

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение высшего образования
«Горно-Алтайский государственный университет»
(ФГБОУ ВО ГАГУ, ГАГУ, Горно-Алтайский государственный университет)

Применение больших данных в экономике рабочая программа дисциплины (модуля)

Закреплена за кафедрой **кафедра экономики, туризма и прикладной информатики**

Учебный план 09.03.03_2023_823.plx
09.03.03 Прикладная информатика
Цифровая экономика

Квалификация **бакалавр**

Форма обучения **очная**

Общая трудоемкость **3 ЗЕТ**

Часов по учебному плану 108
в том числе:
аудиторные занятия 44
самостоятельная работа 54,1
часов на контроль 8,85

Виды контроля в семестрах:
зачеты с оценкой 7

Распределение часов дисциплины по семестрам

Семестр (<Курс>.<Семестр на курсе>)	7 (4.1)		Итого	
	10 5/6			
Неделя				
Вид занятий	УП	РП	УП	РП
Лекции	18	18	18	18
Практические	26	26	26	26
Консультации (для студента)	0,9	0,9	0,9	0,9
Контроль самостоятельной работы при проведении аттестации	0,15	0,15	0,15	0,15
В том числе инт.	16	16	16	16
Итого ауд.	44	44	44	44
Контактная работа	45,05	45,05	45,05	45,05
Сам. работа	54,1	54,1	54,1	54,1
Часы на контроль	8,85	8,85	8,85	8,85
Итого	108	108	108	108

Программу составил(и):

к.э.н., зав. кафедрой, Куттубаева Тосканай Айтмуқановна



Рабочая программа дисциплины

Применение больших данных в экономике

разработана в соответствии с ФГОС:

Федеральный государственный образовательный стандарт высшего образования - бакалавриат по направлению подготовки 09.03.03 Прикладная информатика (приказ Минобрнауки России от 19.09.2017 г. № 922)

составлена на основании учебного плана:

09.03.03 Прикладная информатика

утвержденного учёным советом вуза от 26.12.2022 протокол № 12.

Рабочая программа утверждена на заседании кафедры

кафедра экономики, туризма и прикладной информатики

Протокол от 09.03.2023 протокол № 8

Зав. кафедрой Куттубаева Тосканай Айтмуқановна



Визирование РПД для исполнения в очередном учебном году

Рабочая программа пересмотрена, обсуждена и одобрена для исполнения в 2024-2025 учебном году на заседании кафедры **кафедра экономики, туризма и прикладной информатики**

Протокол от _____ 2024 г. № ____
Зав. кафедрой Куттубаева Тосканай Айтмуқановна

Визирование РПД для исполнения в очередном учебном году

Рабочая программа пересмотрена, обсуждена и одобрена для исполнения в 2025-2026 учебном году на заседании кафедры **кафедра экономики, туризма и прикладной информатики**

Протокол от _____ 2025 г. № ____
Зав. кафедрой Куттубаева Тосканай Айтмуқановна

Визирование РПД для исполнения в очередном учебном году

Рабочая программа пересмотрена, обсуждена и одобрена для исполнения в 2026-2027 учебном году на заседании кафедры **кафедра экономики, туризма и прикладной информатики**

Протокол от _____ 2026 г. № ____
Зав. кафедрой Куттубаева Тосканай Айтмуқановна

Визирование РПД для исполнения в очередном учебном году

Рабочая программа пересмотрена, обсуждена и одобрена для исполнения в 2027-2028 учебном году на заседании кафедры **кафедра экономики, туризма и прикладной информатики**

Протокол от _____ 2027 г. № ____
Зав. кафедрой Куттубаева Тосканай Айтмуқановна

1. ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ	
1.1	<i>Цели:</i> формирование у студентов профессиональной компетенции в области разработки и использования систем обработки и анализа больших массивов данных
1.2	<i>Задачи:</i> Задачи дисциплины – научить производить расчеты с применением технологий анализа больших данных и решать широкий спектр прикладных задач обработки больших наборов данных.

2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ООП	
Цикл (раздел) ООП:	Б1.В.ДВ.04
2.1	Требования к предварительной подготовке обучающегося:
2.1.1	Цифровая экономика
2.1.2	Базы данных
2.2	Дисциплины и практики, для которых освоение данной дисциплины (модуля) необходимо как предшествующее:
2.2.1	Выполнение и защита выпускной квалификационной работы
2.2.2	Преддипломная практика
2.2.3	Интеллектуальные информационные системы

3. КОМПЕТЕНЦИИ ОБУЧАЮЩЕГОСЯ, ФОРМИРУЕМЫЕ В РЕЗУЛЬТАТЕ ОСВОЕНИЯ ДИСЦИПЛИНЫ (МОДУЛЯ)	
ПК-7: Способен собирать детальную информацию для формализации требований пользователей заказчика	
ИД-1.ПК-7: Систематизирует требования пользователей заказчика к ИС	
<p>Знать:</p> <ul style="list-style-type: none"> - сущность, особенности и типы больших данных; - методы анализа и хранения больших объемов данных; - этапы жизненного цикла обработки больших данных; - языки, наиболее приспособленные для обработки и аналитики больших данных; - способы организации хранения и доступа к большим данным. <p>Уметь:</p> <ul style="list-style-type: none"> - выполнять элементы анализа больших данных и интерпретировать результаты; - выбрать подходящий инструмент анализа больших данных, выбрать подходящую технологию хранения больших данных. <p>Владеть:</p> <ul style="list-style-type: none"> - технологиями и методами сбора, обработки и анализа больших данных. 	
ИД-2.ПК-7: Определяет и оценивает виды информации, необходимой для формализации требований пользователей заказчика	
<p>Знать:</p> <ul style="list-style-type: none"> - сущность, особенности и типы больших данных; - методы анализа и хранения больших объемов данных; - этапы жизненного цикла обработки больших данных; - языки, наиболее приспособленные для обработки и аналитики больших данных; - способы организации хранения и доступа к большим данным. <p>Уметь:</p> <ul style="list-style-type: none"> - выполнять элементы анализа больших данных и интерпретировать результаты; - выбрать подходящий инструмент анализа больших данных, выбрать подходящую технологию хранения больших данных. <p>Владеть:</p> <ul style="list-style-type: none"> - технологиями и методами сбора, обработки и анализа больших данных. 	
ИД-3.ПК-7: Осуществляет сбор информации для формализации требований пользователей заказчика	
<p>Знать:</p> <ul style="list-style-type: none"> - сущность, особенности и типы больших данных; - методы анализа и хранения больших объемов данных; - этапы жизненного цикла обработки больших данных; - языки, наиболее приспособленные для обработки и аналитики больших данных; - способы организации хранения и доступа к большим данным. <p>Уметь:</p> <ul style="list-style-type: none"> - выполнять элементы анализа больших данных и интерпретировать результаты; - выбрать подходящий инструмент анализа больших данных, выбрать подходящую технологию хранения больших данных. <p>Владеть:</p> <ul style="list-style-type: none"> - технологиями и методами сбора, обработки и анализа больших данных. 	
ПК-8: Способен проводить описание прикладных процессов и информационного обеспечения решения прикладных задач	

ИД-1.ПК-8: Определяет прикладные задачи, для достижения которых требуется информационное обеспечение
<p>Знать:</p> <ul style="list-style-type: none"> - сущность, особенности и типы больших данных; - методы анализа и хранения больших объемов данных; - этапы жизненного цикла обработки больших данных; - языки, наиболее приспособленные для обработки и аналитики больших данных; - способы организации хранения и доступа к большим данным. <p>Уметь:</p> <ul style="list-style-type: none"> - выполнять элементы анализа больших данных и интерпретировать результаты; - выбрать подходящий инструмент анализа больших данных, выбрать подходящую технологию хранения больших данных. <p>Владеть:</p> <ul style="list-style-type: none"> - технологиями и методами сбора, обработки и анализа больших данных; - навыками работы с программными продуктами по обработке больших данных.
ИД-2.ПК-8: Анализирует и описывает информационные процессы и информационное обеспечение решения прикладных задач
<p>Знать:</p> <ul style="list-style-type: none"> - сущность, особенности и типы больших данных; - методы анализа и хранения больших объемов данных; - этапы жизненного цикла обработки больших данных; - языки, наиболее приспособленные для обработки и аналитики больших данных; - способы организации хранения и доступа к большим данным. <p>Уметь:</p> <ul style="list-style-type: none"> - выполнять элементы анализа больших данных и интерпретировать результаты; - выбрать подходящий инструмент анализа больших данных, выбрать подходящую технологию хранения больших данных. <p>Владеть:</p> <ul style="list-style-type: none"> - технологиями и методами сбора, обработки и анализа больших данных; - навыками работы с программными продуктами по обработке больших данных.
ПК-12: Способен эксплуатировать и сопровождать информационные системы и сервисы
ИД-1.ПК-12: Определяет содержание процесса эксплуатации и сопровождения информационных систем и сервисов
<p>Знать:</p> <ul style="list-style-type: none"> - сущность, особенности и типы больших данных; - методы анализа и хранения больших объемов данных; - этапы жизненного цикла обработки больших данных; - языки, наиболее приспособленные для обработки и аналитики больших данных; - способы организации хранения и доступа к большим данным. <p>Уметь:</p> <ul style="list-style-type: none"> - выбрать подходящий инструмент анализа больших данных, выбрать подходящую технологию хранения больших данных. <p>Владеть:</p> <ul style="list-style-type: none"> - навыками работы с программными продуктами по обработке больших данных.
ИД-2.ПК-12: Эксплуатирует информационные системы и сервисы
<p>Знать:</p> <ul style="list-style-type: none"> - сущность, особенности и типы больших данных; - методы анализа и хранения больших объемов данных; - этапы жизненного цикла обработки больших данных; - языки, наиболее приспособленные для обработки и аналитики больших данных; - способы организации хранения и доступа к большим данным. <p>Уметь:</p> <ul style="list-style-type: none"> - выбрать подходящий инструмент анализа больших данных, выбрать подходящую технологию хранения больших данных. <p>Владеть:</p> <ul style="list-style-type: none"> - навыками работы с программными продуктами по обработке больших данных.
ИД-3.ПК-12: Обеспечивает сопровождение информационных систем и сервисов
<p>Знать:</p> <ul style="list-style-type: none"> - сущность, особенности и типы больших данных; - методы анализа и хранения больших объемов данных; - этапы жизненного цикла обработки больших данных; - языки, наиболее приспособленные для обработки и аналитики больших данных; - способы организации хранения и доступа к большим данным. <p>Уметь:</p> <ul style="list-style-type: none"> - выбрать подходящий инструмент анализа больших данных, выбрать подходящую технологию хранения больших данных. <p>Владеть:</p> <ul style="list-style-type: none"> - навыками работы с программными продуктами по обработке больших данных.

4. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)							
Код занятия	Наименование разделов и тем /вид занятия/	Семестр / Курс	Часов	Компетенции	Литература	Инте вкст.	Примечание
	Раздел 1. Лекционные занятия						
1.1	Тема 1. Определение больших данных. Жизненный цикл больших данных /Лек/	7	2	ИД-1.ПК-7 ИД-2.ПК-7 ИД-1.ПК-8 ИД-2.ПК-8 ИД-1.ПК-12	Л1.1 Л1.2Л2.1	1	
1.2	Тема 2. Основы управления большими данными /Лек/	7	2	ИД-1.ПК-7 ИД-2.ПК-7 ИД-1.ПК-8 ИД-2.ПК-8 ИД-1.ПК-12	Л1.1 Л1.2Л2.1	1	
1.3	Тема 3. Хранилище данных. Технологии хранения больших данных /Лек/	7	2	ИД-1.ПК-7 ИД-2.ПК-7 ИД-1.ПК-8 ИД-2.ПК-8 ИД-1.ПК-12	Л1.1 Л1.2Л2.1	0	
1.4	Тема 4. Архитектура и технологии обработки больших данных /Лек/	7	4	ИД-1.ПК-7 ИД-2.ПК-7 ИД-1.ПК-8 ИД-2.ПК-8 ИД-1.ПК-12	Л1.1 Л1.2Л2.1	0	
1.5	Тема 5. Технологии анализа больших данных. Статистические методы анализа больших данных /Лек/	7	4	ИД-1.ПК-7 ИД-2.ПК-7 ИД-1.ПК-8 ИД-2.ПК-8 ИД-1.ПК-12	Л1.1 Л1.2Л2.1	2	
1.6	Тема 6. Современные программные средства обработки и анализа больших данных /Лек/	7	4	ИД-1.ПК-7 ИД-2.ПК-7 ИД-1.ПК-8 ИД-2.ПК-8 ИД-1.ПК-12	Л1.1 Л1.2Л2.1	0	
	Раздел 2. Практические занятия						
2.1	Практическое занятие 1. Определение больших данных. Жизненный цикл больших данных. Основы управления большими данными Выполнение сквозного проекта по сбору, организации хранения, обработке и анализу больших данных в определенной сфере (по выбору студента). Проект выполняется в малых группах. /Пр/	7	2	ИД-1.ПК-7 ИД-2.ПК-7 ИД-3.ПК-7 ИД-1.ПК-8 ИД-2.ПК-8 ИД-1.ПК-12 ИД-2.ПК-12 ИД-3.ПК-12	Л1.1 Л1.2Л2.1	1	Практическое задание, тест, доклад, вопросы для подготовки к зачету с оценкой

2.2	Практическое занятие 2. Хранилище данных. Технологии хранения больших данных Выполнение сквозного проекта по сбору, организации хранения, обработке и анализу больших данных в определенной сфере (по выбору студента). Проект выполняется в малых группах. /Пр/	7	4	ИД-1.ПК-7 ИД-2.ПК-7 ИД-3.ПК-7 ИД-1.ПК-8 ИД-2.ПК-8 ИД-1.ПК-12 ИД-2.ПК-12 ИД-3.ПК-12	Л1.1 Л1.2Л2.1	1	Практическое задание, тест, доклад, вопросы для подготовки к зачету с оценкой
2.3	Практическое занятие 3. Архитектура системы обработки больших данных. Технологии обработки больших данных Выполнение сквозного проекта по сбору, организации хранения, обработке и анализу больших данных в определенной сфере (по выбору студента). Проект выполняется в малых группах. /Пр/	7	4	ИД-1.ПК-7 ИД-2.ПК-7 ИД-3.ПК-7 ИД-1.ПК-8 ИД-2.ПК-8 ИД-1.ПК-12 ИД-2.ПК-12 ИД-3.ПК-12	Л1.1 Л1.2Л2.1	2	Практическое задание, тест, доклад, вопросы для подготовки к зачету с оценкой
2.4	Практическое занятие 4. Технологии анализа больших данных. Статистические методы анализа больших данных Выполнение сквозного проекта по сбору, организации хранения, обработке и анализу больших данных в определенной сфере (по выбору студента). Проект выполняется в малых группах. /Пр/	7	8	ИД-1.ПК-7 ИД-2.ПК-7 ИД-3.ПК-7 ИД-1.ПК-8 ИД-2.ПК-8 ИД-1.ПК-12 ИД-2.ПК-12 ИД-3.ПК-12	Л1.1 Л1.2Л2.1	4	Практическое задание, тест, доклад, вопросы для подготовки к зачету с оценкой
2.5	Практическое занятие 5. Программные платформы и системы для обработки и анализа больших данных Выполнение сквозного проекта по сбору, организации хранения, обработке и анализу больших данных в определенной сфере (по выбору студента). Проект выполняется в малых группах. /Пр/	7	8	ИД-1.ПК-7 ИД-2.ПК-7 ИД-3.ПК-7 ИД-1.ПК-8 ИД-2.ПК-8 ИД-1.ПК-12 ИД-2.ПК-12 ИД-3.ПК-12	Л1.1 Л1.2Л2.1	4	Практическое задание, тест, доклад, вопросы для подготовки к зачету с оценкой
Раздел 3. Самостоятельная работа							
3.1	Тема 1. Определение больших данных. Жизненный цикл больших данных План самостоятельной работы 1. Самостоятельное изучение основной и дополнительной литературы по теме 2. Подготовка к текущему контролю. 3. Подготовка доклада с презентацией. /Ср/	7	6	ИД-1.ПК-7 ИД-2.ПК-7 ИД-1.ПК-8 ИД-2.ПК-8 ИД-1.ПК-12	Л1.1 Л1.2Л2.1	0	Практическое задание, тест, доклад, вопросы для подготовки к зачету с оценкой
3.2	Тема 2. Основы управления большими данными План самостоятельной работы 1. Самостоятельное изучение основной и дополнительной литературы по теме 2. Подготовка к текущему контролю. 3. Подготовка доклада с презентацией. /Ср/	7	6,1	ИД-1.ПК-7 ИД-2.ПК-7 ИД-1.ПК-8 ИД-2.ПК-8 ИД-1.ПК-12	Л1.1 Л1.2Л2.1	0	Практическое задание, тест, доклад, вопросы для подготовки к зачету с оценкой
3.3	Тема 3. Хранилище данных. Технологии хранения больших данных План самостоятельной работы 1. Самостоятельное изучение основной и дополнительной литературы по теме 2. Подготовка к текущему контролю. 3. Подготовка к практическим работам, формирование отчета по практическим работам. /Ср/	7	8	ИД-1.ПК-7 ИД-2.ПК-7 ИД-1.ПК-8 ИД-2.ПК-8 ИД-1.ПК-12	Л1.1 Л1.2Л2.1	0	Практическое задание, тест, доклад, вопросы для подготовки к зачету с оценкой

3.4	Тема 4. Архитектура и технологии обработки больших данных План самостоятельной работы 1. Самостоятельное изучение основной и дополнительной литературы по теме 2. Подготовка к текущему контролю. 3. Подготовка к практическим работам, формирование отчета по практическим работам. /Ср/	7	10	ИД-1.ПК-7 ИД-2.ПК-7 ИД-1.ПК-8 ИД-2.ПК-8 ИД-1.ПК-12	Л1.1 Л1.2Л2.1	0	Практическое задание, тест, доклад, вопросы для подготовки к зачету с оценкой
3.5	Тема 5. Технологии анализа больших данных. Статистические методы анализа больших данных План самостоятельной работы 1. Самостоятельное изучение основной и дополнительной литературы по теме 2. Подготовка к текущему контролю. 3. Подготовка к практическим работам, формирование отчета по практическим работам. /Ср/	7	12	ИД-1.ПК-7 ИД-2.ПК-7 ИД-1.ПК-8 ИД-2.ПК-8 ИД-1.ПК-12	Л1.1 Л1.2Л2.1	0	Практическое задание, тест, доклад, вопросы для подготовки к зачету с оценкой
3.6	Тема 6. Современные программные средства обработки и анализа больших данных План самостоятельной работы 1. Самостоятельное изучение основной и дополнительной литературы по теме 2. Подготовка к текущему контролю. 3. Подготовка к практическим работам, формирование отчета по практическим работам. /Ср/	7	12	ИД-1.ПК-7 ИД-2.ПК-7 ИД-1.ПК-8 ИД-2.ПК-8 ИД-1.ПК-12	Л1.1 Л1.2Л2.1	0	Практическое задание, тест, доклад, вопросы для подготовки к зачету с оценкой
Раздел 4. Консультации							
4.1	Консультация по дисциплине /Конс/	7	0,9	ИД-1.ПК-7 ИД-2.ПК-7 ИД-3.ПК-7 ИД-1.ПК-8 ИД-2.ПК-8 ИД-1.ПК-12 ИД-2.ПК-12 ИД-3.ПК-12	Л1.1 Л1.2Л2.1	0	
Раздел 5. Промежуточная аттестация (зачёт)							
5.1	Подготовка к зачёту /ЗачётСОц/	7	8,85	ИД-1.ПК-7 ИД-2.ПК-7 ИД-3.ПК-7 ИД-1.ПК-8 ИД-2.ПК-8 ИД-1.ПК-12 ИД-2.ПК-12 ИД-3.ПК-12	Л1.1 Л1.2Л2.1	0	
5.2	Контактная работа /КСРАтт/	7	0,15	ИД-1.ПК-7 ИД-2.ПК-7 ИД-3.ПК-7 ИД-1.ПК-8 ИД-2.ПК-8 ИД-1.ПК-12 ИД-2.ПК-12 ИД-3.ПК-12	Л1.1 Л1.2Л2.1	0	

5. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

5.1. Пояснительная записка

1. Назначение фонда оценочных средств. Оценочные средства предназначены для контроля и оценки образовательных достижений обучающихся, освоивших программу учебной дисциплины "Применение больших данных в экономике".
2. Фонд оценочных средств включает контрольные материалы для проведения текущего контроля в форме заданий для практической работы, тем докладов, тестов для текущего контроля, вопросов к зачету с оценкой для промежуточной аттестации.

5.2. Оценочные средства для текущего контроля

Примерные тесты для входного контроля:

1. Дайте определение Big Data

- а) комплексный набор инструментов обработки структурированных данных колоссальных объемов
- б) комплексный набор подходов, инструментов и методов обработки структурированных и неструктурированных данных колоссальных объемов
- в) комплексный набор методов обработки неструктурированных данных колоссальных объемов
- г) комплексный набор методов обработки структурированных данных колоссальных объемов

2. Отметьте верное понимание Variety в контексте характеристик Big Data...

- а) высокая скорость генерирования данных
- б) разные типы данных в колонках таблиц реляционных СУБД
- в) разнообразие отраслей, являющихся источниками данных
- г) разнообразие типов данных, включающих в себя структурированные, полуструктурированные и неструктурированные

3. Какие из задач решаются Big Data?

- а) мониторинг оборудования
- б) анализ социальных сетей
- в) оптимизация автомобильного движения
- г) все вышеперечисленное

4. Примером применения Big Data не может быть...

- а) зондирование Земли из космоса
- б) родительский контроль
- в) хранение данных клиентов в соцсетях
- г) доступ к крупнейшим библиотекам

5. Какое из нижеперечисленных понятий не относится к перечню необходимых критериев для создания проекта, связанного с Большими данными?

- а) географическое положение
- б) производительность
- в) гибкость анализа
- г) скорость принятия решения

6. Укажите фактор, способствовавший появлению тренда больших данных...

- а) маркетинговые кампании крупных корпораций
- б) повышение издержек на хранение данных
- в) появление новых технологий обработки потоковых данных
- г) выпуск баз данных с обработкой данных в памяти

7. Одна из главных целей Big Data – это...

- а) привлечение все больше пользователей
- б) рост числа обработок данных
- в) снижение издержек операций
- г) таргетирование пользователей

8. Закончите следующее предложение: «С точки зрения машины, информация становится структурированной, если...

- а) машина проинструктирована, каким образом её обрабатывать
- б) информация разделена на части и озаглавлена
- в) информация имеет логическую взаимосвязь внутри себя
- г) машина знает из каких частей состоит информация

9. Выберите неверное высказывание...

- а) большие данные – это данные объёма свыше 1 Тб
- б) проблема больших данных – это проблема, когда при существующих технологиях хранения и обработки существенная обработка данных затруднена или невозможна
- в) большие данные – это тренд в области ИТ, подогреваемый маркетинговыми кампаниями крупных вендоров
- г) большие данные как правило не структурированы

10. До появления Data Mining невозможно было найти связи...

- а) студентов вуза и их успеваемости по предмету
- б) течения COVID-19 и осложнений после заболевания по всей РФ

- в) рекламирования и покупательской активности на сайте веб-магазина
- г) сроков доставки заказа через службу доставки и числом курьеров

Примерные тесты для текущего контроля

1. Выберите технологию потоковой обработки событий в режиме реального времени

- а) Spark Streaming
- б) Apache Hadoop
- в) MapReduce
- г) Apache Kafka

2. Укажите лишний этап построения статистической модели:

- а) сбор и верификация исходных данных;
- б) выбор факторов;
- в) построение модели;
- г) получение оценок;
- д) согласование полученных результатов с заинтересованными лицами;
- е) проверка статистической значимости модели.

3. Кто ввел термин Большие данные?

- а) Клиффорд Линч
- б) Алан Тьюринг
- в) Бьерн Страуструп
- г) Дональд Кнут

4. Какие данные занимают больше мировой памяти относительно остальных?

- а) Structured Data
- б) Unstructured Data
- в) Semi-Structured Data
- г) Quasi-Structured Data

5. Какой метод верификации исходных данных не применяется для верификации данных о стоимости активов?

- а) семантические анализаторы;
- б) матрицы граничных значений;
- в) конверторы отраслевых классификаторов;
- г) наборы решающих правил;
- д) проверка данных с использованием колл-центра;
- е) тестовые и валидационные выборки.

6. Какие нейронные сети лучше подходят для задач поиска аналога исследуемого объекта?

- а) сети Кохонена;
- б) сети встречного распространения;
- в) RBF-сети на радиальных базисных функциях;
- г) любые MLP-нейросети;
- д) все вышеперечисленное.

7. Какие требования к факторам предъявляют классические статистические модели?

- а) значимость;
- б) независимость;
- в) внятная экономическая интерпретация;
- г) все вышеперечисленное.

8. Какая технология машинного обучения реагирует на возникновение новых, не описанных ранее ситуаций, получая данные из внешней среды?

- а) обучение с подкреплением;
- б) обучение с противником;
- в) вероятностное прогнозирование;
- г) распознавание образов.

9. Как не используют выборки из генеральной совокупности аналитики больших данных?

- а) как метод формирования комплексного суждения о генеральной совокупности случайной величины;
- б) как метод тестирования полученных моделей;
- в) как метод верификации исходных данных.

10. Метод главных компонент применяется для решения проблемы:

- а) робастности;
- б) мультиколлинеарности;

- в) гомоскедастичности;
- г) гетероскедастичности.

11. В искусственной нейронной сети типа RBF (сеть с радиальной базисной функцией) применяются следующие виды активационных функций:

- а) радиальная базисная;
- б) линейная;
- в) пороговая;
- г) сигмоидальная.

12. Какая компания создала технологию MapReduce?

- а) Google
- б) Yahoo
- в) EMC
- г) Oracle

13. В искусственной нейронной сети типа MLP (многослойный перцептрон) присутствуют следующие связи:

- а) каждый нейрон связан с каждым нейроном следующего слоя;
- б) каждый нейрон связан с каждым нейроном своего слоя;
- в) каждый нейрон связан с каждым нейроном входного слоя;
- г) каждый нейрон связан с каждым нейроном выходного слоя.

14. В какой последовательности технология MapReduce использует в рабочем процессе задачи-распределители и задачи-редукторы?

- а) последовательно, сначала одни, а затем другие;
- б) параллельно или обе одновременно;
- в) поочередно, одну за другой.

15. Какие функции в MapReduce запускает главный Мастер-контроллер (найдите неверный ответ)?

- а) создание распределителей и редукторов;
- б) назначение задач рабочим процессам;
- в) обработку отказа узла редуктора.

16. В чем состоит стратегия кластеризации?

- а) в объединении близких точек многомерного пространства в один объект (кластер) с усредненными характеристиками;
- б) разделении множества на части с помощью плоскостей;
- с) разделении множества на внутренние, или «свои», точки и внешние, или «чужие», точки.

17. Как ускорить алгоритм иерархической кластеризации для евклидовой метрики?

- а) следует матрицу расстояний между элементами рассчитать только один раз;
- б) следует элементы, объединенные в кластер, вычеркивать из матрицы расстояний;
- в) следует пересчитывать среднее значение для кластера без привлечения исходных координат.

18. В каком году впервые был введен термин Большие данные?

- а) 2002
- б) 2004
- в) 2006
- г) 2008.

19. Как реализуется алгоритм кластеризации потока?

- а) точки потока разбиваются на одинаковые интервалы, в которых хранится информация о кластере;
- б) точки потока разбиваются на интервалы, размеры которых являются степенями двойки;
- в) точки потока разбиваются на интервалы, размеры которых уменьшаются в два раза.

20. Перечислите четыре основные характеристики Big Data:

- а) Virtualization, Volume, Variability, Vehicle;
- б) Variety, Velocity, Volume, Value;
- в) Verification, Volume, Velocity, Visualization;
- г) Video, Value, Variety, Volume.

Критерии оценки:

«отлично» (26 - 28 балла), повышенный уровень - даны верные ответы на 84-100% вопросов

«хорошо» (22- 25 балла), пороговый уровень - даны верные ответы на 66-83% вопросов

«удовлетворительно» (16 - 21 балл), пороговый уровень - даны верные ответы на 50-65% вопросов

«неудовлетворительно» (0 – 15 балла), уровень не сформирован - даны верные ответы на менее 50% вопросов

Практическая работа "Выполнение сквозного проекта по сбору, организации хранения, обработке и анализу больших данных в определенной сфере (по выбору студента)". Проект выполняется в малых группах.

1. Определение цели и задач проекта.
2. Определение типов больших данных. Определение открытых источников данных. Поиск необходимой для решения поставленной задачи информации, критически оценивая надежность различных источников информации на основе больших данных
3. Изучение технологии по сбору больших данных. Осуществление сбора больших данных.
4. Изучение технологии хранения больших данных.
5. Изучение алгоритма, инструментариев и технологии обработки и анализа больших данных. Обработка и анализ больших данных по проекту.
6. Формулирование выводов по результатам анализа больших данных.

Критерии оценки:

Общие требования к выполнению практических работ:

1. изучение теоретического материала;
2. выполнение заданий;
3. ответы на контрольные вопросы.

Форма отчетности:

Практические работы должны оформляться в отдельном файле и содержать:

- номер и название работы;
- цель работы;
- подробное описание хода выполнения заданий;

краткие ответы на контрольные вопросы.

Форма отчетности работ:

Результатом выполнения практических работ является устная защита с предъявлением оформленной работы

«отлично» (91-100%), повышенный уровень

Практическая работа выполнена в полном объеме и правильно. При выполнении задания студент использовал конспекты лекции, основную и дополнительную литературу. Студент выполнил задание самостоятельно.

«хорошо» (75-90%), пороговый уровень

Практическая работа выполнена в полном объеме, но есть отдельные замечания по содержанию задания. При выполнении задания студент использовал конспекты лекции, основную и дополнительную литературу. Студент выполнил задание самостоятельно.

«удовлетворительно» (50-74%), пороговый уровень

Практическая работа выполнена не в полном объеме и есть замечания по содержанию. При выполнении задания студент использовал конспекты лекции, основную литературу.

Студент выполнил задание самостоятельно.

«неудовлетворительно» (менее 50%), уровень не сформирован

Практическая работа выполнена не в полном объеме и с ошибками. Сложно оценить самостоятельность выполнения задания.

5.3. Темы письменных работ (эссе, рефераты, курсовые работы и др.)

Темы докладов:

1. Этапы работы с большими данными. Сбор, консолидация, анализ, принятие решений
2. Инфраструктурные аспекты хранения больших данных. Сети хранения данных
3. Масштабирование систем хранения и обработки больших данных
4. Веб-сервисы доступа к распределенной файловой системы
5. Информационная инфраструктура дата-центров
6. Телекоммуникационная инфраструктура дата-центров
7. Инженерная инфраструктура дата-центра
8. Отказоустойчивость центров хранения и обработки больших данных
9. Стандартизация и нормы систем хранения и обработки больших данных
10. Краудсорсинг больших данных
11. Аналитика и визуализация больших данных с помощью инструмента Grafana
12. Использование облачных вычислений при помощи Apache Spark, и Apache Ignite

Критерии оценки:

Доклад должен иметь следующую структуру:

- титульный лист
- оглавление (в нем последовательно излагаются названия пунктов доклада, указываются страницы, с которых начинается каждый пункт);
- введение (формулирует суть исследуемой проблемы, обосновывается выбор темы, определяются ее значимость и актуальность, указываются цель и задачи доклада, дается характеристика используемой литературы);
- основная часть (каждый раздел ее, доказательно раскрывая отдельную проблему или одну из ее сторон, логически является продолжением предыдущего; в основной части могут быть представлены таблицы, графики, схемы);

- заключение (подводятся итоги или дается обобщенный вывод по теме доклада, предлагаются рекомендации);
- список использованных источников.

«отлично» (4 балла), повышенный уровень

Содержание доклада соответствует требованиям к содержанию и структуре доклада. Тема доклада раскрыта полностью. Выдержан объём, соблюдены требования к внешнему оформлению, изложение материала логично, доступно. На основе представленной презентации формируется полное понимание тематики исследования, раскрыты детали. Презентация полностью соответствует установленным требованиям. При подготовке доклада использованы различные актуальные источники информации. Все ответы на вопросы по докладу при его обсуждении исчерпывающие, аргументированные, корректные. Выступление докладчика полностью соответствует критериям ораторского искусства.

«хорошо» (3 балла), пороговый уровень

Содержание доклада за исключением отдельных моментов, соответствует требованиям к содержанию и структуре доклада. Тема раскрыта в достаточной степени, но не полностью, приведены примеры из практики. Информации представлено недостаточно; в отдельных случаях нарушена логика в изложении материала, не совсем доступно. На основе представленной презентации формируется общее понимание тематики исследования, но не ясны детали. При подготовке доклада использованы различные актуальные источники информации. Не все ответы на вопросы по докладу при его обсуждении исчерпывающие, аргументированные, корректные.

«удовлетворительно» (2 балл), пороговый уровень

Содержание доклада частично соответствует требованиям к содержанию и структуре доклада. Раскрыта малая часть темы; поиск информации проведён поверхностно; в изложении материала отсутствует логика, доступность. На основе представленной презентации не совсем понятна тематика исследования, детали не раскрыты. При подготовке доклада использованы не актуальные источники информации. При обсуждении доклада ответов на вопросы не было, или они не соответствовали заданным вопросам. Выступление докладчика частично соответствует критериям ораторского искусства.

«неудовлетворительно» (0 баллов), уровень не сформирован

Содержание темы не раскрыто, обнаруживается существенное непонимание проблемы.

5.4. Оценочные средства для промежуточной аттестации

1. Понятие Большие данные. Роль цифровой информации в 21 веке.
2. Проблемы анализа и обработки большого объема данных.
3. Базовые принципы обработки больших данных.
4. Жизненный цикл больших данных.
5. Представление наглядной визуализации больших данных (публикация больших данных) (Data Publication)
6. Системы управления потоками данных в системах больших данных
7. Системы хранения больших данных
8. Архивация больших данных (Data Archival)
9. Уничтожение больших данных (Data Purging)
10. Понятие метаданных в больших данных
11. Основные методы критического анализа и основы системного подхода как общенаучного метода изучения больших данных
12. Анализ задач больших данных с использованием основ критического анализа и системного подхода
13. Поиск необходимой для решения поставленной задачи информации, критически оценивая надежность различных источников информации на основе больших данных
14. Оценка требований и анализ контента в больших данных
15. Спецификация системных требований в больших данных
16. Система метаданных в больших данных
17. Сервис и оценка в больших данных
18. Распределенные файловые системы больших данных
19. Распределенные фреймворки
20. Бенчмаркинг в больших данных
21. Серверное программирование в больших данных
22. Планирование больших данных с использованием основ теории вероятностей, математической статистики и эконометрики, в том числе, методов и форм организации статистического наблюдения, методологии первичной обработки статистической информации
23. Планирование больших данных в части экономической информации с использованием следующих методов анализа и обработки: временные ряды, перекрёстные (cross-section) данные, панельные данные; основы регрессионного анализа (линейная модель множественной регрессии); суть метода наименьших квадратов (МНК) и его применение в экономическом анализе; основные методы диагностики (проверки качества) эконометрических моделей
24. Системы развертывания больших данных
25. Интеграция больших данных
26. Информационная безопасность больших данных
27. Машинное обучение на основе больших данных
28. Базы данных NoSQL и новые SQL базы данных
29. Прием больших данных (Data Ingestion)

31. Анализ больших данных (Analysis Layer)
32. Представление результатов больших данных (Consumption Layer)
33. Операторы Map и Reduce в больших данных
34. Оператор Reduce (свертка) в больших данных
35. Оператор Map в больших данных
36. Платформы больших данных
37. Проведение статистических тестов и построение доверительных интервалов, определение статистических свойств полученных оценок при обработке больших данных в реальном масштабе времени
38. Системы управления большими данными
39. Проведение анализа и содержательной интерпретации полученных результатов на основе больших данных с использованием инструментария аналитических платформ

Критерий оценки:

Зачтено «отлично», повышенный уровень

Студентом дан полный, в логической последовательности развернутый ответ на поставленный вопрос, где он продемонстрировал знания предмета в полном объеме учебной программы, достаточно глубоко осмысливает дисциплину, самостоятельно, и исчерпывающе отвечает на дополнительные вопросы, приводит собственные примеры по проблематике поставленного вопроса.

Зачтено, «хорошо», пороговый уровень

Студентом дан развернутый ответ на поставленный вопрос, где студент демонстрирует знания, приобретенные на лекционных и семинарских занятиях, а также полученные посредством изучения обязательных учебных материалов по курсу, дает аргументированные ответы, приводит примеры, в ответе присутствует свободное владение монологической речью, логичность и последовательность ответа. Однако допускается неточность в ответе.

Зачтено, «удовлетворительно», пороговый уровень

Студентом дан ответ, свидетельствующий в основном о знании процессов изучаемой дисциплины, отличающийся недостаточной глубиной и полнотой раскрытия темы, знанием основных вопросов теории, слабо сформированными навыками анализа явлений, процессов, недостаточным умением давать аргументированные ответы и приводить примеры, недостаточно свободным владением монологической речью, логичностью и последовательностью ответа. Допускается несколько ошибок в содержании ответа.

Не зачтено, «неудовлетворительно», уровень не сформирован

Студентом дан ответ, который содержит ряд серьезных неточностей, обнаруживающий незнание процессов изучаемой предметной области, отличающийся неглубоким раскрытием темы, незнанием основных вопросов теории, несформированными навыками анализа явлений, процессов, неумением давать аргументированные ответы, слабым владением монологической речью, отсутствием логичности и последовательности. Выводы поверхностны. Т.е. студент не способен ответить на вопросы даже при дополнительных наводящих вопросах преподавателя.

6. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

6.1. Рекомендуемая литература

6.1.1. Основная литература

	Авторы, составители	Заглавие	Издательство, год	Эл. адрес
Л1.1	Маккинли Уэс, Слинкина А.	Python и анализ данных: практическое издание	Саратов: Профобразование, 2019	http://www.iprbookshop.ru/88752.html
Л1.2	Волкова В.М., Семёнова М.А., Четвертакова [и др.] Е.С.	Программные системы статистического анализа. Обнаружение закономерностей в данных с использованием системы R и языка Python: учебное пособие	Новосибирск: Новосибирский государственный технический университет (НГТУ), 2017	http://www.iprbookshop.ru/91682.html

6.1.2. Дополнительная литература

	Авторы, составители	Заглавие	Издательство, год	Эл. адрес
Л2.1	Мирошников А.И.	Архитектура систем управления базами данных: учебное пособие	Липецк: Липецкий государственный технический университет, ЭБС АСВ, 2018	http://www.iprbookshop.ru/83189.html

6.3.1 Перечень программного обеспечения

6.3.1.1	Kaspersky Endpoint Security для бизнеса СТАНДАРТНЫЙ			
6.3.1.2	MS Office			

6.3.1.3	MS WINDOWS
6.3.1.4	Deductor Academic
6.3.1.5	NVDA
6.3.1.6	Python
6.3.1.7	Python(x,y)
6.3.1.8	R for Windows
6.3.1.9	Яндекс.Браузер
6.3.1.10	LibreOffice
6.3.2 Перечень информационных справочных систем	
6.3.2.1	База данных «Электронная библиотека Горно-Алтайского государственного университета»
6.3.2.2	Электронно-библиотечная система IPRbooks
6.3.2.3	Межвузовская электронная библиотека

7. ОБРАЗОВАТЕЛЬНЫЕ ТЕХНОЛОГИИ	
	презентация
	выполнение практических работ
	работа с программными продуктами
	проблемная лекция
	лекция-визуализация
	метод проектов

8. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)		
Номер аудитории	Назначение	Основное оснащение
317 А2	Компьютерный класс, класс деловых игр, центр (класс) деловых игр, класс имитации деятельности предприятия, лаборатория имитации деятельности предприятия, учебно-тренинговый центр (лаборатория), лаборатория информационно-коммуникативных технологий. Учебная аудитория для проведения занятий лекционного типа, занятий семинарского типа, курсового проектирования (выполнения курсовых работ), групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации. Помещение для самостоятельной работы	Рабочее место преподавателя. Посадочные места обучающихся (по количеству обучающихся). Интерактивная доска с проектором, экран, подключение к интернету, ученическая доска, презентационная трибуна

9. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ (МОДУЛЯ)
<p>Методические указания по самостоятельной работе студента</p> <p>Самостоятельная работа призвана способствовать развитию у студентов творческих навыков, инициативы, умению организовать свое время.</p> <p>При выполнении плана самостоятельной работы студенту необходимо изучить теоретический материал, представленный в рекомендуемой литературе, творчески его переработать и представить его для отчета в форме, рекомендованной планом самостоятельной работы.</p> <p>Самостоятельная работа студента должна быть полной, раскрывающей уровень освоения студентом той или иной темы и грамотно оформленной, показывающей творческий и инициативный подход студента к выполнению задания.</p> <p>В рамках плана самостоятельной работы по данной дисциплине предусмотрены следующие виды самостоятельной работы:</p> <ul style="list-style-type: none"> - подготовка к лабораторным занятиям. Такая подготовка предусматривает самостоятельное изучение основной и дополнительной литературы, лекционного материала и подготовку ответов на вопросы и выполнение заданий по теме лабораторного занятия; - подготовка к зачету предусматривает самостоятельную подготовку студента по экзаменационным вопросам, которые представлены в рабочей программе дисциплины; - подготовка докладов и презентации по отдельным вопросам курса, рекомендованным для самостоятельного изучения. <p>Методические рекомендации по подготовке презентации</p> <p>Компьютерную презентацию, сопровождающую выступление докладчика, удобнее всего подготовить в программе MS</p>

PowerPoint. Презентация как документ представляет собой последовательность сменяющих друг друга слайдов - то есть электронных страничек, занимающих весь экран монитора (без присутствия панелей программы). Количество слайдов адекватно содержанию и продолжительности выступления (например, для 5-минутного выступления рекомендуется использовать не более 10 слайдов). На первом слайде обязательно представляется тема выступления и сведения об авторах. Следующие слайды можно подготовить, используя две различные стратегии их подготовки:

1 стратегия: на слайды выносятся опорный конспект выступления и ключевые слова с тем, чтобы пользоваться ими как планом для выступления. В этом случае к слайдам предъявляются следующие требования: объем текста на слайде – не больше 7 строк; маркированный/нумерованный список содержит не более 7 элементов; отсутствуют знаки пунктуации в конце строк в маркированных и нумерованных списках; значимая информация выделяется с помощью цвета, кегля, эффектов анимации.

2 стратегия: на слайды помещается фактический материал (таблицы, графики, фотографии и пр.), который является уместным и достаточным средством наглядности, помогает в раскрытии стержневой идеи выступления. В этом случае к слайдам предъявляются следующие требования: выбранные средства визуализации информации (таблицы, схемы, графики и т. д.) соответствуют содержанию; использованы иллюстрации хорошего качества (высокого разрешения), с четким изображением. Максимальное количество графической информации на одном слайде – 2 рисунка (фотографии, схемы и т.д.) с текстовыми комментариями (не более 2 строк к каждому). Наиболее важная информация должна располагаться в центре экрана. Основная ошибка при выборе данной стратегии – «соревнование» со своим иллюстративным материалом (аудитории не предоставляется достаточно времени, чтобы воспринять материал на слайдах).

Методические указания по подготовке докладов

Доклад – это развернутое устное изложение какой-либо темы, сделанное публично.

Отличительными признаками доклада являются:

- передача в устной форме информации;
- публичный характер выступления;
- стилевая однородность доклада;
- четкие формулировки и сотрудничество докладчика и аудитории;
- умение в сжатой форме изложить ключевые положения исследуемого вопроса и сделать выводы.

Этапы подготовки доклада:

- выбор темы доклада;
- подбор и изучение наиболее важных учебных, научных работ по данной теме, нормативных правовых актов;
- анализ изученного материала, выделение наиболее значимых для раскрытия темы доклада фактов, мнений;
- составление плана доклада;
- написание текста доклада с соблюдением требований научного стиля.

Структура доклада:

1. Вступление, в котором указываются: тема доклада; цель (основная идея) доклада; связь данной темы с другими темами; краткое перечисление рассматриваемых вопросов; актуальность, проблематика темы, современная оценка предмета изложения; краткий обзор изученной литературы по данной теме и т.п.
2. Основная часть, в которой выступающий должен глубоко раскрыть суть затронутой темы, обычно строится по принципу отчета. Задача основной части - представить достаточно данных для того, чтобы слушатели и заинтересовались темой и захотели ознакомиться с материалами. При этом логическая структура теоретического блока не должны даваться без наглядных пособий, аудио-визуальных и визуальных материалов.
3. Заключение, в котором: подводятся итоги, формулируются выводы; подчеркивается значение рассмотренной проблемы; • выделяются основные проблемы, пути и способы их решения и т.п.