

Министерство науки и высшего образования Российской Федерации
Департамент координации деятельности организаций в сфере
сельскохозяйственных наук
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Волгоградский государственный аграрный университет»

Эколого-мелиоративный
наименование факультета

УТВЕРЖДАЮ
Декан эколого-мелиоративного
наименование факультета

О.А. Корчагина
подпись *инициалы фамилия*

26 октября 2022 г.



ДОКУМЕНТ ПОДПИСАН
ЭЛЕКТРОННОЙ ПОДПИСЬЮ
ВОЛГОГРАДСКИЙ ГАУ

СВЕДЕНИЯ О СЕРТИФИКАТЕ ЭП

Сертификат: 617a770026af82a74a598c23838b44c5
Владелец: Корчагина Ольга Александровна
Действителен: с 06.10.2022 по 06.10.2023

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ
Б1.В.13 Анализ данных и машинное обучение

индекс и наименование дисциплины

Кафедра Информационные системы и
технологии

наименование кафедры

Уровень высшего образования бакалавриат

бакалавриат / специалитет / магистратура

Направление подготовки (специальность) 09.03.03 Прикладная информатика

шифр и наименование направления подготовки

Направленность (профиль) Прикладная информатика в инновационной деятельности

наименование направленности (профиля) программы

Форма обучения очная/заочная

очная / очно-заочная / заочная

Год начала реализации образовательной программы 2022

Волгоград

2022

Автор(ы):

доцент

должность

подпись

С.С. Марченко

инициалы фамилия

Рабочая программа дисциплины согласована с руководителем основной профессиональной образовательной программы высшего образования по направлению подготовки (специальности)

09.03.03 Прикладная информатика

шифр и наименование направления подготовки (специальности)

Прикладная информатика в инновационной деятельности

наименование направленности (профиля) программы

Зав. кафедрой

должность

подпись

О.В. Кочеткова

инициалы фамилия

Рабочая программа дисциплины обсуждена и одобрена на заседании кафедры

Информационные системы и технологии

наименование кафедры

Протокол № _____ от _____ г.

дата

Заведующий кафедрой

подпись

О.В. Кочеткова

инициалы фамилия

Рабочая программа дисциплины обсуждена и одобрена на заседании
методической комиссии эколого-мелиоративного
факультета

наименование факультета

Протокол № _____ от _____ г.

дата

Председатель

методической комиссии факультета

подпись

А.К. Васильев

инициалы фамилия

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

Целью изучения дисциплины «Анализ данных и машинное обучение» является сформировать компетенции обучающегося в области фундаментальных представлений в области современных технологий для обработки и анализа информации и эффективных методов ее обработки с применением современных ЭВМ методами машинного обучения.

Изучение дисциплины направлено на решение следующих задач:

- Получить представление о методах сбора данных из различных источников.
- Овладеть навыком подготовки наборов данных к анализу.
- Освоить специальные библиотеки для анализа данных и машинного обучения.
- Получить представление о методах машинного обучения.
- Получить навыки решения прикладных задач с использованием методов машинного обучения

В результате изучения дисциплины обучающиеся должны приобрести следующие знания, умения, навыки:

Код и наименование компетенции	Код и наименование индикатора достижения компетенции	Планируемые результаты обучения по дисциплине
ПК-1. Способен осуществлять определение первоначальных требований заказчика к ИС и возможности их реализации в ИС на этапе предконтрактных работ	ПК-1.2 Анализирует большие объемы данных на основе методов анализа данных и машинного обучения при обработке информации на этапе предконтрактных работ	Знает методы анализа данных <u>при предконтрактных работах</u>
		Умеет пользоваться специализированными программными инструментами для анализа данных
		Владеет навыками подготовки датасетов для анализа методами машинного обучения
ПК-2. Способен осуществлять определение первоначальных требований заказчика к ИС и возможности их реализации в ИС на этапе предконтрактных работ	ПК-2.10 Разрабатывает архитектуру ИС с учетом методов машинного обучения	Знает методы анализа данных
		Умеет пользоваться специализированными программными инструментами для анализа данных
		Владеет методами машинного обучения

ПК-3. Способен осуществлять организационное и технологическое обеспечение интеграционного тестирования ИС (верификации)	ПК-3.5 Анализирует исходные данные и обладает знаниями в предметной области автоматизации	Знает методы анализа многомерных данных; методы снижения размерности данных и отбора информативных признаков; методы кластеризации; методы классификации; методы регрессионного анализа;
		Умеет пользоваться специализированными программными инструментами для анализа данных
		Владеет методами машинного обучения для решения прикладных задач предметной области

Основными этапами формирования компетенций при изучении дисциплины является последовательное освоение содержательно связанных между собой разделов и тем дисциплины.

2 Место дисциплины в структуре образовательной программы

Дисциплина Анализ данных и машинное обучение (Б1.В.13) относится к дисциплинам обязательной части Блока 1 «Дисциплины (модули)» учебного плана подготовки бакалавров по направлению 09.03.03 «Прикладная информатика» направленность (профиль) «Инжиниринг информационных систем».

Место дисциплины в структуре образовательной программы

Индекс и наименование дисциплины (модуля), практики, участвующих в формировании компетенций	Форма обучения	Курсы обучения*				
		1 курс	2 курс	3 курс	4 курс	5 курс
ПК-1. Способен осуществлять определение первоначальных требований заказчика к ИС и возможности их реализации в ИС на этапе предконтрактных работ						
Б1.О.02 Экономическая теория	Очная	+				
	Заочная	+				
Б1.О.06 Право	Очная		+			
	Заочная		+			
Б1.О.11 Информационные системы и технологии	Очная	+	+			
	Заочная	+	+			
Б1.О.13 Операционные системы	Очная		+			
	Заочная		+			
Б1.О.14 Вычислительные системы, сети	Очная	+				

и телекоммуникации	Заочная	+				
Б1.О.15 Теория систем и системный анализ	Очная	+				
	Заочная	+				
Б1.О.16 Базы данных	Очная		+			
	Заочная		+			
Б1.О.17 Русский язык и психология деловых коммуникаций	Очная		+			
	Заочная			+		
Б1.О.18 Проектирование информационных систем	Очная		+	+		
	Заочная			+	+	
Б1.О.19 Менеджмент	Очная			+		
	Заочная			+		
Б1.О.20 Информационная безопасность	Очная		+			
	Заочная			+		
Б1.О.21 Программная инженерия	Очная			+		
	Заочная			+		
Б1.В.01 Основы киберфизических систем	Очная				+	
	Заочная				+	
Б1.В.02 Математическое и имитационное моделирование	Очная			+		
	Заочная			+		
Б1.В.03 Основы компьютерной электроники	Очная	+				
	Заочная	+				
Б1.В.04 Интернет-программирование	Очная			+		
	Заочная				+	
Б1.В.07 Интеллектуальные информационные системы	Очная		+	+		
	Заочная				+	
Б1.В.09 Реинжиниринг и управление бизнес-процессами	Очная				+	
	Заочная					+
Б1.В.10 Информационные системы бухгалтерского учета	Очная				+	
	Заочная					+
Б1.В.11 Бизнес-инжиниринг	Очная	+				
	Заочная	+				
Б1.В.12 Управление требованиями	Очная		+			
	Заочная			+		
Б1.В.13 Анализ данных и машинное обучение	Очная		+			
	Заочная		+			
Б1.В.14 Автоматизированные системы управления проектами	Очная				+	
	Заочная					+
Б1.В.15 Основы робототехники в агропромышленном комплексе	Очная	+				
	Заочная		+			
Б1.В.16 3D-моделирование и дополненная реальность	Очная				+	
	Заочная					+
Б1.В.17 Цифровые технологии в агропромышленном комплексе	Очная				+	
	Заочная					+

Б1.В.18 Основы проектирования приложений интернета вещей в агропромышленном комплексе	Очная			+	+	
	Заочная					+
Б1.В.19 Основы бухгалтерского учета	Очная		+			
	Заочная			+		
Б1.В.ДВ.01.01 Численные методы	Очная			+		
	Заочная				+	
Б1.В.ДВ.01.02 Математические методы в инженерных и экономических расчетах	Очная			+		
	Заочная				+	
Б1.В.ДВ.02.01 Сетевое администрирование	Очная				+	
	Заочная					+
Б1.В.ДВ.02.02 Компьютерные системы и сети	Очная				+	
	Заочная					+
Б1.В.ДВ.03.01 Компьютерная графика	Очная	+				
	Заочная		+			
Б1.В.ДВ.03.02 Геоинформационные системы	Очная	+				
	Заочная		+			
Б1.В.ДВ.04.01 Электронный документооборот	Очная	+				
	Заочная		+			
Б1.В.ДВ.04.02 Документационное обеспечение информационных систем	Очная	+				
	Заочная		+			
Б2.В.01(П) Технологическая (проектно-технологическая) практика	Очная				+	
	Заочная					+
Б2.В.02(П) Эксплуатационная практика	Очная			+		
	Заочная				+	
Б2.В.03(П) Преддипломная практика	Очная				+	
	Заочная					+
Б3.01(Д) Подготовка к сдаче и сдача государственного экзамена	Очная				+	
	Заочная					+
Б3.02(Д) Выполнение и защита выпускной квалификационной работы	Очная				+	
	Заочная					+
ФТД.01 Моделирование бизнес-процессов	Очная				+	
	Заочная					+
ФТД.02 Информационные системы управления производственной компанией	Очная				+	
	Заочная					+
ПК-2. Способен осуществлять разработку архитектуры ИС						
Б1.О.06 Право	Очная		+			
	Заочная		+			
Б1.О.11 Информационные системы и технологии	Очная	+	+			
	Заочная	+	+			

Б1.О.13 Операционные системы	Очная		+			
	Заочная		+			
Б1.О.14 Вычислительные системы, сети и телекоммуникации	Очная	+				
	Заочная	+				
Б1.О.16 Базы данных	Очная		+			
	Заочная		+			
Б1.О.17 Русский язык и психология деловых коммуникаций	Очная		+			
	Заочная			+		
Б1.О.18 Проектирование информационных систем	Очная		+	+		
	Заочная			+	+	
Б1.О.19 Менеджмент	Очная			+		
	Заочная			+		
Б1.О.20 Информационная безопасность	Очная		+			
	Заочная			+		
Б1.О.21 Программная инженерия	Очная			+		
	Заочная			+		
Б1.В.01 Основы киберфизических систем	Очная				+	
	Заочная				+	
Б1.В.02 Математическое и имитационное моделирование	Очная			+		
	Заочная			+		
Б1.В.03 Основы компьютерной электроники	Очная	+				
	Заочная	+				
Б1.В.04 Интернет-программирование	Очная			+		
	Заочная				+	
Б1.В.06 Системная архитектура информационных систем	Очная			+	+	
	Заочная					+
Б1.В.07 Интеллектуальные информационные системы	Очная		+	+		
	Заочная				+	
Б1.В.10 Информационные системы бухгалтерского учета	Очная				+	
	Заочная					+
Б1.В.13 Анализ данных и машинное обучение	Очная		+			
	Заочная			+		
Б1.В.15 Основы робототехники в агропромышленном комплексе	Очная	+				
	Заочная		+			
Б1.В.17 Цифровые технологии в агропромышленном комплексе	Очная				+	
	Заочная					+
Б1.В.18 Основы проектирования приложений интернета вещей в агропромышленном комплексе	Очная			+	+	
	Заочная					+
Б1.В.19 Основы бухгалтерского учета	Очная		+			
	Заочная			+		
Б1.В.ДВ.01.01 Численные методы	Очная			+		

	Заочная				+	
Б1.В.ДВ.01.02 Математические методы в инженерных и экономических расчетах	Очная			+		
	Заочная				+	
Б1.В.ДВ.02.01 Сетевое администрирование	Очная				+	
	Заочная					+
Б1.В.ДВ.02.02 Компьютерные системы и сети	Очная				+	
	Заочная					+
Б1.В.ДВ.03.01 Компьютерная графика	Очная	+				
	Заочная		+			
Б1.В.ДВ.03.02 Геоинформационные системы	Очная	+				
	Заочная		+			
Б1.В.ДВ.04.01 Электронный документооборот	Очная	+				
	Заочная		+			
Б1.В.ДВ.04.02 Документационное обеспечение информационных систем	Очная	+				
	Заочная		+			
Б2.В.01(П) Технологическая (проектно-технологическая) практика	Очная				+	
	Заочная					+
Б2.В.02(П) Эксплуатационная практика	Очная			+		
	Заочная				+	
Б2.В.03(П) Преддипломная практика	Очная				+	
	Заочная					+
Б3.01(Д) Подготовка к сдаче и сдача государственного экзамена	Очная				+	
	Заочная					+
Б3.02(Д) Выполнение и защита выпускной квалификационной работы	Очная				+	
	Заочная					+

Для успешного освоения дисциплины «Анализ данных и машинное обучение» (Б1.В.13) необходимо обладать знаниями, умениями, навыками, полученными при изучении таких дисциплин, как «Экономическая теория» (Б1.О.02), «Право» (Б1.О.06). Минимальными требованиями к «входным» знаниям, умениям, навыкам, необходимым для изучения данной дисциплины, является удовлетворительное освоение учебной программы по указанным выше дисциплинам. В свою очередь знания, умения, навыки, полученные в ходе изучения дисциплины «Анализ данных и машинное обучение» (Б1.В.13), будут полезными при изучении таких дисциплин как: «Базы данных» (Б1.О.16), «Русский язык и психология деловых коммуникаций» (Б1.О.17), «Проектирование информационных систем» (Б1.О.18), «Менеджмент» (Б1.О.19), «Программная инженерия» (Б1.О.21), «Основы киберфизических систем» (Б1.В.01), «Математическое и имитационное моделирование» (Б1.В.02), «Основы компьютерной электроники» (Б1.В.03), «Интернет-программирование» (Б1.В.04), «Интеллектуальные информационные системы» (Б1.В.07),

«Реинжиниринг и управление бизнес-процессами» (Б1.В.09), «Информационные системы бухгалтерского учета» (Б1.В.10), «Бизнес-инжиниринг» (Б1.В.11), «Управление требованиями» (Б1.В.12), «Автоматизированные системы управления проектами» (Б1.В.14), «Основы робототехники в агропромышленном комплексе» (Б1.В.15), «3D-моделирование и дополненная реальность» (Б1.В.16), «Цифровые технологии в агропромышленном комплексе» (Б1.В.17), «Основы проектирования приложений интернета вещей в агропромышленном комплексе» (Б1.В.18), «Основы бухгалтерского учета» (Б1.В.19), «Численные методы» (Б1.В.ДВ.01.01), «Математические методы в инженерных и экономических расчетах» (Б1.В.ДВ.01.02), «Сетевое администрирование» (Б1.В.ДВ.02.01), «Компьютерные системы и сети» (Б1.В.ДВ.02.02), «Компьютерная графика» (Б1.В.ДВ.03.01), «Геоинформационные системы» (Б1.В.ДВ.03.02), «Электронный документооборот» (Б1.В.ДВ.04.01), «Документационное обеспечение информационных систем» (Б1.В.ДВ.04.02), «Моделирование бизнес-процессов» (ФТД.01), «Информационные системы управления производственной компанией» (ФТД.02), «Системная архитектура информационных систем» (Б1.В.06), прохождении таких практик, как «Технологическая (проектно-технологическая) практика» (Б2.О.03(У)), «Технологическая (проектно-технологическая) практика» (Б2.В.01(П)), «Эксплуатационная практика» (Б2.В.02(П)), «Преддипломная практика» (Б2.В.03(П)), «Эксплуатационная практика» (Б2.О.02(У)), «Преддипломная практика» (Б2.В.03(П)), а также при написании выпускной квалификационной работы «Выполнение и защита выпускной квалификационной работы» (Б3.02(Д)) и сдаче государственного экзамена «Подготовка к сдаче и сдача государственного экзамена» (Б3.01(Д)).

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

Очная форма обучения

Вид учебной работы	Всего часов	Распределение часов по семестрам			
		1	2	3	4
Контактная работа обучающихся с преподавателем (по учебным занятиям), всего**	64	-	-	64	-
Лекционные занятия	16	-	-	16	-
в том числе в форме практической подготовки	-	-	-	-	-
Практические (семинарские) занятия	48	-	-	48	-
в том числе в форме практической подготовки	-	-	-	-	-
Лабораторные занятия	-	-	-	-	-

в том числе в форме практической подготовки	-	-	-	-	-
Самостоятельная работа обучающихся, всего**	80	-	-	80	-
Выполнение курсовой работы	-	-	-	-	-
Выполнение курсового проекта	-	-	-	-	-
Выполнение расчетно-графической работы	-	-	-	-	-
Выполнение реферата	-	-	-	-	-
Самостоятельное изучение разделов и тем	80	-	-	80	-
Промежуточная аттестация***	-	-	-	-	-
Экзамен	-	-	-	-	-
Зачет с оценкой	-	-	-	-	-
Зачет	0	-	-	0	-
Курсовая работа / Курсовой проект	-	-	-	-	-
Общая трудоемкость	часов	144	-	-	144
	зачетных единиц	4	-	-	4

Заочная форма обучения

Вид учебной работы	Всего часов	Распределение часов по сессиям*			
		3	4	5	6
Контактная работа обучающихся с преподавателем (по учебным занятиям), всего**	12	-	-	-	-
Лекционные занятия	4	-	-	-	4
в том числе в форме практической подготовки	-	-	-	-	-
Практические (семинарские) занятия	8	-	-	-	8
в том числе в форме практической подготовки	-	-	-	-	-
Лабораторные занятия	-	-	-	-	-
в том числе в форме практической подготовки	-	-	-	-	-
Самостоятельная работа обучающихся, всего**	128	-	-	-	128
Выполнение курсовой работы	-	-	-	-	-
Выполнение курсового проекта	-	-	-	-	-
Выполнение расчетно-графической работы	-	-	-	-	-
Выполнение реферата	-	-	-	-	-
Выполнение контрольной работы	-	-	-	-	-
Самостоятельное изучение разделов и тем	128	-	-	-	128
Промежуточная аттестация***	-	-	-	-	-
Экзамен	-	-	-	-	-
Зачет с оценкой	-	-	-	-	-
Зачет	4	-	-	-	4
Курсовая работа / Курсовой проект	-	-	-	-	-

Общая трудоемкость	часов	144	-	-	-	144
	зачетных единиц	4	-	-	-	4

4 Содержание дисциплины, структурированное по разделам и темам с указанием отведенного на них количества академических часов и видов учебных занятий

4.1 Тематический план дисциплины

Очная форма обучения

Наименование разделов и тем дисциплины	Контактная работа (по учебным занятиям)						Самостоятельное изучение разделов и тем
	Лекционные занятия	в том числе в форме практической подготовки	Практические (семинарские) занятия	в том числе в форме практической подготовки	Лабораторные занятия	в том числе в форме практической подготовки	
Тема 1. Введение в предметную область.	1	-	-	-	4	-	8
Тема 2. Знакомство со специализированными библиотеками языка программирования Python для научных расчетов и анализа данных. NumPy, SciPy, pandas.	1	-	-	-	4	-	8
Тема 3. Библиотеки для графического анализа данных	1	-	-	-	4	-	8

matplotlib и seaborn.							
Тема 4. Знакомство с различными методами предобработки данных, описательными статистиками и основными способами визуализации данных, методами снижения размерности. Метод главных компонент. Важность нормировки данных. Предобработка данных. Работа с пропущенными значениями, выбросами, нулевыми значениями.	2	-	-	-	4	-	8
Тема 5. Основы машинного обучения и основные типы задач. Классификация задач машинного обучения	1	-	-	-	4	-	8
Тема 6.	2				4		8

Обучение на неразмеченн ых данных. Кластеризац ия. Иерархическ ая кластеризац ия. Метод К- средних, DBSCAN и др. Обзор методов кластеризац ии, реализованн ых в библиотеке sklearn							
Тема 7. Задачи обучения с учителем. Разделение данных на обучающие и тестовые. Нормировка данных. Определение переобученн ости модели. Критерии оценки качества полученных моделей	2				6		8
Тема 8 Постановка задачи регрессии. Линейный регрессионн ый анализ. Отбор	2				6		8

признаков, коллинеарность, влиятельные наблюдения, анализ остатков. Непараметрическая регрессия (ядерное сглаживание). L1 и L2 регуляризация. Метрики качества							
Тема 9. Постановка задачи классификации, обзор основных методов ее решения. Бинарная и многоклассовая классификация. Логистическая регрессия. Решающие деревья. Метрики качества классификации (точность/специфичность, ROC-кривая, площадь под кривой)	2				6		8
Тема 10.	2				6		8

Ансамбли алгоритмов машинного обучения. Агрегирование моделей. Ансамбли решающих деревьев. Метод случайного леса. Градиентный бустинг.							
Итого по дисциплине	16	-	-	-	48	-	80

Заочная форма обучения

Наименование разделов и тем дисциплины	Контактная работа (по учебным занятиям)						Самостоятельное изучение разделов и тем
	Лекционные занятия	в том числе в форме практической подготовки	Практические (семинарские) занятия	в том числе в форме практической подготовки	Лабораторные занятия	в том числе в форме практической подготовки	
Тема 1. Введение в предметную область.	-	-	-	-	-	-	6
Тема 2. Знакомство со специализированными библиотеками языка программирования Python для научных расчетов и	1	-	-	-	2	-	14

анализа данных. NumPy, SciPy, pandas.							
Тема 3. Библиотеки для графическог о анализа данных matplotlib и seaborn.	-	-	-	-	-	-	8
Тема 4. Знакомство с различными методами предобработ ки данных, описательны ми статистикам и и основными способами визуализаци и данных, методами снижения размерности. Метод главных компонент. Важность нормировки данных. Предобработ ка данных. Работа с пропущенны ми значениями, выбросами, нулевыми значениями.	1	-	-	-	-	-	14

Тема 5. Основы машинного обучения и основные типы задач. Классифика ция за-дач машинного обучения	1	-	-	-	2	-	14
Тема 6. Обучение на неразмеченн ых данных. Кластеризац ия. Иерархическ ая кластеризац ия. Метод К- средних, DBSCAN и др. Обзор методов кластеризац ии, реализованн ых в библиотеке sklearn	-						14
Тема 7. Задачи обучения с учителем. Разделение данных на обучающие и тестовые. Нормировка данных. Определение переобученн ости модели. Критерии оценки	1				2		14

качества полученных моделей							
Тема 8 Постановка задачи регрессии. Линейный регрессионн ый анализ. Отбор признаков, коллинеарно сть, влиятельные наблюдения, анализ остатков. Непараметри ческая регрессия (ядерное сглаживание). L1 и L2 регуляриза- ция. Метрики качества	-				2		14
Тема 9. Постановка задачи классификац ии, обзор основных методов ее решения. Бинарная и многоклассо вая классификац ия. Логистическ ая регрессия. Решающие	-						14

деревья. Метрики качества классификац ии (точ- ность/специ фичность, ROC-кривая, площадь под кривой)							
Тема 10. Ансамбли алгоритмов машинного обучения. Агрегирован ие моделей. Ансамбли решающих деревьев. Метод случайного леса. Градиентны й бустинг.	-						14
Итого по дисциплине	4	-	-	-	8	-	128

4.2 Содержание дисциплины

Тема 1. Введение в предметную область. Примеры использования методов машинного обучения для решения прикладных задач. Повторение основ программирования на языке Python.

Тема 2. Знакомство со специализированными библиотеками языка программирования Python для научных расчетов и анализа данных. NumPy, SciPy, pandas.

Тема 3. Библиотеки для графического анализа данных matplotlib и seaborn.

Тема 4. Знакомство с различными методами предобработки данных, описательными статистиками и основными способами визуализации данных, методами снижения размерности. Метод главных компонент. Важность нормировки данных. Предобработка данных. Работа с пропущенными значениями, выбросами, нулевыми значениями.

Тема 5. Основы машинного обучения и основные типы задач. Классификация задач машинного обучения.

Тема 6. Обучение на неразмеченных данных. Кластеризация. Иерархическая кластеризация. Метод К-средних, DBSCAN и др. Обзор методов кластеризации, реализованных в библиотеке sklearn.

Тема 7. Задачи обучения с учителем. Разделение данных на обучающие и тестовые. Нормировка данных. Определение переобученности модели. Критерии оценки качества полученных моделей.

Тема 8. Постановка задачи регрессии. Линейный регрессионный анализ. Отбор признаков, коллинеарность, влиятельные наблюдения, анализ остатков. Непараметрическая регрессия (ядерное сглаживание). L1 и L2 регуляризация. Метрики качества.

Тема 9. Постановка задачи классификации, обзор основных методов ее решения. Бинарная и многоклассовая классификация. Логистическая регрессия. Решающие деревья. Метрики качества классификации (точность/специфичность, ROC-кривая, площадь под кривой).

Тема 10. Ансамбли алгоритмов машинного обучения. Агрегирование моделей. Ансамбли решающих деревьев. Метод случайного леса. Градиентный бустинг.

5 Текущий контроль успеваемости и промежуточная аттестация по дисциплине

Средства и контрольные мероприятия, необходимые для оценки знаний, умений, навыков, приобретенных в результате изучения дисциплины

Наименование разделов и тем дисциплины	Формы оценочных средств текущего контроля	Формы промежуточной аттестации
Тема 1. Введение в предметную область.	Тестирование	Зачет
Тема 2. Знакомство со специализированными библиотеками языка программирования Python для научных расчетов и анализа данных. NumPy, SciPy, pandas.	Тестирование	
Тема 3. Библиотеки для графического анализа данных matplotlib и seaborn.	Контрольная работа	
Тема 4. Знакомство с различными методами предобработки данных, описательными статистиками и основными способами визуализации данных, методами снижения размерности. Метод главных компонент. Важность нормировки данных. Предобработка данных. Работа с пропущенными значениями, выбросами, нулевыми значениями.	Тестирование	

Тема 5. Основы машинного обучения и основные типы задач. Классификация задач машинного обучения	Контрольная работа	
Тема 6. Обучение на неразмеченных данных. Кластеризация. Иерархическая кластеризация. Метод К-средних, DBSCAN и др. Обзор методов кластеризации, реализованных в библиотеке sklearn	Контрольная работа	
Тема 7. Задачи обучения с учителем. Разделение данных на обучающие и тестовые. Нормировка данных. Определение переобученности модели. Критерии оценки качества полученных моделей	Тестирование	
Тема 8 Постановка задачи регрессии. Линейный регрессионный анализ. Отбор признаков, коллинеарность, влиятельные наблюдения, анализ остатков. Непараметрическая регрессия (ядерное сглаживание). L1 и L2 регуляризация. Метрики качества	Тестирование	
Тема 9. Постановка задачи классификации, обзор основных методов ее решения. Бинарная и многоклассовая классификация. Логистическая регрессия. Решающие деревья. Метрики качества классификации (точность/специфичность, ROC-кривая, площадь под кривой)	Тестирование	
Тема 10. Ансамбли алгоритмов машинного обучения. Агрегирование моделей. Ансамбли решающих деревьев. Метод случайного леса. Градиентный бустинг.	Контрольная работа	

Шкала и критерии оценивания знаний, умений, навыков, приобретенных в результате изучения дисциплины

Шкала оценивания	Критерии оценки
На экзамене	

«Отлично»	Обучающийся обнаруживает всестороннее знание учебного материала, выражающееся в полных ответах на поставленные вопросы. Демонстрирует самостоятельность в применении знаний, умений, навыков при выполнении учебных заданий. Усвоил учебную литературу, рекомендованную для изучения дисциплины. Проявляет творческие способности в понимании, изложении и использовании учебного материала. Грамотно излагает свои мысли. В результате обучающийся обнаруживает сформированные и систематические знания, успешное и систематическое умение использовать полученные знания, успешное и систематическое применение навыков. Это подтверждает высокий (продвинутый) уровень достижения планируемых результатов обучения по дисциплине
«Хорошо»	Обучающийся обнаруживает знание учебного материала, однако ответы на поставленные вопросы неполные, но есть дополнения, большая часть материала освоена. Демонстрирует самостоятельность в применении знаний, умений, навыков при выполнении учебных заданий. Усвоил учебную литературу, рекомендованную для изучения дисциплины. Показывает систематический характер знаний учебного материала. Грамотно излагает свои мысли. В результате обучающийся обнаруживает сформированные, но содержащие отдельные пробелы знания, в целом успешное, но содержащее отдельные пробелы умение использовать полученные знания, в целом успешное, но сопровождающееся отдельными ошибками применение навыков. Это подтверждает средний (повышенный) уровень достижения планируемых результатов обучения по дисциплине
«Удовлетворительно»	Обучающийся обнаруживает отдельные пробелы в знаниях учебного материала, неточно давая ответы на поставленные вопросы либо ограничиваясь только дополнениями. Понимает основные понятия и категории дисциплины. Демонстрирует самостоятельность в применении знаний, умений, навыков при выполнении учебных заданий. Знаком с учебной литературой, рекомендованной для изучения дисциплины. В результате обучающийся обнаруживает неполные знания, в целом успешное, но не систематическое умение использовать полученные знания, в целом успешное, но не систематическое применение навыков. Это подтверждает низкий (пороговый) уровень достижения планируемых результатов обучения по дисциплине
«Неудовлетворительно»	Обучающийся обнаруживает существенные пробелы в знаниях учебного материала, поставленные вопросы не раскрыты либо содержание ответа не соответствует сути вопроса. Допускает принципиальные ошибки в трактовке основных понятий и категорий дисциплины. Неспособен самостоятельно продемонстрировать наличие знаний, умений, навыков при

	выполнении учебных заданий. В результате обучающийся обнаруживает фрагментарные знания (отсутствие знаний), фрагментарное умение использовать полученные знания (отсутствие умений), фрагментарное применение навыков (отсутствие навыков). Это подтверждает отсутствие планируемых результатов обучения по дисциплине
--	--

Типовые контрольные задания, соответствующие приведенным формам оценочных средств, необходимые для оценки знаний, умений, навыков, приобретенных в результате изучения дисциплины, а также шкалы и критерии их оценивания как в ходе текущего контроля, так и промежуточной аттестации представлены в виде оценочных материалов по дисциплине отдельным документом.

6 Перечень учебной литературы, необходимой для освоения дисциплины

1. Воронина, В. В. Теория и практика машинного обучения : учебное пособие / В. В. Воронина. — Ульяновск : УлГТУ, 2017. — 290 с. — ISBN 978-5-9795-1712-4. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/165053>
2. Сопов, Е. А. Многокритериальные нейроэволюционные системы в задачах машинного обучения и человеко-машинного взаимодействия : монография / Е. А. Сопов, И. А. Иванов. — Красноярск : СФУ, 2019. — 160 с. — ISBN 978-5-7638-3969-2. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/157729>
3. Беликов, В. В. Применение методов машинного обучения и теории игр при решении задач в области информационной безопасности: Практикум : учебное пособие / В. В. Беликов, С. В. Колесников. — Москва : РТУ МИРЭА, 2022. — 30 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/240047>
4. Котиков, П. Е. Анализ данных : учебно-методическое пособие / П. Е. Котиков. — Санкт-Петербург : СПбГПМУ, 2019. — 48 с. — ISBN 978-5-907184-46-6. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/174498>
5. Кузьмин, В. И. Методы анализа данных : учебное пособие / В. И. Кузьмин, А. Ф. Гадзаов. — 2-е изд., перераб. и доп. — Москва : РТУ МИРЭА, 2020. — 155 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/171433>
6. Вольфсон, М. Б. Анализ данных : учебное пособие / М. Б. Вольфсон. — Санкт-Петербург : СПбГУТ им. М.А. Бонч-Бруевича, 2015. — 81 с. — Текст : электронный // Лань : электронно-библиотечная система. —

URL: <https://e.lanbook.com/book/180254>

7. Вольфсон, М. Б. Анализ данных : методические указания / М. Б. Вольфсон. — Санкт-Петербург : СПбГУТ им. М.А. Бонч-Бруевича, 2013. — 47 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/181497>
8. Новикова, О. А. Анализ данных : учебное пособие / О. А. Новикова, Е. Г. Андрианова. — Москва : РТУ МИРЭА, 2020 — Часть 1 — 2020. — 162 с. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/167597>
9. Низаметдинов, Ш. У. Анализ данных : учебное пособие / Ш. У. Низаметдинов, В. П. Румянцев. — Москва : НИЯУ МИФИ, 2012. — 288 с. — ISBN 978-5-7262-1687-4. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/75847>
10. Приходько, М. А. Математическая статистика и анализ данных / М. А. Приходько, А. В. Приходько. — Омск : Омский ГАУ, 2014. — 60 с. — ISBN 978-5-89764-460-5. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/60684>
11. Клячкин, В. Н. Сборник заданий по статистическим методам анализа данных : учебное пособие / В. Н. Клячкин. — Ульяновск : УлГТУ, 2016. — 123 с. — ISBN 978-5-9795-1582-3. — Текст : электронный // Лань : электронно-библиотечная система. — URL: <https://e.lanbook.com/book/165060>
12. Бруссард, М. Искусственный интеллект: пределы возможного / Мерedit Бруссард ; пер. с англ. - Москва : Альпина нон-фикшн, 2020. - 362 с. - ISBN 978-5-00139-080-0. - Текст : электронный. - URL: <https://znanium.com/catalog/product/1220958>

7 Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины

1. Единое окно доступа к образовательным ресурсам. Каталог учебных продуктов. – Режим доступа: <http://window.edu.ru/>.
2. Веб-сервис для хостинга IT-проектов и их совместной разработки- Режим доступа: <https://github.com>.
3. Программирование на Python. – Режим доступа: <https://smartika.ru/courses/python>
4. Машинное обучение на Python - <https://scikit-learn.org/stable/index.html>

5. Быстрый, мощный, гибкий и простой в использовании инструмент анализа и обработки данных с открытым исходным кодом - <https://pandas.pydata.org/>

6. Высокопроизводительная библиотека с открытым исходным кодом для градиентного бустинга и решающих деревьев - <https://yandex.ru/dev/catboost/>

7. Комплексная платформа машинного обучения с открытым исходным кодом - <https://www.tensorflow.org/>

8 Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине, включая перечень программного обеспечения и информационных справочных систем

Информационные технологии, используемые при осуществлении образовательного процесса по дисциплине:

1. Использование информационно-обучающих (электронные библиотеки), интерактивных (электронная почта) и поисковых (поисковые системы) ресурсов.

2. Использование электронных и информационных ресурсов с текстовой информацией (учебники, учебные пособия, периодические издания, методические материалы), с визуальной информацией (схемы, диаграммы, презентации), с аудио- и видеoinформацией (аудио- и видеозаписи).

Образовательный процесс по дисциплине поддерживается средствами электронной информационно-образовательной среды Университета.

Перечень программного обеспечения и информационных справочных систем:

1. Антивирус Kaspersky Endpoint Security для бизнеса - Стандартный Russian Edition. 500-999 Node 2 year Educational Renewal License - сублиц. договор № КИС-1278-2020 от 24.11.2020 с Компьютерные информационные системы, ООО до 24.11.2022.

2. Система для дистанционного обучения СДО «Прометей 5.0» - договор № 2/ВГАУ/10/20 от 09.10.2020 с Виртуальные технологии в образовании, ООО бессрочно.

9 Методические рекомендации обучающимся по освоению дисциплины

В ходе лекционных занятий обучающимся рекомендуется: 1) вести конспектирование учебного материала; 2) обращать внимание на категории, формулировки, раскрывающие содержание тех или иных явлений и процессов, практические рекомендации по их применению; 3) задавать преподавателю уточняющие вопросы с целью уяснения теоретических положений, разрешения спорных ситуаций; 4) желательно оставить в рабочих конспектах поля, на которых во внеаудиторное время можно сделать пометки из рекомендованной литературы, дополняющие материал

прослушанной лекции, а также подчеркивающие особую важность тех или иных теоретических положений.

На практических (семинарских) занятиях в зависимости от темы занятия выполняется поиск информации по решению соответствующих содержанию дисциплины проблем, выработка индивидуальных или групповых решений, итоговое обсуждение с обменом знаниями, участие в дискуссиях, разбор и описание конкретных ситуаций, командная работа, выполнение индивидуальных заданий.

Самостоятельная работа обучающихся осуществляется в виде изучения литературных источников, работы с лекционным материалом, самостоятельного изучения отдельных тем дисциплины.

Подготовка к контрольным мероприятиям требует от обучающегося не только повторения пройденного материала на аудиторных занятиях, но поиска и анализа материала, выданного на самостоятельное изучение.

Оценка знаний, умений, навыков, характеризующая этапы формирования компетенций в процессе изучения дисциплины, проводится в форме текущего контроля и промежуточной аттестации. Текущий контроль успеваемости осуществляется в течение семестра в ходе повседневной учебной работы, обеспечивая оценивание хода освоения дисциплины. В частности, текущий контроль успеваемости проводится с целью определения уровня освоения обучающимися знаний, оценки формирования у них умений и навыков. Данный вид контроля стимулирует у обучающихся стремление к систематической самостоятельной работе по изучению дисциплины. Текущий контроль успеваемости осуществляется на практических (семинарских) занятиях, а также в ходе индивидуальных консультаций с преподавателем. К оценочным средствам для проведения текущего контроля успеваемости по дисциплине относятся контрольные работы и индивидуальные (домашние) задания.

Промежуточная аттестация обучающихся позволяет определить степень достижения запланированных результатов обучения по дисциплине и проводится в форме экзамена. Данная форма контроля включает в себя задания, позволяющие оценить уровень освоения обучающимися знаний и степень сформированности умений и навыков. Форма проведения экзамена (устная, письменная, тестирование) определяется преподавателем. По результатам экзамена выставляется оценка («отлично» / «хорошо» / «удовлетворительно» / «неудовлетворительно»).

10 Материально-техническая база, необходимая для осуществления образовательного процесса по дисциплине

№ п/п	Наименование учебных аудиторий и помещений	Адрес (местополо жение) учебных аудиторий и	Оснащенность учебных аудиторий и помещений
----------	--	--	--

		помещений	
1	Инновационно-образовательный центр компьютерных технологий для проведения учебных занятий (занятий лекционного типа, занятий семинарского типа, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации) – аудитория 507	400002, Волгоградская обл., г. Волгоград, пр. Университетский, д. 26	комплект учебной мебели, доска меловая, доска мультимедийная, оборудование и технические средства обучения – компьютеры, акустическая система, информационные плакаты
2	Помещения для самостоятельной работы - аудитории 505-508	400002, Волгоградская обл., г. Волгоград, пр. Университетский, д. 26	комплект учебной мебели, рабочие станции, компьютерная техника с возможностью подключения к сети «Интернет» и обеспечением доступа к электронной информационно-образовательной среде Организации
6	Помещение для хранения и профилактического обслуживания учебного оборудования – аудитория 502	400002, Волгоградская обл., г. Волгоград, пр. Университетский, д. 26	комплект мебели, компьютеры