

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

О.Н. Выборнова
«05» мая 2025 г.

УТВЕРЖДАЮ
И.о. заведующего кафедрой
информационной безопасности

В.А. Черкасова
«05» мая 2025 г.

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ(МОДУЛЯ)
«Анализ данных в информационной безопасности»**

Составитель(-и)	Мартьянова А.Е., доцент, к.т.н., доцент кафедры информационной безопасности Барсуков В.А., начальник отдела информационной безопасности Управления корпоративной защиты ООО «Газпром добыча Астрахань»
Согласовано с работодателям	Давидюк Н.В., доцент, к.т.н., заведующий кафедрой «Информационная безопасность», ФГБОУ; Барсуков В.А., начальник отдела информационной безопасности Управления корпоративной защиты ООО «Газпром добыча Астрахань»
Направление подготовки	10.03.01 ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ
Направленность (профиль) ОПОП	«Организация и технологии защиты информации (в сфере информационных и коммуникационных технологий)»
Квалификация (степень)	бакалавр
Форма обучения	очная, очно-заочная
Год приема (курс)	2024
Курс	4 (по очной форме) 5 (по очно-заочной форме)
Семестры	7 (по очной форме) 9 (по очно-заочной форме)

Астрахань, 2025

1. ЦЕЛИ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ(МОДУЛЯ)

1.1. Целью освоения дисциплины «Анализ данных в информационной безопасности» является развитие навыков программирования на языке Python, знакомство с понятиями о сборе, обработке и анализе данных в сфере информационной безопасности, введение в автоматизированные методы работы с данными - машинное обучение и нейронные сети.

1.2. Задачи освоения дисциплины (модуля):

- освоение продвинутых методов исследования взаимосвязей между показателями, характеризующими объекты в социально-экономических исследованиях;
- освоение продвинутых методов распознавания образов и типологизации объектов;
- освоение продвинутых методов оптимизации представления информации об объектах;
- освоение современных пакетов прикладных программ, реализующих алгоритмы многомерного анализа данных;
- приобретение навыков содержательной интерпретации результатов исследования.

2. МЕСТО ДИСЦИПЛИНЫ (МОДУЛЯ) В СТРУКТУРЕ ОПОП

2.1. Учебная дисциплина (модуль) «Анализ данных в информационной безопасности» относится к части плана, формируемой участниками образовательных отношений и осваивается в 7 семестре очной формы обучения и в 9 семестре очно-заочной формы обучения.

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

1. Математика.
2. Информатика.
3. Основы программирования.
4. Вероятностно-статистические методы в анализе данных.

В результате освоения этих дисциплин, студент должен иметь:

Знания:

- основных понятий и методов математического анализа;
- основных понятий и методов аналитической геометрии;
- основных понятий и методов математической статистики и теории вероятности;
- основных понятий и методов линейной алгебры и теории алгебраических систем;
- основных понятий информатики;
- основ программирования.

Умения:

- использовать математические методы и модели для решения прикладных задач;
- использовать языки программирования;
- использовать программные и аппаратные средства персонального компьютера.

Навыки:

- владения методами количественного анализа процессов обработки, поиска и передачи информации;
- поиска информации в глобальной информационной сети Интернет и работы с офисными приложениями (текстовыми процессорами, электронными таблицами, средствами подготовки презентационных материалов, СУБД и т.п.).

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

дисциплиной (модулем):

– Технологии облачных вычислений и виртуализации.

Также дисциплина «Анализ данных в информационной безопасности» поможет студентам при реализации задач преддипломной практики и написанию бакалаврской работы.

3. ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ (МОДУЛЮ)

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

а) профессиональных (ПК):

ПК-2. Способен выполнять работы по установке, настройке и техническому обслуживанию защищенных технических средств обработки информации.

ПК-3. Способен осуществлять внедрение систем защиты информации для обеспечения информационной безопасности автоматизированных систем.

Таблица 1. Декомпозиция результатов обучения

Код компетенции	Код и наименование компетенции	Планируемые результаты обучения по дисциплине (модулю)		
		Знать (1)	Уметь (2)	Владеть (3)
ПК-2	ПК-2.1. Способен выполнять работы по установке, настройке и техническому обслуживанию защищенных технических средств обработки информации.	технические описания и инструкции по эксплуатации технических средств обработки информации в защищенном исполнении, методы контроля защищенности информации от несанкционированного доступа и специальных программных воздействий, порядок аттестации объектов информатизации на соответствие требованиям безопасности информации	проводить настройку защищенных технических средств обработки информации в соответствии с инструкциями по эксплуатации и эксплуатационно-техническими документами; проводить техническое обслуживание защищенных технических средств обработки информации в соответствии с инструкциями по эксплуатации и эксплуатационно-технической документацией	методами защиты информации от несанкционированного доступа и специальных программных воздействий на нее
ПК-3	ПК-3.1. Способен осуществлять внедрение систем защиты информации для обеспечения информационной безопасности автоматизированных систем	основные угрозы безопасности информации и модели нарушителя в автоматизированных системах, содержание эксплуатационной документации автоматизированной системы, типовые средства, методы и	администрировать программные средства системы защиты информации автоматизированных систем, устранять известные уязвимости автоматизированной системы, приводящие к возникновению	методикой анализа структурных и функциональных схем защищенной автоматизированной системы

	протоколы идентификации, аутентификации и авторизации, основные меры по защите информации в автоматизированных системах, нормативные правовые акты в области защиты информации	угроз безопасности информации, применять аналитические и компьютерные модели автоматизированных систем и систем защиты информации, определять параметры настройки программного обеспечения системы защиты информации автоматизированной системы
--	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

4. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

Общая трудоемкость дисциплины в соответствии с учебным планом составляет 4 зачетные единицы (144 часа).

Трудоемкость отдельных видов учебной работы студентов очной, очно-заочной формы обучения приведена в таблице 2.1.

Таблица 2.1. Трудоемкость отдельных видов учебной работы по формам обучения

Вид учебной и внеучебной работы	для очной формы обучения	для очно-заочной формы обучения	для заочной формы обучения
Объем дисциплины в зачетных единицах	4	4	
Объем дисциплины в академических часах	144	144	
Контактная работа обучающихся с преподавателем (всего), в том числе (час.):	55,25	28,25	
- занятия лекционного типа, в том числе:	18	9	
- практическая подготовка (если предусмотрена)	0	0	
- занятия семинарского типа (семинары, практические, лабораторные), в том числе:	36	18	
- практическая подготовка (если предусмотрена)	0	0	
- консультация (предэкзаменационная) ¹	1	1	
- промежуточная аттестация по дисциплине ²	0,25	0,25	

¹ Числовые данные в данной строке соответствуют трудоемкости, указанной в учебном плане в столбце «Конс. (для гр.)»

² Числовые данные в данной строке соответствуют трудоемкости, указанной в учебном плане в столбце

Вид учебной и внеучебной работы	для очной формы обучения	для очно-заочной формы обучения	для заочной формы обучения
Самостоятельная работа обучающихся (час.)	88,75	115,75	
Форма промежуточной аттестации обучающегося (зачет/экзамен), семестр (ы)	экзамен – 7 семестр	экзамен – 9 семестр	

Содержание дисциплины, структурированное по темам (разделам) с указанием отведенного на них количества академических часов и видов учебных занятий и самостоятельной работы представлены в таблице 2.2.

Таблица 2.2. Структура и содержание дисциплины (модуля)
для очной формы обучения

Раздел, тема дисциплины (модуля)	Контактная работа, час.						КР / КП	СР, час.	Итого часов	Форма текущего контроля успеваемости, форма промежуточной аттестации
	Л		ПЗ		ЛР					
	Л	в т.ч. ПП	ПЗ	в т.ч. ПП	ЛР	в т.ч. ПП				
Раздел 1. Основы программирования на языке Python (refresher)	2				4			10	16	Лабораторная работа №1
Раздел 2. Математический аппарат (refresher)	2				4			10	16	Лабораторная работа №2
Раздел 3. Введение в модуль NumPy	2				4			10	16	Лабораторная работа №3
Раздел 4. Основы работы с Pandas	2				4			10	16	Лабораторная работа №4
Раздел 5. Основные этапы анализа данных	2				4			10	16	Лабораторная работа №5
Раздел 6. Визуализация данных. Представление результатов	4				8			20	32	Лабораторная работа №6. Практическая работа №1
Раздел 7. Работа с текстовыми данными. Текстовый анализ	4				8,25			18,75	31	Лабораторная работа №7. Проверочный тест
Консультации										1
Контроль промежуточной аттестации										Экзамен
ИТОГО за семестр:	18				36,25			88,75	143	144

для очно-заочной формы обучения

Раздел, тема дисциплины (модуля)	Контактная работа, час.						КР / КП	СР, час.	Итого часов	Форма текущего контроля успеваемости, форма промежуточно й аттестации
	Л		ПЗ		ЛР					
	Л	в т.ч. ПП	ПЗ	в т.ч. ПП	ЛР	в т.ч. ПП				
Раздел 1. Основы программирования на языке Python (refresher)	1				2			13	16	Лабораторная работа №1
Раздел 2. Математический аппарат (refresher)	1				2			13	16	Лабораторная работа №2
Раздел 3. Введение в модуль NumPy	1				2			13	16	Лабораторная работа №3
Раздел 4. Основы работы с Pandas	1				2			13	16	Лабораторная работа №4
Раздел 5. Основные этапы анализа данных	1				2			13	16	Лабораторная работа №5
Раздел 6. Визуализация данных. Представление результатов	2				4			26	32	Лабораторная работа №6. Практическая работа № 1
Раздел 7. Работа с текстовыми данными. Текстовый анализ	2				4,25			24,75	31	Лабораторная работа №7. Проверочный тест
Консультации									1	
Контроль промежуточной аттестации									Экзамен	
ИТОГО за семестр:	9				18,25			115,75	143	144

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

[При заполнении таблиц 2.2. необходимо учесть следующее:

- заполняются таблицы только по реализуемым формам обучения;
- общий объем часов на каждую тему (раздел) для разных форм обучения должен быть одинаковым;
- практическая подготовка по видам учебных занятий распределяется разработчиком РПД по темам самостоятельно в пределах часов, выделенных в учебном плане на данную дисциплину;
- самостоятельная работа по каждой теме вычисляется как разность между общим объемом часов, выделенных на тему, и количеством часов, выделенных на сумму всех видов контактной работы;
- при подсчете консультаций необходимо учесть, что в случае наличия экзамена по дисциплине проводится одночасовая консультация; разбивать часы на консультации по разделам не нужно;
- при написании курсовой работы на контактную работу с преподавателем отводится 2 часа, объем самостоятельной работы студента на курсовую работу определяется разработчиком; разбивать часы на подготовку курсовой работы по разделам и (или) темам не нужно;

- контроль промежуточной аттестации вносится в соответствующую графу и столбец, разбивать часы на КПА по разделам не нужно.

Далее в данном пункте программы размещается матрица соотношения разделов, тем учебной дисциплины (модуля) и формируемых в них компетенций]

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

Раздел, тема дисциплины (модуля)	Кол-во часов	Код компетенции		Общее количество компетенций
		ПК-2	ПК-3	
Раздел 1. Основы программирования на языке Python (refresher)	16	+	+	2
Раздел 2. Математический аппарат (refresher)	16	+	+	2
Раздел 3. Введение в модуль NumPy	16	+	+	2
Раздел 4. Основы работы с Pandas	16	+	+	2
Раздел 5. Основные этапы анализа данных	16	+	+	2
Раздел 6. Визуализация данных. Представление результатов	32	+	+	2
Раздел 7. Работа с текстовыми данными. Текстовый анализ	31	+	+	2
Консультация	1	+	+	2
ИТОГО	144			

Краткое содержание каждой темы дисциплины (модуля)

Раздел 1. Основы программирования на языке Python (refresher)

Введение в анализ данных на языке Python. Почему Python становится стандартом для работы с большими данными. Прикладные задачи политологов, для решения которых подходит язык Python. Основы программирования на языке Python (refresher): типы данных и методы работы с ними (переменные, листы, словари, кортежи).

Раздел 2. Математический аппарат (refresher)

Математический аппарат для анализа данных: векторы, матрицы, функции и производные. Основы программирования на языке Python: циклы, функции, знание синтаксиса языка.

Раздел 3. Введение в модуль NumPy

Введение в модуль для работы с числовыми данными NumPy (Numerical Python). Особенности типов данных в NumPy. Работа с векторами и матрицами. Вычисление главных статистических метрик с помощью NumPy (среднее, медиана, мода, дисперсия).

Раздел 4. Основы работы с Pandas

Введение в модуль для работы с табличным представлением данных Pandas. Преобразование словарей в табличный формат Pandas, загрузка данных из внешних источников. Особенности фильтрации и обращения к данным. Работа с табличными данными в Pandas на примере.

Раздел 5. Основные этапы анализа данных

Предобработка данных. Разведывательный анализ данных. Статистический анализ данных. Регрессионные модели. Метод наименьших квадратов. Логистическая регрессия. Проведение разведывательного анализа данных на примере.

Раздел 6. Визуализация данных. Представление результатов

Введение в визуализацию данных. Нюансы визуализации данных и принципы человеческого восприятия. Правила создания хороших визуализаций. Создание различных видов визуализаций на синтетических данных и тренировочных наборах данных. Создание интерактивных визуализаций и отчетов с помощью библиотеки Plotly.

Раздел 7. Работа с текстовыми данными. Текстовый анализ

Введение в анализ текста. Особенности подготовки данных для анализа текста. Применение в информационной безопасности.

5. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ПРЕПОДАВАНИЮ И ОСВОЕНИЮ ДИСЦИПЛИНЫ (МОДУЛЯ)

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 темам 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 разделу,
- Заключение,
- рекомендации литературных источников по излагаемым вопросам.

Лабораторные занятия

Лабораторное занятие — целенаправленная форма организации педагогического процесса, направлена 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 в процесс 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 лабораторных занятий должен быть спланирован с расчетом, чтобы за отведенное время они могли быть качественно выполнены большинством учащихся. Лабораторные занятия должны быть так организованы, чтобы студенты ощущали нарастание сложности выполнения заданий, испытывали бы положительные эмоции от переживания собственного успеха в обучении и овладении навыками 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 сформулированными ранее представлениями о порядке и правильности выполнения действий.

В учебном процессе образовательного учреждения выделяются два вида самостоятельной работы:

- 1) аудиторная — выполняется на учебных занятиях, под непосредственным руководством преподавателя и по его заданию (выполнение самостоятельных работ; выполнение контрольных и лабораторных работ; решение задач),
- 2) внеаудиторная — выполняется по заданию преподавателя, но без его непосредственного участия (подготовка к аудиторным занятиям; изучение учебного материала, вынесенного на самостоятельную проработку; выполнение домашних заданий разнообразного 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.

При подготовке к лекционным занятиям необходимо воспользоваться учебно-методической литературой (основной) из п.8.

При подготовке к лабораторным занятиям необходимо воспользоваться учебно-методической литературой (дополнительной) из п.8.

5.2. Указания для обучающихся по освоению дисциплины (модулю)

Лекционные занятия

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 учебниках), а затем посмотреть тему очередной лекции по программе (по плану лекций).

Полезно вести записи (конспекты) лекций: для непонятных вопросов оставлять место при работе над темой лекции с учебными пособиями.

Записи лекций следует вести в отдельной тетради, оставляя место для дополнений во время самостоятельной работы.

При конспектировании лекций выделяйте главы и разделы, параграфы, подчеркивайте основное.

Лабораторные занятия

Лабораторное занятие — наиболее активный вид учебных занятий в вузе. Он предполагает самостоятельную работу над учебными пособиями, основной литературой, открытыми источниками информации.

К каждому лабораторному занятию нужно готовиться. Подготовку следует начинать с повторения теории (по учебному пособию). После этого нужно решать задачи из предложенного домашнего задания.

Самостоятельная работа

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.

Во время самостоятельной работы необходимо воспользоваться учебно-методической литературой из п.8 (основной), (дополнительной), Интернет-ресурсами.

**Таблица 4. Содержание самостоятельной работы обучающихся
для очной формы обучения**

Вопросы, выносимые на самостоятельное изучение	Кол-во часов	Формы работы
Раздел 1. Основы программирования на языке Python (refresher). Подготовка отчета к лабораторной работе 1.	10	Внеаудиторная, изучение учебных пособий
Раздел 2. Математический аппарат (refresher). Подготовка отчета к лабораторной работе 2.	10	Внеаудиторная, изучение учебных пособий
Раздел 3. Введение в модуль NumPy. Подготовка отчета к лабораторной работе 3.	10	Внеаудиторная, изучение учебных пособий
Раздел 4. Основы работы с Pandas. Подготовка отчета к лабораторной работе 4.	10	Внеаудиторная, изучение учебных пособий
Раздел 5. Основные этапы анализа данных. Подготовка отчета к лабораторной работе 5.	10	Внеаудиторная, изучение учебных пособий
Раздел 6. Визуализация данных. Представление результатов. Подготовка отчета к лабораторной работе 6. Подготовка отчета к практической работе 1.	20	Внеаудиторная, изучение учебных пособий
Раздел 7. Работа с текстовыми данными. Текстовый анализ. Подготовка отчета к лабораторной работе 7. Подготовка к проверочному тесту.	18,75	Внеаудиторная, изучение учебных пособий

для очно-заочной формы обучения

Вопросы, выносимые на самостоятельное изучение	Кол-во часов	Формы работы
Раздел 1. Основы программирования на языке Python (refresher). Подготовка отчета к лабораторной работе 1.	13	Внеаудиторная, изучение учебных пособий

Раздел 2. Математический аппарат (refresher). Подготовка отчета к лабораторной работе 2.	13	Внеаудиторная, изучение учебных пособий
Раздел 3. Введение в модуль NumPy. Подготовка отчета к лабораторной работе 3.	13	Внеаудиторная, изучение учебных пособий
Раздел 4. Основы работы с Pandas. Подготовка отчета к лабораторной работе 4.	13	Внеаудиторная, изучение учебных пособий
Раздел 5. Основные этапы анализа данных. Подготовка отчета к лабораторной работе 5.	13	Внеаудиторная, изучение учебных пособий
Раздел 6. Визуализация данных. Представление результатов. Подготовка отчета к лабораторной работе 6. Подготовка отчета к практической работе 1.	26	Внеаудиторная, изучение учебных пособий
Раздел 7. Работа с текстовыми данными. Текстовый анализ. Подготовка отчета к лабораторной работе 7. Подготовка к проверочному тесту.	24,75	Внеаудиторная, изучение учебных пособий

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

Не предусмотрено.

6. ОБРАЗОВАТЕЛЬНЫЕ И ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

При реализации различных видов учебной работы по дисциплине 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 в формах: видеолекций, лекций-презентаций, видеоконференции, собеседования в режиме чат, форума, чата, выполнения виртуальных практических и/или лабораторных работ и др.

Максимальный объем занятий обучающегося с применением электронных образовательных технологий не должен превышать 25%.

Таблица 5. Образовательные технологии, используемые при реализации учебных занятий

Раздел, тема дисциплины (модуля)	Форма учебного занятия		
	Лекция	Практическое занятие, семинар	Лабораторная работа
Раздел 1. Основы программирования на языке Python (refresher)	Обзорная лекция	Не предусмотрено	выполнение лабораторной работы
Раздел 2. Математический аппарат (refresher)	Лекция - презентация	Не предусмотрено	выполнение лабораторной работы
Раздел 3. Введение в модуль NumPy	Лекция - презентация	Не предусмотрено	выполнение лабораторной работы

Раздел 4. Основы работы с Pandas	Лекция - презентация	Не предусмотрено	выполнение лабораторной работы
Раздел 5. Основные этапы анализа данных	Лекция - презентация	Не предусмотрено	выполнение лабораторной работы
Раздел 6. Визуализация данных. Представление результатов	Лекция - презентация	Не предусмотрено	выполнение лабораторной работы
Раздел 7. Работа с текстовыми данными. Текстовый анализ	Обзорная лекция	Не предусмотрено	выполнение лабораторной работы

6.2. Информационные технологии

Название информационной технологии	Темы, разделы дисциплины	Краткое описание применяемой технологии
Использование возможностей Интернета в учебном процессе	1 - 7	Проведение входного, текущего и рейтингового контроля знаний учащихся (в системах дистанционного обучения)
Использование возможностей электронной почты преподавателя	1 - 7	Подготовка к защите отчетов по лабораторным работам
Использование средств представления учебной информации	1 - 7	Использование мультимедийной презентации

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

- - использование возможностей Интернета в учебном процессе (использование информационного сайта преподавателя (рассылка заданий, предоставление выполненных работ, ответы на вопросы, ознакомление учащихся с оценками и т.д.));
- - использование электронных учебников и различных сайтов (например, электронные библиотеки, журналы и т.д.) как источников информации;
- - использование возможностей электронной почты преподавателя;
- - использование средств представления учебной информации (электронных учебных пособий и практикумов, применение новых технологий для проведения очных (традиционных) лекций и семинаров с использованием презентаций и т.д.);
- - 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);
- - использование виртуальной обучающей среды (или системы управления обучением LMS Moodle «Электронное образование») или иных информационных систем, сервисов и мессенджеров.

6.3. Программное обеспечение, 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
Google Chrome	Браузер
Microsoft Office 2013, Microsoft Office Project 2013, Microsoft Office Visio 2013	Офисная программа
7-zip	Архиватор
Microsoft Windows 7 Professional	Операционная система
Kaspersky Endpoint Security	Средство антивирусной защиты

6.3.2. Современные профессиональные базы данных и информационные справочные системы

- 1) Электронный каталог Научной библиотеки АГУ на базе MARK SQL НПО «Информ-систем»: <https://library.asu.edu.ru>.
- 2) Электронный каталог «Научные журналы АГУ»: <http://journal.asu.edu.ru/>.
- 3) Универсальная справочно-информационная полнотекстовая база данных периодических изданий ООО «ИВИС»: <http://dlib.eastview.com/>
- 4) Электронно-библиотечная система eLibrary. <http://elibrary.ru>
- 5) Справочная правовая система КонсультантПлюс: <http://www.consultant.ru>
- 6) Информационно-правовое обеспечение «Система ГАРАНТ»: <http://garant-astrakhan.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 (модулю) «Теория принятия решений и методы оптимизации» проверяется сформированность у обучающихся компетенций, указанных в разделе 3 настоящей программы. Этапность формирования данных компетенций в процессе освоения образовательной программы определяется последовательным освоением дисциплин (модулей) и прохождением практик, а в процессе освоения дисциплины (модуля) – последовательным достижением результатов освоения содержательно связанных между собой разделов, тем.

Таблица 6. Соответствие разделов, тем дисциплины (модуля), результатов обучения по дисциплине (модулю) и оценочных средств

Контролируемые раздел, тема дисциплины (модуля)	Код контролируемой компетенции	Наименование оценочного средства
Раздел 1. Основы программирования на языке	ПК-2, ПК-3	Лабораторная работа 1

Python (refresher)		
Раздел 2. Математический аппарат (refresher)	ПК-2, ПК-3	Лабораторная работа 2
Раздел 3. Введение в модуль NumPy	ПК-2, ПК-3	Лабораторная работа 3
Раздел 4. Основы работы с Pandas	ПК-2, ПК-3	Лабораторная работа 4.
Раздел 5. Основные этапы анализа данных	ПК-2, ПК-3	Лабораторная работа 5
Раздел 6. Визуализация данных. Представление результатов	ПК-2, ПК-3	Лабораторная работа 6
Раздел 7. Работа с текстовыми данными. Текстовый анализ	ПК-2, ПК-3	Лабораторная работа 7. Вопросы к экзамену

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

При решении комплексной ситуационной задачи можно использовать следующие критерии оценки:

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

Шкала оценивания	Критерии оценивания
2 «неудовлетворительно»	не способен правильно выполнить задания

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

Раздел 1. Основы программирования на языке Python (refresher)

Лабораторная работа 1. Основы программирования на языке Python: типы данных и методы работы с ними (переменные, листы, словари, кортежи).

Цель: ознакомиться с основами языка Python, получить умения для выполнения дальнейших лабораторных работ.

Задачи:

- изучить типизацию данных;
- рассмотреть «ветвление» в Python;
- отработать задачи с использованием конструкции «try-except»;

В ней обучающиеся поэтапно проходят все основные аспекта языка, такие как:

1. Типизация данных.
2. Ветвления.
3. Исключения.
4. Переменные, листы, словари, кортежи

Задания для лабораторной работы № 1:

1. Опишите отличия массивов, кортежей, списков и словарей.
2. Приведите пример кода, который записывает/создает текстовый файл, записывает в него две строчки «Hello» и «123», а затем считывает его и выводит его содержимое. Код прокомментируйте.
3. Что такое исключения (ошибки) и как их можно обработать? Что лучше выбрать – написать программу так, чтобы не возникало ошибок, или чтобы ошибки были обработанными?

Раздел 2. Математический аппарат (refresher)

Лабораторная работа 2. Основы программирования на языке Python: циклы, функции, синтаксис языка.

Цель: ознакомиться с основами языка Python: циклы функции, знание синтаксиса языка.

Задачи:

- научиться пользоваться циклами «for» и «while»;
- разобрать функции и пространства имён.

В ней обучающиеся поэтапно проходят все основные аспекта языка, такие как:

1. Пространство имён.
2. Функции.
3. Циклы.

Задания для лабораторной работы № 2:

1. Приведите пример кода, реализующего проверку есть ли словаре запись с каким-то определенным ключом. Например, что контакт есть в телефонной книге.

2. При помощи цикла `for` выведете таблицу умножения для числа 3. Т.е. число 3 должно умножаться на каждое из `[0;9]` чисел и результат – выводится пользователю.

Раздел 3. Введение в модуль NumPy

Лабораторная работа 3. Библиотека NumPy.

Цель: получение представления о функциональности, доступных методах и объектах библиотеки NumPy. Изучение основных принципов практической работы с ними.

Задачи:

1. Объект `ndarray`.
2. Массивы в NumPy.
3. Базовые операции в NumPy.
4. Манипуляции с формой в NumPy.
5. Копии и представления в NumPy.
6. Сохранение массива в файл и чтение из файла в NumPy.

Задания для лабораторной работы № 3:

1. Перечислите наиболее важные атрибуты объектов `ndarray`.
2. Базовая функциональность NumPy – опишите основные доступные функции.
3. Создание массива, заполненного нулями – приведите пример кода.
4. Математические операции между массивами. Представьте код, где массив со значениями от 1 до 9 умножается на константу 2. Должна получиться таблица умножения для числа 2.
5. Слайсы (обрезка массива) – предоставьте код для указанных преподавателем случаев
6. Приведите пример объединения массивов из NumPy.
7. Опишите применение унарных операций к массивам, приведите примеры.
8. Копии и представления при работе с массивами.
9. Сохранение массива в файл и чтение из файла.

Раздел 4. Основы работы с Pandas

Лабораторная работа № 4. Библиотека pandas.

Цель: Получение представления о функциональности, доступных методах и объектах библиотеки pandas. Изучение основных принципов практической работы с ними.

Задачи:

1. Структура данных `Series`.
2. Структура данных `DataFrame`.
3. Доступ к данным в структурах pandas.
4. Добавление элементов в структуры pandas.
5. Обработка отсутствующих данных в pandas.
6. Сохранение объектов в файл и чтение из файла в pandas.

Задания для лабораторной работы № 4:

1. Pandas – основные сведения. Сравнение с функциональностью NumPy.
2. Опишите Конструктор класса `Series`.

3. Приведите пример `Series` и попробуйте: выводить на экран, добавлять элементы, изменять значения.
4. Опишите Конструктор класса `DataFrame`.
5. Повторите пункт 3 для объекта `DataFrame`: выводить на экран, добавлять элементы, изменять значения.
6. Открыть при помощи `pandas` созданный в процессе выполнения работы `dataframe.csv` файл с пропусками или некорректными значениями и исправить их.

Раздел 5. Основные этапы анализа данных

Лабораторная работа № 5. Библиотека `scikit-learn`. Обучение с учителем и без учителя.

Цель занятия:

Получение представления о функциональности, доступных методах и объектах библиотеки `scikit-learn`. Изучение основных принципов практической работы с ними для реализации математических алгоритмов. Рассмотрение основных положений машинного обучения с учителем и без учителя: метод k -средних, линейная регрессия и деревья решений.

Обучение с учителем

Обучение с учителем (англ. `Supervised learning`) – это один из способов машинного обучения, в ходе которого испытуемая система принудительно обучается с помощью примеров «стимул–реакция». К обучению с учителем относятся задачи классификации и регрессии.

Классификация – система группировки объектов исследования или наблюдения в соответствии с их общими признаками. При классификации происходит предсказание признака, множество допустимых значений которого ограничено.

Регрессия – выявление зависимости между случайными переменными и математическим выражением, отражающим связь между зависимой переменной y и независимыми переменными x_i при условии, что это выражение будет иметь статистическую значимость. В отличие от чисто функциональной зависимости $y = f(x_i)$, когда каждому значению независимой переменной x_i соответствует одно определённое значение величины y , при регрессионной связи одному и тому же значению x_i могут соответствовать в зависимости от случая различные значения величины y .

Обучение без учителя

Обучение без учителя (самообучение, спонтанное обучение, англ. `Unsupervised learning`) — один из способов машинного обучения, при котором испытуемая система спонтанно обучается выполнять поставленную задачу без вмешательства со стороны экспериментатора.

Как правило, оно пригодно только для задач, в которых известны описания множества объектов (т.е. существуют обучающие выборки), и требуется обнаружить внутренние взаимосвязи или закономерности, существующие между объектами.

Примерами обучения без учителя являются:

- кластеризация – задача группировки множества объектов на подмножества (кластеры) таким образом, чтобы объекты из одного кластера были более похожи друг на друга, чем на объекты из других кластеров по какому-либо критерию.
- снижение размерности – представление данных в пространстве меньшей размерности с минимальными потерями полезной информации. Обычно в его основе лежит метод главных компонент).
- выявление аномалий – это опознавание во время интеллектуального анализа данных редких данных, событий или наблюдений, которые вызывают подозрения ввиду существенного отличия от большей части данных.

Функциональность Scikit-learn

Библиотека `scikit-learn` реализует следующие основные методы:

1. **Линейные:** модели, задача которых построить разделяющую или аппроксимирующую гиперплоскость (для классификации и регрессии соответственно).
2. **Метрические:** модели, которые вычисляют расстояние по одной из метрик между объектами выборки, и принимают решения в зависимости от этого расстояния (например, метод K-ближайших соседей).
3. **Деревья решений:** обучение моделей, базирующихся на множестве условий, оптимально выбранных для решения задачи.
4. **Ансамблевые методы:** методы, основанные на деревьях решений, которые комбинируют мощь множества деревьев, и таким образом повышают их качество работы, а также позволяют производить отбор признаков (бустинг, бэггинг, случайный лес, мажоритарное голосование).
5. **Нейронные сети:** комплексный нелинейный метод для задач регрессии и классификации.
6. **Метод опорных векторов** (англ. support vector machine, SVM): нелинейный метод, который обучается определять границы принятия решений.
7. **Наивный Байес:** прямое вероятностное моделирование для задач классификации.
8. **Метод главных компонент** (англ. principal component analysis, PCA): линейный метод понижения размерности и отбора признаков.
9. **Стохастическое вложение соседей с t-распределением** (англ. t-distributed Stochastic Neighbor Embedding, t-SNE): нелинейный метод понижения размерности.
10. **K-средних:** самый распространенный метод для кластеризации, требующий на вход число кластеров, по которым должны быть распределены данные.
11. **Кросс-валидация:** метод, при котором для обучения используется весь датасет (в отличие от разбиения на выборки `train/test`), однако обучение происходит многократно, и в качестве валидационной выборки на каждом шаге выступают разные части датасета. Итоговый результат является усреднением полученных результатов.
12. **Поиск по сетке** (англ. Grid Search): метод для нахождения оптимальных гиперпараметров[2] модели путем построения сетки из значений гиперпараметров и последовательного обучения моделей со всеми возможными комбинациями гиперпараметров из сетки.

Задания для лабораторной работы № 5:

1. Классификация – что такое и где применяется.
2. Кластерный анализ – что такое и где применяется, чем отличается от классификации.
3. Линейная регрессия – проведите эксперимент с собственным датасетом (к примеру, подготовьте в Excel датасет случайных значений, лежащих возле какого-либо уравнения).
4. Проведите и проанализируйте собственные эксперименты с K-средних и деревом решений.

Раздел 6. Визуализация данных. Представление результатов

Лабораторная работа № 6. Работа с изображениями и создание различных видов визуализаций

Задание 1. Библиотека TensorFlow Keras. Классификация изображений

Цель: Написание программы для распознавания цифр на базе TensorFlow Keras. Создание нейросети и её обучение по базе данных MNIST. Тестирование MNIST выборкой и пользовательскими картинками.

Создается нейронная сеть, которая классифицирует изображения рукописных цифр. При написании программы используется надстройка «tf.keras», которая является высокоуровневой API для построения и обучения моделей в TensorFlow.

Задачи:

1. Изучение работы с датасетом MNIST.
2. Создание обучающих и тренировочных выборок.
3. Построение модели нейронной сети.
4. Обучение нейронной сети.
5. Проверка точности обученной модели.
6. Предсказание изображений с помощью обученной модели.

Задание 2. Создание различных видов визуализаций на синтетических данных и тренировочных наборах данных

Цель: научить обучающихся основам работы с машинным зрением и показать основные алгоритмы работы с ним.

В последние годы машинное зрение получило большую огласку и вызвало интерес со стороны не только ученых, но и различных инженеров и разработчиков «интеллектуальных» приложений. В лабораторной работе обучающиеся рассмотрят вариант библиотеки OpenCV, написанной на языке «C++» для Python, опробуют некоторый её функционал и протестируют классификатор для распознавания лиц.

Задачи:

1. разобрать импорт и просмотр изображения;
2. разобрать кадрирование;
3. научиться изменять размер изображения;
4. научиться переворачивать изображение;
5. рассмотреть способ преобразование изображения в черно-белое;
6. научиться работать со сглаживанием и размытием;
7. изучить метод распознавания лиц.

Задания для лабораторной работы № 6:

1. Что такое набор данных, или датасет? Для чего может использоваться, как может задаваться (в виде каких типов данных)?
2. Для чего необходимо разбивать датасет на обучающие выборки и валидирующие (тестируемые)? Что дает тестирование?
3. Что такое классы объектов? Зачем им нужны имена? Влияет ли имя класса на обучение, или это дополнительное «удобство» при использовании обученной нейросети?
4. Опишите назначения всех узлов элементов нейросети: входы, синапсы, нейроны, аксоны. Каким узлам соответствуют данные понятия: веса, функции активации?
5. Каким набором вызова функций библиотеки Keras можно определять слои и их параметры (функции активации нейронов на слоях, количество нейронов, тип слоя и прочие параметры)?
6. Подготовьте собственную картинку цифры. Протестируйте на ней классификацию при помощи нейросети.

Практическая работа № 1. Создание интерактивных визуализаций и отчетов с помощью инструмента Plotly

Цель: Получение представления о функциональности, доступных методах и объектах библиотеки Plotly.

Введение в визуализацию данных. Нюансы визуализации данных и принципы человеческого восприятия. Правила создания хороших визуализаций.

Задачи:

1. Рассмотреть особенности подготовки данных для визуализации.
2. Изучить нюансы визуализации данных и принципы человеческого восприятия.
3. Научиться создавать интерактивные визуализации и отчеты с помощью инструмента Plotly.

Раздел 7. Работа с текстовыми данными. Текстовый анализ

Лабораторная работа № 7. Работа с текстовыми данными

Задание1. Модель, классифицирующая отзывы покупателей

Цель: Построение прогностической модели, анализирующей отзывы покупателей о продукции и сортирующую их по двум классам: положительные отзывы и отрицательные. Модель учится на уже классифицированных образцах предсказывать класс других образцов и реализована с помощью scikit-learn.

Задачи:

1. Получение набора образцов.
2. Преобразование текста в числовые векторы признаков.
3. Обучение и оценка модели.

Задание2. Обработка естественного языка

Цель: Научиться предварительно обрабатывать и анализировать содержание текстов, используя библиотеку nltk и метод мешка слов, а также осуществлять перевод текста с помощью библиотеки googletrans.

Задачи:

1. Предварительная оценка текста.
2. Разделение на предложения.

3. Разделение на слова.
4. Перевод в нижний регистр, удаление стоп-слов и знаков пунктуации.
5. Лемматизация.
6. Стемминг.
7. Мешок слов.
8. Определение тональности текста.
9. Перевод и величины достоверности для определения языка.

Задания для лабораторной работы № 7:

1. С какой целью используется разделение текста на предложения и слова? Как называется этот процесс? Приведите пример.
2. С какой целью используется перевод в нижний регистр, удаление стоп-слов и знаков пунктуации? Приведите пример.
3. С какой целью используется лемматизация? Приведите пример.
4. С какой целью используется стемминг? Приведите пример.
5. Опишите принцип метода мешка слов. Приведите пример. С какой целью используется этот метод.
6. Как осуществляется определение тональности текста с помощью библиотеки `nltk`? Приведите примеры.
7. Как осуществляется перевод текста? Приведите примеры.
8. Какие коллизии могут возникать при переводе текстов? Приведите примеры.
9. Как определить величину достоверности для определения языка с которого осуществляется перевод? Приведите примеры.
10. Как определить язык исходного текста? Приведите примеры.
11. Выполнить загрузку из указанного файла так, как это указано в строке кода ниже и осуществить обработку текста.

Проверочный тест

Какой метод используется для предсказания категориальных переменных?

- 1) Линейная регрессия
- 2) Логистическая регрессия
- 3) Кластеризация
- 4) ARIMA

Что такое нормализация данных?

- 1) Процесс удаления выбросов
- 2) Процесс приведения данных к единой шкале
- 3) Процесс кластеризации данных
- 4) Процесс классификации данных

Какой метод используется для предсказания числовых значений?

- 1) Линейная регрессия
- 2) Логистическая регрессия
- 3) Кластеризация
- 4) ARIMA

Выберите верное утверждение:

- 1) Чем ближе к 1 индекс корреляции, тем выше качество модели множественной регрессии.
- 2) Чем ближе к 0 коэффициент детерминации, тем выше качество модели множественной

регрессии.

3) Независимость остатков проверяется с помощью критерия Дарбина – Уотсона.

4) Качество регрессора характеризуется фактом обоснованной зависимости остатков от целевой переменной.

Поставьте в соответствие библиотеку и ее описание

- 1) Scikit-learn
 - 2) NumPy
 - 3) Pandas
 - 4) Seaborn
- a) библиотека для работы с многомерными массивами числовых данных и со сложными математическими операциями.
 - b) библиотека визуализации, основанная на Matplotlib.
 - c) опенсорсная библиотека машинного обучения Python, с широким спектром алгоритмов кластеризации, регрессии и классификации.
 - d) это библиотека с открытым исходным кодом, которая предлагает широкий спектр инструментов для обработки и анализа данных. С ее помощью можно читать данные из широкого спектра источников, таких как CSV, базы данных SQL, файлы JSON и Excel.

Вопросы к экзамену:

1. Язык Python и особенности его стиля программирования. Интерактивный режим Python. Ipython. Jupyter Notebook.
2. Синтаксис и управляющие конструкции языка Python. Переменные, значения и их типы. Типы данных в Python.
3. Встроенные операции и функции. Основные алгоритмические конструкции.
4. Условный оператор. Множественное ветвление.
5. Циклы и счетчики.
6. Определение функций. Параметры и аргументы. Вызовы функций. Оператор возврата. Конструкции *args, **kwargs.
7. Списки, кортежи и словари.
8. Операторы общие для всех типов последовательностей.
9. Специальные операторы и функции для работы со списками. Срезы.
10. Работа со словарями. Методы словарей.
11. Случайные числа. random, randrange, choice.
12. Функции обработки строк. join, replace, split.
13. Стандартная библиотека и pip. Модули и пакеты в Python. Основные стандартные модули.
14. Импортирование модулей. Создание собственных модулей и их импортирование. Специализированные модули и приложения.
15. Файлы и исключения. Работа с внешними источниками данных.
16. Исключения, обработка исключений, вызов исключений (try-except-finally).
17. Утверждения (assert). Открытие, чтение, запись. (open, инструкция with).
18. Работа с текстовыми файлами, xml и csv - файлами.
19. Функциональное программирование. Лямбда-функции.
20. Использование функций map, filter, reduce, zip.
21. Генераторы, декораторы, рекурсия.
22. Модификация функций с помощью декораторов.
23. Итерируемые объекты. Использование генераторов (yield).
24. ООП в Python. Классы, объекты и экземпляры классов. Наследование.
25. Магические методы. Переопределение операторов. Методы классов.

26. Инкапсуляция. Условно частные и строго частные методы.
27. Регулярные выражения. Использование регулярных выражений. Пакет re.
28. Наука о данных и Python. Библиотеки: NumPy, pandas, matplotlib, SciPy.
29. Основы NumPy: массивы и векторные вычисления.
30. Инструменты визуализации данных для Python.
31. Введение в API библиотеки matplotlib.
32. Библиотека pandas. Введение в структуры данных pandas.
33. Объекты Dataframe и Series.
34. Визуализация данных в pandas. Seaborn.
35. Агрегирование данных и групповые операции.
36. Сбор и подготовка данных в Python: извлечение данных с web-страниц (web-scraping). Библиотека BeautifulSoup.
37. Работа с динамическими сайтами с помощью Selenium.
38. Массовый скрепинг с помощью scrapy.
39. Работа со структурированными данными: JSON и XML.
40. Открытые API. Telegram API, VK API.

Таблица 9. Примеры оценочных средств с ключами правильных ответов

№ п/п	Тип задания	Формулировка задания	Правильный ответ	Время выполнения (в минутах)
ПК-2. Способен выполнять работы по установке, настройке и техническому обслуживанию защищенных технических средств обработки информации.				
1)	Задание закрытого типа	Какой метод используется для предсказания категориальных переменных? 1) Линейная регрессия 2) Логистическая регрессия 3) Кластеризация 4) ARIMA	3	2
2)		Что такое нормализация данных? 1) Процесс удаления выбросов 2) Процесс приведения данных к единой шкале 3) Процесс кластеризации данных 4) Процесс классификации данных	2	2
3)		Какой метод используется для предсказания числовых значений? 1) Линейная регрессия 2) Логистическая регрессия 3) Кластеризация 4) ARIMA	1	2
4)		Как связаны между собой AUC и ROC 1. ROC-«кривая ошибок», а AUC – площадь под ней 2. ROC- «кривая правильных ответов» , а AUC- площадь над ней 3. ROC – график, а AUC-оптимальное пороговое значение 4. AUC-«кривая ошибок», а ROC- площадь под ней 5. AUC - «кривая правильных ответов», а ROC - площадь над ней 6. AUC – график, а ROC - оптимальное пороговое значение	1	3
5)		Выберите верное утверждение: 1. Чем ближе к 1 индекс корреляции,	1, 3	3

		<p>тем выше качество модели множественной регрессии.</p> <p>2. Чем ближе к 0 коэффициент детерминации, тем выше качество модели множественной регрессии.</p> <p>3. Независимость остатков проверяется с помощью критерия Дарбина – Уотсона.</p> <p>4. Качество регрессора характеризуется фактом обоснованной зависимости остатков от целевой переменной.</p>		
6)	Задание открытого типа	Что такое точность классификации?	Точность-это доля правильно распознанных экземпляров	5
7)		Какие признаки называются категориальными?	Признаки называются категориальными, если их можно отнести к какой-либо группе, но при этом не важен порядок	6
8)		В чем выражается проблема переобучения?	Модель эффективно работает только с теми данными, на которых была обучена	8
9)		Дайте определение понятию «классификация».	Классификация – это задача отнесения объекта к какому-либо из заранее определенных классов. Относится к типу «обучение с учителем».	8
10)	Задание комбинированного типа	<p>Функция библиотеки Scikit-learn, которая используется для подсчета средней абсолютной ошибки регрессора?</p> <p>а) mean_absolute_error б) predict() в) train_test_split г) accuracy_score д) roc_curve</p>	<p>Функция библиотеки Scikit-learn mean_absolute_error применяется для подсчета средней абсолютной ошибки регрессора.</p>	8
ПК-3. Способен осуществлять внедрение систем защиты информации для обеспечения информационной безопасности автоматизированных систем.				
11)	Задания закрытого типа	<p>Поставьте в соответствие библиотеку и ее описание</p> <p>5) Scikit-learn 6) NumPy 7) Pandas 8) Seaborn</p> <p>е) библиотека для работы с многомерными массивами числовых данных и со сложными</p>	<p>1-с 2-а 3-д 4-б</p>	2

		<p>математическими операциями.</p> <p>f) библиотека визуализации, основанная на Matplotlib.</p> <p>g) опенсорсная библиотека машинного обучения Python, с широким спектром алгоритмов кластеризации, регрессии и классификации.</p> <p>h) это библиотека с открытым исходным кодом, которая предлагает широкий спектр инструментов для обработки и анализа данных. С ее помощью можно читать данные из широкого спектра источников, таких как CSV, базы данных SQL, файлы JSON и Excel.</p>		
12)		<p>С какой целью используется класс OneHotEncoder() библиотеки Scikit Learn?</p> <ol style="list-style-type: none"> 1) Преобразование категориальных признаков в массив чисел. 2) Уменьшение размерности 3) Кодирование категориальных признаков в UTF-8 4) Преобразование количественных признаков в структурированные 	1	2
13)		<p>Под доставкой модели в машинном обучении понимается:</p> <ol style="list-style-type: none"> 1) Отправка модели заказчику 2) Сборка, тестирование и запуск модели в виде ПО с внесенными изменениями 3) Публикация программного кода модели на GitHub 4) Запуск и развертывание модели в виде ПО в фактический продукт 	4	2
14)		<p>Продолжите фразу: "Команда Data Science ...":</p> <ol style="list-style-type: none"> 15) Отвечает за получение и хранение данных заказчика 16) Разрабатывает пользовательский интерфейс для работы с данными 17) Разрабатывает математическое и алгоритмическое решение проблемы 18) Реализует передачу исходных данных разработчикам 	3	2
15)		<p>DevOps сконцентрирован на...</p> <ol style="list-style-type: none"> 1. быстрой и непрерывной поставке работающего программного обеспечения 2. обработке данных, а также возможностях оперативного и безопасного доступа к ним 3. на создании инфраструктуры для корректного хранения, движения и использования данных 	1	2
16)		<p>Опишите функцию библиотеки Scikit-learn, которая используется для разделения множества объектов на обучающее и проверочное множества?</p>	<p>train_test_split.</p> <p>Используется для разделения набора данных на подмножества, чтобы свести к минимуму возможность смещения в</p>	8

			процессе оценки и проверки.	
17)		Опишите функцию библиотеки Scikit-learn, которая используется для определения верности классификации?	accuracy_score. Применяется для оценки качества модели.	8
18)		Опишите функцию библиотеки Scikit-learn, которая используется для анализа ROC-кривых?	roc_curve. Используется для возврата кривой ROC данной модели.	8
19)		Назначение библиотеки pandas	Библиотека pandas предоставляет специальные структуры данных и операции для манипулирования числовыми таблицами и временными рядами: объекты <i>DataFrame</i> и <i>Series</i> .	2
20)	Задание комбинированного типа	Опишите функцию библиотеки Scikit-learn, которая используется для разделения множества объектов на обучающее и проверочное множества? а) mean_absolute_error б) predict() в) train_test_split г) accuracy_score д) roc_curve	Функция библиотеки Scikit-learn train_test_split. Используется для разделения набора данных на подмножества, чтобы свести к минимуму возможность смещения в процессе оценки и проверки.	8

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 (модулю)

Методические рекомендации по выполнению практических работ, проведению экзамена

Отчет по практической работе

Отчет по практической работе представляется в электронном виде. Защита отчета проходит в форме доклада студента по выполненной работе и ответов на вопросы преподавателя. В случае, если оформление отчета и поведение студента во время защиты соответствуют указанным требованиям, студент получает максимальное количество баллов.

Основаниями для снижения количества баллов в диапазоне от max до min являются:

- отсутствие списка использованной литературы,
- небрежное выполнение,

- отсутствие выводов.

Отчет не может быть принят и подлежит доработке в случае:

- отсутствия необходимых разделов,
- отсутствия необходимого графического материала,
- неверных результатов расчета.

В отчете по выполненной практической работе должны быть указаны:

- тема практической работы,
- пакет документов в соответствии с темой практической работы,
- использованная литература.

Критерии оценки по практическим работам:

– оценка «отлично» выставляется обучающемуся, если студент продемонстрировал глубокие знания теоретического материала и умение их применять, обоснованно изложил свои мысли, сделал необходимые выводы и учел основные нормативно-правовые документы по информационной безопасности;

– оценка «хорошо» выставляется обучающемуся, если студент продемонстрировал глубокие знания теоретического материала и умение их применять, обоснованно изложил свои мысли, сделал необходимые выводы и учел основные нормативно-правовые документы по информационной безопасности, допущены некоторые неточности, имеется одна негрубая ошибка.

– оценка «удовлетворительно» выставляется обучающемуся, если студент ответил на вопросы преимущественно верно, имеются затруднения в формулировке выводов, имеются одна или две негрубые ошибки, учтены не все нормативно-правовые документы по информационной безопасности;

– оценка «неудовлетворительно» выставляется обучающемуся, если студент не дал ответы на поставленные вопросы, обоснования неверные, либо дан верный ответ без его обоснования, сделаны грубые ошибки, отсутствуют знания нормативно-правовых документов по информационной безопасности.

Экзамен

Экзамен заключается в письменном ответе на 2 теоретических вопроса и устном собеседовании по каждому теоретическому вопросу.

Основаниями для снижения оценки за теоретический вопрос являются:

- небрежное выполнение;
- неполный ответ;
- наличие мелких неточностей или незначительных искажений фактов;
- неточные объяснения при собеседовании;
- отсутствие ответов на заданные при собеседовании вопросы.

В соответствии с балльно-рейтинговой системой БАРС по дисциплине на экзамен отводится 100 баллов (40 баллов на текущие формы контроля, 10 баллов на бонусы и 50 баллов отводится на экзамен).

Оценивание студентов на экзамене осуществляется в соответствии с требованиями и критериями 100-балльной шкалы. Учитываются как результаты текущего контроля, так и знания, навыки и умения, непосредственно показанные студентами в ходе экзамена.

Критерии оценок на экзамене:

40-50 баллов – студент глубоко понимает пройденный материал, отвечает четко и всесторонне, умеет оценивать факты, самостоятельно рассуждает, отличается способностью обосновать выводы и разьяснять их в логической последовательности.

35-39 баллов – студент глубоко понимает пройденный материал, отвечает четко и всесторонне, умеет оценивать факты, самостоятельно рассуждает, отличается способностью обосновать выводы и разъяснять их в логической последовательности, но допускает отдельные неточности.

25-34 балла – студент глубоко понимает пройденный материал, отвечает четко и всесторонне, умеет оценивать факты, самостоятельно рассуждает, отличается способностью обосновать выводы и разъяснять их в логической последовательности, но допускает некоторые ошибки общего характера.

20-24 балла – студент хорошо понимает пройденный материал, но не может теоретически обосновать некоторые выводы.

15-19 баллов – студент отвечает в основном правильно, но чувствуется механическое заучивание материала. 1

1-14 баллов – в ответе студента имеются существенные недостатки, материал охвачен «половинчато», в рассуждениях допускаются ошибки. 1

0 баллов – ответ студента правилен лишь частично, при разъяснении материала допускаются серьезные ошибки.

6-9 баллов – студент имеет общее представление о теме, но не умеет логически обосновать свои мысли.

1-5 баллов – студент имеет лишь частичное представление о теме. 0 баллов – нет ответа.

Таблица 10. Технологическая карта рейтинговых баллов по дисциплине (модулю)

№ п/п	Контролируемые мероприятия	Количество мероприятий / баллы	Максимальное количество баллов	Срок представления
Основной блок				
1)	<i>Выполнение лабораторной работы</i>	7/6	30	
2)	<i>Выполнение практической работы</i>	1/5	5	
	<i>Проверочный тест</i>	1/5	5	
Всего			40	-
Блок бонусов				
3)	<i>Посещение занятий без пропусков</i>	1	3	
4)	<i>Своевременное выполнение всех заданий</i>	1	3	
5)	<i>Активность студента на занятии</i>	1	4	
Всего			10	-
Дополнительный блок				
6)	<i>Экзамен</i>		50	
Всего			50	-
ИТОГО			100	-

Таблица 11. Система штрафов (для одного занятия)

Показатель	Балл
<i>Опоздание на занятие</i>	- 1
<i>Нарушение учебной дисциплины</i>	- 1
<i>Неготовность к занятию</i>	- 2
<i>Пропуск занятия без уважительной причины</i>	- 2

Таблица 12. Шкала перевода рейтинговых баллов в итоговую оценку за семестр по дисциплине (модулю)

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

При реализации дисциплины (модуля) 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:

2. Кузьмич, Р. И. Модификации метода логического анализа данных для задач классификации : монография / Р. И. Кузьмич, И. С. Масич. — Красноярск : Сибирский федеральный университет, 2018. — 181 с. — ISBN 978-5-7638-3698-1. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <http://www.iprbookshop.ru/84252.html>

3. Воронова, Л. И. Machine Learning: регрессионные методы интеллектуального анализа данных : учебное пособие / Л. И. Воронова, В. И. Воронов. — Москва : Московский технический университет связи и информатики, 2018. — 82 с. — ISBN 2227-8397. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <http://www.iprbookshop.ru/81325.html>

4. Ракитский, А. А. Методы машинного обучения : учебно-методическое пособие / А. А. Ракитский. — Новосибирск : Сибирский государственный университет телекоммуникаций и информатики, 2018. — 32 с. — ISBN 2227-8397. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <http://www.iprbookshop.ru/90591> .

5. Рашка С., Python и машинное обучение: крайне необходимое пособие по новейшей предсказательной аналитике, обязательное для более глубокого понимания методологии машинного обучения / Рашка С. - М. : ДМК Пресс, 2017. - 418 с. - ISBN 978-5-97060-409-0 - Текст : электронный // ЭБС "Консультант студента" : [сайт]. - URL : <https://www.studentlibrary.ru/book/ISBN9785970604090.html>

8.2. Дополнительная литература:

6. Федин, Ф. О. Анализ данных. Часть 1. Подготовка данных к анализу : учебное пособие / Ф. О. Федин, Ф. Ф. Федин. — Москва : Московский городской педагогический университет, 2012. — 204 с. — ISBN 2227-8397. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <http://www.iprbookshop.ru/26444.html>

7. Билл, Фрэнкс Укрошение больших данных : как извлекать знания из массивов информации с помощью глубокой аналитики / Фрэнкс Билл ; перевод А. Баранов. — Москва : Манн, Иванов и Фербер, 2014. — 340 с. — ISBN 978-5-00057-146-0. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <http://www.iprbookshop.ru/39433.html>

8. Флах П., Машинное обучение. Наука и искусство построения алгоритмов, которые извлекают знания из данных / Флах П. - М. : ДМК Пресс, 2015. - 400 с. - ISBN 978-5-97060-273-7 - Текст : электронный // ЭБС "Консультант студента" : [сайт]. - URL : <https://www.studentlibrary.ru/book/ISBN9785970602737.html>

8.3. Интернет-ресурсы, необходимые для освоения дисциплины (модуля)

1. **Электронно-библиотечная система (ЭБС) ООО «Политехресурс» «Консультант студента».** Многопрофильный образовательный ресурс «Консультант студента» является электронной библиотечной системой, предоставляющей доступ через сеть Интернет к учебной литературе и дополнительным материалам, приобретенным на основании прямых договоров с правообладателями. Каталог в настоящее время содержит около 15000 наименований. www.studentlibrary.ru.

2. Электронная библиотека «Астраханский государственный университет» собственной генерации на платформе ЭБС «Электронный Читальный зал – BiblioТех». <https://biblio.asu.edu.ru>

9. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ (МОДУЛЯ)

Для проведения лекционных занятий необходима мультимедийная аудитория, оснащенная компьютерной презентационной техникой.

Для проведения публичной защиты проектов, необходима мультимедийная аудитория с проектором.

Для проведения лабораторных занятий необходима компьютерная аудитория, в которой организован доступ к сети Интернет и установлено программное обеспечение:

10. ОСОБЕННОСТИ РЕАЛИЗАЦИИ ДИСЦИПЛИНЫ (МОДУЛЯ) ПРИ ОБУЧЕНИИ ИНВАЛИДОВ И ЛИЦ С ОГРАНИЧЕННЫМИ ВОЗМОЖНОСТЯМИ ЗДОРОВЬЯ

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. При обучении лиц с ограниченными возможностями здоровья учитываются их индивидуальные психофизические особенности. Обучение инвалидов осуществляется также в соответствии с индивидуальной программой реабилитации инвалида (при наличии).

Для лиц с нарушением слуха возможно предоставление учебной информации в визуальной форме (краткий конспект лекций; тексты заданий, напечатанные увеличенным шрифтом), на аудиторных занятиях допускается присутствие ассистента, а также сурдопереводчиков и тифлосурдопереводчиков. Текущий контроль успеваемости осуществляется в письменной форме: обучающийся письменно отвечает на вопросы, письменно выполняет практические задания. Доклад (реферат) также может быть представлен в письменной форме, при этом требования к содержанию остаются теми же, а требования к качеству изложения материала (понятность, качество речи, взаимодействие с аудиторией и т. д.) заменяются на соответствующие требования, предъявляемые к письменным работам (качество оформления текста и списка литературы, грамотность, наличие иллюстрационных материалов и т.д.). Промежуточная аттестация для лиц с нарушениями слуха проводится в письменной форме, при этом используются общие критерии оценивания. При необходимости время подготовки к ответу может быть увеличено.

Для лиц с нарушением зрения допускается аудиальное предоставление информации, а также использование на аудиторных занятиях звукозаписывающих устройств (диктофонов и т.д.). Допускается присутствие на занятиях ассистента (помощника), оказывающего обучающимся необходимую техническую помощь. Текущий контроль успеваемости осуществляется в устной форме. При проведении промежуточной

аттестации для лиц с нарушением зрения тестирование может быть заменено на устное собеседование по вопросам.

Для лиц с ограниченными возможностями здоровья, имеющих нарушения опорно-двигательного аппарата, на аудиторных занятиях, а также при проведении процедур текущего контроля успеваемости и промежуточной аттестации могут быть предоставлены необходимые технические средства (персональный компьютер, ноутбук или другой гаджет); допускается присутствие ассистента (ассистентов), оказывающего обучающимся необходимую техническую помощь (занять рабочее место, передвигаться по аудитории, прочитать задание, оформить ответ, общаться с преподавателем).