

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ
ФЕДЕРАЦИИ
федеральное государственное автономное образовательное учреждение высшего
образования
"САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
АЭРОКОСМИЧЕСКОГО ПРИБОРОСТРОЕНИЯ"

Кафедра № 82

УТВЕРЖДАЮ

Руководитель образовательной программы

д.э.н., проф.

(должность, уч. степень, звание)

А.С. Будагов

(инициалы, фамилия)

(подпись)

«27» июня 2024 г

Лист согласования рабочей программы дисциплины

Программу составил (а)

доц., к.т.н.

(должность, уч. степень, звание)

(подпись, дата)

В.С. Блюм

(инициалы, фамилия)

Программа одобрена на заседании кафедры № 82

«20» июня 2024 г, протокол № 11

Заведующий кафедрой № 82

д.э.н., проф.

(уч. степень, звание)

(подпись, дата)

А.С. Будагов

(инициалы, фамилия)

Заместитель директора института №8 по методической работе

доц., к.э.н., доц.

(должность, уч. степень, звание)

(подпись, дата)

Л.В. Рудакова

(инициалы, фамилия)

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

«Методы обработки больших данных»
(Наименование дисциплины)

Код направления подготовки/ специальности	09.03.03
Наименование направления подготовки/ специальности	Прикладная информатика
Наименование направленности	Интеллектуальные информационные системы и технологии
Форма обучения	очно-заочная
Год приема	2024

Аннотация

Дисциплина «Методы обработки больших данных» входит в образовательную программу высшего образования – программу бакалавриата по направлению подготовки/ специальности 09.03.03 «Прикладная информатика» направленности «Интеллектуальные информационные системы и технологии». Дисциплина реализуется кафедрой «№82».

Дисциплина нацелена на формирование у выпускника следующих компетенций:

УК-1 «Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач»

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»

ПК-7 «Способность проводить анализ больших данных с использованием существующей в организации методологической и технологической инфраструктуры»

Содержание дисциплины охватывает круг вопросов, связанных с методами решения задач обработки и анализа больших данных, принципы построения современных систем хранения данных, возможности высокопроизводительных вычислительных систем, технологии распределенных вычислений, методы и модели Data Mining.

Преподавание дисциплины предусматривает следующие формы организации учебного процесса: лекции, лабораторные работы и самостоятельную работу.

Программой дисциплины предусмотрены следующие виды контроля: текущий контроль успеваемости, промежуточная аттестация в форме зачета.

Общая трудоемкость освоения дисциплины составляет 3 зачетных единицы, 108 часов.

Язык обучения по дисциплине «русский»

1. Перечень планируемых результатов обучения по дисциплине

1.1. Цели преподавания дисциплины

Формирование у студентов необходимой теоретической базы и практических навыков, которые позволят всесторонне и системно понимать современные проблемы прикладной математики и информатики, проблемы обработки и анализа информации, а также разрабатывать и анализировать концептуальные и теоретические модели при решении научных и прикладных задач в области информационных технологий. Одна из главных проблем современной обработки и анализа данных - рост объемов данных, поэтому вопросам обработки большого объема данных посвящена данная дисциплина. Главная задача курса - сформировать целостное представление о современных проблемах анализа и обработки больших данных, помочь овладеть опытом разработки и анализа концептуальных и теоретических моделей прикладных задач анализа больших данных с применением моделей Data Mining.

1.2. Дисциплина входит в состав части, формируемой участниками образовательных отношений, образовательной программы высшего образования (далее – ОП ВО).

1.3. Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения ОП ВО.

В результате изучения дисциплины обучающийся должен обладать следующими компетенциями или их частями. Компетенции и индикаторы их достижения приведены в таблице 1.

Таблица 1 – Перечень компетенций и индикаторов их достижения

Категория (группа) компетенции	Код и наименование компетенции	Код и наименование индикатора достижения компетенции
Универсальные компетенции	УК-1 Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач	УК-1.3.1 знать методики поиска, сбора и обработки информации, в том числе с использованием информационных технологий, включая интеллектуальные УК-1.У.1 уметь применять методики поиска, сбора и обработки информации, в том числе с использованием искусственного интеллекта УК-1.В.1 владеть навыками критического анализа и синтеза информации, в том числе с помощью цифровых инструментов
Универсальные компетенции	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-2.3.3 знать возможности и ограничения применения цифровых инструментов для решения поставленных задач
Профессиональные компетенции	ПК-7 Способность проводить анализ больших данных с	ПК-7.3.3 знать современные методы и инструментальные средства анализа больших данных

	использованием существующей в организации методологической и технологической инфраструктуры	ПК-7.У.3 уметь формировать предложения по использованию результатов анализа ПК-7.В.3 владеть навыками мониторинга эффективности работы аналитики больших данных
--	---	--

2. Место дисциплины в структуре ОП

Дисциплина может базироваться на знаниях, ранее приобретенных обучающимися при изучении следующих дисциплин:

- « Основы интеллектуального анализа».
- «Технологии программирования».
- « Операционные системы».

Знания, полученные при изучении материала данной дисциплины, имеют как самостоятельное значение, так и могут использоваться при изучении других дисциплин:

- «Интеллектуальные информационные системы»,
- «Интеллектуальный анализ данных»
- «Методы искусственного интеллекта».

3. Объем и трудоемкость дисциплины

Данные об общем объеме дисциплины, трудоемкости отдельных видов учебной работы по дисциплине (и распределение этой трудоемкости по семестрам) представлены в таблице 2.

Таблица 2 – Объем и трудоемкость дисциплины

Вид учебной работы	Всего	Трудоемкость по семестрам
		№9
1	2	3
Общая трудоемкость дисциплины, ЗЕ/ (час)	3/ 108	3/ 108
Из них часов практической подготовки	5	5
Аудиторные занятия, всего час.	34	34
в том числе:		
лекции (Л), (час)	17	17
практические/семинарские занятия (ПЗ), (час)		
лабораторные работы (ЛР), (час)	17	17
курсовой проект (работа) (КП, КР), (час)		
экзамен, (час)		
Самостоятельная работа, всего (час)	74	74
Вид промежуточной аттестации: зачет, дифф. зачет, экзамен (Зачет, Дифф. зач, Экз.**)	Зачет	Зачет

Примечание: ** кандидатский экзамен

4. Содержание дисциплины

4.1. Распределение трудоемкости дисциплины по разделам и видам занятий.

Разделы, темы дисциплины и их трудоемкость приведены в таблице 3.

Таблица 3 – Разделы, темы дисциплины, их трудоемкость

Разделы, темы дисциплины	Лекции (час)	ПЗ (СЗ) (час)	ЛР (час)	КП (час)	СРС (час)
Семестр 9					
Раздел 1. Технологии анализа данных Тема 1.1. Большие данные (Big Data): современные подходы к обработке и хранению. Проблема множественного сравнения данных.	2		2		10
Тема 1.2. Процесс анализа. Общая схема анализа. Извлечение и визуализация данных. Этапы моделирования. Процесс построения моделей. Формы представления данных, типы и виды данных. Представления наборов данных.	3		3		10
Раздел 2. Методы многомерного статистического анализа и анализа нечисловой информации Тема 2.1. Начала корреляционного анализа и регрессионного анализа	4		4		20
Раздел 3. Программные средства для обработки данных и системы Data Mining Тема 3.1. Программные инструменты интеллектуального анализа данных на примере пакета RapidMiner	4		4		20
Раздел 4. Аналитика в больших данных Тема 4.1. Обзор задач и методов машинного обучения и интеллектуального анализа больших данных Тема 4.2. Интерактивное исследование данных	4		4		14
Итого в семестре:	17		17		74
Итого	17	0	17	0	74

Практическая подготовка заключается в непосредственном выполнении обучающимися определенных трудовых функций, связанных с будущей профессиональной деятельностью.

4.2. Содержание разделов и тем лекционных занятий.

Содержание разделов и тем лекционных занятий приведено в таблице 4.

Таблица 4 – Содержание разделов и тем лекционного цикла

Номер раздела	Название и содержание разделов и тем лекционных занятий
1	Раздел 1. Технологии анализа данных Тема 1.1. Большие данные (Big Data): современные подходы к обработке и хранению. Проблема множественного сравнения данных. Тема 1.2. Процесс анализа. Общая схема анализа. Извлечение и визуализация данных. Этапы моделирования. Процесс построения моделей. Формы представления данных, типы и виды данных. Представления наборов данных.
2	Раздел 2. Методы многомерного статистического анализа и

	анализа нечисловой информации Тема 2.1. Начала корреляционного анализа и регрессионного анализа
3	Раздел 3. Программные средства для обработки данных и системы Data Mining Тема 3.1. Программные инструменты интеллектуального анализа данных на примере пакета RapidMiner
4	Раздел 4. Аналитика в больших данных Тема 4.1. Обзор задач и методов машинного обучения и интеллектуального анализа больших данных Тема 4.2. Интерактивное исследование данных

4.3. Практические (семинарские) занятия

Темы практических занятий и их трудоемкость приведены в таблице 5.

Таблица 5 – Практические занятия и их трудоемкость

№ п/п	Темы практических занятий	Формы практических занятий	Трудоемкость, (час)	Из них практической подготовки, (час)	№ раздела дисциплины
Учебным планом не предусмотрено					
Всего					

4.4. Лабораторные занятия

Темы лабораторных занятий и их трудоемкость приведены в таблице 6.

Таблица 6 – Лабораторные занятия и их трудоемкость

№ п/п	Наименование лабораторных работ	Трудоемкость, (час)	Из них практической подготовки, (час)	№ раздела дисциплины
Семестр 9				
1	Начальное знакомство со средой обработки больших данных. Методы подготовки исходных данных	2		1
2	Исследование операций преобразования данных.	3		2
3	Метод деревьев принятия решений для анализа больших данных	4		3
4	Методы регрессии для анализа больших данных.	4		3
5	Методы нейронных сетей для анализа больших данных	4		3
Всего		17		

4.5. Курсовое проектирование/ выполнение курсовой работы

Учебным планом не предусмотрено

4.6. Самостоятельная работа обучающихся

Виды самостоятельной работы и ее трудоемкость приведены в таблице 7.

Таблица 7 – Виды самостоятельной работы и ее трудоемкость

Вид самостоятельной работы	Всего, час	Семестр 9, час
1	2	3
Изучение теоретического материала дисциплины (ТО)	40	40
Курсовое проектирование (КП, КР)		
Расчетно-графические задания (РГЗ)		
Выполнение реферата (Р)	20	20
Подготовка к текущему контролю успеваемости (ТКУ)	4	4
Домашнее задание (ДЗ)		
Контрольные работы заочников (КРЗ)		
Подготовка к промежуточной аттестации (ПА)	10	10
Всего:	74	74

5. Перечень учебно-методического обеспечения

для самостоятельной работы обучающихся по дисциплине (модулю)

Учебно-методические материалы для самостоятельной работы обучающихся указаны в п.п. 7-11.

6. Перечень печатных и электронных учебных изданий

Перечень печатных и электронных учебных изданий приведен в таблице 8.

Таблица 8– Перечень печатных и электронных учебных изданий

Шифр/ URL адрес	Библиографическая ссылка	Количество экземпляров в библиотеке (кроме электронных экземпляров)
DOI: 10.1016/j.procs.2014.05.260	Chernyak E. L., Mirkin B. G. "Refining a Taxonomy by Using Annotated Suffix Trees and Wikipedia Resources" - 2016	20
https://vk.com/wall-54530371_712	Маннинг К.Д., Рагхаван П., Шютце Х. "Введение в информационный поиск" (2017, DjVu)	20

7. Перечень электронных образовательных ресурсов

информационно-телекоммуникационной сети «Интернет»

Перечень электронных образовательных ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины приведен в таблице 9.

Таблица 9 – Перечень электронных образовательных ресурсов информационно-телекоммуникационной сети «Интернет»

URL адрес	Наименование
https://www.coursera.org/	1. Курсы и материалы на платформе Coursera по

	информационным технологиям и поисковым системам.
https://www.udemy.com/	2. Онлайн-курсы на платформе Udey, специализированные на изучении информационно-поисковых систем.
https://openedu.ru/	3. Различные академические ресурсы, такие как статьи и конференции по информационно-поисковым системам, доступные на сайтах различных университетов.

8. Перечень информационных технологий

8.1. Перечень программного обеспечения, используемого при осуществлении образовательного процесса по дисциплине.

Перечень используемого программного обеспечения представлен в таблице 10.

Таблица 10– Перечень программного обеспечения

№ п/п	Наименование
	Не предусмотрено

8.2. Перечень информационно-справочных систем, используемых при осуществлении образовательного процесса по дисциплине

Перечень используемых информационно-справочных систем представлен в таблице 11.

Таблица 11– Перечень информационно-справочных систем

№ п/п	Наименование
	Не предусмотрено

9. Материально-техническая база

Состав материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине, представлен в таблице 12.

Таблица 12 – Состав материально-технической базы

№ п/п	Наименование составной части материально-технической базы	Номер аудитории (при необходимости)
1	Дисплейный класс	14-06 – 14-11 Ленс.
2	Мультимедийная лекционная аудитория	

10. Оценочные средства для проведения промежуточной аттестации

10.1. Состав оценочных средств для проведения промежуточной аттестации обучающихся по дисциплине приведен в таблице 13.

Таблица 13 – Состав оценочных средств для проведения промежуточной аттестации

Вид промежуточной аттестации	Перечень оценочных средств
Зачет	Список вопросов; Тесты; Задачи.

10.2. В качестве критериев оценки уровня сформированности (освоения) компетенций обучающимися применяется 5-балльная шкала оценки сформированности

компетенций, которая приведена в таблице 14. В течение семестра может использоваться 100-балльная шкала модульно-рейтинговой системы Университета, правила использования которой, установлены соответствующим локальным нормативным актом ГУАП.

Таблица 14 – Критерии оценки уровня сформированности компетенций

Оценка компетенции	Характеристика сформированных компетенций
5-балльная шкала	
«отлично» «зачтено»	<ul style="list-style-type: none"> – обучающийся глубоко и всесторонне усвоил программный материал; – уверенно, логично, последовательно и грамотно его излагает; – опираясь на знания основной и дополнительной литературы, тесно привязывает усвоенные научные положения с практической деятельностью направления; – умело обосновывает и аргументирует выдвигаемые им идеи; – делает выводы и обобщения; – свободно владеет системой специализированных понятий.
«хорошо» «зачтено»	<ul style="list-style-type: none"> – обучающийся твердо усвоил программный материал, грамотно и по существу излагает его, опираясь на знания основной литературы; – не допускает существенных неточностей; – увязывает усвоенные знания с практической деятельностью направления; – аргументирует научные положения; – делает выводы и обобщения; – владеет системой специализированных понятий.
«удовлетворительно» «зачтено»	<ul style="list-style-type: none"> – обучающийся усвоил только основной программный материал, по существу излагает его, опираясь на знания только основной литературы; – допускает несущественные ошибки и неточности; – испытывает затруднения в практическом применении знаний направления; – слабо аргументирует научные положения; – затрудняется в формулировании выводов и обобщений; – частично владеет системой специализированных понятий.
«неудовлетворительно» «не зачтено»	<ul style="list-style-type: none"> – обучающийся не усвоил значительной части программного материала; – допускает существенные ошибки и неточности при рассмотрении проблем в конкретном направлении; – испытывает трудности в практическом применении знаний; – не может аргументировать научные положения; – не формулирует выводов и обобщений.

10.3. Типовые контрольные задания или иные материалы.

Вопросы (задачи) для экзамена представлены в таблице 15.

Таблица 15 – Вопросы (задачи) для экзамена

№ п/п	Перечень вопросов (задач) для экзамена	Код индикатора
	Учебным планом не предусмотрено	

Вопросы (задачи) для зачета / дифф. зачета представлены в таблице 16.

Таблица 16 – Вопросы (задачи) для зачета / дифф. зачета

№ п/п	Перечень вопросов (задач) для зачета / дифф. зачета	Код индикатора
1	История возникновения проблемы «больших данных»	УК-1.3.1
2	Причины возникновения проблемы «больших данных» и условия её решения	УК-1.У.1
3	Характеристики больших данных	УК-1.В.1

4	Сравнительные характеристики обычных и больших данных	УК-2.3.3
5	Методика A/B testing.	ПК-7.3.3
6	Методика Ad-hoc GRID	ПК-7.У.3
7	Методики Association rule learning	ПК-7.В.3
8	Методика BOINC-грид	УК-1.3.1
9	Технология Data GRID	УК-1.У.1
10	Методики Data fusion and data integration	УК-1.В.1
11	Когнитивная кластеризация	УК-2.3.3
12	Метод Ensemble learning.	ПК-7.3.3
13	Методика Genetic algorithms	ПК-7.У.3
14	Инфраструктура ГРИД	ПК-7.В.3
15	Вычислительная система со множественным потоком команд и множественным потоком данных	УК-1.3.1
16	Методики распознавания естественного языка человека.	УК-1.У.1
17	Методики анализа связей между узлами в сетях.	УК-1.В.1
18	Численные методы для редизайна сложных систем и процессов	УК-2.3.3
19	Набор методик с элементами самообучения для предсказания поведенческой модели потребителей.	ПК-7.3.3
20	Набор статистических методов для выявления закономерности	ПК-7.У.3
21	Методики оценки настроений потребителей на основе технологии распознавания естественного языка человека	ПК-7.В.3
22	Методики распознавания сигнала на фоне шума и его дальнейшего анализа.	УК-1.3.1
23	Методики на основе технологий машинного обучения, позволяющие выявить функциональные взаимосвязи в анализируемых массивах данных.	УК-1.У.1
24	Методика темпоральной логики анализа повторяющихся с течением времени последовательностей данных.	УК-1.В.1
25	Методы графического представления результатов интеллектуального анализа данных в виде диаграмм или анимированных изображений	УК-2.3.3

Перечень тем для курсового проектирования/выполнения курсовой работы представлены в таблице 17.

Таблица 17 – Перечень тем для курсового проектирования/выполнения курсовой работы

№ п/п	Примерный перечень тем для курсового проектирования/выполнения курсовой работы
	Учебным планом не предусмотрено

Вопросы для проведения промежуточной аттестации в виде тестирования представлены в таблице 18.

Таблица 18 – Примерный перечень вопросов для тестов

№ п/п	Примерный перечень вопросов для тестов	Код индикатора
1	Установите соответствие между методами анализа данных и их описаниями	УК-1.3.1
	Методы анализа данных	Описание
	А. Регрессионный анализ	1. Метод, позволяющий выявить зависимости между переменными.
	В. Кластерный анализ	2. Метод, используемый для группировки объектов по схожести.

	C. Деревья решений	3. Метод, позволяющий визуализировать процесс принятия решений.		
	D. Временные ряды	4. Метод анализа данных, основанный на временных последовательностях.		
2	Установите правильную последовательность этапов обработки больших данных. 1. Сбор данных 2. Хранение данных 3. Предварительная обработка 4. Анализ данных 5. Визуализация результатов		УК-1.3.1	
3	Какой из следующих методов является наиболее подходящим для выявления скрытых паттернов в больших объемах данных? 1. Регрессионный анализ 2. Кластерный анализ 3. Деревья решений 4. Линейная алгебра		УК-1.3.1	
4	Какие из следующих методов можно использовать для обработки больших данных? (Выберите все подходящие варианты) 1. Обработка естественного языка 2. Модели машинного обучения 3. Статистический анализ 4. Ручной ввод данных		УК-1.3.1	
5	Опишите, как системный подход может быть применен для решения задач в области обработки больших данных.		УК-1.3.1	
6	Установите соответствие между методами поиска информации и их характеристиками:		УК-1.У.1	
	Методы поиска	Характеристики		
	1. Поисковые системы	A. Содержат структурированную информацию по конкретным темам		
2. Библиографические базы данных	B. Позволяют находить информацию по ключевым словам			
3. Тематические порталы	C. Предоставляют доступ к библиографическим описаниям публикаций			

	4. Социальные сети	D. Используются для поиска информации в неструктурированных источниках		
7	Расставьте этапы обработки больших данных в правильной последовательности: 1. Сбор данных 2. Очистка и преобразование данных 3. Анализ данных 4. Визуализация результатов 5. Хранение данных		УК-1.У.1	
8	Какой из методов машинного обучения наиболее подходит для прогнозирования временных рядов? 1. Линейная регрессия 2. Логистическая регрессия 3. Дерево решений 4. Кластеризация		УК-1.У.1	
9	Какие методы анализа данных наиболее эффективны для выявления скрытых закономерностей в больших объемах информации?		УК-1.У.1	
10	1. Ассоциативные правила 2. Кластерный анализ 3. Линейная регрессия 4. Нейронные сети 5. Деревья решений Опишите основные этапы процесса обработки больших данных с использованием инструментов искусственного интеллекта.		УК-1.У.1	
11	Установите соответствие между методами поиска информации и их описанием:		УК-1.В.1	
	Методы поиска информации	Описание		
	1. Библиографический поиск	А. Использование ключевых слов для нахождения информации в базах данных		
	2. Интернет-поиск	В. Поиск информации в научных статьях и книгах		

	3. Анкетирование	С. Сбор данных через опросы и анкеты										
	4. Интервью	D. Личное общение с экспертами для получения информации										
12	<p>Установите правильную последовательность этапов критического анализа информации</p> <ol style="list-style-type: none"> 1. : Формулирование вопросов 2. Сбор данных 3. Оценка источников 4. Синтез информации 		УК-1.B.1									
13	<p>Какой из следующих методов является наиболее эффективным для системного анализа информации?</p> <ol style="list-style-type: none"> 1. Сравнительный анализ 2. Статистический анализ 3. SWOT-анализ 4. Метод мозгового штурма 		УК-1.B.1									
14	<p>Какие из следующих методов могут быть использованы для критического анализа информации? (Выберите несколько)</p> <ol style="list-style-type: none"> 1. Анализ данных 2. Системный подход 3. Эмпирическое исследование 4. Логический анализ 		УК-1.B.1									
15	<p>Опишите, как можно использовать системный подход для решения задачи поиска информации</p>		УК-1.B.1									
16	<p>Установите соответствие между задачами и оптимальными способами их решения.</p> <table border="1"> <thead> <tr> <th>Методы поиска информации</th> <th>Описание</th> </tr> </thead> <tbody> <tr> <td>1. Библиографический поиск</td> <td>A. Использование ключевых слов для поиска в базах данных</td> </tr> <tr> <td>2. Анкетирование</td> <td>B. Сбор данных через опросы</td> </tr> <tr> <td>3. Контент-анализ</td> <td>C. Изучение текстов для выявления тем</td> </tr> </tbody> </table>		Методы поиска информации	Описание	1. Библиографический поиск	A. Использование ключевых слов для поиска в базах данных	2. Анкетирование	B. Сбор данных через опросы	3. Контент-анализ	C. Изучение текстов для выявления тем	УК-2.3.3	
Методы поиска информации	Описание											
1. Библиографический поиск	A. Использование ключевых слов для поиска в базах данных											
2. Анкетирование	B. Сбор данных через опросы											
3. Контент-анализ	C. Изучение текстов для выявления тем											

	4. Интервью	D. Личное общение для получения информации												
17	<p>Установите правильную последовательность этапов обработки больших данных</p> <ol style="list-style-type: none"> Сбор данных Оценка достоверности Синтез информации Формулирование выводов 		УК-2.3.3											
18	<p>Какой из следующих методов является наиболее эффективным для анализа больших объемов данных?</p> <ol style="list-style-type: none"> Библиографический поиск Мозговой штурм SWOT-анализ Анкетирование 		УК-2.3.3											
19	<p>Какие из следующих методов способствуют саморазвитию? (выберите все подходящие)</p> <ol style="list-style-type: none"> Сравнительный анализ Статистический анализ Мозговой штурм Контент-анализ 		УК-2.3.3											
20	<p>Опишите, как вы бы определили круг задач для проекта по обработке больших данных с учетом действующих правовых норм и имеющихся ресурсов.</p>		УК-2.3.3											
21	<p>Установите соответствие между методами обработки больших данных и их описанием.</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="width: 50%; text-align: center;">Задача</th> <th style="width: 50%; text-align: center;">Способ решения</th> </tr> </thead> <tbody> <tr> <td>A. Анализ больших данных в маркетинге</td> <td>1. Использование нейронных сетей</td> </tr> <tr> <td>B. Прогнозирование спроса на товар</td> <td>2. Метод предиктивной аналитики</td> </tr> <tr> <td>C. Обработка медицинских данных</td> <td>3. Алгоритмы машинного обучения</td> </tr> <tr> <td>D. Оптимизация логистических процессов</td> <td>4. Алгоритм MapReduce</td> </tr> </tbody> </table>	Задача	Способ решения	A. Анализ больших данных в маркетинге	1. Использование нейронных сетей	B. Прогнозирование спроса на товар	2. Метод предиктивной аналитики	C. Обработка медицинских данных	3. Алгоритмы машинного обучения	D. Оптимизация логистических процессов	4. Алгоритм MapReduce		ПК-7.3.3	
Задача	Способ решения													
A. Анализ больших данных в маркетинге	1. Использование нейронных сетей													
B. Прогнозирование спроса на товар	2. Метод предиктивной аналитики													
C. Обработка медицинских данных	3. Алгоритмы машинного обучения													
D. Оптимизация логистических процессов	4. Алгоритм MapReduce													
22	<p>Установите правильную последовательность этапов анализа больших данных.</p> <ol style="list-style-type: none"> Сбор данных Хранение данных 		ПК-7.3.3											

	<ol style="list-style-type: none"> 3. <input type="text" value="Обработка данных"/> 4. <input type="text" value="Анализ данных"/> 5. <input type="text" value="Визуализация данных"/> 											
23	<p>Какой из перечисленных методов является основным в обработке больших данных?</p> <ol style="list-style-type: none"> 1. <input type="text" value="Ручной анализ"/> 2. <input type="text" value="Статистические методы"/> 3. <input type="text" value="Машинное обучение"/> 4. <input type="text" value="Опросы"/> 	ПК-7.3.3										
24	<p>Какие из следующих методов используются для анализа больших данных? (Выберите все подходящие варианты.)</p> <ol style="list-style-type: none"> 1. <input type="text" value="Чтение книг"/> 2. <input type="text" value="Посещение тренингов"/> 3. <input type="text" value="Общение с единомышленниками"/> 4. <input type="text" value="Игнорирование обратной связи"/> 	ПК-7.3.3										
25	Опишите основные этапы анализа больших данных и их значение в принятии решений.	ПК-7.3.3										
26	<p>Установите соответствие между методами анализа больших данных и их описаниями</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="width: 50%; text-align: center;">Методы обработки</th> <th style="width: 50%; text-align: center;">Описание</th> </tr> </thead> <tbody> <tr> <td>1. Машинное обучение</td> <td>А. Алгоритмы, позволяющие находить закономерности в данных.</td> </tr> <tr> <td>2. Интеллектуальный анализ данных</td> <td>В. Процесс извлечения полезной информации из больших объемов данных.</td> </tr> <tr> <td>3. Нейронные сети</td> <td>С. Модели, имитирующие работу человеческого мозга для решения задач.</td> </tr> <tr> <td>4. MapReduce</td> <td>Д. Алгоритм для обработки больших данных в распределенной среде.</td> </tr> </tbody> </table>	Методы обработки	Описание	1. Машинное обучение	А. Алгоритмы, позволяющие находить закономерности в данных.	2. Интеллектуальный анализ данных	В. Процесс извлечения полезной информации из больших объемов данных.	3. Нейронные сети	С. Модели, имитирующие работу человеческого мозга для решения задач.	4. MapReduce	Д. Алгоритм для обработки больших данных в распределенной среде.	ПК-7.У.3
Методы обработки	Описание											
1. Машинное обучение	А. Алгоритмы, позволяющие находить закономерности в данных.											
2. Интеллектуальный анализ данных	В. Процесс извлечения полезной информации из больших объемов данных.											
3. Нейронные сети	С. Модели, имитирующие работу человеческого мозга для решения задач.											
4. MapReduce	Д. Алгоритм для обработки больших данных в распределенной среде.											
27	<p>Установите правильную последовательность этапов анализа больших данных.</p> <ol style="list-style-type: none"> 1. <input type="text" value="Сбор данных"/> 2. <input type="text" value="Обработка данных"/> 3. <input type="text" value="Анализ данных"/> 4. <input type="text" value="Визуализация результатов"/> 	ПК-7.У.3										
28	Какой из следующих методов является наиболее подходящим для обработки больших объемов неструктурированных данных?	ПК-7.У.3										

	<ol style="list-style-type: none"> 1. Ручной анализ 2. Машинное обучение 3. Статистический анализ 4. Опросы пользователей 												
29	<p>Какие из следующих методов могут быть использованы для анализа больших данных? (Выберите все подходящие варианты)</p> <ol style="list-style-type: none"> 1. Нейронные сети 2. Классические статистические методы 3. Интеллектуальный анализ данных 4. Ручной ввод данных 	ПК-7.У.3											
30	<p>Опишите, как вы можете использовать результаты анализа больших данных для улучшения бизнес-процессов в организации.</p>	ПК-7.У.3											
31	<p>Установите соответствие между методами анализа больших данных и их описаниями.</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="width: 50%; text-align: center;">Методы анализа</th> <th style="width: 50%; text-align: center;">Описание</th> </tr> </thead> <tbody> <tr> <td>1. Машинное обучение</td> <td>А. Используется для прогнозирования и классификации на основе данных.</td> </tr> <tr> <td>2. Анализ временных рядов</td> <td>В. Применяется для анализа данных, собранных в последовательности во времени.</td> </tr> <tr> <td>3. Кластеризация</td> <td>С. Группирует данные по схожести.</td> </tr> <tr> <td>4. Регрессионный анализ</td> <td>Д. Оценивает зависимость одной переменной от другой.</td> </tr> </tbody> </table>		Методы анализа	Описание	1. Машинное обучение	А. Используется для прогнозирования и классификации на основе данных.	2. Анализ временных рядов	В. Применяется для анализа данных, собранных в последовательности во времени.	3. Кластеризация	С. Группирует данные по схожести.	4. Регрессионный анализ	Д. Оценивает зависимость одной переменной от другой.	ПК-7.В.3
	Методы анализа	Описание											
	1. Машинное обучение	А. Используется для прогнозирования и классификации на основе данных.											
	2. Анализ временных рядов	В. Применяется для анализа данных, собранных в последовательности во времени.											
	3. Кластеризация	С. Группирует данные по схожести.											
4. Регрессионный анализ	Д. Оценивает зависимость одной переменной от другой.												
32	<p>Установите правильную последовательность этапов анализа больших данных</p> <ol style="list-style-type: none"> 1. Сбор данных 2. Обработка данных 3. Анализ данных 4. Интерпретация результатов 	ПК-7.В.3											
33	<p>Какой из следующих методов является наиболее подходящим для предсказания будущих значений на основе исторических данных?</p> <ol style="list-style-type: none"> 1. Регрессионный анализ 2. Машинное обучение 3. Кластеризация 4. Анализ временных рядов 	ПК-7.В.3											

34	<p>Какие из следующих методов могут быть использованы для анализа больших данных? (Выберите все подходящие варианты)</p> <ol style="list-style-type: none"> 1. Деревья решений 2. Статистический анализ 3. Системы управления базами данных 4. Глубокое обучение 	ПК-7.В.3
35	Опишите основные этапы процесса анализа больших данных.	ПК-7.В.3

Полный текст вопросов и ключи правильных ответов размещены в приложении к РПД.

Ответ на все тесты считается отличным, если количество баллов находится в диапазоне 27 и выше, хорошим 23-26, удовлетворительным 19-22, неудовлетворительным ниже 19.

Перечень тем контрольных работ по дисциплине обучающихся заочной формы обучения, представлены в таблице 19.

Таблица 19 – Перечень контрольных работ

№ п/п	Перечень контрольных работ
	Не предусмотрено

10.4. Методические материалы, определяющие процедуры оценивания индикаторов, характеризующих этапы формирования компетенций, содержатся в локальных нормативных актах ГУАП, регламентирующих порядок и процедуру проведения текущего контроля успеваемости и промежуточной аттестации обучающихся ГУАП.

11. Методические указания для обучающихся по освоению дисциплины

11.1. Методические указания для обучающихся по освоению лекционного материала.

Основное назначение лекционного материала – логически стройное, системное, глубокое и ясное изложение учебного материала. Назначение современной лекции в рамках дисциплины не в том, чтобы получить всю информацию по теме, а в освоении фундаментальных проблем дисциплины, методов научного познания, новейших достижений научной мысли. В учебном процессе лекция выполняет методологическую, организационную и информационную функции. Лекция раскрывает понятийный аппарат конкретной области знания, её проблемы, дает цельное представление о дисциплине, показывает взаимосвязь с другими дисциплинами.

Планируемые результаты при освоении обучающимися лекционного материала:

- получение современных, целостных, взаимосвязанных знаний, уровень которых определяется целевой установкой к каждой конкретной теме;
- получение опыта творческой работы совместно с преподавателем;
- развитие профессионально-деловых качеств, любви к предмету и самостоятельного творческого мышления.
- появление необходимого интереса, необходимого для самостоятельной работы;
- получение знаний о современном уровне развития науки и техники и о прогнозе их развития на ближайшие годы;

– научиться методически обрабатывать материал (выделять главные мысли и положения, приходить к конкретным выводам, повторять их в различных формулировках);

– получение точного понимания всех необходимых терминов и понятий.

Лекционный материал может сопровождаться демонстрацией слайдов и использованием раздаточного материала при проведении коротких дискуссий об особенностях применения отдельных тематик по дисциплине.

Структура предоставления лекционного материала:

- Лекционный материал представлен в форме презентаций пяти лекций, размещенных в курсе «Методы обработки больших данных» в системе дистанционного обучения ЛМС, <https://lms.guap.ru/new/course/view.php?id=32>
- Дополнительный материал к каждой лекции, представленных в форме файлов Word, размещенных в курсе «Методы обработки больших данных» в системе дистанционного обучения ЛМС.

11.2. Методические указания для обучающихся по участию в семинарах *(если предусмотрено учебным планом по данной дисциплине)*

Основной целью для обучающегося является систематизация и обобщение знаний по изучаемой теме, разделу, формирование умения работать с дополнительными источниками информации, сопоставлять и сравнивать точки зрения, конспектировать прочитанное, высказывать свою точку зрения и т.п. В соответствии с ведущей дидактической целью содержанием семинарских занятий являются узловое, наиболее трудные для понимания и усвоения темы, разделы дисциплины. Спецификой данной формы занятий является совместная работа преподавателя и обучающегося над решением поставленной проблемы, а поиск верного ответа строится на основе чередования индивидуальной и коллективной деятельности.

При подготовке к семинарскому занятию по теме прослушанной лекции необходимо ознакомиться с планом его проведения, с литературой и научными публикациями по теме семинара.

11.3. Методические указания для обучающихся по прохождению практических занятий *(если предусмотрено учебным планом по данной дисциплине)*

Практическое занятие является одной из основных форм организации учебного процесса, заключающаяся в выполнении обучающимися под руководством преподавателя комплекса учебных заданий с целью усвоения научно-теоретических основ учебной дисциплины, приобретения умений и навыков, опыта творческой деятельности.

Целью практического занятия для обучающегося является привитие обучающимся умений и навыков практической деятельности по изучаемой дисциплине.

Планируемые результаты при освоении обучающимся практических занятий:

– закрепление, углубление, расширение и детализация знаний при решении конкретных задач;

– развитие познавательных способностей, самостоятельности мышления, творческой активности;

– овладение новыми методами и методиками изучения конкретной учебной дисциплины;

– выработка способности логического осмысления полученных знаний для выполнения заданий;

– обеспечение рационального сочетания коллективной и индивидуальной форм обучения.

11.4. Методические указания для обучающихся по выполнению лабораторных работ

В ходе выполнения лабораторных работ обучающийся должен углубить и закрепить знания, практические навыки, овладеть современной методикой и техникой эксперимента в соответствии с квалификационной характеристикой обучающегося. Выполнение лабораторных работ состоит из экспериментально-практической, расчетно-аналитической частей и контрольных мероприятий.

Выполнение лабораторных работ обучающимся является неотъемлемой частью изучения дисциплины, определяемой учебным планом, и относится к средствам, обеспечивающим решение следующих основных задач обучающегося:

- приобретение навыков исследования процессов, явлений и объектов, изучаемых в рамках данной дисциплины;
- закрепление, развитие и детализация теоретических знаний, полученных на лекциях;
- получение новой информации по изучаемой дисциплине;
- приобретение навыков самостоятельной работы с лабораторным оборудованием и приборами.

Задание и требования к проведению лабораторных работ

- Представлены на странице курса в <https://lms.guap.ru/new/course/view.php?id=32>

Структура и форма отчета о лабораторной работе

- Представлены на странице курса в <https://lms.guap.ru/new/course/view.php?id=32>

Требования к оформлению отчета о лабораторной работе

Представлены на странице курса в <https://lms.guap.ru/new/course/view.php?id=32>

11.5. Методические указания для обучающихся по прохождению курсового проектирования/выполнения курсовой работы (если предусмотрено учебным планом по данной дисциплине)

Курсовой проект/ работа проводится с целью формирования у обучающихся опыта комплексного решения конкретных задач профессиональной деятельности.

11.6. Методические указания для обучающихся по прохождению самостоятельной работы

В ходе выполнения самостоятельной работы, обучающийся выполняет работу по заданию и при методическом руководстве преподавателя, но без его непосредственного участия.

Для обучающихся по заочной форме обучения, самостоятельная работа может включать в себя контрольную работу.

В процессе выполнения самостоятельной работы, у обучающегося формируется целесообразное планирование рабочего времени, которое позволяет им развивать умения и навыки в усвоении и систематизации приобретаемых знаний, обеспечивает высокий уровень успеваемости в период обучения, помогает получить навыки повышения профессионального уровня.

Методическими материалами, направляющими самостоятельную работу обучающихся являются:

- учебно-методический материал по дисциплине представлен на странице курса в <https://lms.guap.ru/new/course/view.php?id=32>

11.7. Методические указания для обучающихся по прохождению текущего контроля успеваемости.

Текущий контроль успеваемости предусматривает контроль качества знаний обучающихся, осуществляемого в течение семестра с целью оценивания хода освоения дисциплины.

Для проведения текущего контроля успеваемости на каждом лекционном занятии выделяется до 10 минут, а также контроль успеваемости проводится при приеме реферата и лабораторных работ. Результаты текущего контроля успеваемости учитываются при проведении промежуточной аттестации.

11.8. Методические указания для обучающихся по прохождению промежуточной аттестации.

Промежуточная аттестация обучающихся предусматривает оценивание промежуточных и окончательных результатов обучения по дисциплине. Она включает в себя:

– зачет – это форма оценки знаний, полученных обучающимся в ходе изучения учебной дисциплины в целом или промежуточная (по окончании семестра) оценка знаний обучающимся по отдельным разделам дисциплины с аттестационной оценкой «зачтено» или «не зачтено».

Лист внесения изменений в рабочую программу дисциплины

Дата внесения изменений и дополнений. Подпись внесшего изменения	Содержание изменений и дополнений	Дата и № протокола заседания кафедры	Подпись зав. кафедрой