

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

Б1.В.15 Технологии обработки больших данных

наименование дисциплины (модуля) в соответствии с учебным планом

Направление подготовки / специальность

38.03.05 БИЗНЕС-ИНФОРМАТИКА

Направленность (профиль)

38.03.05 БИЗНЕС-ИНФОРМАТИКА

Форма обучения

очная

Год набора

2020

## РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)

Программу составили \_\_\_\_\_

кандидат техн. наук, Доцент, Скуратенко Е.Н.

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

## 1 Цели и задачи изучения дисциплины

### 1.1 Цель преподавания дисциплины

Целью преподавания дисциплины "Технологии обработки больших данных" является формирование знаний, умений и навыков в области обработки больших данных.

### 1.2 Задачи изучения дисциплины

- изучение современных тенденций в области технологий обработки больших данных;
- освоение методов прогнозирования;
- освоение методов регрессионного анализа;
- освоение методов кластерного анализа;
- обзор иных технологий обработки больших данных.

### 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: умение разрабатывать контент и ИТ-сервисы предприятия и интернет-ресурсов</b>	
ПК-16: умение разрабатывать контент и ИТ-сервисы предприятия и интернет-ресурсов	основные принципы обработки больших данных  особенности обработки больших данных  технологии и особенности обработки больших данных разрабатывать контент для обработки больших данных  разрабатывать контент для обработки больших данных разрабатывать контент для обработки больших данных навыками сбора данных для обработки больших данных навыками сбора данных для обработки больших данных навыками обработки больших данных
<b>ПК-19: умение готовить научно-технические отчеты, презентации, научные публикации по результатам выполненных исследований</b>	

ПК-19: умение готовить научно-технические отчеты, презентации, научные публикации по результатам	основные направления развития науки в области технологии обработки больших данных направления развития науки в области технологии
выполненных исследований	обработки больших данных технологии развития науки в области обработки больших данных готовить сообщение о технологиях обработки больших данных готовить доклад и презентацию технологиях обработки больших данных проводить обзор научной литературы о технологиях обработки больших данных методами подготовки сообщений методами подготовки доклада и презентации методами подготовки обзоров научной литературы
<b>ПК-3: выбор рациональных информационных систем и информационно-коммуникативных технологий решения для управления бизнесом</b>	
ПК-3: выбор рациональных информационных систем и информационно-коммуникативных технологий решения для управления бизнесом	основные типы ИС для обработки больших данных основные типы ИС обработки больших данных для управления деятельностью организаций типы ИС и технологии обработки больших данных для управления деятельностью организаций выбирать ИС для управления бизнесом выбирать рациональные ИС для управления бизнесом выбирать рациональные ИС и технологии для управления бизнесом базовыми навыками работы в ИС для обработки больших данных навыками работы в ИС для обработки больших данных навыками работы в ИС для обработки больших данных предприятий

#### 1.4 Особенности реализации дисциплины

Язык реализации дисциплины: Русский.

Дисциплина (модуль) реализуется без применения ЭО и ДОТ.

## 2. Объем дисциплины (модуля)

Вид учебной работы	Всего, зачетных единиц (акад.час)	е
		1
<b>Контактная работа с преподавателем:</b>	<b>1,5 (54)</b>	
занятия лекционного типа	0,5 (18)	
практические занятия	1 (36)	
<b>Самостоятельная работа обучающихся:</b>	<b>0,5 (18)</b>	
курсовое проектирование (КП)	Нет	
курсовая работа (КР)	Нет	
<b>Промежуточная аттестация (Экзамен)</b>	<b>1 (36)</b>	

### 3 Содержание дисциплины (модуля)

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

		Контактная работа, ак. час.							
№ п/п	Модули, темы (разделы) дисциплины	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. Обзор Big-Data. Методы и средства. Используемые программы.</b>									
	1. Big-Data. Инструменты. Технологии. Методы анализа.	4							
	2. Введение. Возможности. Инструменты (Меню. Принцип анализа данных. Основные инструменты и узлы. Возможности построения моделей)			4					
	3. Создание проекта. Настройки источника данных. Типы переменных. Изменение размера выборки. Создание диаграмм. Исследование взаимосвязей между переменными.			4					
	4. Самостоятельная работа с источниками информации, работа с данными, построение моделей							2	
<b>2. Прогнозное моделирование</b>									

1. Область прикладных задач с использованием прогнозного моделирования. Проклятие размерности. Избавление от бесполезных и избыточных входных переменных. Создание обучающих и проверочных данных	2							
2. Прогнозная модель, использующая дерево решений (Создание дерева решений: структура. Алгоритм построения. Поиск разбиений. Прогнозная модель использующая дерево решений: построение, создание правила разбиения. Оптимизация сложности деревьев решений. Оценка качества дерева решений).			4					
3. Самостоятельная работа с источниками информации, работа с данными, построение моделей с использованием ИТ							2	
<b>3. Прогнозное моделирование: работа с регрессионными моделями</b>								
1. Регрессия. Логистическая регрессия. Полиномиальные регрессии. Оценка параметров. Обработка пропущенных значений. Выбор входных переменных. Оптимизация сложности. Интерпретация регрессии.	2							
2. Регрессия. Логистическая регрессия. Полиномиальные регрессии. Оценка параметров. Обработка пропущенных значений. Выбор входных переменных. Оптимизация сложности. Интерпретация регрессии.			2					

3. Регрессии с преобразованными входными переменными. Категориальные переменные в регрессионной модели. Область прикладных задач с использованием прогнозного моделирования. Проклятие размерности. Создание обучающих и проверочных данных	2							
4. Регрессии с преобразованными входными переменными. Категориальные переменные в регрессионной модели. Область прикладных задач с использованием прогнозного моделирования. Проклятие размерности. Создание обучающих и проверочных данных			2					
5. Самостоятельная работа с источниками информации, работа с данными, построение моделей с использованием ИТ							2	
<b>4. Прогнозное моделирование: нейронные сети.</b>								
1. Особенности нейронных сетей. Структура однослойных и многослойных нейронных сетей. Применение нейронных сетей.	2							
2. Обучение нейронной сети, классификация алгоритмов обучения.	2							
3. Прогнозная модель, использующая искусственные нейронные сети(Создание ИНС: структура. Алгоритм обучения т.д.)			4					
4. Самостоятельная работа с источниками информации, работа с данными, построение моделей с использованием ИТ							4	
<b>5. Кластерный анализ</b>								



1. Основные задачи, решаемые с помощью кластерного анализа, сфера применения. ПОДХОДЫ И МЕТОДЫ КЛАСТЕРНОГО АНАЛИЗА	2							
2. Кластерный анализ методом k-средних. Понимание кластера и центра кластера, суть алгоритма. Требования к исходным данным. Определение количества кластеров, критерии оценки качества кластеров. Интерпретация кластеров.	2							
3. Построение модели кластеризации, разбиение объектов в классы в системах распознавания образов "без учителя"			4					
4. Самостоятельная работа с источниками информации, работа с данными, построение моделей с использованием ИТ							4	
<b>6. Payton для анализа данных</b>								
1. Payton для анализа данных			12					
2. Payton для анализа данных							4	
Всего	18		36				18	

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

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

1. Форман Д., Соколова А. Много цифр. Анализ больших данных при помощи Excel: Учебное пособие(Москва: ООО "Альпина Паблишер").
2. Жуков Р.А. Язык программирования Python: практикум: Учебное пособие(Москва: ООО "Научно-издательский центр ИНФРА-М").
3. Редько В. Г., Малинецкий Г. Г. Эволюция, нейронные сети, интеллект: модели и концепции эволюционной кибернетики(Москва: URSS).
4. Емельянов С. В. Искусственный интеллект и принятие решений: Вып. 1 (Москва: URSS).
5. Смолин Д. В. Введение в искусственный интеллект: конспект лекций: методические указания(Москва: Физматлит).
6. Цехановский В. В., Чертовской В. Д. Управление данными: учебник (Санкт-Петербург: Лань).
7. Бродовская Е. В., Домбровская А. Ю. Большие данные в исследовании политических процессов: учебное пособие(Москва: МПГУ).
8. Лейнартас Д. Е. Анализ данных. Математические методы анализа данных: учеб.-метод. пособие для самостоят. работы [для студентов напр. 080500.62 «Бизнес-информатика» очной формы обуч.] (Красноярск: СФУ).
9. Крутиков В. И., Мешечкин В. В. Анализ данных(Кемерово: КемГУ).

### **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. регулярно обновляемый интернет-браузер (Mozilla Firefox, Google Chrome, Yandex Browser, Opera, Internet Explorer, Safari, либо иной);
2. офисный пакет (MS Office, Libre Office, Open Office, либо иной).
3. онлайн компилятор для Python
- 4.

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

1. Сайт библиотеки СФУ. Режим доступа: <http://bik.sfu-kras.ru/>
2. Электронный каталог библиотеки СФУ. Режим доступа: <http://catalog.sfu-kras.ru/>
3. Электронно-библиотечная система Издательства «Лань»  
<http://e.lanbook.com/>
4. ЭБС ЮРАЙТ <http://www.biblio-online.ru/>
5. Электронно-библиотечная система elibrary <https://elibrary.ru>

6. Электронно-библиотечная система ZNANI-UM.COM (ИНФРА-М)  
<http://www.znanium.com/>
- 7.

### **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)**

Лекционные занятия –(лекционная аудитория): рабочее место преподавателя, рабочие места обучающихся, компьютер, активные колонки, проектор.учебные аудитории, позволяющие выступающему (преподавателю, студенту) демонстрировать слайды в форматах pdf, PowerPoint и других графических форматах на экране с одновременным выступлением перед аудиторией;

Практические работы – (компьютерный класс): рабочее место преподавателя. Рабочие места обучающихся оснащены доступом в интернет и пакетом офисных программ.

ПО, используемое в учебном процессе по данной дисциплине:

регулярно обновляемый интернет-браузер (Mozilla Firefox, Google Chrome, Yandex Browser, либо иной);

офисный пакет (MS Office, Libre Office, Open Office, либо иной).

онлайн компилятор для Payton