

Документ подписан простой электронной подписью  
Информация о владельце:  
ФИО: Клочков Юрий Сергеевич  
Должность: **Министр**  
Дата подписания: 28.11.2024 09:29:20  
Уникальный программный ключ:  
4e7c4ea90328ec8e65c5d8058549a2538d7400d1

**МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ**

Федеральное государственное бюджетное образовательное учреждение высшего образования

**«ТЮМЕНСКИЙ ИНДУСТРИАЛЬНЫЙ УНИВЕРСИТЕТ»**

**УТВЕРЖДАЮ**  
Заведующий кафедрой  
Интеллектуальных систем и технологий  
\_\_\_\_\_ О.Ф. Данилов  
«\_\_» \_\_\_\_\_ 2024 г

**РАБОЧАЯ ПРОГРАММА**

дисциплины: **Большие данные**  
направление подготовки: **09.03.02 Информационные системы и технологии**  
направленность (профиль): **Информационные системы и технологии в геологии и нефтегазовой отрасли**  
форма обучения: **очная, заочная**

Рабочая программа рассмотрена  
на заседании кафедры Интеллектуальных систем и технологий

Протокол № \_\_\_\_\_ г.

## 1. Цели и задачи освоения дисциплины

Цель изучения дисциплины состоит в формировании у студентов представлений об общих методологических основах и принципах работы с технологиями больших данных.

Задачи дисциплины:

- сформировать знания о базовые понятия технологии Big Data, методологии и методиках анализа больших данных;
- сформировать умения определять массивы больших данных и организовывать их хранение; применять методы кластерного анализа и алгоритмов деревьев решений;
- сформировать навыки анализа больших данных с использованием инструментов MS-Excel.

## 2. Место дисциплины в структуре ОПОП ВО

Дисциплина относится к части, формируемой участниками образовательных отношений учебного плана. Необходимыми условиями для освоения дисциплины являются:

- знание: базовых понятий статистического анализа данных, баз данных, методов искусственного интеллекта, основные технологии прикладного анализа данных;
- умение: строить различными способами модель данных; применять полученные навыки на практике.
- владение: инструментами MS-Excel.

Содержание дисциплины является логическим продолжением содержания дисциплин «Теория вероятностей и математическая статистика», «Базы данных», «Управление данными», «Системы искусственного интеллекта» и служит основой для освоения дисциплин «Методы и средства поддержки принятия решений в нефтегазовой отрасли», «Нейроинформатика и машинное обучение».

## 3. Результаты обучения по дисциплине

Процесс изучения дисциплины направлен на формирование следующих компетенций:

Таблица 3.1

Код и наименование компетенции	Код и наименование индикатора достижения компетенции (ИДК)	Код и наименование результата обучения по дисциплине
ПКС-5 Способность выполнять работы по обеспечению функционирования баз данных и обеспечению их информационной безопасности	ПКС-5.1 Анализирует программные средства управления базами данных; угрозы и средства защиты их информационной безопасности	З1 Знать: базовые понятия технологии Big Data, технологии хранения больших данных
	ПКС-5.2 Выполняет администрирование баз данных и обеспечение их информационной безопасности	У1 Уметь: организовывать хранение больших данных
ПКС-6 Способность выполнять работы по созданию (модификации) и сопровождению информационных систем	ПКС-6.1 Выполняет анализ и моделирование бизнес-процессов	З2 Знать: задачи и возможности использования анализа больших данных в бизнес-процессах
	ПКС-6.2 Выбирает программные средства для создания информационных систем	У2 Уметь: использовать технологии анализа больших данных, выполнять кластерный анализ и применять методы деревьев решений

#### 4. Объем дисциплины

Общий объем дисциплины составляет 4 зачетных единицы, 144 часов.

Таблица 4.1.

Форма обучения	Курс/ семестр	Аудиторные занятия/контактная работа, час.			Самостоя- тельная ра- бота, час.	Контроль, час	Форма промежу- точной ат- тестации
		Лекции	Практические занятия	Лабораторные занятия			
Очная	4/7	14	-	28	66	36	экзамен
Заочная	5/9	10	-	12	113	9	экзамен

#### 5. Структура и содержание дисциплины

##### 5.1. Структура дисциплины

##### Очная форма обучения (ОФО)

Таблица 5.1.

№ п/п	Структура дисципли- ны		Аудиторные занятия, час.			СРС, час.	Всего, час.	Код ИДК	Оценочные средства
	Номер раздела	Наименование раздела	Л.	Пр.	Лаб.				
1	1	Определение больших дан- ных. Техноло- гии хранения больших дан- ных.	2		4	11	17	ПКС-5.1 ПКС-5.2 ПКС-6.1 ПКС-6.2	Письменная работа
2	2	Технологии анализа больших данных.	4		4	11	19		Письменная работа, лабораторный отчет
3	3	Методы анализа больших данных	6		12	22	40		Письменная работа, лабораторный отчет
4	4	Анализ больших данных в Excel	2		8	22	32		лабораторный отчет
5	Экзамен		-	-	-	36	36		Вопросы к экзамену
Итого:			14		28	102	144		

##### Заочная форма обучения (ЗФО)

Таблица 5.2

№ п/п	Структура дисципли- ны		Аудиторные занятия, час.			СРС, час.	Всего, час.	Код ИДК	Оценочные средства
	Номер раздела	Наименование раздела	Л.	Пр.	Лаб.				
1	1	Определение больших дан- ных. Техноло- гии хранения больших дан- ных.	2		3	30	35	ПКС-5.1 ПКС-5.2 ПКС-6.1 ПКС-6.2	Письменная работа, контрольная работа
2	2	Технологии анализа	2		3	30	35		Письменная работа, лабораторный отчет,

		больших данных.							контрольная работа
3	3	Методы анализа больших данных	4		3	30	37		Письменная работа, лабораторный отчет, контрольная работа
4	4	Анализ больших данных в Excel	2		3	23	28		лабораторный отчет, контрольная работа
5	Экзамен		-	-	-	9	9		Вопросы к экзамену
Итого:			10		12	122	144		

## 5.2. Содержание дисциплины

### 5.2.1. Содержание разделов дисциплины (дидактические единицы).

Раздел 1. «Определение больших данных». Определение больших данных Технологии хранения больших данных. Большие данные (big data) в информационных технологиях. Совокупность подходов, инструментов и методов обработки структурированных и неструктурированных данных огромных объёмов. Средства массово-параллельной обработки неопределённо структурированных данных, NoSQL, алгоритмы MapReduce, программные каркасы и библиотеки проекта Hadoop.

Раздел 2. «Технологии анализа больших данных». Процесс анализа больших данных. Технологии анализа больших данных. Научные проблемы в области больших данных. Методы техники анализа, применимые к большим данным: методы класса Data Mining: обучение ассоциативным правилам (англ. association rule learning), классификация (методы категоризации новых данных на основе принципов, ранее применённых к уже наличествующим данным), кластерный анализ, регрессионный анализ; краудсорсинг

Раздел 3. «Методы анализа данных». Простые методы, Построение деревьев решений. Кластеризация. Прогнозирование. Закон распределения случайной величины. Статистические оценки параметров. Доверительные области. Теория моментов. Корреляционный анализ.

Раздел 4. «Анализ данных в Excel». Программы статистической обработки информации. Представление возможностей пакета Анализ данных в Excel. Применение возможностей Excel для решения задач анализа больших данных.

### 5.2.2. Содержание дисциплины по видам учебных занятий.

#### Лекционные занятия

Таблица 5.2.1

№ п/п	Номер раздела дисциплины	Объем, час.			Тема лекции
		ОФО	ЗФО	ОЗФО	
1	1	2	2	-	Определение больших данных. Технологии хранения больших данных.
2	2	4	2	-	Технологии анализа больших данных.
3	3	6	4	-	Методы анализа больших данных
4	4	2	2	-	Анализ больших данных в Excel
Итого:		14	10	-	

#### Практические занятия

Не предусмотрены учебным планом

## Лабораторные работы

Таблица 5.2.3

№ п/п	Номер раздела дисциплины	Объем, час.			Наименование лабораторной работы
		ОФО	ЗФО	ОЗФО	
1	1,2	8	6	-	Разработка хранилища данных и его основных процессов. Правило OneRule
2	3	12	3	-	Построение деревьев решений. Кластеризация.
3	4	8	3	-	Анализ данных в Excel
Итого:		28	12	-	

## Самостоятельная работа студента

Таблица 5.2.4

№ п/п	Номер раздела дисциплины	Объем, час.			Тема	Вид СРС
		ОФО	ЗФО	ОЗФО		
1	1	11	30	-	Определение больших данных. Технологии хранения больших данных.	Подготовка к письменной работе, контрольная работа
2	2	11	30	-	Технологии анализа больших данных.	Подготовка к письменной работе, контрольная работа
3	3	22	30	-	Методы анализа больших данных	Подготовка/лабораторным работам; Подготовка к письменной работе, контрольная работа
4	4	22	23	-	Анализ больших данных в Excel	Подготовка/лабораторным работам; Подготовка к письменной работе, контрольная работа
	Экзамен	36	9			Подготовка к экзамену
Итого:		102	122	-		

5.2.3. Преподавание дисциплины ведется с применением следующих видов образовательных технологий:

- информационные технологии: использование электронных образовательных ресурсов (электронный конспект, презентация) при подготовке к лекциям, практическим и лабораторным занятиям.

- работа в команде: совместная работа студентов в группе при выполнении лабораторных работ, выполнении групповых домашних заданий.

## 6. Тематика курсовых работ/проектов

Курсовые работы/проекты учебным планом не предусмотрены

## 7. Контрольные работы

7.1. Методические указания для выполнения контрольных работ.

Контрольные работы выполняются самостоятельно в период между сессиями по индивидуальным заданиям. Тематика заданий определяется преподавателем, соответствует разделам дисциплины и сообщается обучающимся не позже, чем за две недели до начала зимней сессии 5 курса. Во время сессии обучающийся должен сдать преподавателю в печатном виде отчет по контрольной работе и устно защитить его.

7.2. Тематика контрольных работ.

Основные темы контрольных работ:

1. Большие данные и бизнес-аналитика
2. Технологии обработки больших данных
3. Обзор решений от производителей ПО в области больших данных
4. Методы и платформы интеграции данных
5. Организации хранилищ больших данных
6. ETL - функции и процессы
7. Технологии многомерных баз данных
8. Миварные базы данных
9. Шаблоны интеграции корпоративных информационных систем
10. Задачи с интенсивным использованием данных
11. Анализ сообществ в социальных сетях
12. Методы и технологии обработки неструктурированных данных
13. Обзор архитектуры Hadoop
14. Нетрадиционные модели данных not only SQL (NOSQL)
15. Представление знаний
16. Консолидация данных
17. Технологии Business Intelligence
18. Параллельные машины баз данных
19. Парадигмы и манифесты баз данных
20. Управление большими данными и анализ потоковых данных
21. Технология MapReduce
22. Технологии Data Mining
23. Современные технологии высокопроизводительных вычислений
24. Грид-системы: проблемы и ограничения
25. Решения ORACLE для работы с большими данными
26. Модели данных для больших данных
27. Роль языков программирования Python и R в аналитике больших данных
28. Интеграция Hadoop с традиционными хранилищами данных
29. Большие данные и машинное обучение
30. Технология HIVE

## 8. Оценка результатов освоения дисциплины

8.1. Критерии оценивания степени полноты и качества освоения компетенций в соответствии с планируемыми результатами обучения приведены в Приложении 1.

8.2. Рейтинговая система оценивания степени полноты и качества освоения компетенций обучающихся очной формы обучения представлена в таблице 8.1.

Таблица 8.1

№ п/п	Виды мероприятий в рамках текущего контроля	Количество баллов
<b>1 текущая аттестация</b>		
	Отчет по лабораторной работе № 1	10
	Письменная работа по теме: Разработка хранилища данных	20
	<b>ИТОГО за первую текущую аттестацию</b>	<b>30</b>
<b>2 текущая аттестация</b>		
	Отчет по лабораторной работе № 2	10
	Письменная работа по теме: Технологии анализа больших данных	20
	<b>ИТОГО за вторую текущую аттестацию</b>	<b>30</b>
<b>3 текущая аттестация</b>		

	Отчет по лабораторной работе № 3	10
	Письменная работа по теме: Кластеризация.	30
	ИТОГО за третью текущую аттестацию	40
	<b>ВСЕГО</b>	<b>100</b>

8.3. Рейтинговая система оценивания степени полноты и качества освоения компетенций обучающихся заочной формы обучения представлена в таблице 8.2.

Таблица 8.2

№ п/п	Виды мероприятий в рамках текущего контроля	Количество баллов
1	Письменная работа 1-3	60
2	Отчет по лабораторной работе 1-3	30
3	Отчет по контрольной работе	10
	<b>ВСЕГО</b>	<b>100</b>

## 9. Учебно-методическое и информационное обеспечение дисциплины

9.1. Перечень рекомендуемой литературы представлен в Приложении 2.

9.2. Современные профессиональные базы данных и информационные справочные системы:

- ЭБС «Издательства Лань»;
- ЭБС «Электронного издательства ЮРАЙТ»;
- Собственная полнотекстовая база (ПБД) БИК ТИУ;
- Научная электронная библиотека «eLIBRARY.RU»;
- ЭБС «IPRbooks»;
- Научно-техническая библиотека ФГБОУ ВО РГУ нефти и газа имени И.М. Губкина;
- Научно-техническая библиотека ФГБОУ ВПО УГНТУ (г. Уфа);
- Научно-техническая библиотека ФГБОУ ВПО УГТУ (г. Ухта);
- ЭБС «Прспект»;
- ЭБС «Консультант студент».

9.3. Лицензионное и свободно распространяемое программное обеспечение, в т.ч. отечественного производства:

1. Microsoft Office Professional Plus;
2. Windows 8 и выше.
3. СУБД MS SQL Server Developer 2017 и выше, MySQL
4. MS Visual Studio Professional 2017 и выше
6. Среда программирования на Python

## 10. Материально-техническое обеспечение дисциплины

Помещения для проведения всех видов работы, предусмотренных учебным планом, укомплектованы необходимым оборудованием и техническими средствами обучения.

Таблица 10.1

### Обеспеченность материально-технических условий реализации ОПОП ВО

№ п/п	Наименование учебной дисциплины	Наименование помещений для проведения учебной дисциплины, в том числе помещения для самостоятельной работы, с указанием перечня основного оборудования, учебно-наглядных пособий	Адрес (местоположение) помещений для проведения учебной дисциплины
1	2	3	4
1	Большие данные	<p>Лекционные занятия: Учебная аудитория для проведения занятий лекционного типа; групповых и индивидуальных консультаций; текущего контроля и промежуточной аттестации, Оснащенность: Учебная мебель: столы, стулья, доска аудиторная. Компьютер в комплекте, проектор, проекционный экран.</p> <p>Практические занятия: Учебная аудитория для проведения занятий семинарского типа (практические занятия); групповых и индивидуальных консультаций; текущего контроля и промежуточной аттестации, Учебная лаборатория. Оснащенность: Учебная мебель: столы, стулья, доска аудиторная, компьютер в комплекте, проектор, проекционный экран, наличие установленных программных комплексов, компьютеры для обучающихся.</p>	<p>625001, Тюменская область, г. Тюмень, ул. Володарского, д.56, ауд.328</p> <p>625001, Тюменская область, г. Тюмень, ул. Володарского, д.56, ауд. 328</p>

№ п/п	Наименование учебной дисциплины	Наименование помещений для проведения учебной дисциплины, в том числе помещения для самостоятельной работы, с указанием перечня основного оборудования, учебно-наглядных пособий	Адрес (местоположение) помещений для проведения учебной дисциплины
1	2	3	4
	Большие данные	Лабораторные занятия: Учебная аудитория для проведения занятий семинарского типа (практические занятия); групповых и индивидуальных консультаций; текущего контроля и промежуточной аттестации, Учебная лаборатория. Оснащенность: Учебная мебель: столы, стулья, доска аудиторная, компьютер в комплекте, проектор, проекционный экран, наличие установленных программных комплексов, компьютеры для обучающихся.	625001, Тюменская область, г. Тюмень, ул. Володарского, д.56, ауд. 328
		Самостоятельная работа: Помещение для самостоятельной работы обучающихся с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду, №355, Учебная мебель: столы, стулья, доска аудиторная. Компьютер в комплекте – 5 шт., проектор – 1 шт., проекционный экран – 1 шт.	625001, Тюменская область, г. Тюмень, ул. Володарского, д.56
		Самостоятельная работа: Помещение для самостоятельной работы обучающихся с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду, Учебная мебель: столы, стулья, доска аудиторная. Компьютер в комплекте – 5 шт.	625001, Тюменская область, г. Тюмень, ул. Володарского, д.56

### 11. Методические указания по организации СРС

Самостоятельная работа заключается в подготовке к практическим, лабораторным занятиям, экзамену.

Лабораторные занятия организуются с использованием интерактивных методов обучения (тренинг, работа в группе). В процессе подготовки к лабораторным занятиям обучающиеся могут прибегать к консультациям преподавателя.

Самостоятельная работа обучающихся также заключается в получении заданий (тем) у преподавателя для индивидуального освоения. Преподаватель на занятии дает рекомендации, необходимые для освоения материала. Обучающиеся должны понимать содержание выполненной работы (знать определения понятий, уметь разъяснить значение и смысл любого термина, используемого в работе и т.п.).

## Планируемые результаты обучения для формирования компетенции и критерии их оценивания

Дисциплина Большие данные

Код, направление подготовки: 09.03.02 Информационные системы и технологии

Направленность (профиль) Информационные системы и технологии в геологии и нефтегазовой отрасли

Код компетенции	Код и наименование результата обучения по дисциплине	Код и наименование результата обучения по дисциплине	Критерии оценивания результатов обучения			
			1-2	3	4	5
ПКС-5 Способность выполнять работы по обеспечению функционирования баз данных и обеспечению их информационной безопасности	ПКС-5.1 Анализирует программные средства управления базами данных; угрозы и средства защиты их информационной безопасности	З1 Знать: базовые понятия технологии Big Data, технологии хранения больших данных	Демонстрирует низкий уровень /отсутствие знаний базовых понятий технологии Big Data, технологий хранения больших данных	Демонстрирует частичное / поверхностное знание базовых понятий технологии Big Data, технологий хранения больших данных	Демонстрирует достаточный уровень знаний базовых понятий технологии Big Data, технологий хранения больших данных	Демонстрирует высокий уровень знаний базовых понятий технологии Big Data, технологий хранения больших данных
	ПКС-5.2 Выполняет администрирование баз данных и обеспечение их информационной безопасности	У1 Уметь: организовывать хранение больших данных	Не умеет: организовывать хранение больших данных	На слабом уровне умеет: организовывать хранение больших данных	На среднем уровне умеет: организовывать хранение больших данных	Профессионально умеет: организовывать хранение больших данных
ПКС-6 Способность выполнять работы по созданию (модификации) и сопровождению информационных систем	ПКС-6.1 Выполняет анализ и моделирование бизнес-процессов	З2 Знать: задачи и возможности использования анализа больших данных в бизнес-процессах	Демонстрирует низкий уровень /отсутствие знаний задач и возможностей использования анализа больших данных в бизнес-процессах	Демонстрирует частичное / поверхностное знание задач и возможностей использования анализа больших данных в бизнес-процессах	Демонстрирует достаточный уровень знаний задач и возможностей использования анализа больших данных в бизнес-процессах	Демонстрирует высокий уровень знаний задач и возможностей использования анализа больших данных в бизнес-процессах
	ПКС-6.2 Выбирает программные средства для создания информационных систем	У2 Уметь: использовать технологии анализа больших данных	Не умеет: использовать технологии анализа больших данных	На слабом уровне умеет: использовать технологии анализа больших данных	На среднем уровне умеет: использовать технологии анализа больших данных	Профессионально умеет: использовать технологии анализа больших данных

**КАРТА**  
**обеспеченности дисциплины учебной и учебно-методической литературой**

Дисциплина Большие данные

Код, направление подготовки 09.03.02 Информационные системы и технологии

Направленность (профиль) Информационные системы и технологии в геологии и нефтегазовой отрасли

№ п/п	Название учебного, учебно-методического издания, автор, издательство, вид издания, год издания	Количество экземпляров в БИК	Контингент обучающихся, использую	Обеспеченность обучающихся литературой	Наличие электронного варианта в ЭБС (+/-)
1	Сьоре, Э. Проектирование и реализация систем управления базами данных / Э. Сьоре ; перевод с английского А. Н. Киселева. — Москва : ДМК Пресс, 2021. — 466 с. — ISBN 978-5-97060-488-5. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <a href="https://e.lanbook.com/book/190718">https://e.lanbook.com/book/190718</a> . — Режим доступа: для авториз. пользователей.	-	25	100	+
2	Фомичева, С. Г. Разработка, проектирование и сопровождение приложений баз данных : учебное пособие / С. Г. Фомичева. — Норильск : ЗГУ им. Н.М. Федоровского, 2021. — 185 с. — ISBN 978-5-89009-744-6. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <a href="https://e.lanbook.com/book/224558">https://e.lanbook.com/book/224558</a> . — Режим доступа: для авториз. пользователей.	-	25	100	+
3	Толстобров, А. П. Управление данными : учебное пособие для вузов / А. П. Толстобров. — 3-е изд., перераб. и доп. — Москва : Издательство Юрайт, 2022. — 272 с. — (Высшее образование). — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <a href="https://urait.ru/bcode/496748">https://urait.ru/bcode/496748</a>	-	25	100	+
4	Постреляционные хранилища данных : учебное пособие для вузов / Ю. П. Парфенов. - М : Издательство Юрайт, 2022. - 121 с. <a href="https://urait.ru/bcode/492609">https://urait.ru/bcode/492609</a>	-	25	100	+