

Министерство науки и высшего образования РФ  
Федеральное государственное автономное образовательное учреждение  
высшего образования  
«СИБИРСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»  
**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)**

Б1.В.ДВ.04.05 Интеллектуальный анализ данных

наименование дисциплины (модуля) в соответствии с учебным планом

Направление подготовки / специальность

09.03.03 Прикладная информатика

Направленность (профиль)

09.03.03 Прикладная информатика

Форма обучения

очная

Год набора

2019

## РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)

Программу составили \_\_\_\_\_

Ассистент, Пахомова Кристина Игоревна

\_\_\_\_\_  
должность, инициалы, фамилия

## 1 Цели и задачи изучения дисциплины

### 1.1 Цель преподавания дисциплины

Целью освоения дисциплины «Введение в анализ больших данных» является теоретическая и практическая подготовка студентов к работе с большими данными. Знания, полученные в результате освоения дисциплины, помогут при сборе и анализе огромных объемов структурированной или неструктурированной информации, при разработке моделей данных и получении новых знаний. Все это необходимо выпускнику, освоившему программу бакалавриата, для решения различных задач практической и научно-исследовательской деятельности.

### 1.2 Задачи изучения дисциплины

Задачи освоения дисциплины:

- приобретение студентами знаний о технологиях подготовки, хранения, обработки и анализа больших данных;
- применение статистических и математических методов для анализа больших объемов информации;
- приобретение практических навыков работы на языке Python.

### 1.3 Перечень планируемых результатов обучения по дисциплине (модулю), соотнесенных с планируемыми результатами освоения образовательной программы

Код и наименование индикатора достижения компетенции	Запланированные результаты обучения по дисциплине
<b>ПК-16: Способен анализировать и оптимизировать прикладные и информационные процессы</b>	
ПК-16.1: Знает прикладные и информационные процессы в организациях	
ПК-16.2: Умеет анализировать и оптимизировать прикладные и информационные процессы в организациях	
ПК-16.3: Владеет навыками анализа и оптимизации прикладных и информационных процессов в организациях	
<b>ПК-7: Способен анализировать информационные потребности пользователей информационных ресурсов</b>	
ПК-7.1: Знает: терминологию и ключевые параметры веб-статистики; основные принципы и методы сбора статистики посещаемости веб-сайтов; популярные сервисы для сбора веб-статистики	

ПК-7.2: Умеет: анализировать	
структурированную и неструктурированную информацию; - владеть популярными сервисами для оценки посещаемости и характеристик аудитории сайтов	
ПК-7.3: Владеет навыками: проведения общего анализа посещаемости сайта; выявления наиболее популярных страниц; выработки решений по наполнению сайта контентом	

#### **1.4 Особенности реализации дисциплины**

Язык реализации дисциплины: Русский.

Дисциплина (модуль) реализуется с применением ЭО и ДОТ

URL-адрес и название электронного обучающего курса: <https://e.sfu-kras.ru/course/view.php?id=21979>.

## 2. Объем дисциплины (модуля)

Вид учебной работы	Всего, зачетных единиц (акад. час)	Сем естр	
		1	2
<b>Контактная работа с преподавателем:</b>	<b>2,5 (90)</b>		
занятия лекционного типа	0,83 (30)		
практические занятия	1,67 (60)		
<b>Самостоятельная работа обучающихся:</b>	<b>4,5 (162)</b>		
курсовое проектирование (КП)	Нет		
курсовая работа (КР)	Нет		

### 3 Содержание дисциплины (модуля)

#### 3.1 Разделы дисциплины и виды занятий (тематический план занятий)

№ п/п		Модули, темы (разделы) дисциплины		Контактная работа, ак. час.							
				Занятия лекционного типа		Занятия семинарского типа				Самостоятельная работа, ак. час.	
						Семинары и/или Практические занятия		Лабораторные работы и/или Практикумы			
Всего	В том числе в ЭИОС	Всего	В том числе в ЭИОС	Всего	В том числе в ЭИОС	Всего	В том числе в ЭИОС				
<b>1. Введение в анализ больших данных</b>											
	1. Основные определения, термины, задачи анализа больших данных.	4									
	2. Data Mining. Когнитивный анализ данных. Обзор источников информации для Big Data.	4									
	3. Методики сбора данных.	2									
	4. Семантический анализ веб-сайтов			8							
	5. Подготовка к проверочной работе по текущему разделу						20				
<b>2. Технологии хранения и обработки больших данных.</b>											
	1. Обзор технологий хранения больших данных. Базы данных. Системы управления базами данных. Модели данных.	2									
	2. Использование реляционного подхода в представлении данных			4							

3. NoSQL базы данных.	2							
4. Использование NOSQL подхода в представлении данных			4					
5. Подготовка к проверочной работе по текущему разделу							20	
<b>3. Статистические методы анализа данных.</b>								
1. Основные понятия математической статистики. Методы анализа данных.	2							
2. Методы обработки и анализа больших данных.			14					
3. Анализ формальных понятий.	2							
4. Реализация алгоритма АФП			6					
5. Подготовка к проверочной работе по текущему разделу							14	
<b>4. OLAP модель представления данных</b>								
1. Технология OLAP	6							
2. Применение OLAP технологии в практических задачах			12					
3. Подготовка к проверочной работе по текущему разделу							54	
<b>5. Современные программные средства анализа больших объемов информации.</b>								
1. Обзор современных популярных программных средства анализа данных; их преимущества и недостатки.	6							
2. Визуализация исходной информации и аналитических данных.			12					
3. Подготовка к проверочной работе по текущему разделу							54	

Bcero	30		60				162	
-------	----	--	----	--	--	--	-----	--



## **4 Учебно-методическое обеспечение дисциплины**

### **4.1 Печатные и электронные издания:**

1. Гмурман В. Е. Теория вероятностей и математическая статистика: учебное пособие для студентов вузов(Москва: Высшая школа).
2. Советов Б. Я., Цехановский В. В., Чертовской В. Д. Базы данных: учебник для прикладного бакалавриата(М.: Юрайт).
3. Кроув Т., Эйвисон Д., Вейнеров О. М. Базы данных в административных информационных системах(М.: Финансы и статистика).
4. Маккинли У. Python и анализ данных(Москва: ДМК Пресс).
5. Коэльо Л. П. Построение систем машинного обучения на языке Python (Москва: ДМК Пресс).
6. Бергер А., Горбач И. Microsoft SQL Server 2005 Analysis Services. OLAP и многомерный анализ данных(СПб.: БХВ - Петербург).
7. Сузи Р. А. Python: Пособие(Санкт-Петербург: Издательство "БХВ-Петербург").
8. Дронов В. А. Django. Практика создания Web-сайтов на Python: Пособие (Санкт-Петербург: Издательство "БХВ-Петербург").
9. Прохоренок Н. А. Python 3 и PyQt 5. Разработка приложений: Пособие (Санкт-Петербург: Издательство "БХВ-Петербург").
10. Митчелл Р. Скрапинг веб-сайтов с помощью Python(Москва: ДМК Пресс).
11. Дюк В., Самойленко А. Data Mining. Учебный курс(Санкт-Петербург: Питер).
12. Голицина О. Л., Максимов Н. В., Попов И. И. Базы данных: учебное пособие для учреждений среднего профессионального образования по специальностям "Автоматизированные системы обработки информации и управления (по отраслям)", "Программное обеспечение вычислительной техники и автоматизированных систем", "Техническое обслуживание средств вычислительной техники и компьютерных сетей"(Москва: Форум).
13. Брянцев И.Н. DATA MINING. Теория и практика: к самостоятельной работе(М.: БДЦ - пресс).
14. Прохоренок Н.А. Python. Самое необходимое: Практическое руководство(Санкт-Петербург: Издательство "БХВ-Петербург").
15. Рашка С. Python и машинное обучение: крайне необходимое пособие по новейшей предсказательной аналитике, обязательное для более глубокого понимания методологии машинного обучения(Москва: ДМК Пресс).

### **4.2 Лицензионное и свободно распространяемое программное обеспечение, в том числе отечественного производства (программное обеспечение, на которое университет имеет лицензию, а также свободно распространяемое программное обеспечение):**

1. Microsoft Office Professional Plus 2007 Russian OLP NL AE,

2. Python (PSFL Open source) (Python Software Foundation License <https://docs.python.org/3/license.html>),
3. PostgreSQL

#### **4.3 Интернет-ресурсы, включая профессиональные базы данных и информационные справочные системы:**

1. не требуется

#### **5 Фонд оценочных средств**

Оценочные средства находятся в приложении к рабочим программам дисциплин.

#### **6 Материально-техническая база, необходимая для осуществления образовательного процесса по дисциплине (модулю)**

Аудитория, оборудованная компьютерами с подключением к сети Интернет и обеспечением доступа в электронную информационно-образовательную среду Университета, доступ к системе виртуальных машин, демонстрационное оборудование:

интерактивная доска обратной проекции; доступ к беспроводной сети WI-FI, маркерная доска.

Занятия организуются с учетом возможности работы обучающихся из числа инвалидов и лиц с ограниченными возможностями в формах, адаптированных к ограничениям их здоровья и восприятия информации в зависимости от нозологии.