

МИНИСТЕРСТВО СЕЛЬСКОГО ХОЗЯЙСТВА РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«КУБАНСКИЙ ГОСУДАРСТВЕННЫЙ АГРАРНЫЙ УНИВЕРСИТЕТ
имени И.Т. ТРУБИЛИНА»

Факультет прикладной информатики
Компьютерных технологий и систем

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)
«БОЛЬШИЕ ДАННЫЕ»**

Уровень высшего образования: бакалавриат

Направление подготовки: 09.03.02 Информационные системы и технологии

Направленность (профиль) подготовки: Создание, модификация и сопровождение информационных систем, администрирование баз данных

Квалификация (степень) выпускника: бакалавр

Форма обучения: очная

Год набора: 2024

Срок получения образования: 4 года

Объем: в зачетных единицах: 6 з.е.
в академических часах: 216 ак.ч.

2024

Разработчики:

Старший преподаватель, кафедры компьютерных технологий и систем Параскевов А.В.

Рабочая программа дисциплины (модуля) составлена в соответствии с требованиями ФГОС ВО по направлению подготовки Направление подготовки: 09.03.02 Информационные системы и технологии, утвержденного приказом Минобрнауки России от 19.09.2017 №926, с учетом трудовых функций профессиональных стандартов: "Специалист по информационным системам", утвержден приказом Минтруда России от 13.07.2023 № 586н; "Администратор баз данных", утвержден приказом Минтруда России от 17.09.2014 № 647н; "Специалист по информационным системам", утвержден приказом Минтруда России от 18.11.2014 № 896н; "Администратор баз данных", утвержден приказом Минтруда России от 27.04.2023 № 408н.

Согласование и утверждение

№	Подразделение или коллегиальный орган	Ответственное лицо	ФИО	Виза	Дата, протокол (при наличии)
1	Компьютерных технологий и систем	Заведующий кафедрой, руководитель подразделения, реализующего ОП	Лукьяненко Т.В.	Согласовано	22.03.2024, № 9

1. Цель и задачи освоения дисциплины (модуля)

Цель освоения дисциплины - изучение математических методов и моделей, используемых в системах обработки и анализа больших данных для поддержки принятия решений, и развитие профессиональных навыков в этой области

Задачи изучения дисциплины:

- формирование теоретических и методологических основ в области анализа неструктурированной информации;
- формирование практических навыков, использования алгоритмов интеллектуального анализа данных;
- формирование навыков проведения сравнительного анализа основных моделей.

2. Планируемые результаты обучения по дисциплине (модулю), соотнесенные с планируемыми результатами освоения образовательной программы

Компетенции, индикаторы и результаты обучения

УК-2 Способен определять круг задач в рамках поставленной цели и выбирать оптимальные способы их решения, исходя из действующих правовых норм, имеющихся ресурсов и ограничений

УК-2.1 Формулирует в рамках поставленной цели проекта совокупность взаимосвязанных задач, обеспечивающих ее достижение.

Определяет ожидаемые результаты решения выделенных задач

Знать:

УК-2.1/Зн1 В рамках поставленной цели проекта совокупность взаимосвязанных задач, обеспечивающих ее достижение, а также ожидаемые результаты решения выделенных задач

Уметь:

УК-2.1/Ум1 Формулировать в рамках поставленной цели проекта совокупность взаимосвязанных задач, обеспечивающих ее достижение. Определять ожидаемые результаты решения выделенных задач

Владеть:

УК-2.1/Нв1 Способностью формулировать в рамках поставленной цели проекта совокупность взаимосвязанных задач, обеспечивающих ее достижение, определять ожидаемые результаты решения выделенных задач

УК-2.2 Проектирует решение конкретной задачи проекта, выбирая оптимальный способ ее решения, исходя из действующих правовых норм и имеющихся ресурсов и ограничений

Знать:

УК-2.2/Зн1 Методы проектирования решения конкретной задачи проекта, выбирая оптимальный способ ее решения, исходя из действующих правовых норм и имеющихся ресурсов и ограничений

Уметь:

УК-2.2/Ум1 Проектировать решение конкретной задачи проекта, выбирая оптимальный способ ее решения, исходя из действующих правовых норм и имеющихся ресурсов и ограничений

Владеть:

УК-2.2/Нв1 Методами проектирования решения конкретной задачи проекта, выбирая оптимальный способ ее решения, исходя из действующих правовых норм и имеющихся ресурсов и ограничений

УК-2.3 Решает конкретные задачи проекта заявленного качества и за установленное время

Знать:

УК-2.3/Зн1 Решение конкретных задач проекта заявленного качества и за установленное время

Уметь:

УК-2.3/Ум1 Решать конкретные задачи проекта заявленного качества и за установленное время

Владеть:

УК-2.3/Нв1 Способностью решать конкретные задачи проекта заявленного качества и за установленное время

УК-2.4 Публично представляет результаты решения конкретной задачи проекта

Знать:

УК-2.4/Зн1 Этапы решения конкретной задачи проекта

Уметь:

УК-2.4/Ум1 Публично представляет результаты решения конкретной задачи проекта

Владеть:

УК-2.4/Нв1 Способностью публично представляет результаты решения конкретной задачи проекта

УК-6 Способен управлять своим временем, выстраивать и реализовывать траекторию саморазвития на основе принципов образования в течение всей жизни

УК-6.1 Применяет знание о своих ресурсах и их пределах (личностных, ситуативных, временных и т.д.) для успешного выполнения порученной работы

Знать:

УК-6.1/Зн1 Информацию о своих ресурсах и их пределах (личностных, ситуативных, временных и т.д.), для успешного выполнения порученной работы

Уметь:

УК-6.1/Ум1 Применить знания о своих ресурсах и их пределах (личностных, ситуативных, временных и т.д.), для успешного выполнения порученной работы

Владеть:

УК-6.1/Нв1 Способностью применить знания о своих ресурсах и их пределах (личностных, ситуативных, временных и т.д.), для успешного выполнения порученной работы

УК-6.2 Понимает важность планирования перспективных целей собственной деятельности с учетом условий, средств, личностных возможностей, этапов карьерного роста, временной перспективы развития деятельности и требований рынка труда

Знать:

УК-6.2/Зн1 Важность планирования перспективных целей собственной деятельности с учетом условий, средств, личностных возможностей, этапов карьерного роста, временной перспективы развития деятельности и требований рынка труда

Уметь:

УК-6.2/Ум1 Понимать важность планирования перспективных целей собственной деятельности с учетом условий, средств, личностных возможностей, этапов карьерного роста, временной перспективы развития деятельности и требований рынка труда

Владеть:

УК-6.2/Нв1 Способностью понимания важности планирования перспективных целей собственной деятельности с учетом условий, средств, личностных возможностей, этапов карьерного роста, временной перспективы развития деятельности и требований рынка труда

УК-6.3 Реализует намеченные цели деятельности с учетом условий, средств, личностных возможностей, этапов карьерного роста, временной перспективы развития деятельности и требований рынка труда

Знать:

УК-6.3/Зн1 Возможность реализации намеченных целей деятельности с учетом условий, средств, личностных возможностей, этапов карьерного роста, временной перспективы развития деятельности и требований рынка труда

Уметь:

УК-6.3/Ум1 Реализовать намеченные цели деятельности с учетом условий, средств, личностных возможностей, этапов карьерного роста, временной перспективы развития деятельности и требований рынка труда

Владеть:

УК-6.3/Нв1 Способами реализации намеченных целей деятельности с учетом условий, средств, личностных возможностей, этапов карьерного роста, временной перспективы развития деятельности и требований рынка труда

УК-6.4 Критически оценивает эффективность использования времени и других ресурсов при решении поставленных задач, а также относительно полученного результата

Знать:

УК-6.4/Зн1 Оценка эффективности использования времени и других ресурсов при решения поставленных задач, а также относительно полученного результата

Уметь:

УК-6.4/Ум1 Критически оценивать эффективность использования времени и других ресурсов при решения поставленных задач, а также относительно полученного результата

Владеть:

УК-6.4/Нв1 Способностью критически оценить эффективность использования времени и других ресурсов при решения поставленных задач, а также относительно полученного результата

УК-6.5 Демонстрирует интерес к учебе и использует предоставляемые возможности для приобретения новых знаний и навыков

Знать:

УК-6.5/Зн1 Предоставляемые возможности для приобретения новых знаний и навыков

Уметь:

УК-6.5/Ум1 Использовать предоставляемые возможности для приобретения новых знаний и навыков

Владеть:

УК-6.5/Нв1 Способностью использовать предоставляемые возможности для приобретения новых знаний и навыков

ОПК-2 Способен использовать современные информационные технологии и программные средства, в том числе отечественного производства, при решении задач профессиональной деятельности

ОПК-2.1 Знает современные информационные технологии и программные средства, в том числе отечественного производства, при решении задач профессиональной деятельности.

Знать:

ОПК-2.1/Зн1 Современные информационные технологии, в том числе отечественного производства

ОПК-2.1/Зн2 Современные программные средства, в том числе отечественного производства.

ОПК-2.2 Умеет выбирать современные информационные технологии и программные средства, в том числе отечественного производства, при решении задач профессиональной деятельности.

Уметь:

ОПК-2.2/Ум1 Умеет выбирать современные информационные, в том числе отечественного производства, при решении задач профессиональной деятельности.

ОПК-2.2/Ум2 Умеет выбирать современные программные средства, в том числе отечественного производства, при решении задач профессиональной деятельности.

ОПК-2.3 Владеет навыками применения современных информационных технологий и программных средств, в том числе отечественного производства, при решении задач профессиональной деятельности.

Владеть:

ОПК-2.3/Нв1 Навыками применения современных информационных технологий, в том числе отечественного производства, при решении задач профессиональной деятельности.

ОПК-2.3/Нв2 Навыками применения современных программных средств, в том числе отечественного производства, при решении задач профессиональной деятельности.

3. Место дисциплины в структуре ОП

Дисциплина (модуль) «Большие данные» относится к обязательной части образовательной программы и изучается в семестре(ах): 5, 6.

В процессе изучения дисциплины студент готовится к видам профессиональной деятельности и решению профессиональных задач, предусмотренных ФГОС ВО и образовательной программой.

4. Объем дисциплины и виды учебной работы

Период обучения	Общая трудоемкость (часы)	Общая трудоемкость (ЗЕТ)	Контактная работа (часы, всего)	Внеаудиторная контактная работа (часы)	Зачет (часы)	Лабораторные занятия (часы)	Лекционные занятия (часы)	Практические занятия (часы)	Самостоятельная работа (часы)	Промежуточная аттестация (часы)
Пятый семестр	72	2	61	1		14	30	16	11	Зачет

Шестой семестр	144	4	47	1		14	16	16	97	Зачет с оценкой
Всего	216	6	108	2		28	46	32	108	

5. Содержание дисциплины

5.1. Разделы, темы дисциплины и виды занятий

(часы промежуточной аттестации не указываются)

Наименование раздела, темы	Всего	Внеаудиторная контактная работа	Лабораторные занятия	Лекционные занятия	Практические занятия	Самостоятельная работа	Планируемые результаты обучения, соответствующие результатам освоения программы
Раздел 1. Раздел 1	8	1	2	2	2	1	УК-2.1
Тема 1.1. Группировка данных	8	1	2	2	2	1	
Раздел 2. Раздел 2	9		2	4	2	1	УК-2.2
Тема 2.1. Обнаружение значимых корреляций	9		2	4	2	1	
Раздел 3. Раздел 3	9		2	4	2	1	УК-2.3
Тема 3.1. Определение отношений между разнородными данными	9		2	4	2	1	
Раздел 4. Раздел 4	10		2	4	2	2	УК-2.4
Тема 4.1. Систематизированные структуры данных	10		2	4	2	2	
Раздел 5. Раздел 5	10		2	4	2	2	УК-6.1
Тема 5.1. Вывод правил для принятия решений	10		2	4	2	2	
Раздел 6. Раздел 6	10		2	4	2	2	УК-6.2
Тема 6.1. Прогнозирование последствий принятых решений	10		2	4	2	2	
Раздел 7. Раздел 7	8		1	4	2	1	УК-6.3
Тема 7.1. Регрессионный анализ	8		1	4	2	1	
Раздел 8. Раздел 8	8		1	4	2	1	УК-6.4
Тема 8.1. Дисперсионный анализ	8		1	4	2	1	
Раздел 9. Раздел 9	19	1	2	2	2	12	УК-6.5
Тема 9.1. Кластерный анализ	19	1	2	2	2	12	
Раздел 10. Раздел 10	18		2	2	2	12	ОПК-2.1
Тема 10.1. Факторный анализ	18		2	2	2	12	
Раздел 11. Раздел 11	18		2	2	2	12	ОПК-2.2
Тема 11.1. Модели распределенных файловых систем	18		2	2	2	12	
Раздел 12. Раздел 12	89		8	10	10	61	ОПК-2.3
Тема 12.1. Поиск подобий в данных	18		2	2	2	12	

Тема 12.2. Анализ потоковых данных	18		2	2	2	12
Тема 12.3. Социально-сетевые графы	18		2	2	2	12
Тема 12.4. Методы кластеризации	17		1	2	2	12
Тема 12.5. Способы применения нейронных сетей	18		1	2	2	13
Итого	216	2	28	46	32	108

5. Содержание разделов, тем дисциплин

Раздел 1. Раздел 1

(Внеаудиторная контактная работа - 1ч.; Лабораторные занятия - 2ч.; Лекционные занятия - 2ч.; Практические занятия - 2ч.; Самостоятельная работа - 1ч.)

Тема 1.1. Группировка данных

(Внеаудиторная контактная работа - 1ч.; Лабораторные занятия - 2ч.; Лекционные занятия - 2ч.; Практические занятия - 2ч.; Самостоятельная работа - 1ч.)

Понятие данных, совокупность инструментов, подходов и методов обработки больших данных

Раздел 2. Раздел 2

(Лабораторные занятия - 2ч.; Лекционные занятия - 4ч.; Практические занятия - 2ч.; Самостоятельная работа - 1ч.)

Тема 2.1. Обнаружение значимых корреляций

(Лабораторные занятия - 2ч.; Лекционные занятия - 4ч.; Практические занятия - 2ч.; Самостоятельная работа - 1ч.)

понятие корреляции, средства массово-параллельной обработки, прогнозная аналитика

Раздел 3. Раздел 3

(Лабораторные занятия - 2ч.; Лекционные занятия - 4ч.; Практические занятия - 2ч.; Самостоятельная работа - 1ч.)

Тема 3.1. Определение отношений между разнородными данными

(Лабораторные занятия - 2ч.; Лекционные занятия - 4ч.; Практические занятия - 2ч.; Самостоятельная работа - 1ч.)

Распознавание объектов, имитационное моделирование, пространственный анализ

Раздел 4. Раздел 4

(Лабораторные занятия - 2ч.; Лекционные занятия - 4ч.; Практические занятия - 2ч.; Самостоятельная работа - 2ч.)

Тема 4.1. Систематизированные структуры данных

(Лабораторные занятия - 2ч.; Лекционные занятия - 4ч.; Практические занятия - 2ч.; Самостоятельная работа - 2ч.)

Анализ временных рядов, коинтеграция, адаптивные и мультипликативные методы прогнозирования

Раздел 5. Раздел 5

(Лабораторные занятия - 2ч.; Лекционные занятия - 4ч.; Практические занятия - 2ч.; Самостоятельная работа - 2ч.)

Тема 5.1. Вывод правил для принятия решений

(Лабораторные занятия - 2ч.; Лекционные занятия - 4ч.; Практические занятия - 2ч.; Самостоятельная работа - 2ч.)

Факторный анализ, ранговые методы, неметрические методы, кластерный анализ

Раздел 6. Раздел 6

(Лабораторные занятия - 2ч.; Лекционные занятия - 4ч.; Практические занятия - 2ч.; Самостоятельная работа - 2ч.)

Тема 6.1. Прогнозирование последствий принятых решений

(Лабораторные занятия - 2ч.; Лекционные занятия - 4ч.; Практические занятия - 2ч.; Самостоятельная работа - 2ч.)

Прогнозная аналитика, нейронные сети, пространственный анализ, сетевой анализ

Раздел 7. Раздел 7

(Лабораторные занятия - 1ч.; Лекционные занятия - 4ч.; Практические занятия - 2ч.; Самостоятельная работа - 1ч.)

Тема 7.1. Регрессионный анализ

(Лабораторные занятия - 1ч.; Лекционные занятия - 4ч.; Практические занятия - 2ч.; Самостоятельная работа - 1ч.)

Понятие и способы применения регрессионного анализа, инструментарий для реализации

Раздел 8. Раздел 8

(Лабораторные занятия - 1ч.; Лекционные занятия - 4ч.; Практические занятия - 2ч.; Самостоятельная работа - 1ч.)

Тема 8.1. Дисперсионный анализ

(Лабораторные занятия - 1ч.; Лекционные занятия - 4ч.; Практические занятия - 2ч.; Самостоятельная работа - 1ч.)

Понятие и способы применения дисперсионного анализа, инструментарий для реализации

Раздел 9. Раздел 9

(Внеаудиторная контактная работа - 1ч.; Лабораторные занятия - 2ч.; Лекционные занятия - 2ч.; Практические занятия - 2ч.; Самостоятельная работа - 12ч.)

Тема 9.1. Кластерный анализ

(Внеаудиторная контактная работа - 1ч.; Лабораторные занятия - 2ч.; Лекционные занятия - 2ч.; Практические занятия - 2ч.; Самостоятельная работа - 12ч.)

Понятие и способы применения кластерного анализа, инструментарий для реализации

Раздел 10. Раздел 10

(Лабораторные занятия - 2ч.; Лекционные занятия - 2ч.; Практические занятия - 2ч.; Самостоятельная работа - 12ч.)

Тема 10.1. Факторный анализ

(Лабораторные занятия - 2ч.; Лекционные занятия - 2ч.; Практические занятия - 2ч.; Самостоятельная работа - 12ч.)

Понятие и способы применения факторного анализа, инструментарий для реализации

Раздел 11. Раздел 11

(Лабораторные занятия - 2ч.; Лекционные занятия - 2ч.; Практические занятия - 2ч.; Самостоятельная работа - 12ч.)

Тема 11.1. Модели распределенных файловых систем

(Лабораторные занятия - 2ч.; Лекционные занятия - 2ч.; Практические занятия - 2ч.; Самостоятельная работа - 12ч.)

Data mining, краудсорсинг, смешение и интеграция данных

Раздел 12. Раздел 12

(Лабораторные занятия - 8ч.; Лекционные занятия - 10ч.; Практические занятия - 10ч.; Самостоятельная работа - 61ч.)

Тема 12.1. Поиск подобий в данных

(Лабораторные занятия - 2ч.; Лекционные занятия - 2ч.; Практические занятия - 2ч.; Самостоятельная работа - 12ч.)

Машинное обучение, оптимизация и генетические алгоритмы

Тема 12.2. Анализ потоковых данных

(Лабораторные занятия - 2ч.; Лекционные занятия - 2ч.; Практические занятия - 2ч.; Самостоятельная работа - 12ч.)

Анализ связей, визуализация аналитических данных

Тема 12.3. Социально-сетевые графы

(Лабораторные занятия - 2ч.; Лекционные занятия - 2ч.; Практические занятия - 2ч.; Самостоятельная работа - 12ч.)

Частые наборы данных, прогнозирование и предвидение в социальных и медиа процессах

Тема 12.4. Методы кластеризации

(Лабораторные занятия - 1ч.; Лекционные занятия - 2ч.; Практические занятия - 2ч.; Самостоятельная работа - 12ч.)

Модели снижения размерности данных, методы машинного обучения большими данными

Тема 12.5. Способы применения нейронных сетей

(Лабораторные занятия - 1ч.; Лекционные занятия - 2ч.; Практические занятия - 2ч.; Самостоятельная работа - 13ч.)

Приложения нейронных сетей, сетевые аналитические модели

6. Оценочные материалы текущего контроля

Раздел 1. Раздел 1

Форма контроля/оценочное средство: Кейс-задание

Вопросы/Задания:

1. Информационное общество и большие данные.

Раздел 2. Раздел 2

Форма контроля/оценочное средство: Кейс-задание

Вопросы/Задания:

1. М. МакЛюэн и роль медиа в больших данных.

Раздел 3. Раздел 3

Форма контроля/оценочное средство: Кейс-задание

Вопросы/Задания:

1. Современные способы хранения и обработки больших данных.

Раздел 4. Раздел 4

Форма контроля/оценочное средство: Кейс-задание

Вопросы/Задания:

1. Современные способы анализа массивов данных

Раздел 5. Раздел 5

Форма контроля/оценочное средство: Кейс-задание

Вопросы/Задания:

1. Современные концепции больших данных

Раздел 6. Раздел 6

Форма контроля/оценочное средство: Кейс-задание

Вопросы/Задания:

1. Блокчейн системы как приложение системы больших данных.

Раздел 7. Раздел 7

Форма контроля/оценочное средство: Кейс-задание

Вопросы/Задания:

1. Структурированные и неструктурированные данные.

Раздел 8. Раздел 8

Форма контроля/оценочное средство: Кейс-задание

Вопросы/Задания:

1. Набор признаков больших данных (VVV-подход)

Раздел 9. Раздел 9

Форма контроля/оценочное средство: Кейс-задание

Вопросы/Задания:

1. Информационно-технологическая структура больших данных

Раздел 10. Раздел 10

Форма контроля/оценочное средство: Кейс-задание

Вопросы/Задания:

1. Анализ стандартных наборов данных с помощью Weka

Раздел 11. Раздел 11

Форма контроля/оценочное средство: Кейс-задание

Вопросы/Задания:

1. Алгоритм k-means.

Раздел 12. Раздел 12

Форма контроля/оценочное средство: Кейс-задание

Вопросы/Задания:

1. Процесс CRISP-DM

7. Оценочные материалы промежуточной аттестации

Пятый семестр, Зачет

*Контролируемые ИДК: УК-2.1 УК-6.1 УК-2.2 УК-6.2 УК-2.3 УК-6.3 УК-2.4 УК-6.4 УК-6.5
ОПК-2.1 ОПК-2.2 ОПК-2.3*

Вопросы/Задания:

1. Понятие «Большие данные». Роль цифровой информации в 21 веке.
2. Виды массивов данных
3. Базовые принципы обработки больших данных.
4. Технологии обработки больших данных: NoSQL, MapReduce, Hadoop, R
5. Технологии Business Intelligence и реляционные системы управления базами данных.
6. Прогнозирование и предвидение: общее и особенное.
7. Виды прогнозов.
8. Общие методы анализа социально-политических и медиа процессов
9. Специальные методы анализа социально-политических и медиа процессов.
10. Предварительный анализ данных.
11. Проверка гипотез о законе распределения случайной величины
12. Статистические оценки параметров. Доверительные области.
13. Теория моментов.
14. Корреляционный анализ.
15. Использование модели множественной линейной регрессии для прогнозирования экономических показателей.
16. Доверительные интервалы для зависимой переменной.
17. Сглаживание временных рядов.
18. Динамические модели с распределенными лагами.
19. Стационарные временные ряды.
20. Тестирование стационарности.

21. Линейная регрессия. Метод наименьших квадратов
22. Примеры использования корреляции и регрессионного анализа в области больших данных
23. Логистическая регрессия.
24. Наивный классификатор Байеса.
25. Алгоритм априори.
26. Алгоритм k-means.
27. Матрица парных расстояний.
28. Виды метрик.
29. Коинтеграция
30. Анализ временных рядов.

Шестой семестр, Зачет с оценкой

*Контролируемые ИДК: УК-2.1 УК-6.1 УК-2.2 УК-6.2 УК-2.3 УК-6.3 УК-2.4 УК-6.4 УК-6.5
ОПК-2.1 ОПК-2.2 ОПК-2.3*

Вопросы/Задания:

1. Понятие «Большие данные». Роль цифровой информации в 21 веке.
2. Виды массивов данных.
3. Базовые принципы обработки больших данных.
4. Технологии обработки больших данных: NoSQL, MapReduce, Hadoop, R.
5. Технологии Business Intelligence и реляционные системы управления базами данных.
6. Прогнозирование и предвидение: общее и особенное.
7. Виды прогнозов.
8. Общие методы анализа социально-политических и медиа процессов.
9. Специальные методы анализа социально-политических и медиа процессов.
10. Предварительный анализ данных.
11. Проверка гипотез о законе распределения случайной величины.
12. Статистические оценки параметров. Доверительные области.

13. Теория моментов.
14. Корреляционный анализ
15. Использование модели множественной линейной регрессии для прогнозирования экономических показателей.
16. Доверительные интервалы для зависимой переменной.
17. Сглаживание временных рядов.
18. Динамические модели с распределенными лагами.
19. Стационарные временные ряды.
20. Тестирование стационарности.
21. Линейная регрессия. Метод наименьших квадратов.
22. Примеры использования корреляции и регрессионного анализа в области больших данных.
23. Логистическая регрессия.
24. Наивный классификатор Байеса.
25. Алгоритм априори.
26. Алгоритм k-means.
27. Матрица парных расстояний.
28. Виды метрик.
29. Коинтеграция.
30. Анализ временных рядов.
31. Адаптивные и мультипликативные методы прогнозирования.
32. Экспоненциальное сглаживание.
33. Авторегрессионные модели.
34. Модели скользящего среднего.
35. Интегрированные процессы.
36. Идентификация авторегрессионной модели скользящего среднего.

37. Прогнозирование с моделями временных рядов.
38. Доверительные интервалы прогноза.
39. Предсказание и прогнозирование социально-экономических прогнозов.
40. Дисперсионный анализ влияния качественных факторов.
41. Ранговые методы.
42. Факторный анализ. Метод главных факторов.
43. Многомерное шкалирование.
44. Классическая модель многомерного шкалирования.
45. Невметрические методы.
46. Кластерный анализ.
47. Дискриминантный анализ.
48. Многомерный статистический анализ.
49. Визуализация стандартных наборов данных с помощью Tableau.
50. Вычисление индекса TF-IDF.
51. Реализация Hadoop.
52. Регуляризация для метода наименьших квадратов.
53. Машина опорных векторов.
54. Тренд больших данных.
55. Определение и примеры больших данных.
56. Основные характеристики Big Data.
57. Принцип Map Reduce.
58. Технологии хранения данных и Big Data.
59. Определение больших данных, ключевые характеристики. Основные виды данных.
60. Роль аналитика по данным (Data Scientist). Ключевые компетенции аналитика.

61. Отличия BI от Data Science.
62. Корреляция и регрессионный анализ. Коэффициент корреляции.
63. Графическое представление. Постановка задачи регрессионного анализа.
64. Линейная регрессия. Метод наименьших квадратов. Привести примеры использования регрессионного анализа.
65. Классификация. Признаковое описание объекта и таблица объект-свойства.
66. Постановка задачи. Отличия задачи классификации от задачи регрессии. Определение модели и алгоритма. Процесс обучения.
67. Проблема переобучения. Регуляризация. Cross validation.
68. Дополнительный вопрос: привести модель в линейной регрессии.
69. Кластеризация. Метрики.
70. Матрица парных расстояний.
71. Постановка задачи кластеризации. Отличие от задачи классификации.
72. Ассоциативные правила. Определение. Достоверность и поддержка.
73. Отличия построения ассоциативного правила от решающего правила задачи классификации.
74. Парадигма Map Reduce. Описать принцип работы. Перечислить слабые и сильные стороны. Обозначить области применимости.
75. Визуализация. Дать определение визуализации. Показать важность визуализации в аналитике больших данных. Привести примеры использования визуализации.
76. «Жизненный цикл» проекта по аналитике больших данных.
77. Типовая архитектура проекта в области больших данных. Перечислить используемые технологии, указать степень вовлеченности каждой из технологий на каждом этапе работы над проектом. Перечислить основные роли исполнителей проекта.
78. Научные проблемы больших данных. Показать значимость проблем, актуальность, связь с областями математики и инженерии.
79. «Песочница» в аналитическом процессе.
80. Использование Hadoop.

8. Материально-техническое и учебно-методическое обеспечение дисциплины

8.1. Перечень основной и дополнительной учебной литературы

Основная литература

1. ПАРАСКЕВОВ А. В. Большие данные: лабораторный практикум / ПАРАСКЕВОВ А. В., Сергеев А. Э., Чемарина А. В.. - Краснодар: КубГАУ, 2023. - 91 с. - Текст: непосредственный.
2. ПАРАСКЕВОВ А. В. Большие данные: учебник / ПАРАСКЕВОВ А. В., Чемарина А. В.. - Краснодар: КубГАУ, 2020. - 173 с. - Текст: электронный. // : [сайт]. - URL: <https://edu.kubsau.ru/mod/resource/view.php?id=10285> (дата обращения: 02.05.2024). - Режим доступа: по подписке
3. ПАРАСКЕВОВ А. В. Большие данные: метод. указания / ПАРАСКЕВОВ А. В.. - Краснодар: КубГАУ, 2020. - 35 с. - Текст: электронный. // : [сайт]. - URL: <https://edu.kubsau.ru/mod/resource/view.php?id=8828> (дата обращения: 21.06.2024). - Режим доступа: по подписке
4. ПАРАСКЕВОВ А.В. Большие данные: учебник / ПАРАСКЕВОВ А.В., Чемарина А. В.. - Краснодар: КубГАУ, 2021. - 172 с. - 978-5-907516-62-5. - Текст: непосредственный.

Дополнительная литература

1. Бродовская,, Е. В. Большие данные в исследовании политических процессов: учебное пособие / Е. В. Бродовская,, А. Ю. Домбровская,. - Большие данные в исследовании политических процессов - Москва: Московский педагогический государственный университет, 2018. - 88 с. - 978-5-4263-0712-4. - Текст: электронный. // IPR SMART: [сайт]. - URL: <https://www.iprbookshop.ru/92872.html> (дата обращения: 20.02.2024). - Режим доступа: по подписке

8.2. Профессиональные базы данных и ресурсы «Интернет», к которым обеспечивается доступ обучающихся

Профессиональные базы данных

1. <http://www.iprbookshop.ru/> - IPRbook

Ресурсы «Интернет»

1. <https://edu.kubsau.ru/> - Образовательный портал КубГАУ

8.3. Программное обеспечение и информационно-справочные системы, используемые при осуществлении образовательного процесса по дисциплине

Информационные технологии, используемые при осуществлении образовательного процесса по дисциплине позволяют:

- обеспечить взаимодействие между участниками образовательного процесса, в том числе синхронное и (или) асинхронное взаимодействие посредством сети «Интернет»;
- фиксировать ход образовательного процесса, результатов промежуточной аттестации по дисциплине и результатов освоения образовательной программы;
- организовать процесс образования путем визуализации изучаемой информации посредством использования презентаций, учебных фильмов;
- контролировать результаты обучения на основе компьютерного тестирования.

Перечень лицензионного программного обеспечения:

- 1 Microsoft Windows - операционная система.
- 2 Microsoft Office (включает Word, Excel, Power Point) - пакет офисных приложений.

Перечень профессиональных баз данных и информационных справочных систем:

- 1 Гарант - правовая, <https://www.garant.ru/>
- 2 Консультант - правовая, <https://www.consultant.ru/>
- 3 Научная электронная библиотека eLibrary - универсальная, <https://elibrary.ru/>

Доступ к сети Интернет, доступ в электронную информационно-образовательную среду университета.

Перечень программного обеспечения

(обновление производится по мере появления новых версий программы)

Не используется.

Перечень информационно-справочных систем

(обновление выполняется еженедельно)

Не используется.

8.4. Специальные помещения, лаборатории и лабораторное оборудование

Университет располагает на праве собственности или ином законном основании материально-техническим обеспечением образовательной деятельности (помещениями и оборудованием) для реализации программы бакалавриата, специалитета, магистратуры по Блоку 1 "Дисциплины (модули)" и Блоку 3 "Государственная итоговая аттестация" в соответствии с учебным планом.

Каждый обучающийся в течение всего периода обучения обеспечен индивидуальным неограниченным доступом к электронной информационно-образовательной среде университета из любой точки, в которой имеется доступ к информационно-телекоммуникационной сети "Интернет", как на территории университета, так и вне его. Условия для функционирования электронной информационно-образовательной среды могут быть созданы с использованием ресурсов иных организаций.

Лекционный зал

451гл

доска ДК11Э2010. - 1 шт.

жалюзи - 32 шт.

Облучатель-рециркулятор воздуха 600 - 1 шт.

парты. - 31 шт.

Проектор Epson EB-X06 - 1 шт.

Экран SACTUS wallscreen CS-PSW-149x265, 265.7x149.4 см, 16:9, настенно-потолочный белый - 1 шт.

Компьютерный класс

205эл

коммутатор - 1 шт.

Компьютер персональный Dell OptiPlex 3050 - 1 шт.

Компьютер персональный IRU Corp 310 i3 3240/4Gb/500Gb/W7Pro64 - 1 шт.

телевизор Samsung LE-46N87BD - 1 шт.

экран настенный - 1 шт.

9. Методические указания по освоению дисциплины (модуля)

Учебная работа по направлению подготовки осуществляется в форме контактной работы с преподавателем, самостоятельной работы обучающегося, текущей и промежуточной аттестаций, иных формах, предлагаемых университетом. Учебный материал дисциплины структурирован и его изучение производится в тематической последовательности. Содержание методических указаний должно соответствовать требованиям Федерального государственного образовательного стандарта и учебных программ по дисциплине. Самостоятельная работа студентов может быть выполнена с помощью материалов, размещенных на портале поддержки Moodle.

Методические указания по формам работы

Лекционные занятия

Передача значительного объема систематизированной информации в устной форме достаточно большой аудитории. Дает возможность экономно и систематично излагать учебный материал. Обучающиеся изучают лекционный материал, размещенный на портале поддержки обучения Moodle.

Лабораторные занятия

Практическое освоение студентами научно-теоретических положений изучаемого предмета, овладение ими техникой экспериментирования в соответствующей отрасли науки. Лабораторные занятия проводятся с использованием методических указаний, размещенных на образовательном портале университета.

Практические занятия

Форма организации обучения, проводимая под руководством преподавателя и служащая для детализации, анализа, расширения, углубления, закрепления, применения (или выполнения разнообразных практических работ, упражнений) и контроля усвоения полученной на лекциях учебной информации. Практические занятия проводятся с использованием учебно-методических изданий, размещенных на образовательном портале университета.

Описание возможностей изучения дисциплины лицами с ОВЗ и инвалидами

Для инвалидов и лиц с ОВЗ может изменяться объём дисциплины (модуля) в часах, выделенных на контактную работу обучающегося с преподавателем (по видам учебных занятий) и на самостоятельную работу обучающегося (при этом не увеличивается количество зачётных единиц, выделенных на освоение дисциплины).

Фонды оценочных средств адаптируются к ограничениям здоровья и восприятия информации обучающимися.

Основные формы представления оценочных средств – в печатной форме или в форме электронного документа.

Формы контроля и оценки результатов обучения инвалидов и лиц с ОВЗ с нарушением зрения:

– устная проверка: дискуссии, тренинги, круглые столы, собеседования, устные коллоквиумы

и др.;

– с использованием компьютера и специального ПО: работа с электронными образовательными ресурсами, тестирование, рефераты, курсовые проекты, дистанционные формы, если позволяет острота зрения - графические работы и др.;

– при возможности письменная проверка с использованием рельефно-точечной системы Брайля, увеличенного шрифта, использование специальных технических средств (тифлотехнических средств): контрольные, графические работы, тестирование, домашние задания, эссе, отчеты и др.

Формы контроля и оценки результатов обучения инвалидов и лиц с ОВЗ с нарушением слуха:

– письменная проверка: контрольные, графические работы, тестирование, домашние задания, эссе, письменные коллоквиумы, отчеты и др.;

– с использованием компьютера: работа с электронными образовательными ресурсами, тестирование, рефераты, курсовые проекты, графические работы, дистанционные формы и др.;

– при возможности устная проверка с использованием специальных технических средств (аудиосредств, средств коммуникации, звукоусиливающей аппаратуры и др.): дискуссии, тренинги, круглые столы, собеседования, устные коллоквиумы и др.

Формы контроля и оценки результатов обучения инвалидов и лиц с ОВЗ с нарушением опорно-двигательного аппарата:

– письменная проверка с использованием специальных технических средств (альтернативных средств ввода, управления компьютером и др.): контрольные, графические работы, тестирование, домашние задания, эссе, письменные коллоквиумы, отчеты и др.;

– устная проверка, с использованием специальных технических средств (средств коммуникаций): дискуссии, тренинги, круглые столы, собеседования, устные коллоквиумы и др.;

– с использованием компьютера и специального ПО (альтернативных средств ввода и управления компьютером и др.): работа с электронными образовательными ресурсами, тестирование, рефераты, курсовые проекты, графические работы, дистанционные формы предпочтительнее обучающимся, ограниченным в передвижении и др.

Адаптация процедуры проведения промежуточной аттестации для инвалидов и лиц с ОВЗ.

В ходе проведения промежуточной аттестации предусмотрено:

– предъявление обучающимся печатных и (или) электронных материалов в формах, адаптированных к ограничениям их здоровья;

– возможность пользоваться индивидуальными устройствами и средствами, позволяющими адаптировать материалы, осуществлять приём и передачу информации с учетом их индивидуальных особенностей;

– увеличение продолжительности проведения аттестации;

– возможность присутствия ассистента и оказания им необходимой помощи (занять рабочее место, передвигаться, прочитать и оформить задание, общаться с преподавателем).

Формы промежуточной аттестации для инвалидов и лиц с ОВЗ должны учитывать индивидуальные и психофизические особенности обучающегося/обучающихся по АОПОП ВО (устно, письменно на бумаге, письменно на компьютере, в форме тестирования и т.п.).

Специальные условия, обеспечиваемые в процессе преподавания дисциплины студентам с нарушениями зрения:

– предоставление образовательного контента в текстовом электронном формате, позволяющем переводить плоскочечную информацию в аудиальную или тактильную форму;

– возможность использовать индивидуальные устройства и средства, позволяющие адаптировать материалы, осуществлять приём и передачу информации с учетом индивидуальных особенностей и состояния здоровья студента;

– предоставление возможности предкурсового ознакомления с содержанием учебной дисциплины и материалом по курсу за счёт размещения информации на корпоративном образовательном портале;

– использование чёткого и увеличенного по размеру шрифта и графических объектов в мультимедийных презентациях;

– использование инструментов «лупа», «прожектор» при работе с интерактивной доской;

- озвучивание визуальной информации, представленной обучающимся в ходе занятий;
- обеспечение раздаточным материалом, дублирующим информацию, выводимую на экран;
- наличие подписей и описания у всех используемых в процессе обучения рисунков и иных графических объектов, что даёт возможность перевести письменный текст в аудиальный;
- обеспечение особого речевого режима преподавания: лекции читаются громко, разборчиво, отчётливо, с паузами между смысловыми блоками информации, обеспечивается интонирование, повторение, акцентирование, профилактика рассеивания внимания;
- минимизация внешнего шума и обеспечение спокойной аудиальной обстановки;
- возможность вести запись учебной информации студентами в удобной для них форме (аудиально, аудиовизуально, на ноутбуке, в виде пометок в заранее подготовленном тексте);
- увеличение доли методов социальной стимуляции (обращение внимания, апелляция к ограничениям по времени, контактные виды работ, групповые задания и др.) на практических и лабораторных занятиях;
- минимизирование заданий, требующих активного использования зрительной памяти и зрительного внимания;
- применение поэтапной системы контроля, более частый контроль выполнения заданий для самостоятельной работы.

Специальные условия, обеспечиваемые в процессе преподавания дисциплины студентам с нарушениями опорно-двигательного аппарата (маломобильные студенты, студенты, имеющие трудности передвижения и патологию верхних конечностей):

- возможность использовать специальное программное обеспечение и специальное оборудование и позволяющее компенсировать двигательное нарушение (коляски, ходунки, трости и др.);
- предоставление возможности предкурсового ознакомления с содержанием учебной дисциплины и материалом по курсу за счёт размещения информации на корпоративном образовательном портале;
- применение дополнительных средств активизации процессов запоминания и повторения;
- опора на определенные и точные понятия;
- использование для иллюстрации конкретных примеров;
- применение вопросов для мониторинга понимания;
- разделение изучаемого материала на небольшие логические блоки;
- увеличение доли конкретного материала и соблюдение принципа от простого к сложному при объяснении материала;
- наличие чёткой системы и алгоритма организации самостоятельных работ и проверки заданий с обязательной корректировкой и комментариями;
- увеличение доли методов социальной стимуляции (обращение внимания, апелляция к ограничениям по времени, контактные виды работ, групповые задания др.);
- обеспечение беспрепятственного доступа в помещения, а также пребывания в них;
- наличие возможности использовать индивидуальные устройства и средства, позволяющие обеспечить реализацию эргономических принципов и комфортное пребывание на месте в течение всего периода учёбы (подставки, специальные подушки и др.).

Специальные условия, обеспечиваемые в процессе преподавания дисциплины студентам с нарушениями слуха (глухие, слабослышащие, позднооглохшие):

- предоставление образовательного контента в текстовом электронном формате, позволяющем переводить аудиальную форму лекции в плоскочечатную информацию;
- наличие возможности использовать индивидуальные звукоусиливающие устройства и сурдотехнические средства, позволяющие осуществлять приём и передачу информации; осуществлять взаимобратный перевод текстовых и аудиофайлов (блокнот для речевого ввода), а также запись и воспроизведение зрительной информации;
- наличие системы заданий, обеспечивающих систематизацию вербального материала, его схематизацию, перевод в таблицы, схемы, опорные тексты, глоссарий;
- наличие наглядного сопровождения изучаемого материала (структурно-логические схемы, таблицы, графики, концентрирующие и обобщающие информацию, опорные конспекты, раздаточный материал);
- наличие чёткой системы и алгоритма организации самостоятельных работ и проверки

заданий с обязательной корректировкой и комментариями;

- обеспечение практики опережающего чтения, когда студенты заранее знакомятся с материалом и выделяют незнакомые и непонятные слова и фрагменты;
- особый речевой режим работы (отказ от длинных фраз и сложных предложений, хорошая артикуляция; четкость изложения, отсутствие лишних слов; повторение фраз без изменения слов и порядка их следования; обеспечение зрительного контакта во время говорения и чуть более медленного темпа речи, использование естественных жестов и мимики);
- чёткое соблюдение алгоритма занятия и заданий для самостоятельной работы (называние темы, постановка цели, сообщение и запись плана, выделение основных понятий и методов их изучения, указание видов деятельности студентов и способов проверки усвоения материала, словарная работа);
- соблюдение требований к предъявляемым учебным текстам (разбивка текста на части; выделение опорных смысловых пунктов; использование наглядных средств);
- минимизация внешних шумов;
- предоставление возможности соотносить вербальный и графический материал; комплексное использование письменных и устных средств коммуникации при работе в группе;
- сочетание на занятиях всех видов речевой деятельности (говорения, слушания, чтения, письма, зрительного восприятия с лица говорящего).

Специальные условия, обеспечиваемые в процессе преподавания дисциплины студентам с прочими видами нарушений (ДЦП с нарушениями речи, заболевания эндокринной, центральной нервной и сердечно-сосудистой систем, онкологические заболевания):

- наличие возможности использовать индивидуальные устройства и средства, позволяющие осуществлять приём и передачу информации;
- наличие системы заданий, обеспечивающих систематизацию вербального материала, его схематизацию, перевод в таблицы, схемы, опорные тексты, глоссарий;
- наличие наглядного сопровождения изучаемого материала;
- наличие чёткой системы и алгоритма организации самостоятельных работ и проверки заданий с обязательной корректировкой и комментариями;
- обеспечение практики опережающего чтения, когда студенты заранее знакомятся с материалом и выделяют незнакомые и непонятные слова и фрагменты;
- предоставление возможности соотносить вербальный и графический материал; комплексное использование письменных и устных средств коммуникации при работе в группе;
- сочетание на занятиях всех видов речевой деятельности (говорения, слушания, чтения, письма, зрительного восприятия с лица говорящего);
- предоставление образовательного контента в текстовом электронном формате;
- предоставление возможности предкурсового ознакомления с содержанием учебной дисциплины и материалом по курсу за счёт размещения информации на корпоративном образовательном портале;
- возможность вести запись учебной информации студентами в удобной для них форме (аудиально, аудиовизуально, в виде пометок в заранее подготовленном тексте);
- применение поэтапной системы контроля, более частый контроль выполнения заданий для самостоятельной работы;
- стимулирование выработки у студентов навыков самоорганизации и самоконтроля;
- наличие пауз для отдыха и смены видов деятельности по ходу занятия.

10. Методические рекомендации по освоению дисциплины (модуля)

Дисциплина "Большие данные" ведётся в соответствии с календарным учебным планом и расписанием занятий по неделям. Темы проведения занятий определяются тематическим планом рабочей программы дисциплины.