	Наименование разделов и тем и их содержание	Вид учебной работы, занятий, контроля	Семестр	Часы в электронной форме	Всего (час.)	Литература	Код (ы) результата(ов) обучения
1.1.	Раздел 1. Визуализация многомерных данных. Задачи визуализации. Способы визуализации. Цель визуализации, задачи визуализации многомерных данных, классификация по визуализируемым объектам. Выбор правильного типа визуализации. Первичный анализ данных с использованием методов визуализации. Библиотеки Python (или R) для решений задач визуализации. Сравнение полученных визуальных образов. Выполнение лабораторной работы № 1.	Лекции	6		4	1-5	OP-2.2.1, OP-2.3.1, OP-2.2.2, OP-2.3.2
		ЛЗ			4		
		СРС			8,15		
2.1.	Раздел 2. Python-библиотеки для визуализации данных в Data Science . Возможности библиотек Matplotlib, Seaborn, Missingno, Altair, Plotly, Vokeh, Pygal, Networkx. Примеры использования перечисленных библиотек. Визуализация многомерных данных с использованием диаграмм Эндрюса. Выполнение лабораторной работы № 2.	Лекции	6		4	1-5	OP-2.2.1, OP-2.3.1, OP-2.2.2, OP-2.3.2
		ЛЗ			4		
		СРС			10		
	Текущий контроль успеваемости: мониторинг сдачи отчетов по лабораторным работам.	Конс.			1,6		OP-2.2.1, OP-2.3.1, OP-2.2.2, OP-2.3.2
3.1.	Раздел 3. Визуализация данных средствами дашбордов. Что из себя представляет дашборд, его свойства, отличие от отчета. Модули дашборда. Виды дашбордов. Основные инструменты: Google Sheets, Яндекс.Метрика, Google Analytics, Google Data Studio, Qlik, Power Bi, Owox Bi. Выполнение лабораторной работы № 3. Выбор темы индивидуального проекта.	Лекции	6		4	1-5	OP-2.2.1, OP-2.3.1, OP-2.2.2, OP-2.3.2, OP-2.2.3, OP-2.3.3
		ЛЗ			4		
		СРС			10		
4.1	Раздел 4. Методы визуализации для решения прикладных задач. Методы визуализации для задач классификации, кластеризации. Визуализация решения задачи временных рядов, демонстрация примеров. Выполнение лабораторной работы № 4. Демонстрация выполнения индивидуальных проектов.	Лекции	6		4	1-5	OP-2.2.1, OP-2.3.1, OP-2.2.2, OP-2.3.2, OP-2.2.3, OP-2.3.3
		ЛЗ			4		
		СРС			10		
	Промежуточная аттестация (по результатам выполнения лабораторных работ (min 70%) и презентации индивидуального проекта -2-3 мин/чел.)	Зачет			0,25		OP-2.2.1, OP-2.3.1, OP-2.2.2, OP-2.3.2, OP-2.2.3, OP-2.3.3

4. Образовательные технологии, учебно-методическое и информационное обеспечение для освоения дисциплины

Освоение дисциплины проходит в классической реализации по схеме «Лекция – лабораторные занятия». На лекции, кроме изложения теоретического материала, преподаватель использует активные формы обучения – взаимодействие с аудиторией: вопрос-ответ, учебная дискуссия, работа в группах, проблемное обучение.

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.

Типовые контрольные задания или иные материалы, необходимые для оценки результатов обучения, характеризующих этапы формирования компетенций, и методические материалы, определяющие процедуры оценивания результатов обучения, приведены в Приложении 1 к рабочей программе «Фонд оценочных средств».

4.1. Рекомендуемая литература и учебно-методическое обеспечение

№ п/п	Авторы составители /	Заглавие	Издательство	Год издания
1	Мастицкий, С.Э., В.К. Шитиков	Статистический анализ и визуализация данных с помощью R URL: https://e.lanbook.com/book/73072	Москва : ДМК Пресс, 2015. — 496 с.	2015
2	Мастицкий, С.Э.	Визуализация данных с помощью ggplot2 https://e.lanbook.com/book/107895	Москва : ДМК Пресс, 2017. — 222 с.	2017
3	Роберт, И.	R в действии. Анализ и визуализация данных в программе R : руководство — URL: https://e.lanbook.com/book/58703	Москва : ДМК Пресс, 2014. — 588 с.	2014
4	Сузи, Р. А.	Язык программирования Python — ISBN 5-9556-0058-2. — Текст : электронный //	Москва : Интернет-Университет Информационных Технологий (ИНТУИТ), 2016. — 350 с.	2015
5.	Маккинни, У.	Python и анализ данных URL: https://e.lanbook.com/book/131721	Москва : ДМК Пресс, 2020. — 540 с.	2020

4.2. Базы данных и информационно-справочные системы, в том числе зарубежные

1. Лань, электронно-библиотечная система
2. IPRBooks, электронно-библиотечная система
3. TOMSK STATE UNIV Electronic Resources

4.3. Перечень лицензионного и программного обеспечения

Программное обеспечение – средства программирования на Python3 и R:

- Python3

- RStudio.

4.4. Оборудование и технические средства обучения

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.

5. Методические указания обучающимся по освоению дисциплины

Методические пособия размещены в локальной сети ИПМКН в каталоге X:\Workspace\Магистратура\ Визуализация многомерных данных.

6. Преподавательский состав, реализующий дисциплину

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

7. Язык преподавания – русский язык.