

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
 Федеральное государственное автономное образовательное учреждение высшего образования
 «НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
 ТОМСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»



РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ
ПРИЕМ 2019 г.
ФОРМА ОБУЧЕНИЯ очная

Методы анализа данных			
Направление подготовки/ специальность	09.04.04 Программная инженерия		
Образовательная программа (направленность (профиль))	Big Data Solutions / Технологии больших данных (Язык обучения: Английский)		
Специализация	Big Data Solutions / Технологии больших данных		
Уровень образования	высшее образование - магистратура		
Курс	1	семестр	2
Трудоемкость в кредитах (зачетных единицах)	3		
Виды учебной деятельности	Временной ресурс		
Контактная (аудиторная) работа, ч	Лекции	16	
	Практические занятия	16	
	Лабораторные занятия	16	
	ВСЕГО	48	
Самостоятельная работа, ч		60	
ИТОГО, ч		108	

Вид промежуточной аттестации	Экзамен	Обеспечивающее подразделение	ОИТ ИШИТР
Заведующий кафедрой- руководитель ОИТ на правах кафедры Руководитель ООП			Шерстнев В.С.
			Савельев А.О.
			Марухина О.В.
Преподаватель			

2020 г.

1. Цели освоения дисциплины

Целями освоения дисциплины является формирование у обучающихся определенного ООП (п. 5.4 Общей характеристики ООП) состава компетенций для подготовки к профессиональной деятельности.

ОПК(У) - 5	Способен разрабатывать и модернизировать программное и аппаратное обеспечение информационных и автоматизированных систем	И.ОПК (У)-5.1	Применяет знания современного программного обеспечения информационных и автоматизированных систем для решения профессиональных задач	ОПК(У) - 5.1В1	Владеет способностью использования языков программирования и инструментальных сред разработки
				ОПК(У) - 5.1У1	Умеет использовать новые и известные методы разработки и модернизации программных систем
				ОПК(У) - 5.1З1	Знает архитектуру современных информационных систем
		И.ОПК (У)-5.2	Осуществляет разработку и модернизацию программного обеспечения информационных и автоматизированных систем	ОПК(У) - 5.2В1	Владеет опытом разработки и тестирования программного обеспечения
				ОПК(У) - 5.2У1	Умеет применить методы и способы эффективного управления разработкой программных средств и проектов
				ОПК(У) - 5.2З1	Знает методы и способы эффективного управления разработкой программных средств и проектов, алгоритмы оптимизации/профилирования запросов
УК(У)-4	Способен применять современные коммуникативные технологии, в том числе на иностранном (-ых) языке (-ах), для академического и профессионального взаимодействия	И.УК (У)-4.2	Составляет академические и (или) профессиональные тексты на иностранном языке	УК(У)-4.2В1	Владеет навыками монологического высказывания на иностранном языке по профилю своей специальности, аргументировано излагая свою позицию и используя вспомогательные средства (таблицы, графики, диаграммы и т.п.)
				УК(У)-4.2У1	Умеет составлять и представлять техническую и научную информацию, используемую в профессиональной деятельности, в виде презентации

2. Место дисциплины (модуля) в структуре ООП

Дисциплина относится к **Вариативной части Междисциплинарного профессионального модуля** учебного плана образовательной программы.

3. Планируемые результаты обучения по дисциплине

После успешного освоения дисциплины будут сформированы результаты обучения:

Планируемые результаты обучения по дисциплине		Индикатор достижения компетенции
Код	Наименование	
РД1	Осуществляет поиск, отбор и систематизацию информации	И.ОПК (У)-5.1
РД2	Определения альтернативные варианты решений в проблемной ситуации	И.ОПК (У)-5.2
РД 3	Осуществлять выбор методов анализа данных в зависимости от природы данных, объема и качества	И.УК (У)-4.2
РД 4	Выполнять обработку и анализ данных, полученных при теоретических и экспериментальных исследованиях	И.УК (У)-4.2

Оценочные мероприятия текущего контроля и промежуточной аттестации представлены в календарном рейтинг-плане дисциплины.

4. Структура и содержание дисциплины

Основные виды учебной деятельности

Разделы дисциплины	Формируемый результат обучения по дисциплине	Виды учебной деятельности	Объем времени, ч.
Раздел 1. Введение в анализ данных. Поиск, очистка	РД1	Лекции	4
		Практические занятия	4
		Лабораторные занятия	4
		Самостоятельная работа	15
Раздел 2. Разнообразие методов анализа данных, выбор методов	РД2	Лекции	4
		Практические занятия	4
		Лабораторные занятия	4
		Самостоятельная работа	15
Раздел 3. Модели классификации, прогнозирования	РД3	Лекции	6
		Практические занятия	4
		Лабораторные занятия	4
		Самостоятельная работа	15
Раздел 4. Решение практических задач	РД4	Лекции	2
		Практические занятия	4
		Лабораторные занятия	4
		Самостоятельная работа	15

Содержание разделов дисциплины:

Раздел 1. Введение в анализ данных. Поиск, очистка

Введение, основы работы с R, RStudio. Создание набора данных. Типы данных. Предварительная обработка данных. Выбор методов в зависимости от типа данных. Методы описательной статистики. Анализ исходных данных на основе использования методов визуализации. Основные типы графиков и соответствующие библиотеки R. Визуализация зависимостей, выявление и анализ «выбросов» в данных, пропущенных значений.

Темы лекций:

- 1. Введение в анализ данных (2 ч.)
- 2. Анализ исходных данных на основе использования методов визуализации (2 ч.)

Темы практических занятий:

- 1. Выбор методов в зависимости от типа данных (2 ч.)
- 2. Методы описательной статистики (2 ч.)

Названия лабораторных работ:

1. Знакомство с R, RStudio. Загрузка и очистка данных.
2. Визуализация зависимостей, выявление и анализ «выбросов» в данных, пропущенных значений.

Раздел 2. Разнообразие методов анализа данных, выбор методов

Основные методы статистической обработки данных: корреляционный анализ, проверка статистической значимости. Визуализация корреляций. Проверка характера распределения данных. Проверка распределения на «нормальность». Основные методы статистической обработки данных: Тесты Стьюдента (для независимых выборок, для зависимых выборок).

Темы лекций:

1. Основные методы статистической обработки данных: корреляционный анализ, проверка статистической значимости (2 ч.)
2. Основные методы статистической обработки данных: Тесты Стьюдента (для независимых выборок, для зависимых выборок) (2 ч.)

Темы практических занятий:

1. Проверка распределения на «нормальность» (2 ч.)
2. Тесты Стьюдента (2 ч.)

Названия лабораторных работ:

1. Корреляционный анализ, проверка статистической значимости, информативность признаков, тесты Стьюдента (для независимых выборок, для зависимых выборок) (4 ч.)

Раздел 3. Модели классификации, прогнозирования

Регрессионный анализ. Оценка параметров модели. Диагностика регрессионной модели. Выбираем лучшую модель. Визуализация результатов. Многомерный регрессионный анализ. Диагностика регрессионной модели. Необычные наблюдения, методы коррекции. Выбираем лучшую модель. Визуализация результатов. Дисперсионный анализ (ANOVA). Оценка параметров модели. Диагностика модели. Визуализация результатов.

Темы лекций:

1. Регрессионный анализ. (2 ч.)
2. Дисперсионный анализ (ANOVA) (4 ч.)

Темы практических занятий:

1. Оценка параметров модели. Диагностика регрессионной модели (2 ч.)
2. Оценка параметров модели. Диагностика дисперсионной модели (2 ч.)

Названия лабораторных работ:

1. Построение регрессионных, дисперсионных моделей (4 ч.)

Раздел 4. Решение практических задач

Обзор и решение задач анализа данных из различных прикладных областей.

Темы лекций:

1. Обзор решения задач анализа данных (примеры решений реальных задач) (2 ч.)

Темы практических занятий:

1. Подготовка данных к анализу, выбор методов анализа (4 ч.).

Названия лабораторных работ:

1. Построение модели анализа данных для решения поставленной задачи (4 ч.).

5. Организация самостоятельной работы студентов

Самостоятельная работа студентов при изучении дисциплины (модуля) предусмотрена в следующих видах и формах:

- Работа с лекционным материалом, поиск и обзор литературы и электронных источников информации по индивидуально заданной проблеме курса;
- Работа в электронном курсе (изучение теоретического материала, выполнение индивидуальных заданий и контролирующих мероприятий и др.);
- Изучение тем, вынесенных на самостоятельную проработку;
- Поиск, анализ, структурирование и презентация информации;
- Подготовка к лабораторным работам, к практическим и семинарским занятиям;
- Работа над междисциплинарным проектом;
- Исследовательская работа и участие в научных студенческих конференциях, семинарах и олимпиадах;
- Анализ научных публикаций по заранее определенной преподавателем теме;
- Подготовка к оценивающим мероприятиям.

6. Учебно-методическое и информационное обеспечение дисциплины

6.1. Учебно-методическое обеспечение

1. Макшанов, А. В.. Технологии интеллектуального анализа данных: учебное пособие [Электронный ресурс] / Макшанов А. В., Журавлев А. Е.. — 2-е изд., стер. — Санкт-Петербург: Лань, 2019. — 212 с. — Книга из коллекции Лань - Информатика. — ISBN 978-5-8114-4493-9.
Схема доступа: <https://e.lanbook.com/book/120063> (контент)
2. Нестеров, С. А.. Основы интеллектуального анализа данных. Лабораторный практикум : учебное пособие [Электронный ресурс] / Нестеров С. А.. — Санкт-Петербург: Лань, 2020. — 40 с. — Книга из коллекции Лань - Информатика. — ISBN 978-5-8114-4509-7.
Схема доступа: <https://e.lanbook.com/book/130181> (контент)
3. Шипунов, А. Б.. Наглядная статистика. Используем R! [Электронный ресурс] / Шипунов А. Б., Балдин Е. М., Волкова П. А., Коробейников А. И.; Назарова С.А., Петров С.В., Суфиянов В.Г.. — Москва: ДМК Пресс, 2014. — 298 с. — Книга из коллекции ДМК Пресс - Информатика. — ISBN 978-5-94074-828-1.
Схема доступа: https://e.lanbook.com/books/element.php?pl1_id=50572 (контент)
4. Роберт, И. . R в действии. Анализ и визуализация данных в программе R [Электронный ресурс] / Роберт И. , Кабаков ; Пер. с англ. Полины Волковой А. — Москва: ДМК Пресс, 2014. — 588 с. — Книга из коллекции ДМК Пресс - Информатика.. — ISBN 978-5-97060-077-1.
Схема доступа: http://e.lanbook.com/books/element.php?pl1_id=58703 (контент)

6.2. Информационное и программное обеспечение

Internet-ресурсы (в т.ч. в среде LMS MOODLE и др. образовательные и библиотечные ресурсы):

1. Электронный курс: <https://stud.lms.tpu.ru/course/view.php?id=2621>
Курс разработан на английском языке для студентов направления 09.04.04 Программная инженерия (образовательная программа Big Data Solutions / Технологии больших данных).
2. <https://rstudio.com/> – официальный сайт для загрузки и установки среды RStudio. Open source.
3. <https://cran.r-project.org/bin/windows/base/> – официальный сайт для загрузки и установки среды R. Open source.

Лицензионное программное обеспечение (в соответствии с **Перечнем лицензионного программного обеспечения ТПУ**):

Для реализации курса специализированного лицензионного программного обеспечения не требуется.

7. Особые требования к материально-техническому обеспечению дисциплины

В учебном процессе используется следующее лабораторное оборудование для практических и лабораторных занятий:

№	Наименование специальных помещений	Наименование оборудования
1.	Аудитория для проведения учебных занятий всех типов, курсового проектирования, консультаций, текущего контроля и промежуточной аттестации (компьютерный класс) Компьютер - 15 шт. 634034, Томская область, г. Томск, Советская улица, 84/3 212	Доска аудиторная настенная - 1 шт.;Шкаф для документов - 1 шт.;Комплект учебной мебели на 10 посадочных мест;Комплект (стол, кресло) - 4 шт.; Компьютер - 15 шт.
	Аудитория для проведения учебных занятий всех типов, курсового проектирования, консультаций, текущего контроля и промежуточной аттестации (лекционная аудитория) 634034 Томская область, Томск, Советская улица, д. 84/3, учебный корпус «Кибернетический центр», аудитория 313.	Комплект оборудования для проведения лекционных занятий: ПК, 1 шт. ,Проектор, 1 шт.

Рабочая программа составлена на основе Общей характеристики образовательной программы по направлению 09.04.04 Программная инженерия/ профиль Big Data Solutions / Технологии больших данных /специализации Big Data Solutions/Технологии больших данных (приема 2019 г., очная форма обучения).

Разработчик(и):

Должность	Подпись	ФИО
Доцент ОИТ ИШИТР		Марухина О.В.

Программа одобрена на заседании Отделения информационных технологий ИШИТР (протокол №13 от 19.06.2019 г.).

Заведующий кафедрой –
руководитель ОИТ на правах кафедры

 / В.С. Шерстнев
подпись

Лист изменений рабочей программы дисциплины:

Учебный год	Содержание /изменение	Обсуждено на заседании ИШИТР (протокол)
2020/2021	1. Обновлено программное обеспечение	от 01.09.2020г. № 19