

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное
образовательное учреждение высшего образования
«Астраханский государственный университет имени В. Н. Татищева»
(Астраханский государственный университет им. В. Н. Татищева)

СОГЛАСОВАНО
Руководитель ОПОП

_____ М.В. Коломина

«8» сентября 2022 г.

УТВЕРЖДАЮ
Зав. кафедрой ПМИ

_____ М.В. Коломина

«8» сентября 2022 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

«Анализ данных»

Составители	Станкевич А. С., к.т.н., доцент ФИТиП, ИТМО Корнеев Г. А., к.т.н., доцент ФИТиП, ИТМО
Направление подготовки / специальность	01.03.02 Прикладная математика и информатика
Направленность (профиль) ОПОП	Программирование и искусственный интеллект
Квалификация (степень)	бакалавр
Форма обучения	очная
Год приёма	2023
Курс	3
Семестр(ы)	6

1. ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ (МОДУЛЯ)

1.1. Целями освоения дисциплины «Анализ данных» является освоение практических аспектов технологий, связанных с хранением, обработкой, подходами к анализу больших объёмов данных.

1.2. Задачи освоения дисциплины:

- изучение вопросов, связанных с хранением и первичной обработкой данных;
- изучение вопросов статистического и машинного обучения;
- формирование практических навыков решения прикладных задач.

2. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОПОП

2.1. Учебная дисциплина «Анализ данных» относится к обязательной части и осваивается в 6 семестре.

2.2. Для изучения данной учебной дисциплины необходимы следующие знания, умения, навыки, формируемые предшествующими учебными дисциплинами:

- Математическая статистика;
- Алгоритмы и структуры данных;
- Технологии программирования.

Знания: основных определений и понятий математической статистики; основных функций и компонентов инструментальных средств проектирования и их практическое воплощение в наиболее развитых программных продуктах.

Умения: решать типовые теоретические и вычислительные задачи;

Навыки: использования и проверки алгебраических соотношений в прикладных математических задачах, ориентировании в основных алгебраических структурах и нахождении связи между ними.

2.3. Последующие учебные дисциплины (модули) и (или) практики, для которых необходимы знания, умения, навыки, формируемые данной учебной дисциплиной (модулем):

- Автоматическое машинное обучение;
- Продвинутое машинное обучение.

3. ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ

Процесс освоения дисциплины направлен на формирование элементов следующих компетенций в соответствии с ФГОС ВО и ОПОП ВО по данному направлению подготовки / специальности:

а) общепрофессиональных (ОПК);

ОПК-3. Способен формулировать, строить и применять математические модели для управления достижением планируемых результатов процессов и объектов профессиональной деятельности на базе знаний математики, программирования и унифицированных пакетов программ.

б) профессиональных (ПК).

ПК-8. Способность понимать, совершенствовать и применять современный математический аппарат.

ПК-14. Способность определять эффективный способ решения прикладных задач с применением информационных технологий и программной инженерии, разрабатывать и внедрять соответствующие программные решения.

ПК-18. Способен разрабатывать и применять методы машинного обучения для решения задач.

ПК-21. Способен осуществлять сбор и подготовку данных для систем искусственного интеллекта.

ПК-23. Способен принимать участие в управлении проектами по созданию и развитию технологий и систем искусственного интеллекта на стадиях их жизненного цикла.

Таблица 1 – Декомпозиция результатов обучения

Код и наименование компетенции	Планируемые результаты обучения по дисциплине (модулю)		
	Знать (1)	Уметь (2)	Владеть (3)
ОПК-3. Способен формулировать, строить и применять математические модели для управления достижением планируемых результатов процессов и объектов профессиональной деятельности на базе знаний математики, программирования и унифицированных пакетов программ.	ИОПК-3.1.1. математические модели, моделирование, методы описания объектов	ИОПК-3.2.1. выявлять и формулировать целевые характеристики описания объекта моделирования, определять методы описания объектов и соответствующие им модели, строить модели объектов и процессов, апробировать и реализовывать математические модели в программной среде, осуществлять их корректировку, применять модели объектов и процессов, оценивать достижение целевых характеристик и показателей, интерпретировать и представлять результаты моделирования процессов и объектов в профессиональной деятельности	ИОПК-3.3.1. навыками описания, построения, применения моделей объектов, оценки и интерпретации результатов моделирования процессов и объектов
ПК-8.1. Владение методами интегрального и дифференциального исчисления одной и нескольких переменных ПК-8.2. Владение методами теории линейных пространств и операторов ПК-8.3. Владение методами функционального анализа для решения сложных задач информатики	ИПК-8.1.1. современный математический аппарат	ИПК-8.2.1. владеть методами функционального анализа для решения сложных задач информатики	ИПК-8.3.1. навыками применения современного математического аппарата
ПК-14.1. Способность проектировать и реализовывать программные решения с применением методов функционального, автоматного и эволюционного программирования	ИПК-14.1.1. методы функционального, автоматного и эволюционного программирования	ИПК-14.2.1. проектировать и реализовывать мобильные и web-приложения	ИПК-14.3.1. навыками определения эффективного способа решения прикладных задач с применением информационных технологий и программной инженерии, разрабатывать и внедрять соответствующие программные решения
ПК-18. ПК-18.1. Проводит анализ требований и определяет необходимые классы задач машинного обучения	ИПК-18.1.1. принципы и методы машинного обучения, типы и классы задач машинного обучения, методо-	ИПК-18.2.1. сопоставить задачам предметной области классы задач машинного обучения	ИПК-18.3.1. методами машинного обучения и статистическими методами анализа данных

	логию ML Ops; статистические методы анализа данных		
ПК-18. ПК-18.2. Принимает участие в оценке, выборе и при необходи- мости разработке методов ма- шинного обучения	ИПК-18.1.2. классические методы и алго- ритмы машин- ного обучения: предиктивные - обучение с учи- телем, обучение без учителя	ИПК-18.2.2. проводить сравни- тельный анализ и осуществлять выбор, настройку при необходи- мости разработку методов и ал- горитмов для решения задач ма- шинного обучения	ИПК-18.3.2. ме- тодами обучение с учителем, обу- чение без учите- ля
ПК-21.1. Осуществляет поиск данных в открытых источниках. Специализированных библиоте- ках и репозиториях ПК-21.2. Выполняет подготовку и разметку структурированных и неструктурированных данных для машинного обучения	ПК-21.1.1. виды представления данных, методы поиска и парсин- га данных уровни пред- ставления дан- ных (ODS, DDL, семантический слой, модель данных) ПК-21.1.2. ос- новные инстру- менты, библио- теки и техноло- гии data Science ПК-21.1.3. мето- ды редукции размерности элементов набо- ра данных и их предварительной статистической обработки раз- метки структу- рированных и неструктуриро- ванных данных	ПК-21.2.1. отделять достоверные источники данных от сомнитель- ных, осуществлять критический отбор данных, проверять их це- лостность и непротиворечивость использовать инструменты и библиотеки для data science для поиска данных в открытых ис- точниках, специализированных библиотеках и репозиториях ПК-21.2.2. выявлять и исклю- чить из массива данных ошибоч- ные данные и выбросы ПК-21.2.3. осуществлять размет- ку структурированных и не- структурированных данных ПК-21.2.4. использовать инстру- менты, библиотеки и технологии Data Science для подготовки и разметки структурированных и неструктурированных данных для машинного обучения ПК-21.2.5. использовать методы и технологии массово- параллельной обработки и ана- лиза данных	ПК-21.3.1. мето- дами поиска и парсинга данных методами редук- ции размерности элементов набо- ра данных и их предварительной статистической обработки раз- метки структу- рированных и неструктуриро- ванных данных,
ПК-23.1 Использует основы управления проектами по созда- нию и развитию технологий и систем искусственного интеллек- та на стадиях их жизненного цикла	ПК-23.1.1. осно- вы управления проектами по созданию и раз- витию техноло- гий и систем ис- кусственного интеллекта на стадиях их жиз- ненного цикла	ПК-23.2.1. управлять проектами по созданию и развитию техно- логий и систем искусственного интеллекта на стадиях их жиз- ненного цикла	ПК-23.3.1. осно- вами управления проектами по созданию и раз- витию техноло- гий и систем искусственного интеллекта на стадиях их жиз- ненного цикла

4. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

Объём дисциплины составляет 4 зачётные единицы, в том числе 72 часа, выделенных на контактную работу обучающихся с преподавателем (из них 36 часов – лекции, 36 часов – лабораторные работы), и 72 часа – на самостоятельную работу обучающихся.

Таблица 2 – Структура и содержание дисциплины

Раздел, тема дисциплины (модуля)	Семестр	Контактная работа (в часах)			Самост. работа		Форма текущего контроля успеваемости, форма промежуточной аттестации
		Л	ПЗ	ЛР	КР	СР	
Хранение и первичная обработка данных	6	18	-	18	-	36	лабораторная работа № 1 контрольная работа № 1 упражнение № 1
Статистическое и машинное обучение	6	18	-	18	-	36	лабораторная работа № 2 контрольная работа № 2 упражнение № 2
Итого		36	-	36	-	72	Экзамен

Примечание: Л – лекция; ПЗ – практическое занятие, семинар; ЛР – лабораторная работа; ПП – практическая подготовка; КР / КП – курсовая работа / курсовой проект; КПА – контроль промежуточной аттестации; КС – консультации; СР – самостоятельная работа

Таблица 3. Матрица соотношения разделов, тем учебной дисциплины и формируемых компетенций

Раздел, тема дисциплины (модуля)	Кол-во часов	Код компетенции						Общее количество компетенций
		ОПК-3	ПК-8	ПК-14	ПК-18	ПК-21	ПК-23	
Хранение и первичная обработка данных	71,75	+	+	+	+	+	+	6
Статистическое и машинное обучение	71,75	+	+	+	+	+	+	6

Краткое содержание каждой темы дисциплины (модуля)

Раздел 1. Хранение и первичная обработка данных

Виды и источники данных. Загрузка и разделение данных. Объединение данных из разных источников. Очистка данных и заполнение пропусков. Контроль диапазонов, Хранение и доступ к данным. Виды баз данных. Реляционные СУБД, Первичная обработка данных. Сглаживание и нормировка данных. Преобразование данных. Визуализация данных. Формы представления количественных и качественных данных. Когнитивная визуализация данных, NoSQL базы данных. Большие данные

Раздел 2. Статистическое и машинное обучение

Элементы математической статистики: точечные и интервальные оценки, проверка гипотез. Методы оценивания некоторых характеристик статистических связей, Регрессии и классификация: линейная, многомерная, полиномиальная, логистическая, наивный Байесовский классификатор, деревья решений и метод ближайших соседей, Метод опорных векторов и нейросети. Задачи машинного обучения (обучение с учителем, без учителя), Композиции: бэггинг, случайный лес, градиентный бустинг.

5. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ПРЕПОДАВАНИЮ И ОСВОЕНИЮ ДИСЦИПЛИНЫ

5.1. Указания для преподавателей по организации и проведению учебных занятий по дисциплине (модулю)

Лекционные занятия

Основной формой реализации теоретического обучения является лекция, которая представляет собой систематическое, последовательное изложение преподавателем-лектором учебного материала теоретического характера. Цель лекции – организация целенаправленной познавательной деятельности студентов по овладению программным материалом учебной дисциплины.

Порядок подготовки лекционного занятия включает в себя выполнение следующих этапов:

- изучение требований программы дисциплины;
- определение целей и задач лекции;
- разработка плана проведения лекции;
- подбор литературы (ознакомление с методической литературой, публикациями периодической печати по теме лекционного занятия);
- отбор необходимого и достаточного по содержанию учебного материала;
- определение методов, приемов и средств поддержания интереса, внимания, стимулирования творческого мышления студентов;
- написание конспекта лекции.

Лекция должна включать следующие разделы:

- формулировку темы лекции;
- указание основных изучаемых разделов или вопросов и предполагаемых затрат времени на их изложение;
- изложение вводной части;
- изложение основной части лекции;
- краткие выводы по каждому из вопросов;
- заключение;
- рекомендации литературных источников по излагаемым вопросам.

Лабораторные занятия

Лабораторное занятие – целенаправленная форма организации педагогического процесса, направленная на углубление научно-теоретических знаний и овладение определенными методами работы, в процессе которых вырабатываются умения и навыки выполнения тех или иных учебных действий в данной сфере науки. Они развивают научное мышление и речь, позволяют проверить знания студентов и выступают как средства оперативной обратной связи.

Правильно организованные лабораторные занятия ориентированы на решение следующих задач:

- обобщение, систематизация, углубление, закрепление полученных на лекциях и в процессе самостоятельной работы теоретических знаний по дисциплине (предмету);
- формирование практических умений и навыков, необходимых в будущей профессиональной деятельности, реализация единства интеллектуальной и практической деятельности;
- выработка при решении поставленных задач таких профессионально значимых качеств, как самостоятельность, ответственность, точность, творческая инициатива.

Состав заданий для лабораторного занятия должен быть спланирован с расчетом, чтобы за отведенное время они могли быть качественно выполнены большинством учащихся.

Лабораторные занятия должны так быть организованы, чтобы студенты ощущали нарастание сложности выполнения заданий, испытывали бы положительные эмоции от переживания собственного успеха в учении, поисками правильных и точных решений.

Самостоятельная работа

Самостоятельная работа – это вид учебной деятельности, которую студент совершает в установленное время и в установленном объеме индивидуально или в группе, без непосредственной помощи преподавателя (но при его контроле), руководствуясь сформированными ранее представлениями о порядке и правильности выполнения действий.

В учебном процессе образовательного учреждения выделяются два вида самостоятельной работы:

- аудиторная – выполняется на учебных занятиях, под непосредственным руководством преподавателя и по его заданию (выполнение самостоятельных работ; выполнение контрольных и практических работ; решение задач);
- внеаудиторная – выполняется по заданию преподавателя, но без его непосредственного участия (подготовка к аудиторным занятиям; изучение учебного материала, вынесенного на самостоятельную проработку; выполнение домашних заданий различного характера; выполнение индивидуальных заданий, направленных на развитие у студентов самостоятельности и инициативы; подготовка к контрольной работе). Внеаудиторные самостоятельные работы представляют собой логическое продолжение аудиторных занятий, проводятся по заданию преподавателя, который инструктирует студентов и устанавливает сроки выполнения задания.

5.2. Указания для обучающихся по освоению дисциплины (модулю)

Лекция

- Лекция – основной вид обучения в вузе.
- В лекции излагаются основные положения теории, ее понятия и законы, приводятся факты, показывающие связь теории с практикой.
- Накануне лекции необходимо повторить содержание предыдущей лекции (а также теорию по изучаемой теме в школьных учебниках геометрии, если эта тема была представлена в них), а затем посмотреть тему очередной лекции по программе (по плану лекций).

Лабораторное занятие

- Лабораторное занятие – наиболее активный вид учебных занятий в вузе. Он предполагает самостоятельную работу над лекциями и учебными пособиями.
- К каждому лабораторному занятию нужно готовиться. Подготовку следует начинать с повторения теории (по записям лекций или по учебному пособию). После этого нужно решать задачи из предложенного домашнего задания.

Организация самостоятельной работы

Самостоятельность в учебной работе способствует развитию заинтересованности студента в изучаемом материале, вырабатывает у него умение и потребность самостоятельно получать знания, что весьма важно для специалиста с высшим образованием. Самостоятельная работа студентов представлена в следующих формах:

- работа с учебной литературой и конспектом лекций с целью подготовки к лабораторным занятиям, составление конспектов тем, выносимых на самостоятельную проработку;
- систематическое выполнение домашних работ.

Таблица 4 – Содержание самостоятельной работы обучающихся

Номер раздела (темы)	Темы/вопросы, выносимые на самостоятельное изучение	Кол-во часов	Форма работы
Раздел 1	Хранение и первичная обработка данных	36	Изучение теоретического материала. Подготовка к лабораторным работам. Решение заданий
Раздел 2	Статистическое и машинное обучение	36	Изучение теоретического материала. Подготовка к лабораторным работам. Решение заданий

5.3. Виды и формы письменных работ, предусмотренных при освоении дисциплины (модуля), выполняемые обучающимися самостоятельно

Экзамен проводится в письменной форме в виде контрольной работы.

Контрольная работа выполняется студентом в день экзамена и включает в себя задания, аналогичные заданиям из текущего контроля успеваемости. Время выполнения ограничено 60-90 минутами. Варианты контрольной работы генерируются автоматически и представляют собой задачи как с закрытым, так и открытым типом вопросов. Контрольная работа включает в себя некоторое количество (5-20) заданий. Задания даются по разделу 1 «Хране-

ние и первичная обработка данных» и разделу 2 – «Статистическое и машинное обучение». Максимальное количество баллов за экзамен – 20.

6. ОБРАЗОВАТЕЛЬНЫЕ И ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

При реализации различных видов учебной работы по дисциплине «Анализ данных» могут использоваться электронное обучение и дистанционные образовательные технологии.

6.1. Образовательные технологии

Учебные занятия по дисциплине могут проводиться с применением информационно-телекоммуникационных сетей при опосредованном (на расстоянии) интерактивном взаимодействии обучающихся и преподавателя в режимах on-line или off-line в формах.

Таблица 5 – Образовательные технологии, используемые при реализации учебных занятий

Раздел, тема дисциплины	Форма учебного занятия		
	Лекция	Практическое занятие, семинар	Лабораторная работа
Хранение и первичная обработка данных	<i>Обзорная лекция</i>	<i>Не предусмотрено</i>	<i>Выполнение лабораторной работы</i>
Статистическое и машинное обучение	<i>Обзорная лекция</i>	<i>Не предусмотрено</i>	<i>Выполнение лабораторной работы</i>

6.2. Информационные технологии

При реализации различных видов учебной и внеучебной работы используются следующие информационные технологии:

- система управления обучением LMS Moodle;
- использование возможностей Интернета в учебном процессе (рассылка заданий, предоставление выполненных работ, ответы на вопросы, ознакомление обучающихся с оценками и т.д.);
- использование электронных учебников и различных сайтов (например, электронные библиотеки, журналы и т.д.) как источник информации;
- использование возможностей электронной почты;
- использование средств представления учебной информации (электронных учебных пособий, применение новых технологий для проведения занятий с использованием презентаций и т.д.);
- использование интерактивных средств взаимодействия участников образовательного процесса (технологии дистанционного или открытого обучения в глобальной сети);
- использование интегрированных образовательных сред, где главной составляющей являются не только применяемые технологии, но и содержательная часть, т.е. информационные ресурсы (доступ к мировым информационным ресурсам, на базе которых строится учебный процесс).

6.3. Современные профессиональные базы данных и информационные справочные системы

6.3.1. Программное обеспечение

Наименование программного обеспечения	Назначение
Adobe Reader	Программа для просмотра электронных документов
Платформа дистанционного обучения LMS Moodle	Виртуальная обучающая среда
Microsoft Office 2013, Microsoft Office Project 2013, Microsoft Office Visio 2013	Пакет офисных программ
7-zip	Архиватор
Microsoft Windows 7 Professional	Операционная система
Kaspersky Endpoint Security	Средство антивирусной защиты
Google Chrome	Браузер
OpenOffice	Пакет офисных программ

6.3.2. Современные профессиональные базы данных и информационные справочные системы

1. Электронная библиотека «Астраханский государственный университет» собственной генерации на платформе ЭБС «Электронный Читальный зал – БиблиоТех». <https://biblio.asu.edu.ru>
2. Электронно-библиотечная система (ЭБС) ООО «Политехресурс» «Консультант студента». www.studentlibrary.ru.
3. Электронная библиотечная система издательства ЮРАЙТ, раздел «Легендарные книги». www.biblio-online.ru
4. Электронный каталог Научной библиотеки АГУ на базе MARK SQL НПО «Информ-систем». <https://library.asu.edu.ru>

7. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ДЛЯ ПРОВЕДЕНИЯ ТЕКУЩЕГО КОНТРОЛЯ И ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ ПО ДИСЦИПЛИНЕ (МОДУЛЮ)

7.1. Паспорт фонда оценочных средств

При проведении текущего контроля и промежуточной аттестации по дисциплине «Анализ данных» проверяется сформированность у обучающихся компетенций, указанных в разделе 3 настоящей программы. Этапность формирования данных компетенций в процессе освоения образовательной программы определяется последовательным освоением дисциплин и прохождением практик, а в процессе освоения дисциплины – последовательным достижением результатов освоения содержательно связанных между собой разделов, тем.

Таблица 6 – Соответствие разделов, тем дисциплины, результатов обучения по дисциплине и оценочных средств

№ п/п	Контролируемые разделы, темы дисциплины (модуля)	Код контролируемой компетенции (компетенций)	Наименование оценочного средства
1	Общие принципы построения сетей, физические и канальные протоколы	ОПК-3, ПК-8, ПК-14, ПК-18, ПК-21, ПК-23	Лабораторные работы, Контрольная работа
2	Протоколы сетевого, транспортного и прикладного уровней	ОПК-3, ПК-8, ПК-14, ПК-18, ПК-21, ПК-23	Лабораторные работы, Контрольная работа

7.2. Описание показателей и критериев оценивания компетенций, описание шкал оценивания

Таблица 7 – Показатели оценивания результатов обучения в виде знаний

Шкала оценивания	Критерии оценивания
5 «отлично»	демонстрирует глубокое знание теоретического материала, умение обоснованно излагать свои мысли по обсуждаемым вопросам, способность полно, правильно и аргументированно отвечать на вопросы, приводить примеры
4 «хорошо»	демонстрирует знание теоретического материала, его последовательное изложение, способность приводить примеры, допускает единичные ошибки, исправляемые после замечания преподавателя
3 «удовлетворительно»	демонстрирует неполное, фрагментарное знание теоретического материала, требующее наводящих вопросов преподавателя, допускает существенные ошибки в его изложении, затрудняется в приведении примеров и формулировке выводов
2 «неудовлетворительно»	демонстрирует существенные пробелы в знании теоретического материала, не способен его изложить и ответить на наводящие вопросы преподавателя, не может привести примеры

Таблица 8 – Показатели оценивания результатов обучения в виде умений и владений

Шкала оценивания	Критерии оценивания
5 «отлично»	демонстрирует способность применять знание теоретического материала при выполнении заданий, последовательно и правильно выполняет задания, умеет обоснованно излагать свои мысли и делать необходимые выводы
4 «хорошо»	демонстрирует способность применять знание теоретического материала при выполнении заданий, последовательно и правильно выполняет задания, умеет обоснованно излагать свои мысли и делать необходимые выводы, допускает единичные ошибки, исправляемые после замечания преподавателя
3 «удовлетворительно»	демонстрирует отдельные, несистематизированные навыки, испытывает затруднения и допускает ошибки при выполнении заданий, выполняет задание по подсказке преподавателя, затрудняется в формулировке выводов
2 «неудовлетворительно»	не способен правильно выполнить задания

7.3. Контрольные задания и иные материалы, необходимые для оценки результатов обучения по дисциплине (модулю)

Типовые контрольные задания, необходимые для оценки достижения запланированных результатов обучения приведены в таблице планирования результатов обучения по дисциплине (БаРС) (Приложение 1)*.

Контроль успеваемости по дисциплине осуществляется с помощью следующих оценочных средств:

Контрольная работа 1

Описание технологии проведения контрольной работы:

Контрольная работа выполняется в форме тестов. После каждой лекции студентам выдаются 1-3 тестовых задания. Срок выполнения теста – не позднее срока освоения дисциплины.

Примеры контрольных заданий для блока 1:

1. Есть исходные данные для анализа следующего вида:

<i>Имя</i>	<i>Возраст</i>	<i>Пол</i>	<i>Образование</i>	<i>Семейное положение</i>
Иннокентий	35	м	высшее	женат
Семен	34	м	высшее	женат
Анфиса	45	ж	среднее	замужем
Светлана	15	ж	начальное	не замужем
Инна	19	ж	среднее	не женат
Михаил	27	м	высшее	разведен
Александр	25	м	высшее	разведен
Екатерина	27	ж	среднее	замужем

Каким шкалам соответствуют значения в каждой из колонок? Ответы выбирать из возможных значений: номинальная, порядковая, интервальная, относительная, дихотомическая (3 балла)

0. Для заданного набора данных (таблица выше) определить распределение в процентах параметра «Образование» (4 балл)

1. Исходные данные:

	2014	2015	2016	2017	2018
Мужчины	13456	11344	12444	14333	12555
Женщины	11333	12323	11212	12121	13121
Пенсионеры	132	153	153	211	242
Студенты	232	345	364	374	435
Пожарники	22	35	38	58	45

Для заданной таблицы данных указать, какие строки можно агрегировать. Ответы выбираются из возможных комбинаций строк (3 балла)

Шкала оценивания и критерии оценки за контрольные работы блока 1:

Минимальное количество баллов — 0 баллов

Максимальное количество баллов — 10 баллов

Оценка снижается при отсутствии развёрнутого решения, наличия доказательных и вычислительных ошибок.

Лабораторная работа 1

Лабораторная работа выполняется в рамках каждого раздела курса с целью усвоения прослушанного студентом теоретического материала. Лабораторная работа реализуется в форме тестовых вопросов внутри лекции.

Объем выполненной работы: каждая лабораторная работа содержит 3-5 вопросов.

Срок сдачи работы указывается преподавателем. Студент получает вопросы и предлагаемые ответы, из которых необходимо выбрать правильные. На выполнение работы дается одна попытка.

Примеры вопросов для лабораторной работы «Хранение и первичная обработка данных»:

1. Укажите шкалы, которые являются дихотомическими:
 - Мнение ('согласен' / 'не могу ответить' / 'не согласен')
 - Пол ('мужской' / 'женский')
 - Возраст ('младше 10 лет' / 'от 10 до 20 лет' / 'от 20 до 30 лет' / 'от 30 до 40 лет' / 'старше 40 лет')
0. Укажите свойства ключа таблицы:
 - Может принимать незаданные значения
 - Может повторяться в разных строках таблицы
 - Должен быть только числовым значением
 - Должен принимать уникальные значения для каждой строки таблицы
 - Может состоять из одного поля, а может из нескольких полей
0. Выберите верные высказывания:
 - Мода всегда больше медианы
 - Мода всегда равна медиане
 - Средневзвешенное значение всегда равно моде
 - Средневзвешенное значение всегда равно медиане
 - Все предыдущие утверждения неверны

Шкала оценивания и критерии оценки (на примере одной лабораторной работы):

Требования	Максимальное количество баллов
1. Логика решения	5
2. Верно произведенные расчеты	5
Итого:	10

Упражнение по блоку 1

После всех лекций блока 1 студентам выдается одно упражнение (1-6 заданий).

Срок выполнения упражнений – не позднее срока освоения дисциплины. Для выполнения задания дается две попытки. За правильное решение со второй попытки количество баллов уменьшается (обычно в два раза). Итоговый балл за упражнения подсчитывается по максимуму из двух попыток целым числом с округлением в большую сторону.

Пример упражнений блока 1:

Студенты получают доступ к таблицам тестовым базам данных (Oracle и MongoDB):

STUDENT_PROFILE(student_number, student_fio, student_birthday),

STUDENT_MARKS (student_number, subject_code, subject_mark, subject_mark_date) STUDENT_PUBLICATIONS(student_number, publication_id, publication_name, publication_issue_date)

Требуется:

1. Выполнить контроль диапазонов отметок (**subject_mark** должен быть в интервале от 1 до 5). Исключить неподходящие записи. (2 балла)
2. Выполнить контроль публикаций студента и дат полученных отметок. Исключить публикации и отметки с сомнительной датой (опубликованные/полученные в возрасте менее 5 лет или с еще не наступившей датой). (2 балла)
3. Объединить данные из нескольких таблиц в одну таблицу **STUDENT_TOTAL** (**student_number, student_fio, mark_5_count, mark_4_count, publication_count**), где **mark_5_count** – количество отметок 5, **mark_4_count** - количество отметок 4, **publication_count** – количество публикаций. (5 баллов)
4. Нормировать значения полей **mark_5_count, mark_4_count** и **publication_count** заданным способом нормировки (с точностью до двух знаков после десятичного разделителя). (5 баллов)
5. Из базы MongoDB получить документ по заданному значению параметра. (4 балла)
6. Вывести в качестве результата сумму нормированных значений **mark_5_count, mark_4_count** и **publication_count**. (4 балла)

Критерии оценивания:

Требования	Максимальное количество баллов
1. Логика решения	10
2. Верно произведенные расчеты	10
Итого:	20

Лабораторная работа 2

Лабораторная работа выполняется в рамках каждого раздела курса с целью усвоения прослушанного студентом теоретического материала. Лабораторная работа реализуется в форме тестовых вопросов внутри лекции.

Объем выполненной работы: каждая лабораторная работа содержит 3-5 вопросов.

Срок сдачи работы указывается преподавателем. Студент получает вопросы и предлагаемые ответы, из которых необходимо выбрать правильные. На выполнение работы дается одна попытка.

Примеры вопросов для лабораторной работы «Статистическое и машинное обучение»:

1. Дискретная случайная величина X принимает значения $\{-1, 0, 1\}$ с вероятностями соответственно $\{0.1, 0.6, 0.3\}$. Ее математическое ожидание равно:
 - 2.2
 - 2.4
 - 0.2
 - 1.2
 - 0.6
2. Логистическая функция задается выражением:
 - $\sigma(x) = \frac{1}{1 + e^{-x}}$
 - $f(s) = \frac{s}{|s| + \alpha}$
 - $f(s) = \frac{s}{|s| + \alpha}$
2. Какие из этих задач являются типичными для машинного обучения без учителя?
 - Оценка вероятности выхода автомобиля из строя
 - Предсказание вероятности покупки автомобиля
 - Группировка изображений по визуальным признакам
 - Рекомендация видеороликов
2. Случайная величина Y задается биномиальным распределением с параметрами $n=10$ и $p=1/5$. Ее дисперсия равна:
 - 5
 - 1/2
 - 4/5
 - 2
 - 8/5
2. Бэггинг – это метод, позволяющий снизить:
 - ошибку разброса
 - ошибку смещения
 - ошибку смещения и ошибку разброса
 - не подходит ни один вариант из приведенных

Шкала оценивания и критерии оценки (на примере одной лабораторной работы):

Требования	Максимальное количество баллов
1. Логика решения	5
2. Верно произведенные расчеты	5

Итого:	10
---------------	-----------

Контрольная работа 2

Описание технологии проведения контрольной работы:

Контрольная работа выполняется в форме тестов

После каждой лекции студентам выдаются 1-3 тестовых задания.

Срок выполнения теста – не позднее срока освоения дисциплины.

Примеры контрольных заданий для блока 2:

На основе анализа некоторого числа писем электронной почты сформирована таблица 1, содержащая информацию о классификации писем на группы «спам» и «не спам», а также суммарное число слов входящих в эти группы. В таблице 2 представлены данные, по уникальным словам, и числу их вхождений в указанные группы.

Таблица 1.

	SPAM	HAM
Emails	19	7
Words	114	42

Таблица 2.

	SPAM	HAM
Credit	21	2
Online	5	2
Bill	0	3
Free	11	11
Cash	3	4
Bonus	23	8
Remove	3	9
Money	12	0
Offer	23	2
Coupon	13	1

Ваша задача найти оценку апостериорного максимума и вероятность для определения наиболее вероятного класса, к которому будет отнесено письмо, содержащее текст:

«Online Credit Offer Access Cash Free Membership»

Правильный ответ может отклоняться от эталонных не более, чем на 2%.

Шкала оценивания и критерии оценки за контрольные работы блока 2:

Минимальное количество баллов — 0 баллов

Максимальное количество баллов— 10 баллов

Каждая правильно решённая задача оценивается в 5 балла. Оценка снижается при отсутствии развёрнутого решения, наличия доказательных и вычислительных ошибок.

Упражнение по блоку 2

После всех лекций блока 2 студентам выдается одно упражнение (1-6 заданий).

Срок выполнения упражнений – не позднее срока освоения дисциплины. Для выполнения задания дается две попытки. За правильное решение со второй попытки количество баллов уменьшается (обычно в два раза). Итоговый балл за упражнения подсчитывается по максимуму из двух попыток целым числом с округлением в большую сторону.

Пример упражнений блока 2:

Для сгенерированного датасета выполнить указанные действия и расчеты.

1. Разделить выборку на тренировочную и тестовую (90% и 10%). Обучить классификатор на основе метода логистической регрессии. (5 баллов)
2. Оценить качество классификации на тренировочной/тестовой выборке. Сделать 10-fold кросс-валидацию и оценить дисперсию. (5 баллов)
3. Построить графики зависимости точности на тренировочном/тестовом наборе для n-fold кросс-валидации для разных значений $n = 2, 5, 8, 10$. (5 баллов)
4. Сделать вывод о сравнительном качестве предсказаний при разных способах оценки точности. (5 баллов)

Критерии оценивания:

Требования	Максимальное количество баллов
1. Логика решения	8
2. Верно произведенные расчеты	12
Итого:	20

Перечень вопросов и заданий, выносимых на экзамен

Экзамен проводится в письменной форме в виде контрольной работы.

Контрольная работа выполняется студентом в день экзамена и включает в себя задания, аналогичные заданиям из текущего контроля успеваемости. Время выполнения ограничено 60-90 минутами. Варианты контрольной работы генерируются автоматически и представляют собой задачи как с закрытым, так и открытым типом вопросов. Контрольная работа включает в себя некоторое количество (5-20) заданий. Задания даются по разделу 1 «Хранение и первичная обработка данных» и разделу 2 – «Статистическое и машинное обучение». Максимальное количество баллов за экзамен — 20.

Раздел 1. Хранение и первичная обработка данных

1. К какому из приведенных типов шкал относится шкала Бофорта? Номинальная, Интервальная, Относительная.
2. Какой оператор языка SQL позволяет добавить (вставить) строки в таблицы: CHANGE INTO, ALTER INTO, MODIFY INTO, ADD INTO, INSERT INTO

Раздел 2. Статистическое и машинное обучение Элементы статистической обработки данных

1. Функция распределения случайной величины ξ задана соотношением

$$F_{\xi}(x) = \begin{cases} 0, & x \leq 0 \\ 1 - e^{-3x}, & x > 0 \end{cases}$$

Какое распределение имеет данная случайная величина?

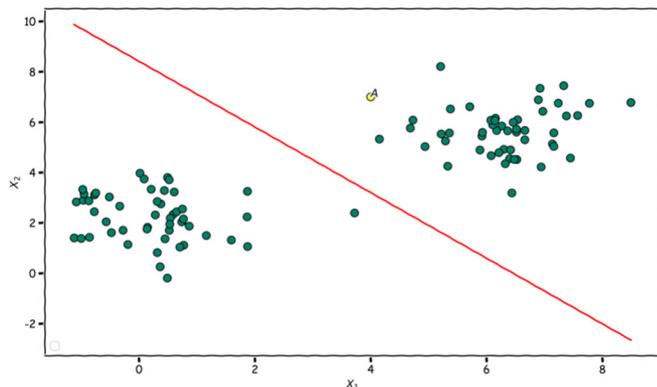
1. Exp_{-3}
2. Exp_3
3. $Exp_{1/3}$

4. $Exp_{-1/3}$ 0. Дана выборка $(1, 2, 1, 4, 2)$.

1. Найдите выборочное среднее.
2. Найдите выборочную медиану.

0. Пусть $\xi \sim Exp_{\lambda}$. Введите выражение для оценки λ , используя выборочное среднее.0. Уравнение линейной регрессии задается соотношением $3 + 4X_1 - 2X_2 = 0$. Предскажите отклик для тестового объекта — $(6, 6)$.

0. Модель логистической регрессии обучена на определение класса +.

 $\Psi = -8.872 + 1.374 \cdot X_1 + 1.055 \cdot X_2$. На рисунке отображены точки, на которых обучена модель, но вам неизвестно какие из них относятся к классу + и -.

Определите, в каком направлении расположен нормальный вектор гиперплоскости. Учитывая, что тестовая точка A с координатами (4,7) относится к классу +, верно ли она классифицирована моделью?

Таблица 9. Примеры оценочных средств с ключами правильных ответов

№ п/п	Тип задания	Формулировка задания	Правильный ответ	Время выполнения (в минутах)
Код и наименование проверяемой компетенции ОПК-3				
1.	Задание закрытого типа	<p><i>Выберите верный ответ.</i></p> <p>Фактором в задаче однофакторного дисперсионного анализа называют ...</p> <ol style="list-style-type: none"> 1. значение измеряемого признака 2. переменную, которая влияет на значение измеряемого признака 3. погрешность измеряемого признака 	2	1-3
2.		<p><i>Выберите верный ответ.</i></p> <p>Откликом в задаче однофакторного дисперсионного анализа называют</p> <ol style="list-style-type: none"> 1. значение измеряемого признака 2. переменную, которая влияет на значение измеряемого признака 3. погрешность измеряемого признака 	1	1-3
3.		<p><i>Выберите верный ответ.</i></p> <p>Уровнем фактора в задаче однофакторного дисперсионного анализа называют</p>	3	1-3

№ п/п	Тип задания	Формулировка задания	Правильный ответ	Время выполнения (в минутах)
		1. значение измеряемого признака 2. переменную, которая влияет на значение измеряемого признака 3. конкретную реализацию фактора		
4.		<i>Выберите верный ответ.</i> Задача однофакторного дисперсионного анализа является обобщением задачи проверки гипотезы об однородности двух выборок против альтернативы о том, что рассматриваемые выборки различаются 1. параметром масштаба 2. параметром сдвига 3. типом распределения	2	1-3
5.		<i>Выберите верный ответ.</i> Количество уровней фактора в задаче однофакторного дисперсионного анализа может быть 1. любым целым числом $a \leq 3$ 2. любым целым числом $a \geq 3$ 3. любым положительным действительным числом	2	1-3
6.	Задание открытого типа	В чем состоит основная (проверяемая) гипотеза в задаче однофакторного дисперсионного анализа?	Неизвестные средние значения наблюдений одинаковы при различных уровнях фактора	3-5
7.		Каким условиям должны удовлетворять погрешности наблюдений в модели однофакторного дисперсионного анализа?	Быть независимыми и иметь одинаковые распределения с нулевым математическим ожиданием	3-5
8.		Какие требования являются необходимыми для применения F-критерия в задаче однофакторного дисперсионного анализа?	Погрешности имеют гауссовское распределение с нулевым математическим ожиданием и одинаковыми дисперсиями	3-5
9.		Что является необходимым условием для применения критерия Краскела-Уоллиса в задаче однофакторного дисперсионного анализа?	Погрешности имеют некоторое непрерывное распределение с нулевым математическим ожиданием	3-5
10.		Что является необходимым условием применения критерия Пейджа в задаче двухфакторного дисперсионного анализа?	Априорная информация о том, что при увеличении уровня главного фактора среднее значение отклика будет увеличиваться	3-5
11.	Задание комбинированного типа	<i>Верно ли утверждение:</i> Позиционное сравнение выявляет соотношение объектов друг к другу, при этом, ключевые слова «колебаться по сравнению», «изменяться в сравнении». Ответ обоснуйте.	Утверждение частично верно. Позиционное сравнение выявляет соотношение объектов друг к другу, при этом, ключевые слова "больше чем", "меньше чем", "равно"	3-5
Код и наименование проверяемой компетенции ПК-8				
12.	Задание закрытого типа	<i>Выберите верный ответ.</i> Для проверки основной гипотезы в задаче двухфакторного дисперсионного анализа применяют F-критерий и ран-	1, 2	1-3

№ п/п	Тип задания	Формулировка задания	Правильный ответ	Время выполнения (в минутах)
		<p>говый критерий Фридмана. Асимптотическая относительная эффективность по Питмену критерия Фридмана по отношению к F-критерию зависит от</p> <ol style="list-style-type: none"> 1. распределения наблюдений 2. количества уровней главного фактора 3. количества уровней мешающего фактора 		
13.		<p><i>Выберите верные ответы.</i></p> <p>Переменная X измерена в количественной шкале. Результаты измерений этой переменной</p> <ol style="list-style-type: none"> 1. можно представить в порядковой шкале измерений 2. можно представить в номинальной шкале измерений 3. нельзя представить в какой-либо другой шкале измерений 	1, 2	1-3
14.		<p><i>Выберите верный ответ.</i></p> <p>Переменная X измерена в порядковой шкале. Результаты измерений этой переменной</p> <ol style="list-style-type: none"> 1. можно представить в количественной шкале измерений 2. можно представить в номинальной шкале измерений 3. нельзя представить в какой-либо другой шкале измерений 	2	1-3
15.		<p><i>Выберите верный ответ.</i></p> <p>Переменная X измерена в номинальной шкале. Результаты измерений этой переменной</p> <ol style="list-style-type: none"> 1. можно представить в количественной шкале измерений 2. можно представить в порядковой шкале измерений 3. нельзя представить в какой-либо другой шкале измерений 	3	1-3
16.		<p><i>Выберите верный ответ.</i></p> <p>Признаки X и Y измерены в номинальной шкале. Какой критерий можно применить для проверки гипотезы о независимости этих признаков?</p> <ol style="list-style-type: none"> 1. ранговый критерий Спирмена 2. ранговый критерий Кендэла 3. критерий хи-квадрат Фишера-Пирсона 4. критерий, основанный на выборочном коэффициенте корреляции 	3	1-3

№ п/п	Тип задания	Формулировка задания	Правильный ответ	Время выполнения (в минутах)
17.	Задание открытого типа	Что является необходимым условием применения F-критерия в задаче двухфакторного дисперсионного анализа?	Погрешности имеют гауссовское распределение с нулевым математическим ожиданием	3-5
18.		Признаки X и Y измерены в количественной шкале. Требуется выяснить, являются ли эти переменные независимыми. Что нужно для того, чтобы решить эту задачу?	(1) преобразовать обе переменные в номинальные переменные и применить критерий хи-квадрат (2) преобразовать обе переменные в порядковые переменные и применить ранговый критерий Спирмена	3-5
19.		В каком случае МНК-оценка параметра θ линейной регрессионной модели совпадает с оценкой максимального правдоподобия параметра θ ?	В том случае, когда шумы имеют гауссовское распределение	3-5
20.		Чем обусловлена проблема мультиколлинеарности в линейной регрессионной модели?	Наличием линейной зависимости между входными (объясняющими) переменными	3-5
21.		К каким последствиям может привести наличие мультиколлинеарности в линейной регрессионной модели?	(1) существенному изменению оценок регрессионных параметров при незначительном изменении матрицы плана (2) большим дисперсиям оценок регрессионных параметров	3-5
Код и наименование проверяемой компетенции				
ПК-14				
22.	Задание закрытого типа	<i>Выберите верный ответ.</i> Переменная X измерена в номинальной шкале, а переменная Y - в количественной шкале. Требуется выяснить, являются ли эти переменные независимыми. Для того чтобы решить эту задачу, можно 1. преобразовать переменную Y в номинальную переменную и применить критерий хи-квадрат 2. преобразовать переменную X в количественную переменную и применить критерий для проверки независимости количественных переменных 3. применить критерий Колмогорова-Смирнова	1	1-3
23.		<i>Выберите верный ответ.</i> В ходе эксперимента получена реализация двумерной выборки. Известно, что первое наблюдение (5;3) , а второе (3;1). Можно сказать, что эти пары 1. согласованы 2. несогласованы 3. коррелированы 4. некоррелированы	1	1-3

№ п/п	Тип задания	Формулировка задания	Правильный ответ	Время выполнения (в минутах)
24.		<p><i>Выберите верный ответ.</i></p> <p>В ходе эксперимента получена реализация двумерной выборки. Известно, что первое наблюдение (7;3), а второе (3;5). Можно сказать, что эти пары</p> <ol style="list-style-type: none"> 1. согласованы 2. несогласованы 3. коррелированы 4. некоррелированы 	2	1-3
25.		<p><i>Выберите верный ответ.</i></p> <p>В ходе эксперимента получена реализация двумерной выборки. Известно, что первое наблюдение (10;3), а второе (3;1). Можно сказать, что эти пары</p> <ol style="list-style-type: none"> 1. согласованы 2. несогласованы 3. коррелированы 4. некоррелированы 	1	1-3
26.		<p><i>Выберите верный ответ.</i></p> <p>Коэффициент конкордации Кендалла $W(m)$, где $m \geq 3$, может принимать значения</p> <ol style="list-style-type: none"> 1. $W(m) \in [0, 1]$ 2. $W(m) \in [0, \infty]$ 3. $W(m) \in [-1, 1]$ 	1	1-3
27.	Задание открытого типа	Какой факт свидетельствует о наличии мультиколлинеарности в линейной регрессионной модели?	<p>(1) в корреляционной матрице входных (объясняющих) переменных есть значения, модуль которых близок к единице</p> <p>(2) матрица $X^T X$ (X - матрица плана) имеет высокий показатель обусловленности $\eta = \lambda_{max} / \lambda_{min}$, λ - собственные значения матрицы $X^T X$</p>	3-5
28.		Что дают экспериментальные исследования?	Критерии оценки обоснованности и приемлемости на практике любых теорий и теоретических предположений	3-5
29.		Что является конечной целью любой обработки экспериментальных данных?	<p>(1) выдвижение гипотез о классе и структуре математической модели</p> <p>(2) выбор возможных методов последующей статистической обработки и их анализ</p>	3-5
30.		Дайте определение понятию <i>математическая модель</i> .	1) приближённое описание какого-либо класса явлений внешнего	3-5

№ п/п	Тип задания	Формулировка задания	Правильный ответ	Время выполнения (в минутах)
			мира, выраженное с помощью математической символики (2) мощный метод познания внешнего мира, его прогнозирования и управления им	
31.		Какие задачи необходимо решить для разработки современной математической модели?	1) объединение нескольких групп измерений (2) выявление статистических связей и взаимовлияния различных измеряемых факторов и результирующих переменных (3) оценка параметров и числовых характеристик наблюдаемых случайных величин или процессов	3-5
Код и наименование проверяемой компетенции ПК-18				
32.	Задание закрытого типа	<i>Выберите верный ответ.</i> Эксперимент является 1. важнейшим средством получения знаний 2. критерием оценки обоснованности принятия решений 3. средством для проведения исследований 4. критерием оценки проведенных исследований	1	1-3
33.		<i>Выберите верные ответы.</i> Вычисленные моменты распределения являются 1. точечными оценками выборочных величин 2. распределительными оценками вычисляемых величин 3. квадратичным отклонением при вычислении точечных оценок 4. дисперсией	1, 2	1-3
34.		<i>Выберите верный ответ.</i> К вычисляемым в результате эксперимента оценкам случайных величин предъявляются следующие требования: 1. состоятельности, несмещенности, эффективности 2. выборочности статичности корреляционности 3. состоятельности, смещенности, островершинности 4. несмещенности, корреляционности, эффективности	1	1-3
35.		<i>Выберите верный ответ.</i> При выборочном наблюдении встречаются ошибки	1	1-3

№ п/п	Тип задания	Формулировка задания	Правильный ответ	Время выполнения (в минутах)
		1. грубые, систематические, случайные 2. грубые, корреляционные, случайные 3. системные, повторяющиеся, смещенные 4. случайные, периодические, ассиметричные		
36.		<i>Выберите верный ответ.</i> Функции аппроксимации... 1. при решении данной функции возникает задача приближенного вычисления значений функций в промежутках между узловыми точками и за их пределами 2. при решении данной функции используются зависимости вида $y(x)$, но число точек этих зависимостей ограничено 3. в общем случае эта функция при каждом обращении к ней возвращает последовательность случайных чисел с равномерным распределением вероятности 4. данная функция задаётся как ранжированная переменная и определяет число случайных чисел	2	1-3
37.	Задание открытого типа	Что относится к <i>грубым ошибкам</i>	(1) просчеты экспериментатора (2) сбой вычислительной техники (3) аномалии в работе измерительных приборов	3-5
38.		Лайте определение понятию <i>итерационные решения.</i>	Повторное возвращение к решению той или иной задачи после получения результатов на последующем этапе обработки	3-5
39.		Дайте определение понятию <i>выборочная оценка.</i>	Случайная величина, точность определения которой и возможные при этом ошибки необходимо контролировать	3-5
40.		Дайте определение понятию <i>грубые ошибки.</i>	Отличающиеся большим отклонением от центра группирования выборки	3-5
41.		Дайте определение понятию <i>систематические ошибки.</i>	Отклонения постоянны при определении каждого члена выборки и зависят от технического уровня измерительной аппаратуры и техники эксперимента	
Код и наименование проверяемой компетенции ПК-21				
42.	Задание закрытого типа	<i>Выберите верные ответы.</i> Функция аппроксимации представлена следующими типами функций: 1. кусочно-линейной 2. сплайновой 3. генератором случайных чисел	1, 2	1-3

№ п/п	Тип задания	Формулировка задания	Правильный ответ	Время выполнения (в минутах)
		4. сплайн – корреляционной		
43.		<p><i>Выберите верные ответы.</i></p> <p>При кусочно-линейной интерполяции</p> <ol style="list-style-type: none"> 1. вычисления дополнительных точек выполняются по линейной зависимости 2. узловые точки соединяются отрезками прямых линий 3. исходная функция заменяется отрезками кубических полиномов, 4. узловые точки проходят через три смежные узловые точки 	1, 2	
44.		<p><i>Выберите верные ответы.</i></p> <p>Функции регрессии бывают:</p> <ol style="list-style-type: none"> 1. линейная, полиномиальная 2. линейная общего вида 3. нелинейная общего вида 	1, 2, 3	1-3
45.		<p><i>Выберите верные ответы.</i></p> <p>При сплайн - аппроксимации</p> <ol style="list-style-type: none"> 1. вычисления дополнительных точек выполняются по линейной зависимости 2. узловые точки соединяются отрезками прямых линий 3. исходная функция заменяется отрезками кубических полиномов, 4. узловые точки проходят через три смежные узловые точки 	3, 4	1-3
46.		<p><i>Выберите верный ответ.</i></p> <p>Линейная регрессия</p> <ol style="list-style-type: none"> 1. описывает отрезок прямой 2. функция, которая содержит коэффициенты полинома n-й степени 3. является линейной комбинацией нескольких функций 4. функция должна быть вектором с символьными выражениями 	1	1-3
47.	Задание открытого типа	Дайте определение <i>понятию случайные ошибки.</i>	Случайными называются неопределенные по величине и знаку ошибки, в появлении каждой из которых не наблюдается какой-либо закономерности. Эти ошибки происходят при любом аналитическом определении.	3-5
48.		Что включает в себя второе правило проведения статистических наблюдений?	В программу наблюдений не стоит включать вопросы, на которые не удастся получить ответы удовлетворительного качества	3-5
49.		Что включает в себя третье правило проведения статистических наблюдений?	В программу наблюдений не должны включаться вопросы, которые могут вызвать недоверие обследуемых субъектов относительно целей проведения	3-5

№ п/п	Тип задания	Формулировка задания	Правильный ответ	Время выполнения (в минутах)
			статистического исследования	
50.		Что включает в себя первое правило проведения статистических наблюдений	Программа статистических наблюдений должна включать только те вопросы, на которые необходимо получить ответы	3-5
51.		Дайте определение понятию <i>полиномиальная регрессия</i> .	(1) функция, которая возвращает вектор, содержащий коэффициенты полинома n -й степени (2) функция наилучшим образом приближается к «облаку» точек с координатами, хранящимися в векторах	3-5
Код и наименование проверяемой компетенции ПК-23				
52.	Задание закрытого типа	<i>Выберите верный ответ.</i> Ряд Фурье используется 1. для периодических функций 2. для равномерных функций 3. для аппроксимации функции	1	1-3
53.		<i>Выберите верный ответ.</i> Как задаются пространственные ориентиры? 1. в виде системы координатных сеток 2. в виде совокупности линий 3. в виде совокупности фигур 4. в виде совокупности точек образующих геометрическую фигуру	1	1-3
54.		<i>Выберите верный ответ.</i> Основное достоинство представления данных в графическом виде. 1. наглядность 2. спорность результатов изображений 3. изображение тенденций измерений	1	1-3
55.		<i>Выберите верные ответы.</i> В графике выделяют следующие основные элементы: 1. пространственные ориентиры 2. графический образ 3. масштабные ориентиры	1, 2, 3	1-3
56.		<i>Выберите верный ответ.</i> Как можно представить явление изучаемое статистикой? 1. в графической форме 2. в виде символической графики 3. в виде символической математики	1	1-3
57.	Задание открытого типа	Дайте определение понятию <i>линейная регрессия общего вида</i> .	Является линейной комбинацией нескольких функций, причем они могут быть и нелинейными	3-5
58.		Дайте определение понятию <i>нелиней-</i>	(1) функция должна быть векто-	

№ п/п	Тип задания	Формулировка задания	Правильный ответ	Время выполнения (в минутах)
		<i>ная регрессия общего вида.</i>	ром с символьными выражениями (2) функция представлена символьными выражениями, которые содержат аналитические выражения для исходной функции и ее производных по всем параметрам	
59.		Дайте характеристику функции предсказания.	(1) обеспечивает высокую точность при монотонных исходных функциях представляемых полиномом невысокой степени. (2) функция, применима к предсказуемым событиям, поведение которых описывается реальной математической зависимостью	3-5
60.		Опишите основную задачу интерполяции.	(1) оценить значение представляемой данными зависимости в промежутках между ее узловыми точками (2) для решения использовать подходящие функции, значения которых в узловых точках совпадают с координатами этих точек	3-5
61.		Что называют графиком в статистике?	(1) наглядное изображение статистических величин и их соотношений при помощи геометрических точек (2) наглядное изображение статистических величин и их соотношений при помощи линий, фигур (3) наглядное изображение статистических величин и их соотношений при помощи географических картосхем	3-5

Полный комплект оценочных материалов по дисциплине (модулю) (фонд оценочных средств) хранится в электронном виде на кафедре, утверждающей рабочую программу дисциплины (модуля), и в Центре мониторинга и аудита качества обучения.

7.4. Методические материалы, определяющие процедуры оценивания результатов обучения по дисциплине

Таблица 10. Технологическая карта рейтинговых баллов по дисциплине

№ п/п	Контролируемые мероприятия	Количество мероприятий / баллы	Максимальное количество баллов	Срок представления
Основной блок				
1.	<i>Выполнение лабораторных работ</i>	2/5	10	Указан в Moodle
2.	<i>Написание контрольных работ</i>	2/12	24	Указан в Moodle
3.	<i>Выполнение упражнений</i>	2/3	6	Указан в Moodle
Всего			40	-
Блок бонусов				
4.	<i>Посещение занятий</i>		10	
Всего			10	-
Дополнительный блок**				
5.	<i>Экзамен</i>			
Всего			50	-

№ п/п	Контролируемые мероприятия	Количество мероприятий / баллы	Максимальное количество баллов	Срок представления
ИТОГО			100	-

Таблица 11. Шкала перевода рейтинговых баллов в итоговую оценку за семестр по дисциплине

Сумма баллов	Оценка по 4-балльной шкале
90–100	5 (отлично)
85–89	4 (хорошо)
75–84	
70–74	
65–69	3 (удовлетворительно)
60–64	
Ниже 60	2 (неудовлетворительно)

При реализации дисциплины в зависимости от уровня подготовленности обучающихся могут быть использованы иные формы, методы контроля и оценочные средства, исходя из конкретной ситуации.

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

8.1. Основная литература

1. Hopcroft J. E., Motwani R., Ullman J. D. Introduction to Automata Theory, Languages, and Computation (3rd Edition). — Addison-Wesley, Boston, MA, USA, 2006. — 750 с.
2. Шень А. Программирование: теоремы и задачи. — М.: МЦНМО, 2014. — 296 с.
3. Шень А., Верещагин Н. Языки и исчисления. — М.: МЦНМО, 2012. — 240 с.
4. Верещагин, Н. К. Колмогоровская сложность и алгоритмическая случайность [Электронный ресурс] / Н. К. Верещагин, В. А. Успенский, А. Шень. — Электрон. дан. — СПб: Лань, 2013. — 575 с. — Режим доступа: <https://e.lanbook.com/book/56395> — Загл. с экрана.

8.2. Учебно-методическое обеспечение для самостоятельной работы обучающихся:

1. Кривцова, И. Е. Основы дискретной математики. Часть 1. Учебное пособие [Электронный ресурс] / И. Е. Кривцова, И. С. Лебедев, А. В. Настека. — Электрон. дан. — СПб: ИТМО, 2016. — 92 с. — Режим доступа: http://books.ifmo.ru/book/1869/osnovy_diskretnoy_matematiki_chast_1_uchebnoe_posobie.htm — Загл. с экрана.

8.3. Дополнительная литература

1. Вики-конспекты. — http://neerc.ifmo.ru/wiki/index.php?title=Заглавная_страница

8.4. Интернет-ресурсы, необходимые для освоения дисциплины (модуля) Электронный каталог Научной библиотеки АГУ на базе MARK SQL НПО «Информ-систем»: <https://library.asu.edu.ru>

1. Корпоративный проект Ассоциации региональных библиотечных консорциумов (АРБИКОН) «Межрегиональная аналитическая роспись статей» (МАРС): <http://mars.arbicon.ru>
2. Единое окно доступа к образовательным ресурсам <http://window.edu.ru>

9. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

Для проведения лекционных занятий:

1. Используется аудитория, оборудованная необходимым количеством столов, стульев, доской маркерной и электронной.
2. Аудитория должна иметь следующие нормы освещенности

- СНиП 23-05-95 «Естественное и искусственное освещение» норма освещенности аудиторий ВУЗов 400 Лк.
 - СанПиН 2.2.1/2.1.1.1278-03 «Гигиенические требования к естественному, искусственному и совмещенному освещению жилых и общественных зданий» пункт 3.3.3. «Общее освещение в помещениях общественных зданий должно быть равномерным».
3. Электронная доска должна быть подключена к сети Интернет.
Для проведения **лабораторных занятий**:
1. Лабораторные занятия проводятся с группами или подгруппами не более 15 человек.
 2. Аудитория должна быть оснащена необходимым количеством столов, стульев, доской маркерной и электронной.
 4. Аудитория должна иметь следующие нормы освещенности
 - СНиП 23-05-95 «Естественное и искусственное освещение» норма освещенности аудиторий ВУЗов 400 Лк.
 - СанПиН 2.2.1/2.1.1.1278-03 «Гигиенические требования к естественному, искусственному и совмещенному освещению жилых и общественных зданий» пункт 3.3.3. «Общее освещение в помещениях общественных зданий должно быть равномерным».
 5. В аудитории должно быть не менее 15 компьютеров, находящихся в исправном состоянии.
 6. Расположение компьютеров в аудитории должно позволять преподавателю подойти к рабочему месту студента.
 7. Компьютеры должны быть соединены локальной сетью со скоростью не менее 1 Гбит/с и подключены к сети Интернет.
 8. Наличие видеокамер.
 9. Компьютеры должны обладать минимальными характеристиками:
 - Объем оперативной памяти 16 Гб
 - Накопитель SDD 500 Гб
 - Процессор 12th Gen Intel(R) Core(TM) i3-12100
 - Видеоадаптер Intel(R) UHD Graphics 730

Рабочая программа дисциплины (модуля) при необходимости может быть адаптирована для обучения (в том числе с применением дистанционных образовательных технологий) лиц с ограниченными возможностями здоровья, инвалидов. Для этого требуется заявление обучающихся, являющихся лицами с ограниченными возможностями здоровья, инвалидами, или их законных представителей и рекомендации психолого-медико-педагогической комиссии. Для инвалидов содержание рабочей программы дисциплины (модуля) может определяться также в соответствии с индивидуальной программой реабилитации инвалида (при наличии).