

Министерство образования и науки Астраханской области
Государственное бюджетное образовательное учреждение
Астраханской области высшего образования
«Астраханский государственный архитектурно-строительный
университет»
(ГБОУ АО ВО «АГАСУ»)

УТВЕРЖДАЮ

И.о. первого проректора



РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Наименование дисциплины

«Технологии анализа данных»

(указывается наименование в соответствии с учебным планом)

По направлению подготовки

09.04.02 «Информационные системы и технологии»

(указывается наименование направления подготовки в соответствии с ФГОС ВО)

Направленность (профиль)

«Искусственный интеллект в проектировании городской среды»

(указывается наименование направленности (профиля) в соответствии с ОПОП)

Кафедра

Системы автоматизированного проектирования и моделирования

Квалификация выпускника *магистр*

Разработчик:

К.т.н., доцент
(занимаемая должность,
учёная степень и учёное звание)


(подпись)

/С.В. Окладникова/
(И. О. Ф.)

Рабочая программа рассмотрена и утверждена на заседании кафедры
«Системы автоматизированного проектирования и моделирования»
протокол № 9 от 22.04.2014 г.

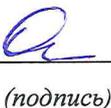
И.о. заведующего кафедрой


(подпись)

/В.В. Соболева /
(И. О. Ф.)

Согласовано:

Председатель МКН «Информационные системы и технологии»
направленность(профиль) «Искусственный интеллект в проектировании городской среды»


(подпись)

/О.М. Шиндурова/
(И.О.Ф)

Начальник УМУ


(подпись)

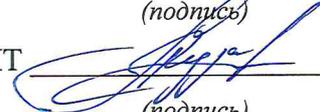
/О.А. Веселов/
(И.О.Ф)

Специалист УМУ


(подпись)

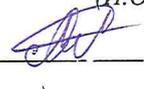
/В.А. Лерин/
(И.О.Ф)

Начальник УИТ


(подпись)

/А.М. Суза/
(И.О.Ф)

Заведующая научной библиотекой


(подпись)

/Л.С. Таврилова/
(И.О.Ф)

Содержание

1. Цель освоения дисциплины	4
2. Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения образовательной программы	4
3. Место дисциплины в структуре ОПОП магистратуры	4
4. Объем дисциплины в зачетных единицах с указанием количества академических часов, выделенных на контактную работу обучающихся с преподавателем (по типам занятий) и на самостоятельную работу обучающихся.....	4
5. Содержание дисциплины, структурированное по разделам с указанием отведенного на них количества академических часов и типов учебных занятий	6
5.1. Разделы дисциплины и трудоемкость по типам учебных занятий и работы обучающихся (в академических часах)	6
5.1.1. Очная форма обучения	6
5.1.2. Заочная форма обучения	6
5.1.3. Очно-заочная форма обучения	7
5.2. Содержание дисциплины, структурированное по разделам.....	8
5.2.1. Содержание лекционных занятий	8
5.2.2. Содержание лабораторных занятий	8
5.2.3. Содержание практических занятий.....	10
5.2.4. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине	10
5.2.5. Темы контрольных работ	11
5.2.6. Темы курсовых проектов/ курсовых работ	11
6. Методические указания для обучающихся по освоению дисциплины	12
7. Образовательные технологии	13
8. Учебно-методическое и информационное обеспечение дисциплины.....	13
8.1. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины.....	13
8.2. Перечень необходимого лицензионного и свободно распространяемого программного обеспечения, в том числе отечественного производства, используемого при осуществлении образовательного процесса по дисциплине	15
8.3. Перечень современных профессиональных баз данных и информационных справочных систем, доступных обучающимся при освоении дисциплины	15
9. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине	15
10. Особенности организации обучения по дисциплине для инвалидов и лиц с ограниченными возможностями здоровья	16

1. Цель освоения дисциплины

Целью освоения дисциплины «Технологии анализа данных» является формирование компетенций у обучающихся в соответствии с требованиями Федерального государственного образовательного стандарта высшего образования по направлению подготовки 09.04.02 «Информационные системы и технологии».

2. Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения образовательной программы

В результате освоения дисциплины обучающийся должен овладеть следующими компетенциями:

ПК-4ИИП - Способен руководить проектами по созданию комплексных систем на основе аналитики больших данных в различных отраслях со стороны заказчика

ПК4.1ИИП – Руководит проектами по построению комплексных систем на основе аналитики больших данных в различных отраслях со стороны заказчика

В результате освоения дисциплины, обучающийся должен овладеть следующими результатами обучения:

знать:

- ПК-4ИИП.1 З-1. Знает методологию и принципы руководства проектами по созданию, поддержке и использованию комплексных систем на основе аналитики больших данных со стороны заказчика
- ПК-4ИИП.1 З-2. Знает специфику сфер и отраслей, для которых реализуется проект по аналитике больших данных

уметь:

- ПК-4ИИП.1 У-1. Умеет решать задачи по руководству коллективной проектной деятельностью для создания, поддержки и использования комплексных систем на основе аналитики больших данных со стороны заказчика
- ПК-4ИИП.1 У-2. Умеет выявлять небольшие по масштабу проекты аналитики, которые потенциально могут представлять интерес для ряда подразделений / служб или для организации в целом
- ПК-4ИИП.1 У-3. Умеет выявлять области деловой деятельности, которые потенциально могут получить отдачу от аналитики

3. Место дисциплины в структуре ОПОП магистратуры

Дисциплина Б1.В.02 «Технологии анализа данных» реализуется в рамках Блока 1 «Дисциплины (модули)»: части формируемой участниками образовательных отношений..

Дисциплина базируется на основах обучения, полученных в рамках изучения дисциплин: «Специальные главы математики», «Прикладной искусственный интеллект (базовый уровень)».

4. Объем дисциплины в зачетных единицах с указанием количества академических часов, выделенных на контактную работу обучающихся с преподавателем (по типам учебных занятий) и на самостоятельную работу обучающихся

Форма обучения	Очная	Заочная
1	2	3
Трудоемкость в зачетных единицах:	2 семестр – 2 з.е.; всего – 2 з.е. 3 семестр – 3 з.е.; всего – 3 з.е.	1 семестр – 2 з.е.; всего – 2 з.е. 2 семестр – 3 з.е.; всего – 3 з.е.
Лекции (Л)	2 семестр – 14 часов; всего – 14 часов	1 семестр – 4 часа; 2 семестр – 4 часа; всего – 8 часов
Лабораторные занятия (ЛЗ)	2 семестр – 14 часов; 3 семестр – 42 часа; всего – 56 часов;	1 семестр – 8 часов; 2 семестр – 10 часов; всего – 18 часов;
Практические занятия (ПЗ)	учебным планом <i>не предусмотрены</i>	учебным планом <i>не предусмотрены</i>
Самостоятельная работа (СР)	2 семестр – 44 часа; 3 семестр – 66 часов; всего – 110 часов;	1 семестр – 60 часов; 2 семестр – 94 часа; всего – 154 часа
Форма текущего контроля:		
Контрольная работа	3 семестр	1 семестр 2 семестр
Форма промежуточной аттестации:		
Экзамены	3 семестр	2 семестр
Зачет	2 семестр	1 семестр
Зачет с оценкой	учебным планом <i>не предусмотрены</i>	учебным планом <i>не предусмотрены</i>
Курсовая работа	учебным планом <i>не предусмотрены</i>	учебным планом <i>не предусмотрены</i>
Курсовой проект	учебным планом <i>не предусмотрены</i>	учебным планом <i>не предусмотрены</i>

5. Содержание дисциплины, структурированное по разделам с указанием отведенного на них количества академических часов и типов учебных занятий

5.1. Разделы дисциплины и трудоемкость по типам учебных занятий и работы обучающихся (в академических часах)

5.1.1. Очная форма обучения

№ п/п	Раздел дисциплины. (по семестрам)	Всего часов на раздел	Семестр	Распределение трудоемкости раздела (в часах) по типам учебных занятий и работы обучающихся				Форма текущего контроля и промежуточной аттестации
				контактная			СР	
				Л	ЛЗ	ПЗ		
1	2	3	4	5	6	7	8	9
1.	Раздел 1. Введение в большие данные	22	2	4	4	-	14	контрольная работа зачет
2.	Раздел 2. Этапы обработки больших данных	22	2	4	4	-	14	
3.	Раздел 3. Статистические методы анализа больших данных	28	2	6	6	-	16	
Итого за 2 семестр:		72		14	14	-	44	
4.	Раздел 4. Визуализация больших данных	32	3	-	10	-	22	контрольная работа экзамен
5.	Раздел 5. Хранение и обработка больших данных	36	3	-	14	-	22	
6.	Раздел 6. Современные программные средства анализа больших данных	40	3	-	18	-	22	
Итого за 3 семестр:		108		-	42	-	66	
Итого:		180		14	56	-	110	зачет, экзамен

5.1.2. Заочная форма обучения

№ п/п	Раздел дисциплины. (по семестрам)	Всего часов на раздел	Семестр	Распределение трудоемкости раздела (в часах) по типам учебных занятий и работы обучающихся				Форма текущего контроля и промежуточной аттестации
				контактная			СР	
				Л	ЛЗ	ПЗ		
1	2	3	4	5	6	7	8	9
1.	Раздел 1. Введение в большие данные	22	1	1	2	-	19	контрольная работа, зачет
2.	Раздел 2. Этапы обработки больших данных	22	1	1	2	-	19	

3.	Раздел 3. Статистические методы анализа больших данных	28	1	2	4	-	22	контрольная работа, экзамен
Итого за 2 семестр:		72	72	4	8	-	60	
4.	Раздел 4. Визуализация больших данных	32	2	1	2	-	29	
5.	Раздел 5. Хранение и обработка больших данных	36	2	1	4	-	31	
6.	Раздел 6. Современные программные средства анализа больших данных	40	2	2	4	-	34	
Итого за 3 семестр:		108	108	4	10	-	94	
Итого:		180	180	8	18	-	154	зачет, экзамен

5.1.3. Очно-заочная форма обучения

«ОПОП не предусмотрено»

5.2. Содержание дисциплины, структурированное по разделам

5.2.1. Содержание лекционных занятий

№	Наименование раздела дисциплины	Содержание
1	2	3
1.	Раздел 1. Введение в большие данные	<p>Понятие больших данных. Классификация больших данных. Источники больших данных. Компоненты экосистемы больших данных.</p> <p><i>Методология и принципы руководства проектами по созданию, поддержке и использованию комплексных систем на основе аналитики больших данных со стороны заказчика.</i></p> <p><i>Специфика сфер и отраслей, для которых реализуется проект по аналитике больших данных</i></p>
2.	Раздел 2. Этапы обработки больших данных	<p>Обработка, прием, сбор, анализ данных. Представление результатов. Технологии хранения и обработки больших данных.</p> <p><i>Методология и принципы руководства проектами по созданию, поддержке и использованию комплексных систем на основе аналитики больших данных со стороны заказчика.</i></p> <p><i>Специфика сфер и отраслей, для которых реализуется проект по аналитике больших данных</i></p>
3.	Раздел 3. Статистические методы анализа больших данных	<p>Методы анализа больших данных. Методы класса Data mining.</p> <p><i>Методология и принципы руководства проектами по созданию, поддержке и использованию комплексных систем на основе аналитики больших данных со стороны заказчика.</i></p> <p><i>Специфика сфер и отраслей, для которых реализуется проект по аналитике больших данных</i></p>

5.2.2. Содержание лабораторных занятий

№	Наименование раздела дисциплины	Содержание
1	2	3
1.	Раздел 1. Введение в большие данные	<p>Входное тестирование</p> <p>Лабораторная работа 1. Технологии сбора открытых данных</p> <p><i>Задачи по руководству коллективной проектной деятельностью для создания, поддержки и использования комплексных систем на основе аналитики больших данных со стороны заказчика</i></p> <p><i>Небольшие по масштабу проекты аналитики, которые потенциально могут представлять интерес для ряда подразделений / служб или для организации в целом</i></p> <p><i>Области деловой деятельности, которые потенциально могут получить отдачу от аналитики</i></p>

2.	Раздел 2. Этапы обработки больших данных	<p>Лабораторная работа 2. Предобработка и классификация данных в Python</p> <p><i>Задачи по руководству коллективной проектной деятельностью для создания, поддержки и использования комплексных систем на основе аналитики больших данных со стороны заказчика</i></p> <p><i>Небольшие по масштабу проекты аналитики, которые потенциально могут представлять интерес для ряда подразделений / служб или для организации в целом</i></p> <p><i>Области деловой деятельности, которые потенциально могут получить отдачу от аналитики</i></p>
3.	Раздел 3. Статистические методы анализа больших данных	<p>Лабораторная работа 3. Исследовательский анализ данных</p> <p><i>Задачи по руководству коллективной проектной деятельностью для создания, поддержки и использования комплексных систем на основе аналитики больших данных со стороны заказчика</i></p> <p><i>Небольшие по масштабу проекты аналитики, которые потенциально могут представлять интерес для ряда подразделений / служб или для организации в целом</i></p> <p><i>Области деловой деятельности, которые потенциально могут получить отдачу от аналитики</i></p>
4.	Раздел 4. Визуализация больших данных	<p>Лабораторная работа 4. Инструменты интеграции и визуализации данных в Python</p> <p><i>Задачи по руководству коллективной проектной деятельностью для создания, поддержки и использования комплексных систем на основе аналитики больших данных со стороны заказчика</i></p> <p><i>Небольшие по масштабу проекты аналитики, которые потенциально могут представлять интерес для ряда подразделений / служб или для организации в целом</i></p> <p><i>Области деловой деятельности, которые потенциально могут получить отдачу от аналитики</i></p>
5.	Раздел 5. Хранение и обработка больших данных	<p>Лабораторная работа 5.</p> <p>Основы NoSQL СУБД MONGODB</p> <p><i>Задачи по руководству коллективной проектной деятельностью для создания, поддержки и использования комплексных систем на основе аналитики больших данных со стороны заказчика</i></p> <p><i>Небольшие по масштабу проекты аналитики, которые потенциально могут представлять интерес для ряда подразделений / служб или</i></p>

		<i>для организации в целом Области деловой деятельности, которые потенциально могут получить отдачу от аналитики</i>
6.	Раздел 6. Современные программные средства анализа больших данных	Лабораторная работа 6. Анализ данных по технологии MapReduce <i>Задачи по руководству коллективной проектной деятельностью для создания, поддержки и использования комплексных систем на основе аналитики больших данных со стороны заказчика</i> <i>Небольшие по масштабу проекты аналитики, которые потенциально могут представлять интерес для ряда подразделений / служб или для организации в целом</i> <i>Области деловой деятельности, которые потенциально могут получить отдачу от аналитики</i>

5.2.3. Содержание практических занятий

Учебным планом не предусмотрены

5.2.4. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине

Очная форма обучения

№	Наименование раздела дисциплины	Содержание	Учебно-методическое обеспечение
1	2	3	4
1	Раздел 1. Введение в большие данные	Изучение теоретического материала по рекомендованной в рабочей программе литературе и методического материала, размещенного на образовательном портале АГАСУ. Подготовка к контрольной работе, итоговому тестированию, зачету, экзамену.	[1,2,7,10,14-17,18]
2	Раздел 2. Этапы обработки больших данных		[3,4,7,10,11,14-17,19]
3	Раздел 3. Статистические методы анализа больших данных		[5,6,7,10,12,13,14-17,19]
4	Раздел 4. Визуализация больших данных		[3,7,10,11,14-17,19]
5	Раздел 5. Хранение и обработка больших данных		[8,9]
6	Раздел 6. Современные программные средства анализа больших данных		[8,9]

Заочная форма обучения

№	Наименование раздела дисциплины	Содержание	Учебно-методическое обеспечение
1	2	3	4
1	Раздел 1. Введение в большие данные	Изучение теоретического материала по рекомендованной	[1,2,7,10,14-17,18]

№	Наименование раздела дисциплины	Содержание	Учебно-методическое обеспечение
1	2	3	4
2	Раздел 2. Этапы обработки больших данных	в рабочей программе литературе и методического материала, размещенного на образовательном портале АГАСУ. Подготовка к контрольной работе, итоговому тестированию, зачету, экзамену.	[3,4,7,10,11,14-17, 19]
3	Раздел 3. Статистические методы анализа больших данных		[5,6,7,10,12,13,14-17, 19]
4	Раздел 4. Визуализация больших данных		[3,7,10,11,14-17, 19]
5	Раздел 5. Хранение и обработка больших данных		[8,9]
6	Раздел 6. Современные программные средства анализа больших данных		[8,9]

5.2.5. Темы контрольных работ

1. Статистические методы анализа больших данных
2. Хранение и обработка больших данных

5.2.6. Темы курсовых проектов/ курсовых работ

Учебным планом не предусмотрены.

6. Методические указания для обучающихся по освоению дисциплины

Организация деятельности студента
<p><u>Лекция.</u></p> <p>В ходе лекционных занятий необходимо вести конспектирование учебного материала, обращать внимание на категории, формулировки, раскрывающие содержание тех или иных явлений и процессов, научные выводы и практические рекомендации. Необходимо задавать преподавателю уточняющие вопросы с целью уяснения теоретических положений, разрешения спорных ситуаций. Целесообразно дорабатывать свой конспект лекции, делая в нем соответствующие записи из литературы, рекомендованной преподавателем и предусмотренной учебной программой.</p>
<p><u>Лабораторное занятие.</u></p> <p>Целью лабораторных занятий является углубление и закрепление теоретических знаний, полученных студентами на лекциях и в процессе самостоятельного изучения учебного материала, а, следовательно, формирование у них определенных умений и навыков. На лабораторных занятиях студент вначале знакомится с содержанием работы, пользуясь электронными методическими материалами, размещенными на образовательном портале АГАСУ, затем выполняет задание и показывает результаты преподавателю. Лабораторные работы, выполняются студентом самостоятельно, возникающие при их выполнении проблемы разрешаются в рамках учебного времени и индивидуальных и групповых консультаций.</p>
<p><u>Самостоятельная работа.</u></p> <p>Самостоятельная работа студента над усвоением учебного материала по учебной дисциплине может выполняться в помещениях для самостоятельной работы, а также в домашних условиях. Содержание самостоятельной работы студента определяется учебной программой дисциплины, методическими материалами, заданиями и указаниями преподавателя. Самостоятельная работа в аудиторное время включает работу с учебно-методическим материалом, выполнение заданий лабораторных работ.</p> <p>Самостоятельная работа во внеаудиторное время может состоять из:</p> <ul style="list-style-type: none">– повторения теоретического материала;– подготовки к лабораторным занятиям;– изучения учебной и научной литературы;– подготовка к итоговому тестированию;– самопроверка изученного учебного материала.
<p><u>Контрольная работа.</u></p> <p>Теоретическая и практическая части контрольной работы выполняются по установленным темам (вариантам) с использованием практических материалов, полученных на практических (лабораторных) занятиях и при прохождении практики. К каждой теме контрольной работы рекомендуется примерный перечень основных вопросов, список необходимой литературы. Необходимо изучить литературу, рекомендуемую для выполнения контрольной работы. Чтобы полнее раскрыть тему, следует использовать дополнительные источники и материалы. Инструкция по выполнению контрольной работы находится в методических материалах по дисциплине</p>
<p><u>Подготовка к зачету/экзамену.</u></p> <p>Подготовка студентов к зачету/ экзамену включает:</p> <ul style="list-style-type: none">– самостоятельная работа в течение семестра;– непосредственная подготовка в дни, предшествующие зачету/экзамену.

7. Образовательные технологии

Перечень образовательных технологий, используемых при изучении дисциплины «Технологии анализа данных».

Традиционные образовательные технологии

Дисциплина «Технологии анализа данных» проводится с использованием традиционных образовательных технологий, ориентирующихся на организацию образовательного процесса, предполагающую прямую трансляцию знаний от преподавателя к студенту (преимущественно на основе объяснительно-иллюстративных методов обучения), учебная деятельность студента носит в таких условиях, как правило, репродуктивный характер. Формы учебных занятий с использованием традиционных технологий:

Лекция – последовательное изложение материала в дисциплинарной логике, осуществляемое преимущественно вербальными средствами (монолог преподавателя).

Лабораторное занятие – занятие, посвященное освоению конкретных умений и навыков по предложенному алгоритму.

Интерактивные технологии

По дисциплине «Технологии анализа данных» лекционные занятия проводятся с использованием следующих интерактивных технологий:

Лекция-визуализация представляет собой визуальную форму подачи лекционного материала средствами ТСО или аудиовидеотехники (видео-лекция). Чтение такой лекции сводится к развернутому или краткому комментированию просматриваемых визуальных материалов (в виде схем, таблиц, графов, графиков, моделей). Лекция-визуализация помогает студентам преобразовывать лекционный материал в визуальную форму, что способствует формированию у них профессионального мышления за счет систематизации и выделения наиболее значимых, существенных элементов.

По дисциплине «Технологии анализа данных» лабораторные занятия проводятся с использованием следующих интерактивных технологий:

Работа в малых группах – это одна из самых популярных стратегий, так как она дает всем обучающимся (в том числе и стеснительным) возможность участвовать в работе, практиковать навыки сотрудничества, межличностного общения (в частности, умение активно слушать, вырабатывать общее мнение, разрешать возникающие разногласия). Все это часто бывает невозможно в большом коллективе.

Разработка проекта (метод проектов) – организация обучения, при которой учащиеся приобретают знания в процессе планирования и выполнения практических заданий-проектов.

8. Учебно-методическое и информационное обеспечение дисциплины

8.1. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины

а) основная учебная литература:

1. Черемухин А. Д. Большие данные: учебное пособие / А. Д. Черемухин. – Москва: Ай Пи Ар Медиа, 2023. – 782 с. – ISBN 978-5-4497-2138-9. – Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. – URL: <https://www.iprbookshop.ru/129721.html>
2. Целых А. Н. Применение временных рядов для анализа больших данных: учебное по-

- собрание по курсу «Математические методы анализа больших данных» / А. Н. Целых, В. С. Васильев, Э. М. Котов. – Ростов-на-Дону, Таганрог: Издательство Южного федерального университета, 2021. – 84 с. – ISBN 978-5-9275-3983-3. – Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. – URL: <https://www.iprbookshop.ru/121929.html>
3. Целых А.Н. Современные методы прикладной информатики в задачах анализа данных: учебное пособие по курсу «Методы интеллектуального анализа данных» / Целых А.Н., Целых А.А., Котов Э.М. – Ростов-на-Дону, Таганрог: Издательство Южного федерального университета, 2021. – 130 с. – ISBN 978-5-9275-3783-9. – Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. – URL: <https://www.iprbookshop.ru/117165.html>
 4. Сузи Р. А. Язык программирования Python : учебное пособие / Р. А. Сузи. — 3-е изд. — Москва : Интернет-Университет Информационных Технологий (ИНТУИТ), Ай Пи Ар Медиа, 2020. — 350 с. - Режим доступа: <https://www.iprbookshop.ru/97589.html>

б) дополнительная учебная литература:

5. Железнов М.М. Методы и технологии обработки больших данных: учебно-методическое пособие / Железнов М.М. – Москва: МИСИ-МГСУ, ЭБС АСВ, 2020. – 46 с. – ISBN 978-5-7264-2193-3. – Текст : электронный // Цифровой образовательный ресурс IPR SMART : [сайт]. – URL: <https://www.iprbookshop.ru/101802.html>
6. Воронов В.И. Data Mining - технологии обработки больших данных: учебное пособие / Воронов В.И., Воронова Л.И., Усачев В.А. – Москва: Московский технический университет связи и информатики, 2018. – 47 с. – Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. – URL: <https://www.iprbookshop.ru/81324.html>
7. Адлер Ю. П. Статистическое управление процессами. «Большие данные»: учебное пособие / Ю. П. Адлер, Е. А. Черных. – Москва: Издательский Дом МИСиС, 2016. – 52 с. – ISBN 978-5-87623-969-3. – Текст: электронный // Цифровой образовательный ресурс IPR SMART: [сайт]. – URL: <https://www.iprbookshop.ru/64199.html>
8. Шелудько В. М. Язык программирования высокого уровня Python. Функции, структуры данных, дополнительные модули : учебное пособие / В. М. Шелудько. — Ростов-на-Дону, Таганрог : Издательство Южного федерального университета, 2017. — 107 с. - Режим доступа: <https://www.iprbookshop.ru/87530.html>
9. Программные системы статистического анализа. Обнаружение закономерностей в данных с использованием системы R и языка Python : учебное пособие / В. М. Волкова, М. А. Семёнова, Е. С. Четвертакова, С. С. Вожов. — Новосибирск: Новосибирский государственный технический университет, 2017. — 74 с. - Режим доступа: <https://www.iprbookshop.ru/91682.html>
10. Маккинли Уэс. Python и анализ данных / Уэс Маккинли; перевод А. Слинкина. — 2-е изд. — Саратов : Профобразование, 2019. — 482 с. - Режим доступа: <https://www.iprbookshop.ru/88752.html>
11. Глебов В. И. Практикум по математической статистике. Проверка гипотез с использованием Excel, MatCalc, R и Python : учебное пособие / В. И. Глебов, С. Я. Криволапов. — Москва : Прометей, 2019. — 86 с. - Режим доступа: <https://www.iprbookshop.ru/94504.html>

в) перечень учебно-методического обеспечения:

12. ЭОР «Анализ данных». РЭУ им. Плеханова.
<http://moodle.aucu.ru/mod/page/view.php?id=149075>

г) перечень онлайн курсов:

13. ЭОР «Основы Python для анализа данных». Академии Data Diving.
<https://stepik.org/lesson/976082/step/1?unit=983112>

8.2. Перечень необходимого лицензионного и свободно распространяемого программного обеспечения, в том числе отечественного производства, используемого при осуществлении образовательного процесса по дисциплине

- 7-Zip;
- Adobe Acrobat Reader DC;
- Apache Open Office;
- VLC media player;
- Kaspersky Endpoint Security
- Yandex browser
- Python 3.12.4
- MongoDB

8.3. Перечень современных профессиональных баз данных и информационных справочных систем, доступных обучающимся при освоении дисциплины

1. Электронная информационно-образовательная среда Университета: образовательный портал (<http://moodle.aucu.ru>)
2. Электронно-библиотечная система «Университетская библиотека» (<https://biblioclub.ru/>)
3. Электронно-библиотечная система «IPRbooks» (www.iprbookshop.ru)
4. Научная электронная библиотека (<http://www.elibrary.ru/>)

9. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине

№ п/п	Наименование специальных помещений и помещений для самостоятельной работы	Оснащенность специальных помещений и помещений для самостоятельной работы
1	Учебные аудитории для проведения учебных занятий: 414056, г. Астрахань, ул. Татищева, 18, аудитории № 204, 207, 209, 211	<p>№204 Комплект учебной мебели Учебно-наглядные пособия Стационарный мультимедийный комплект Доступ к информационно – телекоммуникационной сети «Интернет»</p>
		<p>№207 Комплект учебной мебели Компьютеры - 15 шт. Стационарный мультимедийный комплект Доступ к информационно – телекоммуникационной сети «Интернет»</p>
		<p>№209 Комплект учебной мебели Компьютеры - 15 шт. Стационарный мультимедийный комплект Графические планшеты – 16 шт. Источник бесперебойного питания –</p>

№ п/п	Наименование специальных помещений и помещений для самостоятельной работы	Оснащенность специальных помещений и помещений для самостоятельной работы
		1 шт. Доступ к информационно – телекоммуникационной сети «Интернет» №211 Комплект учебной мебели Компьютеры - 15 шт. Стационарный мультимедийный комплект Доступ к информационно – телекоммуникационной сети «Интернет»
2	Помещения для самостоятельной работы: 414056, г. Астрахань, ул. Татищева, 22а, аудитории №201,203 414056, г. Астрахань, ул. Татищева, 18а, библиотека, читальный зал	№201 Комплект учебной мебели Компьютеры – 8 шт. Доступ к информационно – телекоммуникационной сети «Интернет» №203 Комплект учебной мебели Компьютеры – 8 шт. Доступ к информационно – телекоммуникационной сети «Интернет» библиотека, читальный зал Комплект учебной мебели Компьютеры - 4 шт. Доступ к информационно – телекоммуникационной сети «Интернет»

10. Особенности организации обучения по дисциплине для инвалидов и лиц с ограниченными возможностями здоровья

Для обучающихся из числа инвалидов и лиц с ограниченными возможностями здоровья на основании письменного заявления дисциплина «Технологии анализа данных» реализуется с учетом особенностей психофизического развития, индивидуальных возможностей и состояния здоровья (далее – индивидуальных особенностей).

Аннотация
к рабочей программе дисциплины «Технологии анализа данных»
по направлению подготовки **09.04.02 «Информационные системы и технологии»**
направленность (профиль) **«Искусственный интеллект в проектировании городской среды»**

Общая трудоемкость дисциплины составляет 5 зачетных единиц.

Форма промежуточной аттестации: зачет, экзамен.

Целью освоения дисциплины «Технологии анализ данных» является формирование компетенций обучающихся в соответствии с требованиями Федерального государственного образовательного стандарта высшего образования по направлению подготовки 09.04.02 «Информационные системы и технологии».

Учебная дисциплина Б1.В.02 «Технологии анализа данных» входит в **Блок 1 «Дисциплины (модули)» части формируемой участниками образовательных отношений**. Для освоения дисциплины необходимы знания, полученные при изучении следующих дисциплин: «Специальные главы математики», «Прикладной искусственный интеллект (базовый уровень)».

Краткое содержание дисциплины

- Раздел 1. Введение в большие данные
- Раздел 2. Этапы обработки больших данных
- Раздел 3. Статистические методы анализа больших данных
- Раздел 4. Визуализация больших данных
- Раздел 5. Хранение и обработка больших данных
- Раздел 6. Современные программные средства анализа больших данных

И.о. заведующего кафедрой



(подпись)

/В.В. Соболева/
И.О.Ф

РЕЦЕНЗИЯ
на рабочую программу, оценочные и методические материалы
Б1.В.02 «Технологии анализа данных»
(наименование дисциплины с указанием блока)

ОПОП ВО по направлению подготовки
09.04.02 «Информационные системы и технологии»
направленность (профиль)
«Искусственный интеллект в проектировании городской среды»
по программе магистратуры

Соболева В.В. (далее по тексту рецензент), проведена рецензия рабочей программы, оценочных и методических материалов по дисциплине «Технологии анализа данных» ОПОП ВО по направлению подготовки 09.04.02 «Информационные системы и технологии», направленность (профиль) «Искусственный интеллект в проектировании городской среды», по программе магистратуры, разработанной в ГБОУ АО ВО "Астраханский государственный архитектурно-строительный университет", на кафедре САПРиМ (разработчик – Окладникова С.В.).

Рассмотрев представленные на рецензию материалы, рецензент пришел к следующим выводам:

Предъявленная рабочая программа учебной дисциплины «Технологии анализа данных» (далее по тексту Программа) соответствует требованиям ФГОС ВО по направлению подготовки 09.04.02 «Информационные системы и технологии», утвержденного приказом Министерства образования и науки Российской Федерации от 19.09.2017г. №917, редакция с изменениями № 1456 от 26.11.2020г., с изменениями и дополнениями от 26.11.2020 г., 08.02.2021 г. и зарегистрированного в Минюсте России от 16.10.2017г, №48550.

Представленная в Программе актуальность учебной дисциплины в рамках реализации ОПОП ВО не подлежит сомнению – дисциплина относится к Блок 1 «Дисциплины (модули)»: части формируемой участниками образовательных отношений.

Представленные в Программе цели учебной дисциплины соответствуют требованиям ФГОС ВО по направлению подготовки 09.04.02 «Информационные системы и технологии», направленность (профиль) «Искусственный интеллект в проектировании городской среды».

В соответствии с Программой за дисциплиной «Технологии анализа данных» закреплена 1 компетенция, которая реализуется в объявленных требованиях.

Предложенные в Программе индикаторы компетенций в категориях знать, уметь, иметь практический опыт отражают специфику и содержание дисциплины, а представленные в ОММ показатели и критерии оценивания компетенций по дисциплине на различных этапах их формирования, а также шкалы оценивания позволяют определить степень достижения заявленных результатов, т.е. уровень освоения обучающимися соответствующих компетенций в рамках дисциплины.

Учебная дисциплина «Технологии анализа данных» взаимосвязана с другими дисциплинами ОПОП ВО по направлению подготовки 09.04.02 «Информационные системы и технологии», направленность (профиль) «Искусственный интеллект в проектировании городской среды» и возможность дублирования в содержании не выявлена.

Представленная Программа предполагает использование современных образовательных технологий при реализации различных видов учебной работы. Формы образовательных технологий соответствуют специфике дисциплины.

Представленные и описанные в Программе формы текущей оценки знаний соответствуют специфике дисциплины и требованиям к выпускникам.

Промежуточная аттестация знаний магистратуры, предусмотренная Программой, осуществляется в форме **зачета, экзамена**. Формы оценки знаний, представленные в Рабочей программе, соответствуют специфике дисциплины и требованиям к выпускникам.

Учебно-методическое обеспечение дисциплины представлено основной, дополнительной литературой, интернет-ресурсами и соответствует требованиям ФГОС ВО

направления подготовки 09.04.02 «Информационные системы и технологии», направленность (профиль) «Искусственный интеллект в проектировании городской среды».

Материально-техническое обеспечение соответствует требованиям ФГОС ВО направления подготовки 09.04.02 «Информационные системы и технологии», направленность (профиль) «Искусственный интеллект в проектировании городской среды» и специфике дисциплины «Технологии анализа данных» и обеспечивает использование современных образовательных, в том числе интерактивных методов обучения.

Представленные на рецензию оценочные и методические материалы направления подготовки 09.04.02 «Информационные системы и технологии», разработаны в соответствии с нормативными документами, представленными в Программе. Оценочные и методические материалы по дисциплине «Технологии анализа данных» предназначены для текущего контроля и промежуточной аттестации и представляет собой совокупность разработанных кафедрой САПРиМ материалов для установления уровня и качества достижения обучающимися результатов обучения.

Задачами оценочных и методических материалов является контроль и управление процессом освоения обучающимися компетенций, заявленных в образовательной программе по направлению подготовки 09.04.02 «Информационные системы и технологии», направленность (профиль) «Искусственный интеллект в проектировании городской среды».

Оценочные и методические материалы по дисциплине «Технологии анализа данных» представлены: перечнем материалов текущего контроля и промежуточной аттестации.

Данные материалы позволяют в полной мере оценить результаты обучения по дисциплине «Технологии анализа данных» в АГАСУ, а также оценить степень сформированной компетенций.

ОБЩИЕ ВЫВОДЫ

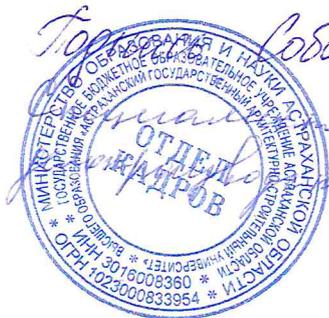
На основании проведенной рецензии можно сделать заключение, что характер, структура, содержание рабочей программы, оценочных и методических материалов дисциплины «Технологии анализа данных» ОПОП ВО по направлению подготовки 09.04.02 «Информационные системы и технологии», направленность (профиль) «Искусственный интеллект в проектировании городской среды», по программе *магистратуры*, разработанная к.т.н., доц. Окладниковой С.В. соответствует требованиям ФГОС ВО, современным требованиям отрасли, рынка труда, профессиональных стандартов направлению подготовки 09.04.02 «Информационные системы и технологии», направленность (профиль) «Искусственный интеллект в проектировании городской среды» и могут быть рекомендованы к использованию.

Рецензент:

Соболева Вера Владимировна,
кандидат педагогических наук,
и.о. заведующего кафедрой
«Системы автоматизированного
проектирования и моделирования
(САПРиМ)» ГБОУ АО ВО «АГАСУ».


(подпись)

/В.В. Соболева/
(И.О.Ф.)



Соболевой В.В. заверяю.

по договору

№ 10/2019 от 10.08.2019 г. с г-ном Д.В. Странуновым

РЕЦЕНЗИЯ
на рабочую программу, оценочные и методические материалы
Б1.В.02 «Технологии анализа данных»
(наименование дисциплины с указанием блока)

ОПОП ВО по направлению подготовки
09.04.02 «Информационные системы и технологии»
направленность (профиль)
«Искусственный интеллект в проектировании городской среды»
по программе магистратуры

Хоменко Т.В. (далее по тексту рецензент), проведена рецензия рабочей программы, оценочных и методических материалов по дисциплине «Технологии анализа данных» ОПОП ВО по направлению подготовки 09.04.02 «Информационные системы и технологии», направленность (профиль) «Искусственный интеллект в проектировании городской среды», по программе магистратуры, разработанной в ГБОУ АО ВО "Астраханский государственный архитектурно-строительный университет", на кафедре САПРиМ (разработчик – Окладникова С.В.).

Рассмотрев представленные на рецензию материалы, рецензент пришел к следующим выводам:

Предъявленная рабочая программа учебной дисциплины «Технологии анализа данных» (далее по тексту Программа) соответствует требованиям ФГОС ВО по направлению подготовки 09.04.02 «Информационные системы и технологии», утвержденного приказом Министерства образования и науки Российской Федерации от 19.09.2017г. №917, редакция с изменениями № 1456 от 26.11.2020г., с изменениями и дополнениями от 26.11.2020 г., 08.02.2021 г. и зарегистрированного в Минюсте России от 16.10.2017г, №48550.

Представленная в Программе актуальность учебной дисциплины в рамках реализации ОПОП ВО не подлежит сомнению – дисциплина относится к Блок 1 «Дисциплины (модули)»: части формируемой участниками образовательных отношений.

Представленные в Программе цели учебной дисциплины соответствуют требованиям ФГОС ВО по направлению подготовки 09.04.02 «Информационные системы и технологии», направленность (профиль) «Искусственный интеллект в проектировании городской среды».

В соответствии с Программой за дисциплиной «Технологии анализа данных» закреплена 1 компетенция, которая реализуется в объявленных требованиях.

Предложенные в Программе индикаторы компетенций в категориях знать, уметь, иметь практический опыт отражают специфику и содержание дисциплины, а представленные в ОММ показатели и критерии оценивания компетенций по дисциплине на различных этапах их формирования, а также шкалы оценивания позволяют определить степень достижения заявленных результатов, т.е. уровень освоения обучающимися соответствующих компетенций в рамках дисциплины.

Учебная дисциплина «Технологии анализа данных» взаимосвязана с другими дисциплинами ОПОП ВО по направлению подготовки 09.04.02 «Информационные системы и технологии», направленность (профиль) «Искусственный интеллект в проектировании городской среды» и возможность дублирования в содержании не выявлена.

Представленная Программа предполагает использование современных образовательных технологий при реализации различных видов учебной работы. Формы образовательных технологий соответствуют специфике дисциплины.

Представленные и описанные в Программе формы текущей оценки знаний соответствуют специфике дисциплины и требованиям к выпускникам.

Промежуточная аттестация знаний магистратуры, предусмотренная Программой, осуществляется в форме **зачета, экзамена**. Формы оценки знаний, представленные в Рабочей программе, соответствуют специфике дисциплины и требованиям к выпускникам.

Учебно-методическое обеспечение дисциплины представлено основной, дополнительной литературой, интернет-ресурсами и соответствует требованиям ФГОС ВО

направления подготовки 09.04.02 «Информационные системы и технологии», направленность (профиль) «Искусственный интеллект в проектировании городской среды».

Материально-техническое обеспечение соответствует требованиям ФГОС ВО направления подготовки 09.04.02 «Информационные системы и технологии», направленность (профиль) «Искусственный интеллект в проектировании городской среды» и специфике дисциплины «Технологии анализа данных» и обеспечивает использование современных образовательных, в том числе интерактивных методов обучения.

Представленные на рецензию оценочные и методические материалы по направлению подготовки 09.04.02 «Информационные системы и технологии», разработаны в соответствии с нормативными документами, представленными в Программе. Оценочные и методические материалы по дисциплине «Технологии анализа данных» предназначены для текущего контроля и промежуточной аттестации и представляет собой совокупность разработанных кафедрой САПРиМ материалов для установления уровня и качества достижения обучающимися результатов обучения.

Задачами оценочных и методических материалов является контроль и управление процессом освоения обучающимися компетенций, заявленных в образовательной программе направления подготовки 09.04.02 «Информационные системы и технологии», направленность (профиль) «Искусственный интеллект в проектировании городской среды».

Оценочные и методические материалы по дисциплине «Технологии анализа данных» представлены: перечнем материалов текущего контроля и промежуточной аттестации.

Данные материалы позволяют в полной мере оценить результаты обучения по дисциплине «Технологии анализа данных» в АГАСУ, а также оценить степень сформированной компетенций.

ОБЩИЕ ВЫВОДЫ

На основании проведенной рецензии можно сделать заключение, что характер, структура, содержание рабочей программы, оценочных и методических материалов дисциплины «Технологии анализа данных» ОПОП ВО по направлению подготовки 09.04.02 «Информационные системы и технологии», направленность (профиль) «Искусственный интеллект в проектировании городской среды», по программе *магистратуры*, разработанная к.т.н., доц. Окладниковой С.В. соответствует требованиям ФГОС ВО, современным требованиям отрасли, рынка труда, профессиональных стандартов направлению подготовки 09.04.02 «Информационные системы и технологии», направленность (профиль) «Искусственный интеллект в проектировании городской среды» и могут быть рекомендованы к использованию.

Рецензент:

Хоменко Татьяна Владимировна,
доктор технических наук, доцент
зав. кафедрой «Автоматизированные
системы обработки информации и
управления (АСОИУ)» ФГБОУ ВО
«Астраханский государственный
технический университет»


(подпись)

/ Т.В. Хоменко/
(И.О.Ф.)



Министерство образования и науки Астраханской области
Государственное бюджетное образовательное учреждение
Астраханской области высшего образования
«Астраханский государственный архитектурно-строительный
университет»
(ГБОУ АО ВО «АГАСУ»)

УТВЕРЖДАЮ



И.о. первого проректора

/ С.П. Стрелков /

(подпись)

И. О. Ф.

2024 г.

ОЦЕНОЧНЫЕ И МЕТОДИЧЕСКИЕ МАТЕРИАЛЫ

Наименование дисциплины

«Технологии анализа данных»

(указывается наименование в соответствии с учебным планом)

По направлению подготовки

09.04.02 «Информационные системы и технологии»

(указывается наименование направления подготовки в соответствии с ФГОС ВО)

Направленность (профиль)

«Искусственный интеллект в проектировании городской среды»

(указывается наименование направленности (профиля) в соответствии с ОПОП)

Кафедра

Системы автоматизированного проектирования и моделирования

Квалификация выпускника *магистр*

Разработчик:

к.т.н., доцент

(занимаемая должность,
учёная степень и учёное звание)



(подпись)

/С.В. Окладникова/

(И. О. Ф.)

Оценочные и методические материалы рассмотрены и утверждены на заседании кафедры
«Системы автоматизированного проектирования и моделирования»

Протокол № 9 от 22.04 2024 г.

И.о. заведующего кафедрой САПРиМ



(подпись)

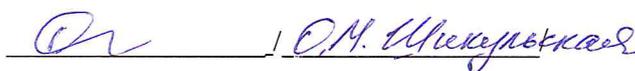
/Соболева В.В./

(И. О. Ф.)

Согласовано:

Председатель МКН «Информационные системы и технологии»

направленность(профиль) «Искусственный интеллект в проектировании городской среды»



Начальник УМУ



(подпись)



(И. О. Ф.)

Специалист УМУ



(подпись)



(И. О. Ф.)

Содержание

1. Оценочные и методические материалы для проведения текущего контроля успеваемости и промежуточной аттестации обучающихся по дисциплине	4
1.1. Перечень компетенций с указанием этапов их формирования в процессе освоения образовательной программы	4
1.2. Описание показателей и критериев оценивания компетенций на различных этапах их формирования, описание шкал оценивания.	5
1.2.1. Перечень оценочных средств текущего контроля успеваемости	5
1.2.2. Описание показателей и критериев оценивания компетенций по дисциплине на различных этапах их формирования, описание шкал оценивания	5
1.2.3. Шкала оценивания	8
2. Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков, характеризующих этапы формирования компетенций в процессе освоения образовательной программы	9
2.1. Зачет, экзамен	9
2.2. Защита лабораторной работы.....	10
2.3. Контрольная работа.....	10
2.4. Тест	11
3. Перечень и характеристики процедуры оценивания знаний, умений, навыков, характеризующих этапы формирования компетенций	12
<i>Приложение 1</i>	13
<i>Приложение 2</i>	15
<i>Приложение 3</i>	16
<i>Приложение 4</i>	19
<i>Приложение 5</i>	20
<i>Приложение 6</i>	22

1. Оценочные и методические материалы для проведения текущего контроля успеваемости и промежуточной аттестации обучающихся по дисциплине

Оценочные и методические материалы являются неотъемлемой частью рабочей программы дисциплины (далее РПД) и представлены в виде отдельного документа.

1.1. Перечень компетенций с указанием этапов их формирования в процессе освоения образовательной программы

Индекс и формулировка компетенции		Индикаторы достижений компетенций, установленные ОПОП	Номер раздела дисциплины (в соответствии с п.5.1 РПД)						Формы контроля с конкретизацией задания	
			1	2	3	4	5	6		
1		2	3	4	5	6	7	8	9	
ПК-4ИИП. Способен руководить проектами по созданию комплексных систем на основе аналитики больших данных в различных отраслях со стороны заказчика	ПК4.1ИИП – Руководит проектами по построению комплексных систем на основе аналитики больших данных в различных отраслях со стороны заказчика	Знать:								
		Знает методологию и принципы руководства проектами по созданию, поддержке и использованию комплексных систем на основе аналитики больших данных со стороны заказчика (ПК-4.1ИИП. З-1.)	X	X	X					Вопросы к зачету [1-15] Вопросы к экзамену [1-14] Итоговое тестирование [1-10] Вопросы к зачету [16-30] Вопросы к экзамену [15-28] Итоговое тестирование [11-20]
		Знает специфику сфер и отраслей, для которых реализуется проект по аналитике больших данных (ПК-4.1ИИП. З-2.)	X	X	X					
		Уметь:								
		Умеет решать задачи по руководству коллективной проектной деятельностью для создания, поддержки и использования комплексных систем на основе аналитики больших данных со стороны заказчика (ПК-4.1ИИП. У-1.)	X	X	X	X	X	X		Контрольная работа [задание 1] Отчет по выполнению лабораторной работы Контрольная работа [задание 2]
		Умеет выявлять небольшие по масштабу проекты аналитики, которые потенциально могут представлять интерес для ряда подразделений / служб или для организации в целом (ПК-4.1ИИП. У-2.)	X	X	X	X	X	X		Отчет по выполнению лабораторной работы Контрольная работа [задание 3]
Умеет выявлять области деловой деятельности, которые потенциально могут получить отдачу от аналитики (ПК-4.1ИИП. У-3.)	X	X	X	X	X	X		Отчет по выполнению лабораторной работы		

1.2. Описание показателей и критериев оценивания компетенций на различных этапах их формирования, описание шкал оценивания.

1.2.1. Перечень оценочных средств текущего контроля успеваемости

Наименование оценочного средства	Краткая характеристика оценочного средства	Представление оценочного средства в фонде
Лабораторная работа	Средство, позволяющее оценить умение и владение обучающегося излагать суть поставленной задачи, самостоятельно применять стандартные методы решения поставленной задачи с использованием имеющейся лабораторно-практической базы, проводить анализ полученного результата работы. Рекомендуется для оценки умений и владений студентов	Темы лабораторной работ и требования к защите
Контрольная работа	Средство проверки умений применять полученные знания для решения задач определенного типа по теме или разделу	Комплект контрольных заданий по вариантам
Тест	Система стандартизированных заданий, позволяющая автоматизировать процедуру измерения уровня знаний и умений обучающегося	Фонд тестовых заданий

1.2.2. Описание показателей и критериев оценивания компетенций по дисциплине на различных этапах их формирования, описание шкал оценивания

Компетенция, этапы освоения компетенции	Планируемые результаты обучения	Показатели и критерии оценивания результатов обучения			
		Ниже порогового уровня (неудовлетворительно)	Пороговый уровень (удовлетворительно)	Продвинутый уровень (хорошо)	Высокий уровень (отлично)
1	2	3	4	5	6
ПК-4ИИП. Способен руководить проектами по созданию комплексных систем на основе аналитики больших данных в различных	Знать: Знает методологию и принципы руководства проектами по созданию, поддержке и использованию комплексных систем на основе	Не знает методологию и принципы руководства проектами по созданию, поддержке и использованию комплексных систем на основе аналитики больших данных со стороны заказчика. Не знает специфику сфер и отраслей, для которых реализуется	Знает методологию и принципы руководства проектами по созданию, поддержке и использованию комплексных систем на основе аналитики больших данных со стороны заказчика. Знает специфику сфер и отраслей, для	Знает методологию и принципы руководства проектами по созданию, поддержке и использованию комплексных систем на основе аналитики	Знает методологию и принципы руководства проектами по созданию, поддержке и использованию комплексных систем на основе аналитики

Компетенция, этапы освоения компетенции		Планируемые результаты обучения	Показатели и критерии оценивания результатов обучения			
			Ниже порогового уровня (неудовлетворительно)	Пороговый уровень (удовлетворительно)	Продвинутый уровень (хорошо)	Высокий уровень (отлично)
1		2	3	4	5	6
отраслях со стороны заказчика		аналитики больших данных со стороны заказчика (ПК-4.1ИИП. 3-1.) Знает специфику сфер и отраслей, для которых реализуется проект по аналитике больших данных (ПК-4.1ИИП. 3-2.)	проект по аналитике больших данных.	которых реализуется проект по аналитике больших данных.	больших данных со стороны заказчика. Знает специфику сфер и отраслей, для которых реализуется проект по аналитике больших данных.	больших данных со стороны заказчика. Знает специфику сфер и отраслей, для которых реализуется проект по аналитике больших данных.
		Уметь: Умеет решать задачи по руководству коллективной проектной деятельностью для создания, поддержки и использования комплексных систем на основе аналитики больших данных со	Не умеет решать задачи по руководству коллективной проектной деятельностью для создания, поддержки и использования комплексных систем на основе аналитики больших данных со стороны заказчика.	Умеет решать задачи по руководству коллективной проектной деятельностью для создания, поддержки и использования комплексных систем на основе аналитики больших данных со стороны заказчика.	Умеет решать задачи по руководству коллективной проектной деятельностью для создания, поддержки и использования комплексных систем на основе аналитики больших данных со	Умеет решать задачи по руководству коллективной проектной деятельностью для создания, поддержки и использования комплексных систем на основе аналитики больших данных со

Компетенция, этапы освоения компетенции		Планируемые результаты обучения	Показатели и критерии оценивания результатов обучения			
			Ниже порогового уровня (неудовлетворительно)	Пороговый уровень (удовлетворительно)	Продвинутый уровень (хорошо)	Высокий уровень (отлично)
1		2	3	4	5	6
		стороны заказчика (ПК- 4.1ИИП. У-1.)	Не умеет выявлять		стороны заказчика.	стороны заказчика.

Компетенция, этапы освоения компетенции		Планируемые результаты обучения	Показатели и критерии оценивания результатов обучения			
			Ниже порогового уровня (неудовлетворительно)	Пороговый уровень (удовлетворительно)	Продвинутый уровень (хорошо)	Высокий уровень (отлично)
1		2	3	4	5	6
ПК4.1ИИП – Руководит проектами по построению комплексных систем на основе аналитики больших данных в различных отраслях со стороны заказчика		Умеет выявлять небольшие по масштабу проекты аналитики, которые потенциально могут представлять интерес для ряда подразделений / служб или для организации в целом (ПК-4.1ИИП. У-2.) Умеет выявлять области деловой деятельности, которые потенциально могут получить отдачу от аналитики (ПК-4.1ИИП. У-3.)	небольшие по масштабу проекты аналитики, которые потенциально могут представлять интерес для ряда подразделений / служб или для организации в целом. Не умеет выявлять области деловой деятельности, которые потенциально могут получить отдачу от аналитики.	Умеет выявлять небольшие по масштабу проекты аналитики, которые потенциально могут представлять интерес для ряда подразделений / служб или для организации в целом. Умеет выявлять области деловой деятельности, которые потенциально могут получить отдачу от аналитики.	Умеет выявлять небольшие по масштабу проекты аналитики, которые потенциально могут представлять интерес для ряда подразделений / служб или для организации в целом. Умеет выявлять области деловой деятельности, которые потенциально могут получить отдачу от аналитики.	Умеет выявлять небольшие по масштабу проекты аналитики, которые потенциально могут представлять интерес для ряда подразделений / служб или для организации в целом. Умеет выявлять области деловой деятельности, которые потенциально могут получить отдачу от аналитики.

1.2.3. Шкала оценивания

Уровень достижений	Отметка в 5-бальной шкале
высокий	«5»(отлично)
продвинутый	«4»(хорошо)
пороговый	«3»(удовлетворительно)
ниже порогового	«2»(неудовлетворительно)

2. Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков, характеризующих этапы формирования компетенций в процессе освоения образовательной программы

ТИПОВЫЕ ЗАДАНИЯ ДЛЯ ПРОВЕДЕНИЯ ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ

2.1. Зачет, экзамен

а) типовые вопросы к зачету и экзамену (см. приложение 1);

б) критерии оценки.

При оценке знаний на зачете и экзамене учитывается:

1. Уровень сформированности компетенций.
2. Уровень усвоения теоретических положений дисциплины, правильность формулировки основных понятий и закономерностей.
3. Уровень знания фактического материала в объеме программы.
4. Логика, структура и грамотность изложения вопроса.
5. Умение связать теорию с практикой.
6. Умение делать обобщения, выводы.

№ п/п	Оценка	Критерии оценки
1	отлично	Ответы на поставленные вопросы излагаются логично, последовательно и не требуют дополнительных пояснений. Полно раскрываются причинно-следственные связи между явлениями и событиями. Делаются обоснованные выводы. Демонстрируются глубокие знания базовых нормативно-правовых актов. Соблюдаются нормы литературной речи.
2	хорошо	Ответы на поставленные вопросы излагаются систематизировано и последовательно. Базовые нормативно-правовые акты используются, но в недостаточном объеме. Материал излагается уверенно. Раскрыты причинно-следственные связи между явлениями и событиями. Демонстрируется умение анализировать материал, однако не все выводы носят аргументированный и доказательный характер. Соблюдаются нормы литературной речи.
3	удовлетворительно	Допускаются нарушения в последовательности изложения. Имеются упоминания об отдельных базовых нормативно-правовых актах. Неполно раскрываются причинно-следственные связи между явлениями и событиями. Демонстрируются поверхностные знания вопроса, с трудом решаются конкретные задачи. Имеются затруднения с выводами. Допускаются нарушения норм литературной речи.
4	неудовлетворительно	Материал излагается непоследовательно, сбивчиво, не представляет определенной системы знаний по дисциплине. Не раскрываются причинно-следственные связи между явлениями и событиями. Не проводится анализ. Выводы отсутствуют. Ответы на дополнительные вопросы отсутствуют. Имеются заметные нарушения норм литературной речи.
5	зачтено	Выставляется при соответствии параметрам экзаменационной шкалы на уровнях «отлично», «хорошо», «удовлетворительно».

№ п/п	Оценка	Критерии оценки
6	не зачтено	Выставляется при соответствии параметрам экзаменационной шкалы на уровне «неудовлетворительно».

ТИПОВЫЕ ЗАДАНИЯ ДЛЯ ПРОВЕДЕНИЯ ТЕКУЩЕГО КОНТРОЛЯ

2.2. Защита лабораторной работы

- а) типовые задания лабораторной работ (см. приложение 2);
 б) критерии оценки.

При оценке знаний на защите лабораторной работы учитывается:

1. Уровень сформированности компетенций.
2. Уровень усвоения теоретических положений дисциплины, правильность формулировки основных понятий и закономерностей.
3. Уровень знания фактического материала в объеме программы.
4. Логика, структура и грамотность изложения вопроса.
5. Умение связать теорию с практикой.
6. Умение делать обобщения, выводы.

№ п/п	Оценка	Критерии оценки
1	2	3
1	отлично	Студент правильно называет метод исследования, правильно называет прибор, правильно демонстрирует методику исследования /измерения, правильно оценивает результат.
2	хорошо	Студент правильно называет метод исследования, правильно называет прибор, допускает единичные ошибки в демонстрации методики исследования /измерения и оценке его результатов
3	удовлетворительно	Студент неправильно называет метод исследования, но при этом дает правильное название прибора. Допускает множественные ошибки в демонстрации методики исследования /измерения и оценке его результатов
4	неудовлетворительно	Студент неправильно называет метод исследования, дает неправильное название прибора. Не может продемонстрировать методику исследования /измерения, а также оценить результат
5	зачтено	Выставляется при соответствии параметрам экзаменационной шкалы на уровнях «отлично», «хорошо», «удовлетворительно».
6	не зачтено	Выставляется при соответствии параметрам экзаменационной шкалы на уровне «неудовлетворительно».

2.3. Контрольная работа

- а) варианты контрольной работы (см. приложение 3);
 б) критерии оценки.

При оценке знаний контрольной работы учитывается:

1. Уровень сформированности компетенций.
2. Уровень усвоения теоретических положений дисциплины, правильность формулировки основных понятий и закономерностей.
3. Уровень знания фактического материала в объеме программы.
4. Логика, структура и грамотность изложения вопроса.
5. Умение связать теорию с практикой.
6. Умение делать обобщения, выводы.

№ п/п	Оценка	Критерии оценки
1	2	3
1	отлично	Студент правильно называет метод исследования, правильно называет прибор, правильно демонстрирует методику исследования /измерения, правильно оценивает результат.
2	хорошо	Студент правильно называет метод исследования, правильно называет прибор, допускает единичные ошибки в демонстрации методики исследования /измерения и оценке его результатов
3	удовлетворительно	Студент неправильно называет метод исследования, но при этом дает правильное название прибора. Допускает множественные ошибки в демонстрации методики исследования /измерения и оценке его результатов
4	неудовлетворительно	Студент неправильно называет метод исследования, дает неправильное название прибора. Не может продемонстрировать методику исследования /измерения, а также оценить результат
5	зачтено	Выставляется при соответствии параметрам экзаменационной шкалы на уровнях «отлично», «хорошо», «удовлетворительно».
6	не зачтено	Выставляется при соответствии параметрам экзаменационной шкалы на уровне «неудовлетворительно».

2.4. Тест

- а) *типовой комплект заданий для входного тестирования (Приложение 4)*
типовой комплект заданий для итогового тестирования (Приложение 5)
- б) *критерии оценки.*

При оценке знаний по результатам тестов учитывается:

1. Уровень сформированности компетенций.
2. Уровень усвоения теоретических положений дисциплины, правильность формулировки основных понятий и закономерностей.
3. Уровень знания фактического материала в объеме программы.
4. Логика, структура и грамотность изложения вопроса.
5. Умение связать теорию с практикой.
6. Умение делать обобщения, выводы.

№ п/п	Оценка	Критерии оценки
1	2	3
1	отлично	если выполнены следующие условия: - даны правильные ответы не менее чем на 90% вопросов теста, исключая вопросы, на которые студент должен дать свободный ответ; - на все вопросы, предполагающие свободный ответ, студент дал правильный и полный ответ.
2	хорошо	если выполнены следующие условия: - даны правильные ответы не менее чем на 75% вопросов теста, исключая вопросы, на которые студент должен дать свободный ответ; - на все вопросы, предполагающие свободный ответ, студент дал правильный ответ, но допустил незначительные ошибки и не показал необходимой полноты.
3	удовлетворительно	если выполнены следующие условия: - даны правильные ответы не менее чем на 50% вопросов теста, исключая вопросы, на которые студент должен дать свободный ответ;

№ п/п	Оценка	Критерии оценки
1	2	3
		-- на все вопросы, предполагающие свободный ответ, студент дал непротиворечивый ответ, или при ответе допустил значительные неточности и не показал полноты.
4	неудовлетворительно	если студентом не выполнены условия, предполагающие оценку «удовлетворительно».
5	зачтено	Выставляется при соответствии параметрам экзаменационной шкалы на уровнях «отлично», «хорошо», «удовлетворительно».
6	не зачтено	Выставляется при соответствии параметрам экзаменационной шкалы на уровне «неудовлетворительно».

3. Перечень и характеристики процедуры оценивания знаний, умений, навыков, характеризующих этапы формирования компетенций

Процедура проведения текущего контроля успеваемости и промежуточной аттестации обучающихся по дисциплине регламентируется локальным нормативным актом.

Перечень и характеристика процедур текущего контроля и промежуточной аттестации по дисциплине

№	Наименование оценочного средства	Периодичность и способ проведения процедуры оценивания	Виды вставляемых оценок	Форма учета
1.	Зачет	Раз в семестр, по окончании изучения дисциплины	По пятибалльной шкале	Ведомость, зачетная книжка, портфолио
2	Экзамен	Раз в семестр, по окончании изучения дисциплины	По пятибалльной шкале	Ведомость, зачетная книжка, портфолио
3	Лабораторная работа	Систематически на занятиях	По пятибалльной шкале	Лабораторная тетрадь, журнал успеваемости преподавателя
4	Контрольная работа	Раз в семестр по окончании изучения дисциплины	По пятибалльной шкале	Журнал успеваемости преподавателя
5	Тест	Входное тестирование по дисциплине – вначале изучения дисциплины (в начале семестра) Итоговое тестирование – по окончании изучения дисциплины	По пятибалльной шкале	Лист результатов компьютерного тестирования, журнал успеваемости преподавателя, ведомость, зачетная книжка, портфолио

**Типовые вопросы к зачету
по дисциплине «Технологии анализа данных»**

Знать - ПК-4.1ИИП. 3-1

1. Какие инструментальные средства используются для организации рабочего места специалиста Data Science?
2. Какие библиотеки Python используются для работы в области анализа данных? Дайте краткую характеристику каждой библиотеке.
3. Почему при реализации систем анализа данных широкое распространение получили библиотеки Python?
4. Поясните принцип разбиения набора данных на обучающую и тестовую выборку.
5. Поясните особенности основных методов метрической классификации: метод ближайшего соседа, метод k ближайших соседей.
6. Поясните принцип построения дерева решений.
7. Почему при реализации линейной модели регрессии нет необходимости выполнять масштабирование признаков?
8. Почему при реализации модели линейной регрессии в качестве функции потерь используется квадратичное отклонение, а не модуль отклонения?
9. Поясните назначение метода predict класса LinearRegression.
10. Поясните назначение метода plot и scatter класса pyplot.
11. Что такое кластерный анализ?
12. Перечислите известные методы кластерного анализа.
13. Перечислите классы и функции Python, которые задействованы при реализации кластерного анализа.
14. Опишите принцип определения оптимального количества кластеров.
15. Опишите принципиальные отличия методов регрессии, кластеризации и классификации.

Знать - ПК-4.1ИИП. 3-2

16. Какие проблемы при импорте данных могут возникнуть, какие способы их решения могут быть использованы?
17. Есть ли в наборе данных количественная дискретная переменная?
18. Есть ли в данных категориальные переменные? Как преобразовать их в числовые векторы?
19. Есть ли в данных количественные переменные? Какой у них тип?
20. Как построить срез данных по условию? Как провести сортировку набора данных по одному из столбцов?
21. С помощью параметров описательной статистики опишите одну из количественных переменных и одну из категориальных переменных.
22. Есть ли в количественной переменной выбросы? Какое значение является выбросом и почему?
23. По какой категориальной переменной можно группировать данные? Какие группировки представляют интерес и почему?
24. Какие визуализации построены для количественных переменных? Какие методы библиотек на языке Python и с какими входными параметрами при этом использовались? Какие особенности распределения данных они позволяют наблюдать?
25. Какие визуализации построены для категориальных переменных? Какие методы библиотек на языке Python и с какими входными параметрами при этом использовались? Какие особенности распределения данных они позволяют наблюдать?
26. Какие визуализации построены для пар переменных? Какие методы библиотек на языке Python и с какими входными параметрами при этом использовались? Какие особенности распределения данных они позволяют наблюдать?

27. Какие гипотезы нужно протестировать для ответа на исследовательский вопрос? Альтернативная гипотеза, сформулированная вами, является односторонней или двусторонней?
28. Какой статистический критерий необходимо использовать при тестировании гипотез?
29. Какой уровень значимости был выбран? Что позволяет определить уровень значимости при тестировании гипотез? Как уровень значимости влияет на вывод?
30. Какой методы на языке Python использовались для тестирования гипотез? Какой содержательный вывод можно сделать по результатам тестирования?

**Типовые вопросы к экзамену
по дисциплине «Технологии анализа данных»**

Знать - ПК-4.1ИИП. 3-1

1. Что означает термин NoSQL? Какие преимущества предоставляют NoSQL базы данных в сравнении с реляционными базами данных?
2. Какими особенностями обладает MongoDB? Сколькими способами можно произвести установку MongoDB? Кратко опишите эти способы.
3. Какие существуют способы взаимодействия с БД Mongo? Существует ли возможность взаимодействовать с БД Mongo посредством WEB?
4. Какими особенностями обладает Mongo Explorer? Опишите интерфейс приложения Mongo Explorer.
5. Дайте определения терминам JSON и BSON. На каких структурах строится JSON? Какие форма представления данных используются в JSON?
6. Возможно ли в MongoDB использовать массив документов? Существует ли возможность в MongoDB использовать вложенные документы?
7. Какие способы создания БД присутствуют в MongoDB?
8. Для чего используется индекс в БД Mongo? Как создается и удаляется индекс в MongoDB? Перечислите опции создания индекса? Какие типы индексов можно создать в MongoDB?
9. Каким образом можно вставить в массив несколько элементов? Приведите пример запроса.
10. Для чего используются модификаторы массивов в запросах?
11. Какие способы обновления данных в массиве вы знаете? Приведите примеры.
12. Для чего используются регулярные выражения? Какие метасимволы используются при составлении регулярных выражений?
13. Можно ли к метасимволам применять какие-либо модификаторы, если можно, то какие?
14. Какими правилами необходимо руководствоваться при составлении регулярных выражений?

Знать - ПК-4.1ИИП. 3-2

15. Что означает термин MapReduce?
16. Из каких шагов состоит работа MapReduce?
17. Какими преимуществами обладает MapReduce по сравнению с обычными вычислениями?
18. Опишите работу с MapReduce в MongoDB.
19. Какие утилиты используются для резервирования и восстановления баз данных?
20. Какие параметры запуска имеет утилита «mongodump.exe»?
21. Какие параметры запуска имеет утилита «mongorestore.exe»?
22. Приведите синтаксис команды для добавления пользователя в БД.
23. Каким образом производится смена пароля для пользователя?
24. Перечислите роли, которые можно назначать пользователям в БД.
25. Что означает термин MapReduce?
26. Из каких шагов состоит работа MapReduce?
27. В чем особенности выполнения mapReduce над распределенными коллекциями?
28. Что такое шардинг? Опишите преимущества и недостатки различных типов ключей, используемых для шардинга?

**Типовые задания по выполнению лабораторных работ
по дисциплине «Технологии анализа данных»**

Уметь – ПК-4.1ИИП. У-1, ПК-4.1ИИП. У-2, ПК-4.1ИИП. У-3

Лабораторная работа 1. «Технологии сбора открытых данных»

Цель: изучение структуры источников открытых данных, способов хранения и представления данных.

Задание 1.

структуру и организацию данных на Портале открытых данных РФ (data.gov.ru):

- 1) Выберите интересующую Вас тематику исследования. По выбранной тематике подберите набор данных в формате CSV. Изучите паспорт и структуру набора данных.
- 2) Загрузите выбранный набор данных на компьютер и импортируйте его в документ MS Excel (Лист 1).
- 3) Проведите анализ набора данных: определите количество записей и полей в наборе данных и т.п.

Задание 2.

Исследуйте наборы данных на портале World Bank Open Data (data.worldbank.org):

- 1) Найдите произвольный набор данных по тематике, выбранной в Задании 1.
- 2) Загрузите на компьютер найденный набор данных и импортируйте его в документ MS Excel (Лист 2).
- 3) На основе набора данных подготовьте выборку, содержащую значения показателя за все годы для трёх произвольно выбранных стран мира.
- 4) На основе подготовленной выборки постройте график, иллюстрирующий изменение показателя со временем для трёх стран мира.

Задание 3.

Постройте модель получения данных в режиме реального времени с веб-ресурсов:

- 1) Найдите веб-ресурс, который не реже чем с ежедневной периодичностью обновляет данные (например, курс валюты, прогноз погоды, онлайн-табло, прайс-лист и т.п.).
- 2) Загрузите используя Power Query на отдельные Листы документа MS Excel данные не менее чем за 5 периодов (например, за 5 дней/недель/месяцев/лет). Установите период обновления данных.
- 3) Выполните объединение запросов. Для сводных данных постройте спарклайны, показывающие тенденцию изменения данных.

Задание 4.

Подготовьте отчет о выполнении Заданий 1 – 3, в котором:

- 1) Дайте краткую характеристику источников данных и размещенных на них наборов данных.
- 2) Опишите основные этапы выполнения заданий и полученные результаты.

Лабораторная работа 2. «Предобработка и классификация данных в Python»

Цель: ознакомиться с методами предобработки данных из библиотеки Scikit Learn

Задание 1:

- 1) Выполнить загрузку данных из социальной сети VK (в соответствии с вариантом), используя Python vk_api. Выбрать сообщество и скачать все посты на странице этого сообщества. Изучить служебную информацию в загруженном файле.
- 2) Подсчитать слова в загруженном текстовом документе, используя алгоритм WordCount.

- 3) Решить задачу кластеризации текстовых данных. Рассмотреть реализацию алгоритма Алгоритм Incremental PCA.
- 4) Из социальной сети VK выполнить загрузку фотографий.
- 5) Выполнить кластеризацию цветов на изображениях. Изучить кластеризацию изображений на основе доминирующих цветов.
- 6) Построить и визуализировать граф связей друзей пользователя.

Задание 2:

- 1) Выполнить загрузку табличных данных в формате CSV из открытых источников.
- 2) Изучите возможности библиотеки Pandas для обработки датафреймов (выбор строк, столбцов, подсчета количества статистик, заполнение пропусков данных, нахождение дубликатов, изменение типов данных, категоризация данных и др.). Выполните необходимые преобразования.

Лабораторная работа 3. «Исследовательский анализ данных»

Цель: изучить технологии статистического анализа данных на основе библиотеки Pandas

Задание:

Исследовать набор данных методами: а) методом K-ближайших соседей, б) линейной регрессии, в) дерева решений, г) случайного леса.

Лабораторная работа 4. Инструменты интеграции и визуализации данных в Python

Цель: изучение программных средств для интеграции и визуализации данных

Задание:

- 1) В СУБД PostgreSQL создайте тестовую БД
- 2) Используя psycopg2 осуществите связь Python и БД в PostgreSQL
- 3) Создайте в Python новую таблицу и заполните ее тестовыми данными
- 4) Используя функцию SELECT в Python создайте не менее 3 различных запросов из различных таблиц
- 5) По результатам запросов из п.4. построить графики, используя библиотеку Matplotlib.

Лабораторная работа 5. «Основы NoSQL СУБД MONGODB»

Цель: сформировать навыки моделирования данных в MongoDB

Задание:

- 1) Установить MongoDB
- 2) Провести подключение к тестовой базе данных
- 3) Сформировать набор тестовых данных в БД с использованием командной строки
- 4) Извлечь тестовые данные с использованием командной строки
- 5) Добавьте тестовые данные в БД с использованием Mongo Explorer
- 6) Создать БД в любой реляционной СУБД, представить ее в нереляционном виде, записанную в JSON
- 7) Создать коллекцию тестовых документов (не менее 10). Документы коллекции должны содержать, по крайней мере, один массив, и один вложенный документ
- 8) Создайте составной и полнотекстовый индексы для документов в вашей коллекции.
- 9) Создайте запрос, производящий замену документа.
- 10) Создайте запрос, производящий обновление любого поля документа и запрос, добавляющий в него новые поля.
- 11) Создайте запрос для получения всех документов коллекции и определите, используется ли для него индекс

- 12) Создайте не менее 10 уникальных запросов (2-3 простых запросов для выборки данных из БД, 3-4 сложных запроса с каждым из перечисленных модификаторов, 3-4 запроса с использованием методов сортировки, ограничения и пропуска данных)
- 13) Создайте не менее 10 уникальных запросов для вставки данных в массив, производящие обновление данных в массиве: как по позиции элемента в массиве, так и по его значению, удаляющие элементы из массива: по позиции элемента в массиве и по его значению.
- 14) Создайте не менее 10 уникальных запросов с использованием регулярных выражений для поиска: документа начинающегося с определенной последовательности символов, документов содержащих определенную последовательность символов, документов со сложным именем (шаблонное имя)

Лабораторная работа 6. «Анализ данных по технологии MapReduce»

Цель: изучить модель распределенных вычислений MapReduce

Задание:

- 1) Создайте коллекцию документов для обработки её с помощью MapReduce
- 2) Наполните коллекцию документами
- 3) Произведите обработку коллекции с использованием модели распределенных вычислений MapReduce
- 4) Получите диагностическую информацию о БД и содержащихся в ней коллекциях
- 5) Создайте резервную копию данных БД
- 6) Восстановите БД из резервной копии
- 7) Создайте для БД несколько пользователей, имеющих различные роли

**Типовые задания по выполнению контрольной работы
по дисциплине «Технологии анализа данных»**

Уметь – ПК-4.1ИИП. У-1

Задание 1. Импортируйте набор данных в соответствии с вариантом. Проведите разведывательный анализ данных. Поставьте гипотезу. Проведите необходимые преобразования данных. На основании поставленной гипотезы выберите модели для предсказания. Обоснуйте выбор. Постройте выбранные модели, выберите лучшую. Постройте ансамбль моделей. Улучшились ли предсказания? Оформите результаты в виде отчета.

Уметь – ПК-4.1ИИП. У-2

Задание 2. Соберите данные для анализа с сайта Федеральной службы государственной статистики. Проведите предварительную обработку данных. Поставьте задачу классификации. Проведите необходимые преобразования данных. Постройте модель. Оцените, насколько хорошо модель справляется с поставленной задачей. Опишите результат, назовите возможные причины плохой работы модели (в случае, если результаты неудовлетворительные).

Уметь – ПК-4.1ИИП. У-3

Задание 3. Вы хотите узнать на сайте Центробанка РФ курсы доллара с 1 по 31 января 2021 года, на ваш запрос появилась такая ссылка в браузере: http://www.cbr.ru/scripts/XML_dynamic.asp?date_req1=01/01/2021&date_req2=31/01/2021&VAL_NM_RQ=R01235 Это текст, который содержит интересующие вас данные. Какие? Охарактеризуйте их. Как называется данная выгрузка?

Типовой комплект заданий для тестов

Полный комплект тестовых материалов по входному тестированию размещен на образовательном портале «АГАСУ»

Типовые тесты для входного тестирования

1. Объект статистического наблюдения - это
 - а) единица наблюдения;
 - б) статистическая совокупность;
 - в) единица статистической совокупности;
 - г) отчетная единица.
2. Субъект, от которого поступают данные в ходе статистического наблюдения, называется:
 - а) единица наблюдения;
 - б) единица статистической совокупности;
 - в) отчетная единица.
3. Перечень признаков (или вопросов), подлежащих регистрации в процессе наблюдения, называется:
 - а) статистический формуляр;
 - б) программа наблюдения;
 - в) инструментарий наблюдения.
4. Расхождение между расчетными значениями и действительным значением изучаемых величин называется:
 - а) ошибкой наблюдения;
 - б) ошибкой регистрации;
 - в) ошибкой репрезентативности.
5. К дискретным признакам группировок относится:
 - а) заработная плата рабочих;
 - б) численность населения страны;
 - в) число членов семьи;
 - г) прибыль банка.
6. К атрибутивным признакам группировок относится:
 - а) прибыль предприятия;
 - б) пол человека;
 - в) разряд работы;
 - г) размер обуви.
7. Основанием группировки может быть:
 - а) количественный признак;
 - б) качественный признак;
 - в) как количественный признак, так и качественный признак.
8. По характеру разработки сказуемого различают статистические таблицы:
 - а) простые;
 - б) сложные;
 - в) перечневые;
 - г) моногамные.
9. Соотношение одного и того же абсолютного показателя, характеризующего разные объекты, называется относительной величиной:
 - а) динамики;
 - б) структуры;
 - в) координации;
 - г) интенсивности;
 - д) сравнения.

10. Изменится ли средняя величина, если все веса увеличить в два раза:
- а) увеличиться в 2 раза;
 - б) не изменится;
 - в) изменится в 4 раза.
 - г) уменьшится в 2 раза
11. Можно ли для одного соотношения вычислять среднюю различными формами:
- а) возможно;
 - б) нельзя;
 - в) можно, если наложит дополнительные условия;
 - г) можно для определенных совокупностей
12. Средняя величина признака равна 20, а коэффициент вариации – 25%. Дисперсия признака равна:
- а) 20;
 - б) 25;
 - в) 125;
 - г) 45.
13. Чему равен коэффициент вариации, если средняя величина признака равна 22 рубля, а дисперсия 36:
- а) 25,0%
 - б) 27,3%
 - в) 33,5%
 - г) 63,6%
14. Коэффициент детерминации представляет собой долю:
- а) межгрупповой дисперсии в общей
 - б) межгрупповой дисперсии в остаточной
 - в) дисперсии теоретических значений в общей дисперсии
15. Средний уровень интервального ряда динамики с равными временными промежутками исчисляется по формуле средней:
- а) арифметической простой;
 - б) арифметической взвешенной;
 - в) гармонической простой;
 - г) гармонической взвешенной.
16. Под выборочным наблюдением понимают:
- а) обследование наиболее крупных единиц изучаемой совокупности
 - б) сплошное наблюдение всех единиц совокупности
 - в) несплошное наблюдение части единиц совокупности, отобранных случайным способом
 - г) несплошное наблюдение части единиц совокупности
17. Выборочный метод наблюдения основан на:
- а) случайном отборе единиц совокупности;
 - б) обследовании самых существенных единиц совокупности;
 - в) обследовании отдельных единиц совокупности, обычно представителей каких-либо новых типов явлений;
 - г) изучении всех единиц совокупности.
18. Средняя ошибка выборки зависит от:
- а) доверительной вероятности утверждения;
 - б) вариации значений признаков выборочной совокупности;
 - в) значения модального интервала

Типовой комплект заданий для тестов

Полный комплект тестовых материалов по итоговому тестированию размещен на образовательном портале «АГАСУ»

Типовые тесты для итогового тестирования

Знать - ПК-4.1ИИП. 3-1

1. Укажите виды данных согласно их классификации по способу получения информации
 - a) собственные и приобретенные
 - b) «data-activ» и «dark-data»
 - c) сырые и производные
2. Выберите верные определения категории «большие данные»:
 - d) ряд распределения, состоящий из чисел, каждое из которых больше чем 1000
 - e) информационный массив, который невозможно обработать и хранить с помощью традиционных инструментов или на одном компьютере
 - f) методы анализа структурированных и неструктурированных данных для решения конкретных социально-экономических задач
 - g) категории вычислительных технологий для обработки, хранения и передачи больших наборов данных
3. Среди представленных видов источников больших данных отметьте те, которые относятся к классификации «по типам данных»
 - a) аудио-, видеозаписи
 - b) текстовые
 - c) изображения
 - d) генерируемые устройством
 - e) открытые
4. Характеристика больших данных Velocity (скорость накопления и обработки) заключается
 - a) в разнообразии типов получаемых данных
 - a) в формировании неструктурированных массивов информации
 - b) в возможности производить сбор и обработку данных в режиме реального времени
5. Первая модель для описания характеристик больших данных называлась
 - a) 7V
 - b) 12V
 - c) 3V
 - d) 5V
6. Какое соотношение «тестовая: обучающая» наиболее оптимально?
 - a) 20:80,
 - b) 50:50,
 - c) 25:75,
 - d) 5:95,
 - e) 40:30?
7. Какой код лучше использовать при загрузке данных из csv-файла?
 - a) `dataset = read_csv("data.csv")`
 - b) `dataset = import("data.csv")`
 - c) `dataset = read.csv("data.csv")`
 - d) `dataset = import.csv("data.csv")`
 - e) `dataset = read_xls("data.csv")`
8. Какая библиотека python предназначена для управления наборами данных:
 - a) `numpy`,
 - b) `pandas`,
 - c) `sklearn`,

- d) `opencv`,
 - e) `matplotlib`?
9. Какая стратегия является нежелательной при обработке пропусков в данных?
- a) замена пропущенных значений в столбце медианным значением по данному столбцу;
 - b) удаление строк, содержащих пропуски в данных;
 - c) замена пропущенных значений в столбце средним арифметическим значением по данному столбцу;
 - d) замена пропущенных значений в столбце наиболее часто встречающимся значением по данному столбцу;
10. Какой код лучше использовать при загрузке данных из csv-файла?
- a) `dataset = read_csv("data.csv")`
 - b) `dataset = import("data.csv")`
 - c) `dataset = read.csv("data.csv")`
 - d) `dataset = import.csv("data.csv")`
 - e) `dataset = read_xls("data.csv")`

Знать - ПК-4.1ИИП. 3-2

11. Выберите примеры генерации больших данных:
- a) записи видеорегистраторов ГИБДД о дорожно-транспортной ситуации
 - b) запись телефонных разговоров согласно Федеральному закону от 6 июля 2016 г. N 374-ФЗ "О внесении изменений в Федеральный закон "О противодействии терроризму" и отдельные законодательные акты Российской Федерации в части установления дополнительных мер противодействия терроризму и обеспечения общественной безопасности"
 - c) отслеживание маршрутов автомобильного транспорта системой ГЛОНАСС
 - d) сбор результатов контрольных работ школьным учителем
12. С позиции предприятия способы монетизации больших данных можно объединить в две группы:
- a) отраслевая и секторальная
 - b) краткосрочная и долгосрочная
 - c) внешняя и внутренняя
13. Направления оптимизации бизнеса, основанные на использовании больших данных:
- a) персонализированный маркетинг
 - b) скоринговые системы
 - c) реконструирование и ремонт офиса предприятия
 - d) снижение простоев грузового транспорта
14. В России большие данные наиболее активно аккумулируются, анализируются и используются в таких сферах деятельности как
- a) торговля
 - b) телекоммуникации
 - c) хореографическое искусство
 - d) рыболовческое хозяйство
 - e) банковский сектор
15. Участниками рынка больших данных, призванными помочь клиентам извлекать знания из накопленного массива больших данных, являются
- a) поставщики специальной инфраструктуры
 - b) датамайнеры или разработчики алгоритмов
 - c) IT-консультанты
16. Барьерами для российского рынка больших данных являются
- a) проблемы качества данных
 - b) нехватка опыта внедрений
 - c) высокая стоимость решений и отсутствие быстрых результатов

- d) защита персональных данных
 - e) невозможность обработки данных на одном компьютере
17. Основные направления монетизации больших данных
- a) Публикация данных (пресс-релизы, отчеты)
 - b) Сбор и продажа данных (социальные сети, поисковые системы)
 - c) Продукты, использующие решение на данных (беспилотные автомобили)
 - d) Извлечение пользы из данных (рекомендательная система)
18. Сбор и хранение неструктурированных разнообразных данных, возможность обработки сразу всех доступных данных в их исходном виде, анализ в реальном времени – это особенности
- a) работы с визуализированной информацией
 - b) работы с массивами больших данных
 - c) процесса работы с традиционными данными
19. Информационные массивы, собираемые и обрабатываемые согласно законодательно-нормативным актам, подлежащие размещению в интернете или свободной публикации, называются
- a) закрытые данные
 - b) поисковые системы
 - c) поисковые системы
20. Data Driven-подход подразумевает
- a) анализ клиентской базы и оптимизация персональных предложений
 - b) принятие решений основано на анализе данных, а не интуиции и личном опыте
 - c) повышение эффективности за счет снижения издержек и минимизации рисков