МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное бюджетное образовательное учреждение высшего образования
«КУБАНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»
Факультет компьютерных технологий и прикладной математики

УТВЕРЖДАЮ
Проректор по учебной работе, качеству образования – первый проректор
_____ Хагуров Т.А.
подпись

«28» мая 2021 г.

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ
Б1.В.01«Машинное обучение»

Направление подготовки 02.03.03 Математическое обеспечение и администрирование информационных систем

Направленность (профиль) Технология программирования

Форма обучения очная

Квалификация бакалавр

Краснодар 2021

ДОКУМЕНТ ПОДПИСАН
ЭЛЕКТРОННОЙ ПОДПИСЬЮ

Сертификат:
4E23D21DF7CC6DF6D9D04D6AEAE74F342DA725D9
Владелец: Хагуров Темыр Айтечевич
Действителен: с 13.05.2021 до 13.08.2022

Рабочая программа дисциплины «Машинное обучение» составлена в соответствии с федеральным государственным образовательным стандартом высшего образования (ФГОС ВО) по направлению подготовки 02.03.03Математическое обеспечение и администрирование информационных систем.

Программу составил(и):

Харченко Анна Владимировна, старший преподаватель _____

И.О. Фамилия, должность, ученая степень, ученое звание подпись

Рабочая программа дисциплины «Машинное обучение» утверждена на заседании кафедры информационных технологий протокол №15 от «20» мая 2021 г.

Заведующий кафедрой (разработчика)

В. В. Подколзин _____

подпись

Рабочая программа обсуждена на заседании кафедры информационных технологий протокол №15 от «20» мая 2021 г.

Заведующий кафедрой (выпускающей)

В. В. Подколзин _____

подпись

Утверждена на заседании учебно-методической комиссии факультета компьютерных технологий и прикладной математики протокол №1 от «21» мая 2021 г.

Председатель УМК факультета

А. В. Коваленко _____

подпись

Рецензенты:

Бегларян М. Е., зав. кафедрой социально-гуманитарных и естественнонаучных дисциплин СКФ ФГБОУВО «Российский государственный университет правосудия», канд. физ.-мат. наук, доцент

Рубцов Сергей Евгеньевич, кандидат физико-математических наук, доцент кафедры математического моделирования ФГБГОУ «КубГУ»

1 Цели и задачи изучения дисциплины (модуля)

1.1 Цель освоения дисциплины

Целью освоения дисциплины являются формирование у студентов теоретических знаний и практических навыков по основам машинного обучения, овладение студентами инструментарием, моделями и методами машинного обучения, а также приобретение навыков исследователя данных (data scientist) и разработчика математических моделей, методов и алгоритмов анализа данных.

Содержательное наполнение дисциплины обусловлено общими задачами в подготовке бакалавра.

Научной основой для построения программы данной дисциплины является теоретико-прагматический подход в обучении.

1.2 Задачи дисциплины

– изучение ключевых понятий, целей и задач использования машинного обучения;

– изучение методологических основ применения алгоритмов машинного обучения.

– изучение методов машинного обучения, соответствующих различным исследовательским задачам, интерпретирование полученных результатов;

– умение визуализировать результаты работы алгоритмов машинного обучения, выбирать;

– приобретение опыта по применению методов машинного обучения, построения и оценки качества моделей.

1.3 Место дисциплины (модуля) в структуре образовательной программы

Дисциплина «Машинное обучение» относится к «Часть, формируемая участниками образовательных отношений» Блока 1 «Дисциплины (модули)» учебного плана.

Входными знаниями для освоения данной дисциплины являются знания, умения и опыт, накопленный студентами в процессе изучения дисциплин «Фундаментальные дискретные модели», «Основы программирования», «Методы программирования». Обучающийся должен:

- иметь базовые навыки в написании программ на процедурных и объектно-ориентированных языках;
- быть знакомым с наиболее часто встречающимися структурами данных, уметь ими пользоваться и знать внутреннюю организацию.

1.4 Перечень планируемых результатов обучения по дисциплине (модулю), соотнесенных с планируемыми результатами освоения образовательной программы

Изучение данной учебной дисциплины направлено на формирование у обучающихся следующих компетенций:

ПК-4	**Способен применять современные информационные технологии при проектировании, реализации, оценке качества и анализа эффективности программного обеспечения для решения задач в различных предметных областях**
Знать	ИПК-4.1 (D/03.6 Зн.2) Типовые решения, библиотеки программных модулей, шаблоны, классы объектов, информационные технологии при проектировании, реализации, оценке качества и анализа эффективности программного обеспечения для решения задач в различных предметных областях ИПК-4.2 (D/03.6 Зн.3) Методы и средства, современные информационные технологии проектирования программного обеспечения для решения задач в различных предметных областях

ИПК-4.5 (С/16.6 Зн.2) Инструменты, методы и современные информационные технологии проектирования и дизайна ИС
ИПК-4.7 (С/16.6 Зн.5) Предметная область автоматизации при решении задач в различных предметных областях
ИПК-4.17 (А/01.5 Зн.1) Цели и задачи проводимых исследований и разработок при решении задач в различных предметных областях
ИПК-4.18 (А/01.5 Зн.2) Методы анализа и обобщения отечественного и международного опыта области информационных технологий
ИПК-4.20 (А/01.5 Зн.4) Методы проведения экспериментов и наблюдений, обобщения и обработки результатов работы программного обеспечения
ИПК-4.21 (А/01.5 Др.1 Зн.) Деятельность, направленная на решение задач аналитического характера, предполагающих выбор и многообразие актуальных способов решения задач с применением современных информационных технологий при проектировании, реализации, оценке качества и анализа эффективности программного обеспечения

Уметь ИПК-4.26 (А/01.5 У.3) Применять современные методы анализа научно-технической информации

Владеть ИПК-4.30 (А/01.5 Тд.2) Сбор, обработка, анализ и обобщение передового отечественного и международного опыта в области современных информационных технологий
ИПК-4.31 (А/01.5 Тд.3) Сбор, обработка, анализ и обобщение результатов экспериментов и исследований в области современных информационных технологий

ПК-7 Способен использовать современные методы разработки и реализации конкретных алгоритмов математических моделей на базе языков программирования и пакетов прикладных программ моделирования

Знать ИПК-7.1 (D/03.6 Зн.1) Принципы построения архитектуры программного обеспечения и виды архитектуры программного обеспечения, современные методы разработки и реализации конкретных алгоритмов математических моделей
ИПК-7.2 (D/03.6 Зн.2) Типовые решения, библиотеки программных модулей, шаблоны, классы объектов, используемые при разработке и реализации конкретных алгоритмов математических моделей на базе языков программирования и пакетов прикладных программ моделирования
ИПК-7.14 (А/01.5 Зн.2) Методы анализа и обобщения отечественного и международного опыта в области разработки и реализации алгоритмов математических моделей на базе языков программирования и пакетов прикладных программ моделирования
ИПК-7.15 (А/01.5 Др.1 Зн.) Деятельность, направленная на решение задач аналитического характера, предполагающих выбор и многообразие актуальных способов решения задач математического моделирования на базе языков программирования и пакетов прикладных программ моделирования

Уметь ИПК-7.16 (D/03.6 У.1) Использовать существующие типовые решения и шаблоны, современные методы разработки и реализации конкретных алгоритмов математических моделей на базе языков программирования и пакетов прикладных программ моделирования
ИПК-7.17 (D/03.6 У.2) Применять методы и средства проектирования программного обеспечения, структур данных, баз данных, программных

интерфейсов при реализации конкретных алгоритмов математических моделей

Владеть ИПК-7.21 (А/01.5 Тд.3) Сбор, обработка, анализ и обобщение результатов экспериментов и исследований в области знаний алгоритмизации математических моделей на базе языков программирования и пакетов прикладных программ моделирования

Результаты обучения по дисциплине достигаются в рамках осуществления всех видов контактной и самостоятельной работы обучающихся в соответствии с утвержденным учебным планом.

Индикаторы достижения компетенций считаются сформированными при достижении соответствующих им результатов обучения.

2. Структура и содержание дисциплины

2.1 Распределение трудоёмкости дисциплины по видам работ

Общая трудоёмкость дисциплины составляет 4 зач. ед. (144часов), их распределение по видам работ представлено в таблице

Вид учебной работы		**Всего часов**	**Семестры (часы)**					
			7					
Контактная работа, в том числе:		**78,3**	**78,3**					
Аудиторные занятия (всего):		**68**	**68**					
Занятия лекционного типа		**34**	34					
Лабораторные занятия		**34**	34					
Занятия семинарского типа (семинары, практические занятия)								
Иная контактная работа:		**10,3**	**10,3**					
Контроль самостоятельной работы (КСР)		**10**	10					
Промежуточная аттестация (ИКР)		**0,3**	0,3					
Самостоятельная работа, в том числе:		**30**	**30**					
Проработка учебного (теоретического) материала		**10**	**10**					
Выполнение индивидуальных заданий (подготовка сообщений, презентаций)		**16**	**16**					
Подготовка к текущему контролю		**4**	**4**					
Контроль:		**35,7**	**35,7**					
Подготовка к экзамену		**35,7**	35,7					
Общая трудоемкость	**час.**	**144**	**144**					
	в том числе контактная работа	**78,3**	**78,3**					
	зач. ед	**4**	**4**					

2.2 Структура дисциплины

Распределение видов учебной работы и их трудоемкости по разделам дисциплины.

Разделы (темы) дисциплины, изучаемые в 7 семестре

№	Наименование разделов (тем)	Количество часов				
		Всего	Аудиторная работа			Внеаудиторная работа
			Л	ПЗ	ЛР	СРС
1	2	3	4	5	6	7
1.	Введение в машинное обучение		6		6	4
2.	Линейные модели регрессии		6		6	6
3.	Логистическая регрессия		4		4	4
4.	Классификация		6		6	4
5.	Кластеризация		4		4	4
6.	Деревья решений		4		4	4
7.	Ансамблевые методы		4		4	4
ИТОГО по разделам дисциплины			**34**		**34**	**30**
Контроль самостоятельной работы (КСР)		10				
Промежуточная аттестация (ИКР)		0,3				
Подготовка к текущему контролю		35,7				
Общая трудоемкость по дисциплине		**144**				

Примечание: Л – лекции, ПЗ – практические занятия/семинары, ЛР – лабораторные занятия, СРС – самостоятельная работа студента

2.3 Содержание разделов (тем) дисциплины

2.3.1 Занятия лекционного типа

№	Наименование раздела (темы)	Содержание раздела (темы)	Форма текущего контроля
1	2	3	4
1.	Введение в машинное обучение	Введение в предметную область. Примеры использования методов машинного обучения для решения прикладных задач. .	К, Т
2.	Введение в машинное обучение	Знакомство со специализированными библиотеками языка программирования Py thon для научных расчетов и анализа данных. NumPy, SciPy, pandas.	К, Т
3.	Введение в машинное обучение	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	К, Т
4.	Линейные модели регрессии	Постановка задачи регрессии. Линейный регрессионный анализ.	К, Т
5.	Линейные модели регрессии	тбор признаков, коллинеарность, влиятельные наблюдения, анализ остатков.	К, Т
6.	Линейные модели регрессии	Непараметрическая регрессия (ядерное сглаживание). L1 и L2 регуляризация. Метрики качества.	К, Т
7.	Логистическая регрессия	Разделение данных на обучающие и тестовые. Нормировка данных.	К, Т
8.	Логистическая регрессия	Определение переобученности модели. Критерии оценки качества полученных моделей.	К, Т

№	Наименование раздела (темы)	Содержание раздела (темы)	Форма текущего контроля
1	2	3	4
9.	Классификация	Постановка задачи классификации, обзор основных методов ее решения.	К, Т
10.	Классификация	Бинарная и многоклассовая классификация. Логистическая регрессия.	К, Т
11.	Классификация	Метрики качества классификации (точность/специфичность, ROC-кривая, площадь под кривой).	К, Т
12.	Деревья решений	Структура деревьев решений. Виды разделяющих функций.	К, Т
13.	Деревья решений	Обучения дерева решений. Алгоритм Random Forest.	К, Т
14.	Кластеризация	Обучение на неразмеченных данных. Кластеризация. Иерархическая кластеризация. Метод K-средних, DBSCAN и др.	К, Т
15.	Кластеризация	Обзор методов кластеризации, реализованных в библиотеке sklearn.	К, Т
16.	Ансамблевые методы	Ансамбли алгоритмов машинного обучения. Агрегирование моделей	К, Т
17.	Ансамблевые методы	. Ансамбли решающих деревьев. Метод случайного леса. Градиентный бустинг.	К, Т

Примечание: ЛР – отчет/защита лабораторной работы, КП - выполнение курсового проекта, КР - курсовой работы, РГЗ - расчетно-графического задания, Р - написание реферата, Э - эссе, К - коллоквиум, Т – тестирование, РЗ – решение задач.

2.3.2 Занятия семинарского типа

Не предусмотрены

Примечание: ЛР – отчет/защита лабораторной работы, КП - выполнение курсового проекта, КР - курсовой работы, РГЗ - расчетно-графического задания, Р - написание реферата, Э - эссе, К - коллоквиум, Т – тестирование, РЗ – решение задач.

2.3.3 Лабораторные занятия

№	Наименование раздела (темы)	Наименование лабораторных работ	Форма текущего контроля
1	2	3	4
1.	Введение в машинное обучение	Примеры использования методов машинного обучения для практических задач. Краткий обзор синтаксиса языка Python. Встроенные операции и функции, типы и структуры данных.	*РЗ*
2.	Введение в машинное обучение	Библиотеки NumPy и SciPy. Матрицы. Разреженные матрицы. Индексирование, срезы. Объединение массивов. Библиотека pandas. Запросы к таблицам: выборка строк/столбцов по заданным критериям. Модификация элементов таблицы. Добавление строк/столбцов.	*РЗ*
3.	Введение в машинное обучение	Группировка и агрегирование. Объединение таблиц (различные виды join). Многомерные данные: мультииндексы. Операции stack-unstack. Построение сводных таблиц (pivottables).	*РЗ*

№	Наименование раздела (темы)	Наименование лабораторных работ	Форма текущего контроля
1	2	3	4
4.	Линейные модели регрессии	Описательные статистики. Обзор библиотек matplotlib, seaborn, bokeh. Базовые типы визуализации данных. Знакомство с библиотекой scikit-learn (sklearn).	*РЗ*
5.	Линейные модели регрессии	Предобработка данных. Метод главных компонент. Работа с пропущенными значениями	*РЗ*
6.	Линейные модели регрессии	Полиномиальная регрессия. Смещение и дисперсия. Гребневая регрессия.	*РЗ*
7.	Логистическая регрессия	Целевая функция логистической регрессии. Регуляризация логистической регрессии.	*РЗ*
8.	Логистическая регрессия	Определения целевой метрики качества. Сравнение различных метрик качества моделей. Работа с несбалансированными наборами	*РЗ*
9.	Классификация	Реализация регрессионных и классификационных моделей с помощью sklearn.	*РЗ*
10.	Классификация	Работа с синтетическими данными. Реализация метода градиентного спуска.	*РЗ*
11.	Классификация	Реализация моделей на основе метода k-ближайших соседей метода опорных векторов	*РЗ*
12.	Деревья решений	Правила и анализ качества (точность, полнота). Анализ с помощью ROC кривой. Алгоритм построения деревьев решений.	*РЗ*
13.	Деревья решений	. Критерий информационного выигрыша и критерий Джини. Леса решающих деревьев.	*РЗ*
14.	Кластеризация	Алгоритмы кластеризации с фиксированным количеством кластеров.	*РЗ*
15.	Кластеризация	. Алгоритмы кластеризации по плотности. Иерархическая кластеризация.	*РЗ*
16.	Ансамблевые методы	Реализация моделей с помощью метода градиентного бустинга, метода случайного леса. Блендинг и стеккинг.	*РЗ*
17.	Ансамблевые методы	Реализация моделей с помощью метода градиентного бустинга, метода случайного леса. Методы отбора признаков. Оптимизация гиперпараметров	*РЗ*

Примечание: ЛР – отчет/защита лабораторной работы, КП - выполнение курсового проекта, КР - курсовой работы, РГЗ - расчетно-графического задания, Р - написание реферата, Э - эссе, К - коллоквиум, Т – тестирование, РЗ – решение задач.

2.3.4 Примерная тематика курсовых работ (проектов)

Не предусмотрены

2.4 Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине (модулю)

№	Вид СРС	Перечень учебно-методического обеспечения дисциплины по выполнению самостоятельной работы
1	2	3
1	Изучение теоретического материала	Методические указания по организации самостоятельной работы студентов, утвержденные кафедрой информационных технологий, протокол №1 от 30.08.2019
2	Решение задач	Методические указания по организации самостоятельной работы студентов, утвержденные кафедрой информационных технологий, протокол №1 от 30.08.2019

Учебно-методические материалы для самостоятельной работы обучающихся из числа инвалидов и лиц с ограниченными возможностями здоровья (ОВЗ) предоставляются в формах, адаптированных к ограничениям их здоровья и восприятия информации:

Для лиц с нарушениями зрения:

– в печатной форме увеличенным шрифтом,

– в форме электронного документа,

– в форме аудиофайла,

– в печатной форме на языке Брайля.

Для лиц с нарушениями слуха:

– в печатной форме,

– в форме электронного документа.

Для лиц с нарушениями опорно-двигательного аппарата:

– в печатной форме,

– в форме электронного документа,

– в форме аудиофайла.

Данный перечень может быть конкретизирован в зависимости от контингента обучающихся.

3. Образовательные технологии

В соответствии с требованиями ФГОС в программа дисциплины предусматривает использование в учебном процессе следующих образовательные технологии: чтение лекций с использованием мультимедийных технологий; метод малых групп, разбор практических задач и кейсов.

При обучении используются следующие образовательные технологии:

– Технология коммуникативного обучения – направлена на формирование коммуникативной компетентности студентов, которая является базовой, необходимой для адаптации к современным условиям межкультурной коммуникации.

– Технология разноуровневого (дифференцированного) обучения – предполагает осуществление познавательной деятельности студентов с учётом их индивидуальных способностей, возможностей и интересов, поощряя их реализовывать свой творческий потенциал. Создание и использование диагностических тестов является неотъемлемой частью данной технологии.

– Технология модульного обучения – предусматривает деление содержания дисциплины на достаточно автономные разделы (модули), интегрированные в общий курс.

– Информационно-коммуникационные технологии (ИКТ) - расширяют рамки образовательного процесса, повышая его практическую направленность, способствуют интенсификации самостоятельной работы учащихся и повышению познавательной активности. В рамках ИКТ выделяются 2 вида технологий:

– Технология использования компьютерных программ – позволяет эффективно дополнить процесс обучения языку на всех уровнях.

– Интернет-технологии – предоставляют широкие возможности для поиска информации, разработки научных проектов, ведения научных исследований.

– Технология индивидуализации обучения – помогает реализовывать личностно-ориентированный подход, учитывая индивидуальные особенности и потребности учащихся.

– Проектная технология – ориентирована на моделирование социального взаимодействия учащихся с целью решения задачи, которая определяется в рамках профессиональной подготовки, выделяя ту или иную предметную область.

– Технология обучения в сотрудничестве – реализует идею взаимного обучения, осуществляя как индивидуальную, так и коллективную ответственность за решение учебных задач.

– Игровая технология – позволяет развивать навыки рассмотрения ряда возможных способов решения проблем, активизируя мышление студентов и раскрывая личностный потенциал каждого учащегося.

– Технология развития критического мышления – способствует формированию разносторонней личности, способной критически относиться к информации, умению отбирать информацию для решения поставленной задачи.

Комплексное использование в учебном процессе всех вышеназванных технологий стимулируют личностную, интеллектуальную активность, развивают познавательные процессы, способствуют формированию компетенций, которыми должен обладать будущий специалист.

Основные виды интерактивных образовательных технологий включают в себя:

– работа в малых группах (команде) - совместная деятельность студентов в группе под руководством лидера, направленная на решение общей задачи путём творческого сложения результатов индивидуальной работы членов команды с делением полномочий и ответственности;

– проектная технология - индивидуальная или коллективная деятельность по отбору, распределению и систематизации материала по определенной теме, в результате которой составляется проект;

– анализ конкретных ситуаций - анализ реальных проблемных ситуаций, имевших место в соответствующей области профессиональной деятельности, и поиск вариантов лучших решений;

– развитие критического мышления – образовательная деятельность, направленная на развитие у студентов разумного, рефлексивного мышления, способного выдвинуть новые идеи и увидеть новые возможности.

Подход разбора конкретных задач и ситуаций широко используется как преподавателем, так и студентами во время лекций, лабораторных занятий и анализа результатов самостоятельной работы. Это обусловлено тем, что при исследовании и решении каждой конкретной задачи имеется, как правило, несколько методов, а это требует разбора и оценки целой совокупности конкретных ситуаций.

Семестр	**Вид занятия**	**Используемые интерактивные образовательные технологии**	**количество интерактивных часов**
	ЛР	Практические занятия в режимах взаимодействия «преподаватель – студент» и «студент – студент»	16
Итого			**16**

Примечание: Л – лекции, ПЗ – практические занятия/семинары, ЛР – лабораторные занятия, СРС – самостоятельная работа студента

Темы, задания и вопросы для самостоятельной работы призваны сформировать навыки поиска информации, умения самостоятельно расширять и углублять знания, полученные в ходе лекционных и практических занятий.

Подход разбора конкретных ситуаций широко используется как преподавателем, так и студентами при проведении анализа результатов самостоятельной работы.

Для лиц с ограниченными возможностями здоровья предусмотрена организация консультаций с использованием электронной почты.

Для лиц с нарушениями зрения:

– в печатной форме увеличенным шрифтом,

– в форме электронного документа.

Для лиц с нарушениями слуха:

– в печатной форме,

– в форме электронного документа.

Для лиц с нарушениями опорно-двигательного аппарата:

– в печатной форме,

– в форме электронного документа.

Для лиц с ограниченными возможностями здоровья предусмотрена организация консультаций с использованием электронной почты.

Данный перечень может быть конкретизирован в зависимости от контингента обучающихся.

4. Оценочные и методические материалы

4.1 Оценочные средства для текущего контроля успеваемости и промежуточной аттестации

Оценочные средства предназначены для контроля и оценки образовательных достижений обучающихся, освоивших программу учебной дисциплины «название дисциплины».

Оценочные средства включает контрольные материалы для проведения **текущего контроля** в форме тестовых заданий, разноуровневых заданий, и **промежуточной аттестации** в форме вопросов и заданий к экзамену.

Оценочные средства для инвалидов и лиц с ограниченными возможностями здоровья выбираются с учетом их индивидуальных психофизических особенностей.

– при необходимости инвалидам и лицам с ограниченными возможностями здоровья предоставляется дополнительное время для подготовки ответа на экзамене;

– при проведении процедуры оценивания результатов обучения инвалидов и лиц с ограниченными возможностями здоровья предусматривается использование технических средств, необходимых им в связи с их индивидуальными особенностями;

– при необходимости для обучающихся с ограниченными возможностями здоровья и инвалидов процедура оценивания результатов обучения по дисциплине может проводиться в несколько этапов.

Процедура оценивания результатов обучения инвалидов и лиц с ограниченными возможностями здоровья по дисциплине (модулю) предусматривает предоставление информации в формах, адаптированных к ограничениям их здоровья и восприятия информации:

Для лиц с нарушениями зрения:

– в печатной форме увеличенным шрифтом,

– в форме электронного документа.

Для лиц с нарушениями слуха:

– в печатной форме,

– в форме электронного документа.

Для лиц с нарушениями опорно-двигательного аппарата:

– в печатной форме,

– в форме электронного документа.

Данный перечень может быть конкретизирован в зависимости от контингента обучающихся.

Структура оценочных средств для текущей и промежуточной аттестации

№ п/п	Контролируемые разделы (темы) дисциплины*	Код контролируемой компетенции (или ее части)	Наименование оценочного средства	
			Текущий контроль	Промежуточная аттестация

1	Введение в машинное обучение	**ПК-4** ИПК-4.1 (D/03.6 Зн.2) ИПК-4.2 (D/03.6 Зн.3) ИПК-4.5 (C/16.6 Зн.2) ИПК-4.7 (C/16.6 Зн.5) ИПК-4.17 (A/01.5 Зн.1) ИПК-4.18 (A/01.5 Зн.2) ИПК-4.20 (A/01.5 Зн.4) ИПК-4.21 (A/01.5 Др.1 Зн.) ИПК-4.26 (A/01.5 У.3) ИПК-4.30 (A/01.5 Тд.2) ИПК-4.31 (A/01.5 Тд.3) **ПК-7** ИПК-7.1 (D/03.6 Зн.1) ИПК-7.2 (D/03.6 Зн.2) ИПК-7.14 (A/01.5 Зн.2) ИПК-7.15 (A/01.5 Др.1 Зн.) Уметь ИПК-7.16 (D/03.6 У.1) ИПК-7.17 (D/03.6 У.2) ИПК-7.21 (A/01.5 Тд.3)	*Типовые тестовые задания 1-5*	*Вопрос на экзамене 1-4*
2	Линейные модели регрессии	**ПК-4** ИПК-4.1 (D/03.6 Зн.2) ИПК-4.2 (D/03.6 Зн.3) ИПК-4.5 (C/16.6 Зн.2) ИПК-4.7 (C/16.6 Зн.5) ИПК-4.17 (A/01.5 Зн.1) ИПК-4.18 (A/01.5 Зн.2) ИПК-4.20 (A/01.5 Зн.4) ИПК-4.21 (A/01.5 Др.1 Зн.) ИПК-4.26 (A/01.5 У.3) ИПК-4.30 (A/01.5 Тд.2) ИПК-4.31 (A/01.5 Тд.3) **ПК-7** ИПК-7.1 (D/03.6 Зн.1) ИПК-7.2 (D/03.6 Зн.2) ИПК-7.14 (A/01.5 Зн.2) ИПК-7.15 (A/01.5 Др.1 Зн.) Уметь ИПК-7.16 (D/03.6 У.1) ИПК-7.17 (D/03.6 У.2) ИПК-7.21 (A/01.5 Тд.3)	*Типовые тестовые задания 6-8* *Типовые контрольные задания 1-2*	*Вопрос на экзамене 5-7*
3	Логистическая регрессия	**ПК-4** ИПК-4.1 (D/03.6 Зн.2) ИПК-4.2 (D/03.6 Зн.3) ИПК-4.5 (C/16.6 Зн.2) ИПК-4.7 (C/16.6 Зн.5) ИПК-4.17 (A/01.5 Зн.1) ИПК-4.18 (A/01.5 Зн.2) ИПК-4.20 (A/01.5 Зн.4) ИПК-4.21 (A/01.5 Др.1 Зн.) ИПК-4.26 (A/01.5 У.3) ИПК-4.30 (A/01.5 Тд.2) ИПК-4.31 (A/01.5 Тд.3) **ПК-7** ИПК-7.1 (D/03.6 Зн.1) ИПК-7.2 (D/03.6 Зн.2) ИПК-7.14 (A/01.5 Зн.2) ИПК-7.15 (A/01.5 Др.1 Зн.) Уметь ИПК-7.16 (D/03.6 У.1) ИПК-7.17 (D/03.6 У.2) ИПК-7.21 (A/01.5 Тд.3)	*Тестовые вопросы 9-11* *Типовые контрольные задания 3*	*Вопрос на экзамене 8*

4	Классификация	**ПК-4** ИПК-4.1 (D/03.6 Зн.2) ИПК-4.2 (D/03.6 Зн.3) ИПК-4.5 (C/16.6 Зн.2) ИПК-4.7 (C/16.6 Зн.5) ИПК-4.17 (A/01.5 Зн.1) ИПК-4.18 (A/01.5 Зн.2) ИПК-4.20 (A/01.5 Зн.4) ИПК-4.21 (A/01.5 Др.1 Зн.) ИПК-4.26 (A/01.5 У.3) ИПК-4.30 (A/01.5 Тд.2) ИПК-4.31 (A/01.5 Тд.3) **ПК-7** ИПК-7.1 (D/03.6 Зн.1) ИПК-7.2 (D/03.6 Зн.2) ИПК-7.14 (A/01.5 Зн.2) ИПК-7.15 (A/01.5 Др.1 Зн.) Уметь ИПК-7.16 (D/03.6 У.1) ИПК-7.17 (D/03.6 У.2) ИПК-7.21 (A/01.5 Тд.3)	*Типовые тестовые задания 12-14* *Типовые контрольные задания 4-6*	*Вопрос на экзамене 9-13*
5	Деревья решений	**ПК-4** ИПК-4.1 (D/03.6 Зн.2) ИПК-4.2 (D/03.6 Зн.3) ИПК-4.5 (C/16.6 Зн.2) ИПК-4.7 (C/16.6 Зн.5) ИПК-4.17 (A/01.5 Зн.1) ИПК-4.18 (A/01.5 Зн.2) ИПК-4.20 (A/01.5 Зн.4) ИПК-4.21 (A/01.5 Др.1 Зн.) ИПК-4.26 (A/01.5 У.3) ИПК-4.30 (A/01.5 Тд.2) ИПК-4.31 (A/01.5 Тд.3) **ПК-7** ИПК-7.1 (D/03.6 Зн.1) ИПК-7.2 (D/03.6 Зн.2) ИПК-7.14 (A/01.5 Зн.2) ИПК-7.15 (A/01.5 Др.1 Зн.) Уметь ИПК-7.16 (D/03.6 У.1) ИПК-7.17 (D/03.6 У.2) ИПК-7.21 (A/01.5 Тд.3)	*Типовые тестовые задания 15* *Типовые контрольные задания 7*	*Вопрос на экзамене 15-17*
6	Кластеризация	**ПК-4** ИПК-4.1 (D/03.6 Зн.2) ИПК-4.2 (D/03.6 Зн.3) ИПК-4.5 (C/16.6 Зн.2) ИПК-4.7 (C/16.6 Зн.5) ИПК-4.17 (A/01.5 Зн.1) ИПК-4.18 (A/01.5 Зн.2) ИПК-4.20 (A/01.5 Зн.4) ИПК-4.21 (A/01.5 Др.1 Зн.) ИПК-4.26 (A/01.5 У.3) ИПК-4.30 (A/01.5 Тд.2) ИПК-4.31 (A/01.5 Тд.3) **ПК-7** ИПК-7.1 (D/03.6 Зн.1) ИПК-7.2 (D/03.6 Зн.2) ИПК-7.14 (A/01.5 Зн.2) ИПК-7.15 (A/01.5 Др.1 Зн.) Уметь ИПК-7.16 (D/03.6 У.1) ИПК-7.17 (D/03.6 У.2) ИПК-7.21 (A/01.5 Тд.3)	*Типовые тестовые задания 16 -18* *Типовые контрольные задания 8-9*	*Вопрос на экзамене 14*

7	Ансамблевые методы	**ПК-4** ИПК-4.1 (D/03.6 Зн.2) ИПК-4.2 (D/03.6 Зн.3) ИПК-4.5 (C/16.6 Зн.2) ИПК-4.7 (C/16.6 Зн.5) ИПК-4.17 (A/01.5 Зн.1) ИПК-4.18 (A/01.5 Зн.2) ИПК-4.20 (A/01.5 Зн.4) ИПК-4.21 (A/01.5 Др.1 Зн.) ИПК-4.26 (A/01.5 У.3) ИПК-4.30 (A/01.5 Тд.2) ИПК-4.31 (A/01.5 Тд.3) **ПК-7** ИПК-7.1 (D/03.6 Зн.1) ИПК-7.2 (D/03.6 Зн.2) ИПК-7.14 (A/01.5 Зн.2) ИПК-7.15 (A/01.5 Др.1 Зн.) Уметь ИПК-7.16 (D/03.6 У.1) ИПК-7.17 (D/03.6 У.2) ИПК-7.21 (A/01.5 Тд.3)	*Типовые тестовые задания 19-21* *Типовые контрольные задания 10*	*Вопрос на экзамене 18-20*

Показатели, критерии и шкала оценки сформированных компетенций

Соответствие **<u>пороговому уровню</u>** освоения компетенций планируемым результатам обучения и критериям их оценивания (оценка: **удовлетворительно /зачтено**):

ПК-4 **Способен применять современные информационные технологии при проектировании, реализации, оценке качества и анализа эффективности программного обеспечения для решения задач в различных предметных областях**

Знать ИПК-4.1 (D/03.6 Зн.2) Типовые решения, библиотеки программных модулей, шаблоны, классы объектов, информационные технологии при проектировании, реализации, оценке качества и анализа эффективности программного обеспечения для решения задач в различных предметных областях с помощью языка программирования Python

ИПК-4.2 (D/03.6 Зн.3) Методы и средства, современные информационные технологии проектирования программного обеспечения для решения задач в различных предметных областях

ИПК-4.5 (C/16.6 Зн.2) Инструменты, методы и современные информационные технологии проектирования и дизайна ИС, средства разработки программ PyCharm, Google Collab

ИПК-4.7 (C/16.6 Зн.5) Предметная область автоматизации при решении задач в различных предметных областях

ИПК-4.17 (A/01.5 Зн.1) Цели и задачи проводимых исследований и разработок при решении задач в различных предметных областях

ИПК-4.18 (A/01.5 Зн.2) Методы анализа и обобщения отечественного и международного опыта области информационных технологий

ИПК-4.20 (A/01.5 Зн.4) Методы проведения экспериментов и наблюдений, обобщения и обработки результатов работы программного обеспечения

ИПК-4.21 (A/01.5 Др.1 Зн.) Деятельность, направленная на решение задач аналитического характера, предполагающих выбор и многообразие актуальных способов решения задач с применением современных

информационных технологий при проектировании, реализации, оценке качества и анализа эффективности программного обеспечения

Уметь ИПК-4.26 (А/01.5 У.3) Применять современные методы анализа научно-технической информации

Владеть ИПК-4.30 (А/01.5 Тд.2) Сбор, обработка, анализ и обобщение передового отечественного и международного опыта в области современных информационных технологий, в том числе языка программирования Python и средств разработки программ PyCharm, Google Collab

ИПК-4.31 (А/01.5 Тд.3) Сбор, обработка, анализ и обобщение результатов экспериментов и исследований в области современных информационных технологий, в том числе языка программирования Python и средств разработки программ PyCharm, Google Collab

ПК-7 Способен использовать современные методы разработки и реализации конкретных алгоритмов математических моделей на базе языков программирования и пакетов прикладных программ моделирования

Знать ИПК-7.1 (D/03.6 Зн.1) Принципы построения архитектуры программного обеспечения и виды архитектуры программного обеспечения, современные методы разработки и реализации конкретных алгоритмов математических моделей на языке программирования Python

ИПК-7.2 (D/03.6 Зн.2) Типовые решения, библиотеки программных модулей, шаблоны, классы объектов, используемые при разработке и реализации конкретных алгоритмов математических моделей на базе языков программирования, в том числе Python и пакетов прикладных программ моделирования

ИПК-7.14 (А/01.5 Зн.2) Методы анализа и обобщения отечественного и международного опыта в области разработки и реализации алгоритмов математических моделей на базе языка программирования Python и пакетов прикладных программ моделирования numpy, sklearn

ИПК-7.15 (А/01.5 Др.1 Зн.) Деятельность, направленная на решение задач аналитического характера, предполагающих выбор и многообразие актуальных способов решения задач математического моделирования на базе языка программирования Python и пакетов прикладных программ моделирования numpy, sklearn

Уметь ИПК-7.16 (D/03.6 У.1) Использовать существующие типовые решения и шаблоны, современные методы разработки и реализации конкретных алгоритмов математических моделей на базе языка программирования Python и пакетов прикладных программ моделирования numpy, sklearn

ИПК-7.17 (D/03.6 У.2) Применять методы и средства проектирования программного обеспечения, структур данных, баз данных, программных интерфейсов при реализации конкретных алгоритмов математических моделей на базе языка программирования Python

Владеть ИПК-7.21 (А/01.5 Тд.3) Сбор, обработка, анализ и обобщение результатов экспериментов и исследований в области знаний алгоритмизации математических моделей на базе языка программирования Python и пакетов прикладных программ моделирования numpy, sklearn

Соответствие **базовому уровню** освоения компетенций планируемым результатам обучения и критериям их оценивания (оценка: **хорошо /зачтено**):

ПК-4 **Способен применять современные информационные технологии при проектировании, реализации, оценке качества и анализа эффективности программного обеспечения для решения задач в различных предметных областях**

Знать ИПК-4.1 (D/03.6 Зн.2) Современные типовые решения, библиотеки программных модулей, шаблоны, классы объектов, информационные технологии при проектировании, реализации, оценке качества и анализа эффективности программного обеспечения для решения задач в различных предметных областях с помощью языка программирования Python

ИПК-4.2 (D/03.6 Зн.3) Методы и средства, современные информационные технологии проектирования программного обеспечения для решения задач в различных предметных областях

ИПК-4.5 (C/16.6 Зн.2) Современные инструменты, методы и современные информационные технологии проектирования и дизайна ИС, средства разработки программ PyCharm, Google Collab

ИПК-4.7 (C/16.6 Зн.5) Предметная область автоматизации при решении задач в различных предметных областях

ИПК-4.17 (A/01.5 Зн.1) Цели и задачи проводимых исследований и разработок при решении задач в различных предметных областях

ИПК-4.18 (A/01.5 Зн.2) Новейшие методы анализа и обобщения отечественного и международного опыта области информационных технологий

ИПК-4.20 (A/01.5 Зн.4) Современные методы проведения экспериментов и наблюдений, обобщения и обработки результатов работы программного обеспечения

ИПК-4.21 (A/01.5 Др.1 Зн.) Деятельность, направленная на решение задач аналитического характера, предполагающих выбор и многообразие актуальных способов решения задач с применением современных информационных технологий при проектировании, реализации, оценке качества и анализа эффективности программного обеспечения

Уметь ИПК-4.26 (A/01.5 У.3) Применять новейшие современные методы анализа научно-технической информации

Владеть ИПК-4.30 (A/01.5 Тд.2) Сбор, обработка, анализ и обобщение передового отечественного и международного опыта в области современных информационных технологий, в том числе языка программирования Python и средств разработки программ PyCharm, Google Collab

ИПК-4.31 (A/01.5 Тд.3) Сбор, обработка, анализ и обобщение результатов экспериментов и исследований в области современных информационных технологий, в том числе языка программирования Python и средств разработки программ PyCharm, Google Collab

ПК-7 **Способен использовать современные методы разработки и реализации конкретных алгоритмов математических моделей на базе языков программирования и пакетов прикладных программ моделирования**

Знать ИПК-7.1 (D/03.6 Зн.1) Современные принципы построения архитектуры программного обеспечения и виды архитектуры программного обеспечения, современные методы разработки и реализации конкретных алгоритмов математических моделей на языке программирования Python

ИПК-7.2 (D/03.6 Зн.2) Современные типовые решения, библиотеки программных модулей, шаблоны, классы объектов, используемые при разработке и реализации конкретных алгоритмов математических моделей на базе языков программирования, в том числе Python и пакетов прикладных программ моделирования

ИПК-7.14 (A/01.5 Зн.2) Новейшие методы анализа и обобщения отечественного и международного опыта в области разработки и реализации алгоритмов математических моделей на базе языка программирования Python и пакетов прикладных программ моделирования numpy, sklearn

ИПК-7.15 (A/01.5 Др.1 Зн.) Деятельность, направленная на решение задач аналитического характера, предполагающих выбор и многообразие актуальных способов решения задач математического моделирования на базе языка программирования Python и пакетов прикладных программ моделирования numpy, sklearn

Уметь ИПК-7.16 (D/03.6 У.1) Эффективно использовать существующие типовые решения и шаблоны, современные методы разработки и реализации конкретных алгоритмов математических моделей на базе языка программирования Python и пакетов прикладных программ моделирования numpy, sklearn

ИПК-7.17 (D/03.6 У.2) Грамотно применять методы и средства проектирования программного обеспечения, структур данных, баз данных, программных интерфейсов при реализации конкретных алгоритмов математических моделей на базе языка программирования Python

Владеть ИПК-7.21 (A/01.5 Тд.3) Сбор, обработка, анализ и обобщение результатов экспериментов и исследований в области знаний алгоритмизации математических моделей на базе языка программирования Python и пакетов прикладных программ моделирования numpy, sklearn

Соответствие **<u>продвинутому уровню</u>** освоения компетенций планируемым результатам обучения и критериям их оценивания (оценка: **отлично /зачтено**):

Знать ИПК-4.1 (D/03.6 Зн.2) Современные типовые решения, библиотеки программных модулей, шаблоны, классы объектов, информационные технологии при проектировании, реализации, оценке качества и анализа эффективности программного обеспечения для решения задач в различных предметных областях с помощью языка программирования Python на высоком профессиональном уровне

ИПК-4.2 (D/03.6 Зн.3) Методы и средства, современные информационные технологии проектирования программного обеспечения для решения задач в различных предметных областях

ИПК-4.5 (C/16.6 Зн.2) Современные инструменты, методы и современные информационные технологии проектирования и дизайна ИС, средства разработки программ PyCharm, Google Collab

ИПК-4.7 (C/16.6 Зн.5) Предметная область автоматизации при решении задач в различных предметных областях

ИПК-4.17 (A/01.5 Зн.1) Цели и задачи проводимых исследований и разработок при решении задач в различных предметных областях

ИПК-4.18 (A/01.5 Зн.2) Новейшие методы анализа и обобщения отечественного и международного опыта области информационных технологий на высоком профессиональном уровне

ИПК-4.20 (А/01.5 Зн.4) Современные методы проведения экспериментов и наблюдений, обобщения и обработки результатов работы программного обеспечения

ИПК-4.21 (А/01.5 Др.1 Зн.) Эффективная деятельность, направленная на решение задач аналитического характера, предполагающих выбор и многообразие актуальных способов решения задач с применением современных информационных технологий при проектировании, реализации, оценке качества и анализа эффективности программного обеспечения

Уметь ИПК-4.26 (А/01.5 У.3) Применять новейшие современные методы анализа научно-технической информации

Владеть ИПК-4.30 (А/01.5 Тд.2) Сбор, обработка, анализ и обобщение передового отечественного и международного опыта в области современных информационных технологий, в том числе языка программирования Python и средств разработки программ PyCharm, Google Collab на высоком профессиональном уровне

ИПК-4.31 (А/01.5 Тд.3) Сбор, обработка, анализ и обобщение результатов экспериментов и исследований в области современных информационных технологий, в том числе языка программирования Python и средств разработки программ PyCharm, Google Collab на высоком профессиональном уровне

ПК-7 Способен использовать современные методы разработки и реализации конкретных алгоритмов математических моделей на базе языков программирования и пакетов прикладных программ моделирования

Знать ИПК-7.1 (D/03.6 Зн.1) Современные принципы построения архитектуры программного обеспечения и виды архитектуры программного обеспечения, современные методы разработки и реализации конкретных алгоритмов математических моделей на языке программирования Python на высоком профессиональном уровне

ИПК-7.2 (D/03.6 Зн.2) Современные типовые решения, библиотеки программных модулей, шаблоны, классы объектов, используемые при разработке и реализации конкретных алгоритмов математических моделей на базе языков программирования, в том числе Python и пакетов прикладных программ моделирования

ИПК-7.14 (А/01.5 Зн.2) Новейшие методы анализа и обобщения отечественного и международного опыта в области разработки и реализации алгоритмов математических моделей на базе языка программирования Python и пакетов прикладных программ моделирования numpy, sklearn на высоком профессиональном уровне

ИПК-7.15 (А/01.5 Др.1 Зн.) Деятельность, направленная на решение задач аналитического характера, предполагающих выбор и многообразие актуальных способов решения задач математического моделирования на базе языка программирования Python и пакетов прикладных программ моделирования numpy, sklearn

Уметь ИПК-7.16 (D/03.6 У.1) Эффективно использовать существующие типовые решения и шаблоны, современные методы разработки и реализации конкретных алгоритмов математических моделей на базе языка программирования Python и пакетов прикладных программ моделирования numpy, sklearn на высоком профессиональном уровне

ИПК-7.17 (D/03.6 У.2) Грамотно применять методы и средства проектирования программного обеспечения, структур данных, баз данных, программных интерфейсов при реализации конкретных алгоритмов математических моделей на базе языка программирования Python

Владеть ИПК-7.21 (А/01.5 Тд.3) Сбор, обработка, анализ и обобщение результатов экспериментов и исследований в области знаний алгоритмизации математических моделей на базе языка программирования Python и пакетов прикладных программ моделирования numpy, sklearn на высоком профессиональном уровне

Типовые контрольные задания или иные материалы, необходимые для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций в процессе освоения образовательной программы

Типовые тестовые задания

1) Метеоролог Вася хочет построить модель, прогнозирующую температуру воздуха на завтра. К какому типу относится данная задача?

- Кластеризация
- Классификация
- Ранжирование
- Регрессия

2) Доктор Петя хочет построить модель, прогнозирующую возможность развития осложнений у пациента после операции - все ли будет благополучно в течение нескольких следующих месяцев или нет. К какому типу относится данная задача?

- Ранжирование
- Регрессия
- Кластеризация
- Классификация

3) Астроном Витя хочет построить модель, которая сможет разбить известные науке звезды на группы по их характеристикам, чтобы лучше изучить их особенности. К какому типу относится данная задача?

- Классификация
- Регрессия
- Кластеризация
- Ранжирование

4) В задаче какого типа в обучающей выборке для объектов НЕТ ответов?

- Регрессия
- Классификация
- Кластеризация
- Во всех этих задачах в обучающей выборке есть ответы

5) Выберите все верные утверждения:

- Модель машинного обучения, по сути, является отображением пространства ответов в пространство объектов
- Функционал ошибки показывает, насколько плохое качество имеют данные, используемые для решения задачи
- Процесс обучения модели заключается в минимизации функционала ошибки
- Элементами обучающей выборки являются объекты, характеристики которых являются значениями признаков

6) Рассмотрим признак "Образовательная программа" при анализе данных по студентам университета. Этот признак может принимать три значения: "Экономика", "Математика", "Философия". Воспользуемся one-hot кодированием и заменим этот признак на три бинарных, которые будут соответствовать категориям в том порядке, в котором они перечислены выше. Как будет закодирован признак со значением "Философия"?

- (0, 0, 1)
- (0, 1, 0)

- (1, 0, 0)

7) Чему будет равен корень из среднеквадратичной ошибки для набора из 3 наблюдений, где отклонение предсказания линейной регрессии от реальных значений равны: -1, 2, 2?

- 0
- 3
- 2

8) Предположим, что мы строим модель предсказания роста по возрасту и весу человека. Модель с каким коэффициентами вероятнее всего переобучилась?

- (возраст) + 0.5 * (вес)
- 1402325.3 * (возраст) + -1404370.5 (вес)
- (возраст) + 0.33 (вес)

9) Чем стохастический градиентный спуск (SGD) лучше обычного градиентного спуска? Выберите все подходящие ответы.

- Один шаг в SGD быстрее, поэтому в целом этот метод может быстрее выдать решение.
- В SGD гарантируется, что на каждой итерации уменьшается ошибка модели.
- В SGD гарантируется, что будет найден глобальный минимум.
- Один шаг в SGD точнее, чем в обычном градиентном спуске, поэтому требуется меньше шагов для получения решения.

10) Заполните пропуск: «Для выпуклой функции .... указывает сторону наискорейшего убывания»

- Изохрона
- Линия уровня
- Антиградиент
- Градиент

11) Что может являться критерием останова в градиентном спуске?

- Норма градиента на текущем шаге
- Норма разницы весов на соседних шагах алгоритма
- Сумма элементов вектора градиента
- Число шагов алгоритма

12) Рассмотрим два объекта: у первого отступ линейного классификатора равен -10, у второго -1000. Выберите верные утверждения про то, как соотносятся эти два объекта:

- Первый объект находится дальше от разделяющей поверхности, чем второй
- Второй объект находится дальше от разделяющей поверхности, чем первый -
- Модель больше уверена в своём ответе на первом объекте, чем на втором
- Модель больше уверена в своём ответе на втором объекте, чем на первом -

13) *Рассмотрим пользователя социальной сети как объект в задаче машинного обучения. Что из перечисленного является задачей классификации?*

- Предсказание заработной платы пользователя
- Предсказание пола пользователя
- Предсказание профессии пользователя
- Предсказание, какой пост пользователь сделает следующим

14) *Чем задача классификации с пересекающимися классами (П) отличается от задачи классификации с непересекающимися классами (Н)?*

- В задаче П один объект может относиться к нескольким классам одновременно, а в задаче Н один объект относится ровно к одному классу
- В задаче П один объект может относиться либо к одному классу, либо ни к какому классу, а в задаче Н один объект относится ровно к одному классу
- В задаче П один признак может использоваться для предсказания нескольких классов, а в задаче Н один признак используется для предсказания только одного класса
- В задаче П для объектов обучающей выборки известны метки классов, а в задаче Н - нет

15) *Выберите, какой из приведенных алгоритмов предсказания стоимости квартиры является примером решающего дерева:*

- - Если площадь больше 100 кв. метров, то 4 млн руб., иначе если этаж первый, то 2.3 млн. рублей, иначе 3.15 млн. рублей
- 400 тыс. рублей. за каждый квадратный метр и 27 тыс. руб. за каждый этаж
- Использовать стоимость квартиры, наиболее похожей на текущую, из продаваемых ранее
- Продавать за 3 млн. рублей

16) *Метод главных компонент… (один правильный ответ)*

- Строит новые признаки как сложные нелинейные функции от исходных признаков
- Строит новые признаки как линейные функции от исходных признаков
- Отбирает самые важные признаки из исходных

- Предсказывает класс объекта на основе главных компонент

17) *Выберите верные утверждения про задачу кластеризации:*
- В задаче кластеризации сложно придумать универсальный критерий качества, говорящий, хорошая или плохая получилась кластеризация, что усложняет решение этой задачи
- На практике данные, как правило, не разделяются на четкие кластеры, что усложняет решение задачи кластеризации
- Метод k-Means - самый эффективный метод кластеризации, способный находить качественную кластеризацию практически любых данных
- Задача кластеризации предполагает выделение самых важных признаков и удаление всех остальных признаков

18) *Какие преимущества дает отбор признаков?*
- Чем меньше признаков, тем более сложные зависимости может восстановить алгоритм
- Чем меньше признаков, тем быстрее работает алгоритм
- Отбор признаков уменьшает объем данных и затраты на их хранение
- Отбор неинформативных признаков может повысить качество работы алгоритма

19) *Ансамлевые методы обычно делают более качественные предсказания, чем отдельные алгоритмы (у ансамблей ниже ошибка на тестовой выборке). А какой показатель всегда ухудшается при ансамблировании?*
- *Ошибка на обучающей выборке*
- Скорость выполнения предсказаний
- Количество данных, необходимое для обучения
- Число используемых алгоритмом признаков

20) *Алгоритм бэггинг подразумевает выбор случайных частей данных, обучение алгоритма (например, решающего дерева) на каждой части и составление ансамбля из обученных алгоритмов. Для чего нужно обучать алгоритмы на разных частях данных?*
- Для того, чтобы обученные алгоритмы выполняли несовпадающие (различные) предсказания
- Для того, чтобы ускорить обучение
- Для того, чтобы выделить важные для обучения объекты
- Без этого невозможно обучить решающее дерево

21) *Вы обучили три модели для предсказания суммы кредита, которая потребуется клиенту, и выполняете предсказание для нового клиента: линейная модель предсказывает 50 тысяч, решающее дерево - 30 тысяч, метод ближайшего соседа – 20 тысяч. Какое предсказание выполнит ансамбль трех алгоритмов, если вес линейной модели 0.4, вес решающего дерева 0.4, а вес метода ближайшего соседа - 0.2?*
- 33.3 тысяч
- 30 тысяч
- 36 тысяч
- 38 тысяч

Типовые контрольные задания

1. Реализуйте алгоритм линейной регрессии, и полиномиальной регрессии (для датасета noisysine – степеней от 2 до 5, для датасета hydrodynamics – степени 2) без регуляризации.
2. Реализуйте алгоритм гребневой регрессии и найдите оптимальный параметр регуляризации.
3. Реализуйте алгоритм логистической регрессии со стохастическим градиентным спуском, обучите его на датасете spambase_old (train) и проверьте на датасете spambase_new (val). Получите ROC кривые для вариантов без нормировки и с нормировкой признаков.
4. Реализуйте алгоритм kNN классификации по k ближайшим соседям, используя простое евклидовое расстояние.
5. Примените метод SVM (например, из библиотеки sklearn) для датасета blobs2. Визуализируйте результат (разбиение плоскости и опорные вектора) при разных вариантах ядер (линейное; полиномиальное степеней 2,3,5; RBF).
6. Реализуйте мультиклассовую классификацию с softmax в качестве решающей функции и кросс-энтропией в качестве функции потерь и обучите на подготовленном датасете mnist.
7. Реализуйте алгоритмы построения дерева с критерием информационного выигрыша и критерием Джини и определению класса по мажоритарному классу в листе. Найдите оптимальную глубину дерева в обоих случаях (в отрезке 2-10).
8. Реализуйте алгоритм k-means для кластеризации на 2-4 кластера.

9. Реализуйте алгоритм DBSCAN, найдите параметры для кластеризации на 4 кластера.
10. Примените ансамбевые методы для решения задачи классификации на подготовленном датасете mnist.

Зачетно-экзаменационные материалы для промежуточной аттестации (экзамен/зачет)

Вопросы для подготовки к экзамену

1. Препроцессинг. Масштабирование. Нормировка. Полиномиальные признаки. One-hot encoding.
2. Типы обучения: с учителем, без учителя, с подкреплением, с частичным участием учителя, активное обучение.
3. Смещение и дисперсия (bias and variance). Понятие средней гипотезы.
4. Ошибка внутри и вне выборки. Ошибка обобщения. Неравенство Хёфдинга. Валидация и кросс-валидация.
5. Линейная регрессия.
6. Полиномиальная регрессия.
7. Гребневая регрессия.
8. Логистическая регрессия. Градиентный спуск.
9. Байесовский классификатор. Типы оценки распределений признаков (Gaussian, Bernoulli, Multinomial). EM алгоритм.
10. Метод опорных векторов. Постановка задачи. Формулировка и решение двойственной задачи. Типы опорных векторов. Ядра.
11. Пороговые условия. Эффективность по Парето.
12. Presicion-Recall и ROC кривые. AUC.
13. Метрические классификаторы. kNN. WkNN. Отбор эталонов. DROP5. Kdtree.
14. Кластеризация. kMeans, MeanShift, DBSCAN, Affinity Propagation.
15. Деревья решений. Информационный выигрыш, критерий Джини.
16. Регуляризация деревьев. Небрежные решающие деревья.
17. Бустинг деревьев решений.
18. Ансамблевые методы. Soft and Hard Voting. Bagging.
19. Случайные леса. AdaBoost.
20. Ансамблевые методы регрессии. RANSAC. Theil-Sen. Huber.

4.2 Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций

Методические рекомендации, определяющие процедуры оценивания тестов:

Тест проводится онлайн в системе Moodle или Google Docs и ограничен по времени. На сдачу теста дается две попытки. Тест считается успешно пройденным если студент правильно ответил на 70% вопросов.

Методические рекомендации, определяющие процедуры оценивания выполнения контрольных заданий:

Задание считается выполненным при выполнении следующих условий:
– предоставлен исходный код в среде PyCharm, Google Collab
– продемонстрирована работоспособность программы

– студент понимает исходный код и отвечает на вопросы по его организации.

Методические рекомендации к сдаче экзамена

Процедура промежуточной аттестации проходит в соответствии с Положением о текущем контроле и промежуточной аттестации обучающихся ФГБОУ ВО «КубГУ».

Текущий контроль успеваемости студентов проводится в целях совершенствования и непрерывного контроля качества образовательного процесса, проверки усвоения учебного материала, активизации самостоятельной работы студентов, стимулирования их учебной работы, обеспечения эффективности образовательного процесса, предупреждения рисков отчисления студентов.

Текущий контроль знаний студентов осуществляется постоянно в течение всего семестра.

Виды текущего контроля: устный (письменный) опрос на занятиях; проверка выполнения домашних заданий; проведение контрольных работ; оценка активности студента на занятии.

При отсутствии зачетной книжки у студента экзаменатор не имеет права принимать у него зачет/экзамен. Такой студент считается не явившимся на зачет/экзамен. В исключительных случаях, на основании распоряжения декана преподаватель может допустить студента к зачету/экзамену при наличии документа, удостоверяющего личность.

В целях объективного оценивания знаний во время проведения зачетов и экзаменов не допускается наличие у студентов посторонних предметов и технических устройств.

Студенту, использующему в ходе экзамена неразрешенные источники и средства получения информации, выставляется неудовлетворительная оценка, и он удаляется из аудитории.

Во время экзамена студенты могут пользоваться утвержденной рабочей программой учебной дисциплины, которая должна быть в наличии на экзамене, а также с разрешения экзаменатора справочной литературой и другими пособиями.

Студенты, нарушающие правила поведения при проведении зачетов и экзаменов, могут быть незамедлительно удалены из аудитории, к ним могут быть применены меры дисциплинарного воздействия.

На зачете/экзамене могут присутствовать ректор, проректор по учебной работе, декан факультета, заведующий кафедрой, которая обеспечивает учебный процесс по данной дисциплине. Присутствие на экзаменах и зачетах посторонних лиц без разрешения ректора или проректора по учебной работе не допускается.

После прослушивания лекции рекомендуется выполнить упражнения, приводимые в лекции для самостоятельной работы, а также выполнить на компьютере с использованием среды Python задачи, приводимые в лекции в качестве примеров.

При самостоятельной работе студентов необходимо изучить литературу, приведенную в перечнях, для осмысления вводимых понятий, анализа предложенных подходов и методов разработки параллельных программ. Разрабатывая решение новой

задачи, студент должен уметь выбрать методы решения задачи с учетом целевой аппаратной платформы, проводить отладку и профилирование программы на языке Python.

В качестве систем программирования для решения задач и изучения методов и алгоритмов, приведенных в лекциях, рекомендуется использовать на практических занятиях и при самостоятельной работе стандартную реализацию языка Python в связке со средой разработки PyCharm, Google Collab.

Методические рекомендации, определяющие процедуры оценивания на экзамене:

Критерии оценивания и шкала оценки экзамена

Оценка	**Критерии выставления оценки**
«Отлично»	Дан развернутый ответ на поставленные вопросы. Материал изложен последовательно. Имеются логичные и аргументированные выводы. В течении семестра студент правильно решил 80 % задач, имеет представление как решать остальные задачи
«Хорошо»	Дан развернутый ответ на поставленные вопросы. Материал изложен в целом последовательно. Имеются логичные и аргументированные выводы. В течении семестра студент правильно решил 70 % задач, имеет представление как решать остальные задачи
«Удовлетворительно»	Ответ на вопрос не является полным. Материал изложен непоследовательно. Выводы не аргументированы. В течении семестра студент правильно решил 60 % задач, имеет представление как решать остальные задачи
«Неудовлетворительно»	Обучающийся не знает значительной части программного материала, допускает существенные ошибки. В течении семестра студент правильно решил менее 70 % задач и/или не имеет представление как решать остальные задачи

Оценочные средства для инвалидов и лиц с ограниченными возможностями здоровья выбираются с учетом их индивидуальных психофизических особенностей.

– при необходимости инвалидам и лицам с ограниченными возможностями здоровья предоставляется дополнительное время для подготовки ответа на экзамене;

– при проведении процедуры оценивания результатов обучения инвалидов и лиц с ограниченными возможностями здоровья предусматривается использование технических средств, необходимых им в связи с их индивидуальными особенностями;

– при необходимости для обучающихся с ограниченными возможностями здоровья и инвалидов процедура оценивания результатов обучения по дисциплине может проводиться в несколько этапов.

Процедура оценивания результатов обучения инвалидов и лиц с ограниченными возможностями здоровья по дисциплине предусматривает предоставление информации в формах, адаптированных к ограничениям их здоровья и восприятия информации:

Для лиц с нарушениями зрения:

– в печатной форме увеличенным шрифтом,

– в форме электронного документа.

Для лиц с нарушениями слуха:

– в печатной форме,

– в форме электронного документа.

Для лиц с нарушениями опорно-двигательного аппарата:

– в печатной форме,

– в форме электронного документа.

Данный перечень может быть конкретизирован в зависимости от контингента обучающихся.

5. Перечень учебной литературы, информационных ресурсов и технологий

5.1 Основная литература:

1. Миркин, Б. Г. Введение в анализ данных : учебник и практикум / Б. Г. Миркин. — Москва : Издательство Юрайт, 2020. — 174 с. — (Высшее образование). — ISBN 978-5-9916-5009-0. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: https://urait.ru/bcode/450262

2. Анализ данных : учебник для вузов / В. С. Мхитарян [и др.] ; под редакцией В. С. Мхитаряна. — Москва : Издательство Юрайт, 2021. — 490 с. — (Высшее образование). — ISBN 978-5-534-00616-2. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: https://urait.ru/bcode/469022

3. Келлехер, Д. Наука о данных: базовый курс : [16+] / Д. Келлехер, Б. Тирни ; науч. ред. З. Мамедьяров ; пер. с англ. М. Белоголовского. – Москва : Альпина Паблишер, 2020. – 224 с. : схем., табл. – Режим доступа: по подписке. – URL: https://biblioclub.ru/index.php?page=book&id=598235 (дата обращения: 22.08.2021). – ISBN 978-5-9614-3170-4. – Текст : электронный.

Для освоения дисциплины инвалидами и лицами с ограниченными возможностями здоровья имеются издания в электронном виде в электронно-библиотечных системах *«Лань» и «Юрайт».*

5.2 Дополнительная литература:

1. Маккинли У. (пер. с англ. СлинкинА.А.), Python и анализ данных // Издательство “ДМК Пресс”, 2015, 482 с.

2. Мюллер А., Гвидо С. Введение в машинное обучение с помощью Python. Руководство для специалистов по работе с данными // Вильямс, 2017, 480 с.

3. Программные системы статистического анализа: обнаружение закономерностей в данных с использованием системы R и языка Python : [16+] / В. М. Волкова, М. А. Семенова, Е. С. Четвертакова, С. С. Вожов. – Новосибирск : Новосибирский государственный технический университет, 2017. – 74 с. : ил., табл. – Режим доступа: по подписке. – URL: https://biblioclub.ru/index.php?page=book&id=576496 (дата обращения: 02.07.2021). – Библиогр.: с. 48. – ISBN 978-5-7782-3183-2. – Текст : электронный.

5.3. Периодические издания:

1. Базы данных компании «Ист Вью» http://dlib.eastview.com
2. Электронная библиотека GREBENNIKON.RU https://grebennikon.ru/

5.4. Интернет-ресурсы, в том числе современные профессиональные базы данных и информационные справочные системы

Электронно-библиотечные системы (ЭБС):

1. ЭБС «ЮРАЙТ» https://urait.ru/
2. ЭБС «УНИВЕРСИТЕТСКАЯ БИБЛИОТЕКА ОНЛАЙН» www.biblioclub.ru
3. ЭБС «BOOK.ru» https://www.book.ru
4. ЭБС «ZNANIUM.COM» www.znanium.com
5. ЭБС «ЛАНЬ» https://e.lanbook.com

5.5.Профессиональные базы данных:

1. Web of Science (WoS) http://webofscience.com/
2. Scopus http://www.scopus.com/
3. ScienceDirect www.sciencedirect.com
4. Журналы издательства Wiley https://onlinelibrary.wiley.com/
5. Научная электронная библиотека (НЭБ) http://www.elibrary.ru/
6. Полнотекстовые архивы ведущих западных научных журналов на Российской платформе научных журналов НЭИКОН http://archive.neicon.ru
7. Национальная электронная библиотека (доступ к Электронной библиотеке диссертаций Российской государственной библиотеки (РГБ) https://rusneb.ru/
8. Президентская библиотека им. Б.Н. Ельцина https://www.prlib.ru/
9. Электронная коллекция Оксфордского Российского Фонда https://ebookcentral.proquest.com/lib/kubanstate/home.action
10. Springer Journals https://link.springer.com/
11. Nature Journals https://www.nature.com/siteindex/index.html
12. Springer Nature Protocols and Methods https://experiments.springernature.com/sources/springer-protocols
13. Springer Materials http://materials.springer.com/
14. zbMath https://zbmath.org/
15. Nano Database https://nano.nature.com/
16. Springer eBooks: https://link.springer.com/
17. "Лекториум ТВ" http://www.lektorium.tv/
18. Университетская информационная система РОССИЯ http://uisrussia.msu.ru
19. Система организации конкурсов по исследованию данных, а также социальная сеть специалистов по обработке данных и машинному обучению https://www.kaggle.com
20. Портал открытых данных Российской Федерации https://data.gov.ru/
21. Финам.ру — финансовый портал: новости фондового рынка ценных бумаг и экономики, прогнозы и анализ. Инвестиционная компания Финам https://www.finam.ru

5.6. Информационные справочные системы:

1. Консультант Плюс - справочная правовая система (доступ по локальной сети с компьютеров библиотеки)

5.7.Ресурсы свободного доступа:

1. Американская патентная база данных http://www.uspto.gov/patft/
2. Полные тексты канадских диссертаций http://www.nlc-bnc.ca/thesescanada/
3. КиберЛенинка (http://cyberleninka.ru/);
4. Министерство науки и высшего образования Российской Федерации https://www.minobrnauki.gov.ru/;
5. Федеральный портал "Российское образование" http://www.edu.ru/;
6. Информационная система "Единое окно доступа к образовательным ресурсам" http://window.edu.ru/;
7. Единая коллекция цифровых образовательных ресурсов http://school-collection.edu.ru/ .
8. Федеральный центр информационно-образовательных ресурсов (http://fcior.edu.ru/);
9. Проект Государственного института русского языка имени А.С. Пушкина "Образование на русском" https://pushkininstitute.ru/;
10. Справочно-информационный портал "Русский язык" http://gramota.ru/;
11. Служба тематических толковых словарей http://www.glossary.ru/;
12. Словари и энциклопедии http://dic.academic.ru/;
13. Образовательный портал "Учеба" http://www.ucheba.com/;
14. Законопроект "Об образовании в Российской Федерации". Вопросы и ответы http://xn--273--84d1f.xn--p1ai/voprosy_i_otvety

5.8.Собственные электронные образовательные и информационные ресурсы КубГУ:

1. Среда модульного динамического обучения http://moodle.kubsu.ru
2. База учебных планов, учебно-методических комплексов, публикаций и конференций http://mschool.kubsu.ru/
3. Библиотека информационных ресурсов кафедры информационных образовательных технологий http://mschool.kubsu.ru;
4. Электронный архив документов КубГУ http://docspace.kubsu.ru/
5. Электронные образовательные ресурсы кафедры информационных систем и технологий в образовании КубГУ и научно-методического журнала "ШКОЛЬНЫЕ ГОДЫ" http://icdau.kubsu.ru/

6. Методические указания для обучающихся по освоению дисциплины (модуля)

В освоении дисциплины инвалидами и лицами с ограниченными возможностями здоровья большое значение имеет индивидуальная учебная работа (консультации) – дополнительное разъяснение учебного материала.

Индивидуальные консультации по предмету являются важным фактором, способствующим индивидуализации обучения и установлению воспитательного контакта между преподавателем и обучающимся инвалидом или лицом с ограниченными возможностями здоровья.

7. Материально-техническое обеспечение по дисциплине (модулю)

По всем видам учебной деятельности в рамках дисциплины используются аудитории, кабинеты и лаборатории, оснащенные необходимым специализированным и лабораторным оборудованием.

№	Вид работ	Наименование учебной аудитории, ее оснащенность оборудованием и техническими средствами обучения
1.	Лекционные занятия	Аудитория, укомплектованная специализированной мебелью и техническими средствами обучения
2.	Лабораторные занятия	Аудитория, укомплектованная специализированной мебелью и техническими средствами обучения, компьютерами, проектором, программным обеспечением
3.	Практические занятия	Аудитория, укомплектованная специализированной мебелью и техническими средствами обучения
4.	Групповые (индивидуальные) консультации	Аудитория, укомплектованная специализированной мебелью и техническими средствами обучения, компьютерами, программным обеспечением
5.	Текущий контроль, промежуточная аттестация	Аудитория, укомплектованная специализированной мебелью и техническими средствами обучения, компьютерами, программным обеспечением
6.	Самостоятельная работа	Кабинет для самостоятельной работы, оснащенный компьютерной техникой с возможностью подключения к сети «Интернет», программой экранного увеличения и обеспеченный доступом в электронную информационно-образовательную среду университета.

Примечание: Конкретизация аудиторий и их оснащение определяется ОПОП.