

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ



Частное учреждение высшего образования
«Высшая школа предпринимательства (институт)»
(ЧУВО «ВШП»)

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ
Б1.В.13 «Основы машинного обучения»

Направление подготовки: 09.03.02 Информационные системы и технологии

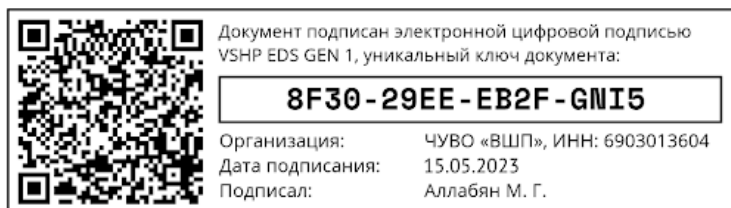
Направленность (профиль) программы бакалавриата
«Информационные системы в экономике»

ОДОБРЕНО

Ученым советом ЧУВО «ВШП»

Протокол заседания

№01-02/23 от 15 мая 2023 г.



Тверь, 2023

Рабочая программа учебной дисциплины **Б1.В.13 Основы машинного обучения**, компонента основной профессиональной образовательной программы высшего образования — программы бакалавриата по направлению подготовки **09.03.02 Информационные системы и технологии** направленность (профиль) **«Информационные системы в экономике»**, направлена на обеспечение у обучающегося способности осуществлять профессиональную деятельность в соответствующей области и сферах профессиональной деятельности, в том числе на их практическую подготовку с учётом рабочей программы воспитания и календарного плана воспитательной работы Частном учреждении высшего образования **«Высшая школа предпринимательства (институт)»** (далее — **ЧУВО «ВШП»**).

1. ОБЛАСТЬ ПРИМЕНЕНИЯ И НОРМАТИВНЫЕ ССЫЛКИ

Настоящая рабочая программа учебной дисциплины устанавливает требования к результатам обучения студента и определяет содержание и виды учебных занятий и отчетности.

Программа предназначена для преподавателей и студентов направления подготовки 09.03.02 Информационные системы и технологии.

Программа учебной дисциплины разработана в соответствии с ФГОС ВО, утвержденного приказом Минобрнауки России от 19.09.2017 № 926 «Об утверждении федерального государственного образовательного стандарта высшего образования - бакалавриата по направлению подготовки 09.03.02 Информационные системы и технологии», основной профессиональной образовательной программой высшего образования по направлению подготовки 09.03.02 Информационные системы и технологии, направленность (профиль) Информационные системы в экономике.

2. ЦЕЛИ И ЗАДАЧИ ДИСЦИПЛИНЫ

Целью изучения дисциплины «Основы машинного обучения» является ознакомление студентов с теоретическими основами и основными принципами машинного обучения и формирование у них практических навыков работы с данными и решения прикладных задач анализа данных.

Для этого в рамках дисциплины решаются следующие задачи:

- изучить основные методы машинного обучения;
- изучить существующие программные библиотеки машинного обучения;
 - научиться самостоятельно реализовывать методы машинного обучения в виде программ;
- научиться применять методы машинного обучения для решения прикладных задач.

3. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОПОП ВО

Дисциплина является компонентом части, формируемой участниками образовательных отношений Блока 1 основной профессиональной образовательной программы высшего образования по направлению подготовки 09.03.02 Информационные системы и технологии, направленность (профиль) — Информационные системы в экономике.

4. ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОБУЧЕНИЯ ПО УЧЕБНОЙ ДИСЦИПЛИНЕ В РАМКАХ ПЛАНИРУЕМЫХ РЕЗУЛЬТАТОВ ОСВОЕНИЯ ОСНОВНОЙ ПРОФЕССИОНАЛЬНОЙ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

Процесс освоения учебной дисциплины направлен на формирование у обучающихся следующих компетенций.

В результате освоения учебной дисциплины обучающийся должен демонстрировать следующие результаты обучения: УК-1, УК-2, ПК-1, ПК-2, ПК-5, ПК-6.

Таблица 1. Результаты обучения

Код компетенции	Наименование компетенции	Индекс и наименование индикатора содержания компетенции	Дескрипторы – основные признаки освоения (показатели достижения результата)
УК-1	Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход	УК-1.1 Способен осуществлять поиск и критический анализ информации	Знать: <ul style="list-style-type: none">- Основные источники информации в области машинного обучения.- Методы критического анализа научных публикаций и статей. Уметь:

	для решения поставленных задач		<ul style="list-style-type: none"> - Осуществлять поиск актуальной информации по теме машинного обучения. - Анализировать и оценивать достоверность и значимость найденной информации. <p>Владеть:</p> <ul style="list-style-type: none"> - Навыками поиска информации в специализированных базах данных и научных журналах. - Методами критического анализа и синтеза информации.
		УК-1.2 Способен применять системный подход для решения поставленных задач	<p>Знать:</p> <ul style="list-style-type: none"> - Принципы системного подхода и его значение в машинном обучении. - Основные этапы разработки и внедрения систем машинного обучения. <p>Уметь:</p> <ul style="list-style-type: none"> - Применять системный подход при разработке моделей машинного обучения. - Интегрировать различные компоненты системы для решения комплексных задач. <p>Владеть:</p> <ul style="list-style-type: none"> - Методами системного анализа и проектирования в контексте машинного обучения. - Навыками разработки комплексных решений с использованием машинного обучения.
УК-2	Способен определять круг задач в рамках поставленной цели и выбирать оптимальные способы их решения, исходя из действующих правовых норм, имеющихся ресурсов и ограничений	УК-2.1 Способен определять круг задач в рамках поставленной цели	<p>Знать:</p> <ul style="list-style-type: none"> - Методы декомпозиции задач в области машинного обучения. - Основные типы задач машинного обучения (регрессия, классификация, кластеризация и т.д.). <p>Уметь:</p> <ul style="list-style-type: none"> - Определять задачи, вытекающие из поставленной цели проекта машинного обучения. - Формулировать задачи и подзадачи для разработки моделей машинного обучения. <p>Владеть:</p> <ul style="list-style-type: none"> - Навыками структурирования и декомпозиции задач. - Методами определения ключевых задач и их приоритизации.
		УК-2.2 Способен выбирать оптимальные способы решения задач, исходя из правовых норм, ресурсов и ограничений	<p>Знать:</p> <ul style="list-style-type: none"> - Основные методы и алгоритмы машинного обучения. - Правовые нормы и ограничения, связанные с обработкой данных и разработкой моделей машинного обучения. <p>Уметь:</p> <ul style="list-style-type: none"> - Выбирать методы и алгоритмы для решения конкретных задач машинного обучения. - Оценивать ресурсы и ограничения, влияющие на выбор решений.

			<p>Владеть:</p> <ul style="list-style-type: none"> - Навыками оценки и выбора оптимальных методов и алгоритмов. - Методами обоснования выбора решений с учетом правовых норм и ресурсов.
ПК-1	Разработка и отладка программного кода	<p>ПК-1.1 Способен разрабатывать программный код</p>	<p>Знать:</p> <ul style="list-style-type: none"> - Основные языки программирования, используемые в машинном обучении (Python, R). - Структуру и синтаксис языков программирования. <p>Уметь:</p> <ul style="list-style-type: none"> - Писать программный код для реализации алгоритмов машинного обучения. - Использовать библиотеки и фреймворки для машинного обучения (scikit-learn, TensorFlow, PyTorch). <p>Владеть:</p> <ul style="list-style-type: none"> - Навыками разработки кода для решения задач машинного обучения. - Методами отладки и оптимизации программного кода.
		<p>ПК-1.2 Способен отлаживать программный код</p>	<p>Знать:</p> <ul style="list-style-type: none"> - Основные методы и инструменты отладки программного кода. - Принципы тестирования и проверки корректности программного обеспечения. <p>Уметь:</p> <ul style="list-style-type: none"> - Отлаживать код для выявления и устранения ошибок. - Использовать инструменты для автоматизированного тестирования. <p>Владеть:</p> <ul style="list-style-type: none"> - Навыками эффективной отладки кода. - Методами тестирования и обеспечения качества программного обеспечения.
ПК-2	Проверка работоспособности и рефакторинг кода программного обеспечения	<p>ПК-2.1 Способен проверять работоспособность программного кода</p>	<p>Знать:</p> <ul style="list-style-type: none"> - Методы и инструменты тестирования программного обеспечения. - Принципы обеспечения надежности и стабильности кода. <p>Уметь:</p> <ul style="list-style-type: none"> - Проводить тестирование и верификацию программного кода. - Оценивать работоспособность и производительность кода. <p>Владеть:</p> <ul style="list-style-type: none"> - Навыками тестирования и верификации программного обеспечения. - Методами анализа и улучшения производительности кода.
		<p>ПК-2.2 Способен рефакторить программный код</p>	<p>Знать:</p> <ul style="list-style-type: none"> - Принципы рефакторинга и оптимизации кода. - Стандарты и лучшие практики написания чистого и поддерживаемого кода. <p>Уметь:</p>

			<ul style="list-style-type: none"> - Проводить рефакторинг для улучшения структуры и читаемости кода. - Оптимизировать код для повышения его производительности. <p>Владеть:</p> <ul style="list-style-type: none"> - Навыками рефакторинга и оптимизации кода. - Методами обеспечения читаемости и поддерживаемости программного кода.
ПК-5	Анализ экономических данных и применение методов экономического моделирования для обоснования управленческих решений	ПК-5.1 Способен проводить анализ экономических данных	<p>Знать:</p> <ul style="list-style-type: none"> - Методы сбора и анализа экономических данных. - Основные экономические показатели и их интерпретацию. <p>Уметь:</p> <ul style="list-style-type: none"> - Проводить сбор и предварительную обработку экономических данных. - Анализировать экономические данные с использованием методов машинного обучения. <p>Владеть:</p> <ul style="list-style-type: none"> - Навыками анализа и интерпретации экономических данных. - Методами применения машинного обучения для анализа экономических данных.
		ПК-5.2 Способен применять методы экономического моделирования	<p>Знать:</p> <ul style="list-style-type: none"> - Основные методы экономического моделирования. - Принципы построения экономических моделей с использованием машинного обучения. <p>Уметь:</p> <ul style="list-style-type: none"> - Разрабатывать и оценивать экономические модели. - Применять методы машинного обучения для решения задач экономического моделирования. <p>Владеть:</p> <ul style="list-style-type: none"> - Навыками разработки и применения экономических моделей. - Методами использования машинного обучения для экономического моделирования.
ПК-6	Применение математических и статистических методов для анализа данных и принятия решений в условиях неопределенности.	ПК-6.1 Способен применять математические методы анализа данных	<p>Знать:</p> <ul style="list-style-type: none"> - Основные математические методы, используемые в анализе данных. - Принципы и методы машинного обучения для анализа данных. <p>Уметь:</p> <ul style="list-style-type: none"> - Применять математические методы для анализа данных. - Использовать алгоритмы машинного обучения для анализа данных. <p>Владеть:</p> <ul style="list-style-type: none"> - Навыками применения математических методов и алгоритмов машинного обучения для анализа данных. - Методами интерпретации результатов анализа данных.

		ПК-6.2 Способен применять статистические методы для анализа данных	Знать: - Основные статистические методы, используемые в анализе данных. - Принципы статистического анализа и интерпретации данных. Уметь: - Применять статистические методы для анализа данных. - Использовать статистические методы в машинном обучении для построения моделей. Владеть: - Навыками применения статистических методов для анализа данных. - Методами использования статистических методов в контексте машинного обучения.
--	--	---	--

5. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

5.1. Объем дисциплины в зачетных единицах с указанием количества академических часов

Общая трудоемкость учебной дисциплины составляет 4 зачетные единицы, 144 часа, включая все формы контактной и самостоятельной работы обучающихся.

Объем дисциплины по учебному плану составляет –
4 зачётных единицы = 144 академических часа.

Контактная работа обучающегося (студенты) с научно-педагогическим работником организации (всего) - 50 академических часов,

в том числе:

Лекционные занятия (Лек.) - 18 академических часов,

Практические занятия (Пр.) - 30 академических часов,

Консультации (Конс.) - 2 академических часа.

Самостоятельная работа обучающегося (студента):

Самостоятельная работа (СР) - 58 академический час,

Текущий контроль успеваемости

и промежуточно-заочная аттестация обучающегося (студента):

Часы на контроль - 36 академических часов.

Таблица 2. Объём дисциплины

№ п/п	Раздел/тема дисциплины	Семестр/курс	Виды учебной деятельности, включая самостоятельную работу обучающихся (студентов), и трудоёмкость (в ак. часах)				Коды формируемых компетенций
			Виды учебных занятий по дисциплине			Самостоятельная работа	
			Лек.	Пр.	Конс.		
1	Тема 1. Введение в машинное обучение	9 семестр/ 5 курс	1	2		4	УК-1.1, УК-1.2
2	Тема 2. Основы статистики и вероятности для машинного обучения	9 семестр/ 5 курс	1	2		4	ПК-6.1, ПК-6.2

3	Тема 3. Предобработка данных	9 семестр/ 5 курс	2	2		4	УК-2.1, УК-2.2
4	Тема 4. Линейные модели в машинном обучении	9 семестр/ 5 курс	2	2		4	ПК-1.1, ПК-1.2
5	Тема 5. Деревья решений и ансамблевые методы	9 семестр/ 5 курс	1	2		4	ПК-1.1, ПК-1.2
6	Тема 6. Методы снижения размерности	9 семестр/ 5 курс	1	2		4	ПК-6.1, ПК-6.2
7	Тема 7. Кластеризация и методы обучения без учителя	9 семестр/ 5 курс	2	2		4	ПК-6.1, ПК-6.2
8	Тема 8. Нейронные сети и глубокое обучение	9 семестр/ 5 курс	1	2		4	ПК-1.1, ПК-2.2
9	Тема 9: Оценка моделей и перекрестная проверка	9 семестр/ 5 курс	2	2		4	ПК-2.1, ПК-2.2
10	Тема 10: Применение машинного обучения в экономике	9 семестр/ 5 курс	2	4		4	ПК-5.1, ПК-5.2
11	Тема 11: Практическое использование библиотек и инструментов для машинного обучения	9 семестр/ 5 курс	2	2		4	ПК-1.1, ПК-1.2
12	Тема 12: Этика и правовые аспекты машинного обучения	9 семестр/ 5 курс	1	2		4	УК-1.2, УК-2.2
13	Тема 13: Итоговое занятие и подготовка к экзамену	9 семестр/ 5 курс		4	2	8	УК-1.1, УК-1.2, УК-2.1, УК-2.2, ПК-1.1, ПК-1.2, ПК-2.1, ПК-2.2, ПК-5.1, ПК-5.2, ПК-6.1, ПК-6.2
ИТОГО аудиторных часов/СР:		9 семестр/ 5 курс	48 ак. часов			60 ак. часа	-
Часы на контроль		9 семестр/ 5 курс	36 ак. час (форма промежуточной аттестации – экзамен – 9 семестр)				
ВСЕГО ак. часов:		9 семестр/ 5 курс	144 академических часа				

5.2. Тематическое содержание дисциплины

* количество академических часов и виды занятий представлены в таблице № 2.

Тема 1: Введение в машинное обучение (УК-1.1, УК-1.2)

Содержание: Основные понятия и принципы машинного обучения. Классификация методов

машинного обучения: обучение с учителем, обучение без учителя, обучение с подкреплением. Примеры применения машинного обучения в различных отраслях.

Тема 2: Основы статистики и вероятности для машинного обучения (ПК-6.1, ПК-6.2)

Содержание: Основные понятия статистики и теории вероятностей. Дискретные и непрерывные случайные величины. Распределения вероятностей. Проверка гипотез и доверительные интервалы. Роль статистики в машинном обучении.

Тема 3: Предобработка данных (УК-2.1, УК-2.2)

Содержание: Методы очистки, нормализации и трансформации данных. Работа с пропущенными значениями. Принципы выбора признаков и их влияние на результаты моделей. Использование библиотек для предобработки данных (pandas, sklearn).

Тема 4: Линейные модели в машинном обучении (ПК-1.1, ПК-1.2)

Содержание: Линейная регрессия и классификация. Математическое описание и интерпретация линейных моделей. Градиентный спуск и его вариации. Применение линейных моделей к реальным данным.

Тема 5: Деревья решений и ансамблевые методы (ПК-1.1, ПК-1.2)

Содержание: Основы деревьев решений. Преимущества и недостатки деревьев решений. Ансамблевые методы: бэггинг, бустинг и случайные леса. Применение ансамблевых методов для улучшения точности моделей.

Тема 6: Методы снижения размерности (ПК-6.1, ПК-6.2)

Содержание: Принципы и методы снижения размерности данных. Главные компоненты анализа (PCA). Линейный дискриминантный анализ (LDA). Применение методов снижения размерности в задачах визуализации и предварительного анализа данных.

Тема 7: Кластеризация и методы обучения без учителя (ПК-6.1, ПК-6.2)

Содержание: Основные алгоритмы кластеризации: k-means, иерархическая кластеризация, DBSCAN. Методы оценки качества кластеризации. Примеры применения кластеризации в различных отраслях.

Тема 8: Нейронные сети и глубокое обучение (ПК-1.1, ПК-2.2)

Содержание: Основы нейронных сетей. Архитектуры и виды нейронных сетей: многослойные перцептроны, сверточные нейронные сети, рекуррентные нейронные сети. Основы обучения глубоких моделей. Применение глубокого обучения в реальных задачах.

Тема 9: Оценка моделей и перекрестная проверка (ПК-2.1, ПК-2.2)

Содержание: Методы оценки качества моделей: точность, полнота, F-мера, ROC-кривая. Принципы перекрестной проверки. Методы борьбы с переобучением. Применение методов оценки моделей на практике.

Тема 10: Применение машинного обучения в экономике (ПК-5.1, ПК-5.2)

Содержание: Примеры использования машинного обучения для анализа экономических данных. Прогнозирование временных рядов. Модели для оценки кредитного риска. Анализ рыночных данных и инвестиционных стратегий с помощью машинного обучения.

Тема 11: Практическое использование библиотек и инструментов для машинного обучения (ПК-1.1, ПК-1.2)

Содержание: Обзор популярных библиотек и инструментов для машинного обучения (scikit-learn, TensorFlow, Keras, PyTorch). Примеры использования библиотек для решения типовых задач машинного обучения. Практические задания по разработке и оценке моделей.

Тема 12: Этика и правовые аспекты машинного обучения (УК-1.2, УК-2.2)

Содержание: Основные этические вопросы в области машинного обучения. Влияние машинного обучения на общество и экономику. Правовые аспекты обработки данных и использования моделей машинного обучения. Принципы ответственного использования технологий машинного обучения.

Тема 13: Итоговое занятие и подготовка к экзамену (УК-1.1, УК-1.2, УК-2.1, УК-2.2, ПК-1.1, ПК-1.2, ПК-2.1, ПК-2.2, ПК-5.1, ПК-5.2, ПК-6.1, ПК-6.2)

Содержание: Повторение и обобщение пройденного материала. Решение типовых задач и вопросов для подготовки к экзамену. Обсуждение практических аспектов использования методов машинного обучения.

5.2.1 Содержание практических занятий

Таблица 3

№ п/п	Наименование темы (раздела) дисциплины	Содержание практического занятия
1	Практическое занятие 1: Введение в машинное обучение	Задание: Установить и настроить среды разработки (Anaconda, Jupyter Notebook). Создать и выполнить первый проект на Python с использованием библиотеки scikit-learn. Цель: Ознакомиться с инструментами и средами разработки для машинного обучения.
2	Практическое занятие 2: Основы статистики и вероятности для машинного обучения	Задание: Выполнить расчет основных статистических показателей (среднее, медиана, дисперсия) и построить распределения данных с использованием библиотеки pandas. Цель: Понять роль статистики в машинном обучении и научиться использовать основные статистические методы.
3	Практическое занятие 3: Предобработка данных	Задание: Провести очистку и нормализацию набора данных. Обработать пропущенные значения и выполнить преобразование признаков. Цель: Научиться применять методы предобработки данных для подготовки данных к моделированию.
4	Практическое занятие 4: Линейные модели в машинном обучении	Задание: Реализовать линейную регрессию на практике с использованием scikit-learn. Построить модель, провести обучение и оценить качество модели. Цель: Понять основы линейных моделей и их применение в машинном обучении.
5	Практическое занятие 5: Деревья решений и ансамблевые методы	Задание: Реализовать деревья решений и случайные леса на практике. Сравнить результаты одиночных деревьев решений и ансамблевых методов. Цель: Освоить методы деревьев решений и ансамблей для повышения точности моделей.
6	Практическое занятие 6: Методы снижения размерности	Задание: Применить PCA и LDA для снижения размерности данных и визуализации результатов. Сравнить влияние снижения размерности на точность моделей. Цель: Научиться применять методы снижения размерности и интерпретировать результаты.
7	Практическое занятие 7: Кластеризация и методы обучения без учителя	Задание: Реализовать алгоритмы кластеризации (k-means, иерархическая кластеризация). Оценить качество кластеризации и визуализировать результаты. Цель: Понять принципы и методы кластеризации данных.
8	Практическое занятие 8: Нейронные сети и глубокое обучение	Задание: Создать простую нейронную сеть с использованием TensorFlow и Keras. Обучить модель на наборе данных и оценить ее производительность. Цель: Ознакомиться с основами нейронных сетей и глубокого обучения.
9	Практическое занятие 9: Оценка моделей и перекрестная проверка	Задание: Реализовать методы оценки моделей (точность, полнота, F-мера). Провести перекрестную проверку для оценки стабильности модели. Цель: Научиться применять методы оценки моделей и предотвращения переобучения.
10	Практическое занятие 10: Применение	Задание: Разработать модель для прогнозирования временных рядов с использованием экономических данных. Применить методы машинного

	машинного обучения в экономике	обучения для анализа рыночных данных. Цель: Понять применение машинного обучения в анализе экономических данных.
1 1	Практическое занятие 11: Использование библиотек и инструментов для машинного обучения	Задание: Использовать библиотеки scikit-learn, TensorFlow, Keras для решения различных задач машинного обучения. Реализовать и оценить модели на практике. Цель: Освоить использование популярных библиотек и инструментов для машинного обучения.
1 2	Практическое занятие 12: Этика и правовые аспекты машинного обучения	Задание: Обсудить этические и правовые вопросы в области машинного обучения. Рассмотреть реальные кейсы и предложить решения для ответственного использования технологий машинного обучения. Цель: Понять важность этических и правовых аспектов в машинном обучении.
1 3	Практическое занятие 13: Итоговое занятие и подготовка к экзамену	Задание: Повторение пройденного материала, решение типовых задач и вопросов. Обсуждение практических аспектов использования методов машинного обучения. Цель: Подготовиться к экзамену, закрепляя изученный материал и практические навыки.

5.2.2 Содержание самостоятельной работы

Таблица 4

№ п/п	Наименование темы (раздела) дисциплины	Содержание самостоятельной работы	Форма контроля
1	Самостоятельная работа 1: Введение в машинное обучение	Задание: Прочитать рекомендованные материалы по основам машинного обучения, подготовить краткий обзор. Цель: Понять основные принципы и классификацию методов машинного обучения.	Ответы на контрольные вопросы, доклад
2	Самостоятельная работа 2: Основы статистики и вероятности для машинного обучения	Задание: Изучить материалы по основам статистики и теории вероятностей, выполнить дополнительные упражнения по расчету статистических показателей. Цель: Освоить основные статистические методы и их применение в машинном обучении. Форма контроля: Ответы на контрольные вопросы, выполнение практического задания.	Ответы на контрольные вопросы, выполнение практического задания.
3	Самостоятельная работа 3: Предобработка данных	Задание: Изучить методы очистки и нормализации данных, провести дополнительное исследование по трансформации данных. Цель: Научиться преобразовывать данные для улучшения качества моделей.	Выполнение практического задания.
4	Самостоятельная работа 4: Линейные модели в машинном обучении	Задание: Прочитать материалы по линейной регрессии и классификации, выполнить упражнения по построению линейных моделей. Цель: Понять математическое описание и применение линейных моделей.	Ответы на контрольные вопросы, выполнение практического задания.
5	Самостоятельная работа 5: Деревья решений и ансамблевые методы	Задание: Изучить материалы по деревьям решений и ансамблевым методам, выполнить упражнения по реализации и оценке этих моделей. Цель: Освоить методы деревьев решений и ансамблей.	Выполнение практического задания.
6	Самостоятельная работа 6: Методы снижения размерности	Задание: Прочитать материалы по методам снижения размерности, провести дополнительные исследования по PCA и LDA. Цель: Научиться применять методы снижения размерности для анализа данных.	Выполнение практического задания.
7	Самостоятельная работа 7: Кластеризация и методы обучения без учителя	Задание: Изучить материалы по алгоритмам кластеризации, выполнить упражнения по реализации методов кластеризации. Цель: Освоить основные методы обучения без учителя.	Выполнение практического задания.
8	Самостоятельная	Задание: Прочитать материалы по основам нейронных	Выполнение

	работа 8: Нейронные сети и глубокое обучение	сетей и глубокого обучения, выполнить упражнения по созданию и обучению нейронных сетей. Цель: Ознакомиться с нейронными сетями и методами глубокого обучения.	практического задания.
9	Самостоятельная работа 9: Оценка моделей и перекрестная проверка	Задание: Изучить методы оценки моделей и перекрестной проверки, провести дополнительное исследование по методам борьбы с переобучением. Цель: Научиться оценивать качество моделей и предотвращать переобучение.	Выполнение практического задания.
10	Самостоятельная работа 10: Применение машинного обучения в экономике	Задание: Прочитать материалы по применению машинного обучения в анализе экономических данных, выполнить упражнения по анализу временных рядов и оценке кредитного риска. Цель: Понять применение машинного обучения для решения экономических задач.	Выполнение практического задания.
11	Самостоятельная работа 11: Использование библиотек и инструментов для машинного обучения	Задание: Изучить документацию по библиотекам scikit-learn, TensorFlow, Keras и выполнить упражнения по их применению для решения типовых задач машинного обучения. Цель: Освоить использование популярных библиотек для машинного обучения.	Ответы на контрольные вопросы, выполнение практического задания.
12	Самостоятельная работа 12: Этика и правовые аспекты машинного обучения	Задание: Прочитать материалы по этическим и правовым вопросам в области машинного обучения, подготовить доклад по одному из кейсов. Цель: Понять этические и правовые аспекты машинного обучения.	Тест.
13	Самостоятельная работа 13: Итоговое занятие и подготовка к экзамену	Задание: Повторение пройденного материала, решение типовых задач и вопросов. Обсуждение практических аспектов использования методов машинного обучения. Цель: Подготовиться к экзамену, закрепляя изученный материал и практические навыки.	Экзамен.

6. Оценочные материалы по дисциплине

Оценочные материалы находятся в документе «Оценочные материалы по дисциплине «Основы машинного обучения»».

7. Методические материалы для обучающихся по освоению дисциплины (модуля)

А) Рекомендации обучающемуся (студенту) по работе с конспектом после лекции

Какими бы замечательными качествами в области методики ни обладал лектор, какое бы большое значение на занятиях ни уделял лекции слушатель, глубокое понимание материала достигается только путем самостоятельной работы над ним. Самостоятельную работу следует начинать с доработки конспекта, желательно в тот же день, пока время не стерло содержание лекции из памяти (через 10 часов после лекции в памяти остается не более 30-40 % материала). С целью доработки необходимо в первую очередь прочитать записи, восстановить текст в памяти, а также исправить описки, расшифровать не принятые ранее сокращения, заполнить пропущенные места, понять текст, вникнуть в его смысл. Далее прочитать материал по рекомендуемой литературе, разрешая в ходе чтения, возникшие ранее затруднения, вопросы, а также дополнения и исправляя свои записи. Записи должны быть наглядными, для чего следует применять различные способы выделений. В ходе доработки конспекта углубляются, расширяются и закрепляются знания, а также дополняется, исправляется и совершенствуется конспект. Подготовленный конспект и рекомендуемая литература используется при подготовке к практическому занятию. Подготовка сводится к внимательному прочтению учебного материала, к выводу с карандашом в руках всех утверждений и формул, к решению примеров, задач, к ответам на вопросы, предложенные в конце лекции преподавателем или помещенные в рекомендуемой литературе. Примеры, задачи, вопросы по теме являются материалом самоконтроля. Непременным условием глубокого усвоения учебного материала является знание основ, на которых строится изложение материала. Обычно преподаватель напоминает, какой

ранее изученный материал и в какой степени требуется подготовить к очередному занятию. Эта рекомендация, как и требование систематической и серьезной работы над всем лекционным курсом, подлежит безусловному выполнению. Потери логической связи как внутри темы, так и между ними приводит к негативным последствиям: материал учебной дисциплины перестает основательно восприниматься, а творческий труд подменяется утомленным переписыванием. Обращение к ранее изученному материалу не только помогает восстановить в памяти известные положения, выводы, но и приводит разрозненные знания в систему, углубляет и расширяет их. Каждый возврат к старому материалу позволяет найти в нем что-то новое, переосмыслить его с иных позиций, определить для него наиболее подходящее место в уже имеющейся системе знаний. Неоднократное обращение к пройденному материалу является наиболее рациональной формой приобретения и закрепления знаний. Очень полезным, но, к сожалению, еще мало используемым в практике самостоятельной работы, является предварительное ознакомление с учебным материалом. Даже краткое, беглое знакомство с материалом очередной лекции дает многое. Обучающиеся (студенты) получают общее представление о её содержании и структуре, о главных и второстепенных вопросах, о терминах и определениях. Все это облегчает работу на лекции и делает ее целеустремленной.

Б) Рекомендации обучающемуся (студенту) по подготовке к занятиям семинарского типа

Обучающийся (студент) должен чётко уяснить, что именно с лекции начинается его подготовка к лабораторному/ практическому/ семинарскому/ методическому/ клиническому практическому занятию. Вместе с тем, лекция лишь организует мыслительную деятельность, но не обеспечивает глубину усвоения программного материала. При подготовке к такому виду занятий можно выделить 2 этапа:

1-й - организационный,

2-й - закрепление и углубление теоретических знаний.

На первом этапе обучающийся (студент) планирует свою самостоятельную работу, которая включает:

- уяснение задания на самостоятельную работу;

- подбор рекомендованной литературы;

- составление плана работы, в котором определяются основные пункты предстоящей подготовки.

Составление плана дисциплинирует и повышает организованность в работе. Второй этап включает непосредственную подготовку обучающегося (студента) к занятию. Начинать надо с изучения рекомендованной литературы. Необходимо помнить, что на лекции обычно рассматривается не весь материал, а только его часть. Остальная его часть восполняется в процессе самостоятельной работы. В связи с этим работа с рекомендованной литературой обязательна. Особое внимание при этом необходимо обратить на содержание основных положений и выводов, объяснение явлений и фактов, уяснение практического приложения рассматриваемых теоретических вопросов. В процессе этой работы обучающийся (студент) должен стремиться понять и запомнить основные положения рассматриваемого материала, примеры, поясняющие его, а также разобраться в иллюстративном материале. Заканчивать подготовку следует составлением плана (перечня основных пунктов) по изучаемому материалу (вопросу). Такой план позволяет составить концентрированное, сжатое представление по изучаемым вопросам. В процессе подготовки к семинарскому занятию рекомендуется взаимное обсуждение материала, во время которого закрепляются знания, а также приобретается практика в изложении и разъяснении полученных знаний, развивается речь. При необходимости следует обращаться за консультацией к преподавателю. Идя на консультацию, необходимо хорошо продумать вопросы, которые требуют разъяснения. В начале семинарского занятия обучающиеся (студента) под руководством преподавателя более глубоко осмысливают теоретические положения по теме занятия, раскрывают и объясняют основные явления и факты. В процессе творческого обсуждения и дискуссии вырабатываются умения и навыки использовать приобретенные знания для решения практических задач.

В) Рекомендации по самостоятельной работе обучающегося (студента) над изучаемым материалом

Успешное освоение данного курса базируется на рациональном сочетании нескольких видов учебной деятельности - лекций, семинарских занятий, самостоятельной работы. При этом самостоятельную работу следует рассматривать одним из главных звеньев полноценного высшего образования, на которую отводится значительная часть учебного времени.

Самостоятельная работа студентов складывается из следующих составляющих:

- работа с основной и дополнительной литературой, с материалами интернета и конспектами лекций;
- внеаудиторная подготовка к контрольным работам, выполнение докладов, рефератов и курсовых работ;
- выполнение самостоятельных практических работ;
- подготовка к экзаменам (зачетам) непосредственно перед ними.

Для правильной организации работы необходимо учитывать порядок изучения разделов курса, находящихся в строгой логической последовательности. Поэтому хорошее усвоение одной части дисциплины является предпосылкой для успешного перехода к следующей. Задания, проблемные вопросы, предложенные для изучения дисциплины, в том числе и для самостоятельного выполнения, носят междисциплинарный характер и базируются, прежде всего, на причинно-следственных связях между компонентами окружающего нас мира. В течение семестра, необходимо подготовить рефераты (проекты) с использованием рекомендуемой основной и дополнительной литературы и сдать рефераты для проверки преподавателю. Важным составляющим в изучении данного курса является решение ситуационных задач и работа над проблемно-аналитическими заданиями, что предполагает знание соответствующей научной терминологии и т.д.

Для лучшего запоминания материала целесообразно использовать индивидуальные особенности и разные виды памяти: зрительную, слуховую, ассоциативную. Успешному запоминанию также способствует приведение ярких свидетельств и наглядных примеров. Учебный материал должен постоянно повторяться и закрепляться.

При выполнении докладов, творческих, информационных, исследовательских проектов особое внимание следует обращать на подбор источников информации и методику работы с ними.

Для успешной сдачи экзамена (зачета) рекомендуется соблюдать следующие правила:

1. Подготовка к экзамену (зачету) должна проводиться систематически, в течение всего семестра.
2. Интенсивная подготовка должна начаться не позднее, чем за месяц до экзамена.
3. Время непосредственно перед экзаменом (зачетом) лучше использовать таким образом, чтобы оставить последний день свободным для повторения курса в целом, для систематизации материала и доработки отдельных вопросов.

На экзамене высокую оценку получают студенты, использующие данные, полученные в процессе выполнения самостоятельных работ, а также использующие собственные выводы на основе изученного материала.

Учитывая значительный объем теоретического материала, студентам рекомендуется регулярное посещение и подробное конспектирование лекций.

8. ОСОБЕННОСТИ РЕАЛИЗАЦИИ ДИСЦИПЛИНЫ ДЛЯ ИНВАЛИДОВ И ЛИЦ С ОГРАНИЧЕННЫМИ ВОЗМОЖНОСТЯМИ ЗДОРОВЬЯ

Обучающимся с ограниченными возможностями здоровья предоставляются специальные учебники и учебные пособия, иная учебная литература, специальные технические средства обучения коллективного и индивидуального пользования, предоставление услуг ассистента (помощника), оказывающего обучающимся необходимую техническую помощь, а также услуги сурдопереводчиков и тифлосурдопереводчиков.

- а) для слабовидящих:

- на промежуточной аттестации присутствует ассистент, оказывающий студенту необходимую техническую помощь с учетом индивидуальных особенностей (он помогает занять рабочее место, передвигаться, прочитать и оформить задание, в том числе записывая под диктовку);

- задания для выполнения, а также инструкция о порядке проведения промежуточной аттестации оформляются увеличенным шрифтом;

- задания для выполнения на промежуточной аттестации зачитываются ассистентом;

- письменные задания выполняются на бумаге, надиктовываются ассистенту;

- обеспечивается индивидуальное равномерное освещение не менее 300 люкс;

- студенту для выполнения задания при необходимости предоставляется увеличивающее устройство;

в) для глухих и слабослышащих:

- на промежуточной аттестации присутствует ассистент, оказывающий студенту необходимую техническую помощь с учетом индивидуальных особенностей (он помогает занять рабочее место, передвигаться, прочитать и оформить задание, в том числе записывая под диктовку);

- промежуточно-заочная аттестация проводится в письменной форме;

- обеспечивается наличие звукоусиливающей аппаратуры коллективного пользования, при необходимости поступающим предоставляется звукоусиливающая аппаратура индивидуального пользования;

- по желанию студента промежуточно-заочная аттестация может проводиться в письменной форме;

д) для лиц с нарушениями опорно-двигательного аппарата (тяжелыми нарушениями двигательных функций верхних конечностей или отсутствием верхних конечностей):

- письменные задания выполняются на компьютере со специализированным программным обеспечением или надиктовываются ассистенту;

- по желанию студента промежуточно-заочная аттестация проводится в устной форме.

Примечание:

а) Для обучающегося (бакалавра), осваивающего учебную дисциплину, обязательный компонент основной профессиональной образовательной программы высшего образования — программы бакалавриата по направлению подготовки **09.03.02 Информационные системы и технологии (направленность (профиль) «Информационные системы в экономике»), форма обучения — очно-заочная), одобренной на заседании Учёного совета образовательной организации, утверждённой ректором Частного образовательного учреждения высшего образования «Высшая школа предпринимательства», по индивидуальному учебному плану (при наличии факта зачисления в образовательную организацию такого обучающегося (бакалавра)), Институт:**

- разрабатывает, согласовывает с участниками образовательных отношений и утверждает в установленном порядке согласно соответствующему локальному нормативному акту **индивидуальный учебный план** конкретного обучающегося (бакалавра) (*учебный план, обеспечивающий освоение конкретной основной образовательной программы высшего образования на основе индивидуализации её содержания с учётом особенностей и образовательных потребностей конкретного обучающегося (бакалавра)*);

- устанавливает для конкретного обучающегося (бакалавра) по индивидуальному учебному плану **одинаковые дидактические единицы** — элементы содержания учебного материала, изложенного в виде утверждённой в установленном образовательной организацией порядке согласно соответствующему локальному нормативному акту рабочей программы учебной дисциплины, обязательного компонента разработанной и реализуемой Институтom основной профессиональной образовательной программы высшего образования — программы бакалавриата по направлению подготовки **09.03.02 Информационные системы и технологии** (направленность (профиль) «Информационные системы в экономике»), форма обучения — очно-заочная), как и для обучающегося (бакалавра), осваивающего основную образовательную

программу высшего образования в учебной группе;

- определяет в индивидуальном учебном плане конкретного обучающегося (бакалавра) **объём учебной дисциплины** с указанием количества академических часов/ ЗЕТ, выделенных на его контактную работу (групповую и (или) индивидуальную работу) с руководящими и (или) научно-педагогическими работниками, реализующими основную образовательную программу высшего образования;

- определяет в индивидуальном учебном плане конкретного обучающегося (бакалавра) количество академических часов/ ЗЕТ по учебной дисциплине, выделенных на его самостоятельную работу (*при необходимости*).

б) Для обучающегося (бакалавра) с ограниченными возможностями здоровья и инвалида, осваивающего учебную дисциплину, обязательный компонент основной профессиональной образовательной программы высшего образования — программы бакалавриата по направлению подготовки **09.03.02 Информационные системы и технологии** (направленность (профиль) «Информационные системы в экономике»), форма обучения — очно-заочная), одобренной на заседании Учёного совета образовательной организации, утверждённой ректором Частного образовательного учреждения высшего образования «Высшая школа предпринимательства», (*при наличии факта зачисления в образовательную организацию такого обучающегося (бакалавра) с учётом конкретной (конкретных) нозологии (нозологий)*), **Институт:**

- разрабатывает, согласовывает с участниками образовательных отношений и утверждает в установленном порядке согласно соответствующему локальному нормативному акту **индивидуальный учебный план** конкретного обучающегося (бакалавра) с ограниченными возможностями здоровья/ инвалида (*при наличии факта зачисления в образовательную организацию такого обучающегося (бакалавра) с учётом конкретной (конкретных) нозологии (нозологий)*) (учебный план, обеспечивающий освоение конкретной основной образовательной программы высшего образования на основе индивидуализации её содержания с учётом особенностей и образовательных потребностей конкретного обучающегося (бакалавра));

- устанавливает для конкретного обучающегося (бакалавра) с ограниченными возможностями здоровья содержание образования (**одинаковые дидактические единицы** — элементы содержания учебного материала, как и для обучающего (бакалавра), осваивающего основную образовательную программу высшего образования в учебной группе) и условия организации обучения, изложенного в виде утверждённой в установленном Институте порядке согласно соответствующему локальному нормативному акту рабочей программы учебной дисциплины, обязательного компонента разработанной и реализуемой им адаптированной основной профессиональной образовательной программы высшего образования - программы бакалавриата по направлению подготовки **09.03.02 Информационные системы и технологии** (направленность (профиль) «Информационные системы в экономике»), форма обучения — очно-заочная), а для инвалидов также в соответствии с индивидуальной программой реабилитации инвалида (для конкретного обучающегося (бакалавра) с ограниченными возможностями здоровья/ инвалида (*при наличии факта зачисления в образовательную организацию такого обучающегося (бакалавра) с учётом конкретной (конкретных) нозологии (нозологий)*));

- определяет в индивидуальном учебном плане конкретного обучающегося (бакалавра) с ограниченными возможностями здоровья/ инвалида (*при наличии факта зачисления такого обучающегося (бакалавра) с учётом конкретной (конкретных) нозологии (нозологий)*) **объём учебной дисциплины** с указанием количества академических часов/ ЗЕТ, выделенных на его контактную работу (групповую и (или) индивидуальную работу) с руководящими и (или) научно-педагогическими работниками, реализующими основную образовательную программу высшего образования;

- определяет в индивидуальном учебном плане конкретного обучающегося (бакалавра) с ограниченными возможностями здоровья/ инвалида (*при наличии факта зачисления в образовательную организацию такого обучающегося (бакалавра) с учётом конкретной*

(конкретных) нозологии (нозологий)) количество академических часов/ ЗЕТ по учебной дисциплине, выделенных на его самостоятельную работу (при необходимости).

9. УЧЕБНО-МЕТОДИЧЕСКОЕ И ИНФОРМАЦИОННОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

9.1 Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины

Основная литература:

1. Бишоп К.М., Распознавание образов и машинное обучение / К.М. Бишоп ; пер. с англ. и ред. Д.А. Ключин. - М.: Вильямс, 2020. - 960 с. - ISBN 978-5-4461-0877-0.
2. Гудфеллоу, И. Глубокое обучение / И. Гудфеллоу, Й. Бенжио, А. Курвилль. - М.: Вильямс, 2018. - 800 с.
3. Платонов А.В., Машинное обучение [Электронный ресурс] : учебное пособие для вузов / А.В. Платонов. - М. : Издательство Юрайт, 2022. - 85 с. - (Высшее образование) - ISBN 978-5-534-15561-7.
4. Хастие Т., Основы статистического обучения. Интеллектуальный анализ данных, логический вывод и прогнозирование / Т. Хастие, Р. Тибширани, Дж. Фридман ; пер. с англ. и ред. Д.А. Ключин. - Вильямс, 2020. - 768 с. - ISBN 978-5-907144-42-2

Дополнительная литература:

1. Мэрфи К., Вероятностное машинное обучение. Введение / К. Мэрфи ; пер. А.А. Слинкин. - ДМК-Пресс, 2022. - 990 с. - ISBN 978-5-93700-119-1.
2. Алпайдин Э., Введение в машинное обучение / Э. Алпайдин. - 4-е изд. - М.: Вильямс, 2020. - 640 с.

9.2 Используемое программное обеспечение (комплект лицензионного и свободно распространяемого программного обеспечения, в том числе отечественного производства в соответствии с п.4.3.2. ФГОС ВО 09.03.02):

1. Microsoft Windows 11 Pro или аналогичная ОС, включая дистрибутивы Linux, например Debian, Ubuntu, OpenSuse, в том числе отечественного производства, например ОС Astra Linux Common Edition (Разработчик: АО «НПО РусБИТех»), ОС «РОСА» (Разработчик: «НТИЦ ИТ РОСА»).
2. Microsoft Office 365 или аналогичный офисный пакет, например OpenOffice, LibreOffice, ONLYOFFICE, в том числе отечественного производства, например МойОфис (Разработчик: ООО «НОВЫЕ ОБЛАЧНЫЕ ТЕХНОЛОГИИ»).
3. Adobe Reader или аналогичный просмотрщик PDF, например Okular, Foxit Reader, в том числе отечественного производства, например Окуляр ГОСТ (Разработчик: ООО «Лаборатория 50»).
4. Google Chrome или аналогичный веб-браузер, например Microsoft Edge, Mozilla Firefox, в том числе отечественного производства, например Яндекс.Браузер (Разработчик: ООО «ЯНДЕКС»).
5. Microsoft Visual Studio Code или аналогичная IDE, например Sublime Text, Eclipse, в том числе отечественного производства
6. PyCharm / IntelliJ IDEA / CLion / WebStorm либо аналогичная IDE полного стека, в том числе отечественного производства
7. MySQL CE 8.0 / MySQL Workbench или аналогичные СУБД, например MS SQL, PostgreSQL, в том числе отечественного производства
8. Android Studio или аналогичная IDE для разработки мобильных приложений, в том числе отечественного производства
9. Figma или аналогичное ПО для подготовки макетов, например Penpot, Lunacy, в том числе отечественного производства

9.3 Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины (модуля) (в соответствии с п.4.3.4. ФГОС ВО 09.03.02)

1. Электронно-библиотечная система BOOK.RU [Электронный ресурс]. - Режим доступа: <https://book.ru/>
2. Habrahabr: Машинное обучение [Электронный ресурс]. - Режим доступа: https://habr.com/ru/hub/machine_learning/
3. Kaggle: Основы машинного обучения [Электронный ресурс]. - Режим доступа: <https://www.kaggle.com/learn/intro-to-machine-learning>
4. Machine Learning Mastery: Machine Learning Algorithms [Электронный ресурс]. - Режим доступа: <https://machinelearningmastery.com/start-here/>

9.4 Базы данных, информационно-справочные и поисковые системы (в соответствии с п.4.3.4. ФГОС ВО 09.03.02)

1. КонсультантПлюс: справочно-поисковая система [Электронный ресурс]. - <http://www.consultant.ru>
2. Мировая цифровая библиотека: <http://wdl.org/ru>
3. Научная электронная библиотека «Scopus»: <https://www.scopus.com>
4. Научная электронная библиотека ScienceDirect: <http://www.sciencedirect.com>
5. Научная электронная библиотека «eLIBRARY»: <https://elibrary.ru>
6. Портал «Гуманитарное образование» <http://www.humanities.edu.ru>
7. Федеральный портал «Российское образование» <http://www.edu.ru>
8. Федеральное хранилище «Единая коллекция цифровых образовательных ресурсов» <http://school-collection.edu.ru>
9. Поисковые системы Yandex, Rambler и др.
10. Электронная библиотека Российской Государственной Библиотеки (РГБ): <http://elibrary.rsl.ru>
11. Электронно-библиотечная система <http://www.sciteclibrary.ru>

10. Материально-техническое обеспечение дисциплины

Помещения для проведения всех видов работы, предусмотренных учебным планом, укомплектованы необходимым оборудованием и техническими средствами обучения.

Наименование помещений для проведения всех видов учебной деятельности, предусмотренной учебным планом, в том числе помещения для самостоятельной работы, с указанием перечня основного оборудования, учебно-наглядных пособий и используемого программного обеспечения	Адрес (местоположение) помещений для проведения всех видов учебной деятельности, предусмотренной учебным планом (в случае реализации образовательной программы в сетевой форме дополнительно указывается наименование организации, с которой заключен договор)
Специализированная многофункциональная учебная аудитория для проведения учебных занятий лекционного типа, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, в том числе, для организации практической подготовки обучающихся, с перечнем основного оборудования: <ul style="list-style-type: none">- Столы для обучающихся;- Стулья для обучающихся;	170001, Россия, город Тверь, улица Спартака, дом 26а

<ul style="list-style-type: none"> - Стол педагогического работника; - Стул педагогического работника; - Компьютеры с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду лицензиата; - Маркерная или меловая доска; - Проектор. 	
<p>Специализированная многофункциональная учебная аудитория для проведения учебных занятий семинарского типа, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, в том числе, для организации практической подготовки обучающихся, с перечнем основного оборудования:</p> <ul style="list-style-type: none"> - Столы для обучающихся; - Стулья для обучающихся; - Стол педагогического работника; - Стул педагогического работника; - Компьютеры с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду лицензиата; - Маркерная или меловая доска; - Проектор. 	<p>170001, Россия, город Тверь, улица Спартака, дом 26а</p>
<p>Помещение для самостоятельной работы обучающихся с перечнем основного оборудования:</p> <ul style="list-style-type: none"> - Столы для обучающихся; - Стулья для обучающихся; - Стол педагогического работника; - Стул педагогического работника; - Компьютеры с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду лицензиата; - Маркерная или меловая доска; - Проектор. 	<p>170001, Россия, город Тверь, улица Спартака, дом 26а</p>
<p>Помещение для практических занятий на персональных компьютерах:</p> <ul style="list-style-type: none"> - Столы для обучающихся; - Стулья для обучающихся; - Стол педагогического работника; - Стул педагогического работника; - Компьютеры с возможностью подключения к сети «Интернет» и 	<p>170001, Россия, город Тверь, улица Спартака, дом 26а</p>

<p>обеспечением доступа в электронную информационно-образовательную среду лицензиата;</p> <ul style="list-style-type: none">- Ноутбуки с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду лицензиата;- Маркерная или меловая доска;- Проектор.	
---	--

**МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ**



**Частное учреждение высшего образования
«Высшая школа предпринимательства (институт)»
(ЧУВО «ВШП»)**

**ОЦЕНОЧНЫЕ МАТЕРИАЛЫ
по дисциплине
Б1.В.13 «Основы машинного обучения»**

**Направление подготовки: 09.03.02 Информационные системы и технологии
Направленность (профиль) программы бакалавриата
«Информационные системы в экономике»**

ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОБУЧЕНИЯ ПО УЧЕБНОЙ ДИСЦИПЛИНЕ В РАМКАХ ПЛАНИРУЕМЫХ РЕЗУЛЬТАТОВ ОСВОЕНИЯ ОСНОВНОЙ ПРОФЕССИОНАЛЬНОЙ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

Процесс освоения учебной дисциплины направлен на формирование у обучающихся следующих компетенций.

В результате освоения учебной дисциплины обучающийся должен демонстрировать следующие результаты обучения: УК-1, УК-2, ПК-1, ПК-2, ПК-5, ПК-6.

Код компетенции	Наименование компетенции	Индекс и наименование индикатора содержания компетенции	Дескрипторы – основные признаки освоения (показатели достижения результата)
УК-1	Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач	УК-1.1 Способен осуществлять поиск и критический анализ информации	Знать: <ul style="list-style-type: none"> - Основные источники информации в области машинного обучения. - Методы критического анализа научных публикаций и статей. Уметь: <ul style="list-style-type: none"> - Осуществлять поиск актуальной информации по теме машинного обучения. - Анализировать и оценивать достоверность и значимость найденной информации. Владеть: <ul style="list-style-type: none"> - Навыками поиска информации в специализированных базах данных и научных журналах. - Методами критического анализа и синтеза информации.
		УК-1.2 Способен применять системный подход для решения поставленных задач	Знать: <ul style="list-style-type: none"> - Принципы системного подхода и его значение в машинном обучении. - Основные этапы разработки и внедрения систем машинного обучения. Уметь: <ul style="list-style-type: none"> - Применять системный подход при разработке моделей машинного обучения. - Интегрировать различные компоненты системы для решения комплексных задач. Владеть: <ul style="list-style-type: none"> - Методами системного анализа и проектирования в контексте машинного обучения. - Навыками разработки комплексных решений с использованием машинного обучения.
УК-2	Способен определять круг задач в рамках поставленной цели и выбирать оптимальные способы их решения, исходя из действующих правовых норм,	УК-2.1 Способен определять круг задач в рамках поставленной цели	Знать: <ul style="list-style-type: none"> - Методы декомпозиции задач в области машинного обучения. - Основные типы задач машинного обучения (регрессия, классификация, кластеризация и т.д.). Уметь: <ul style="list-style-type: none"> - Определять задачи, вытекающие из поставленной цели проекта машинного обучения.

	имеющихся ресурсов и ограничений		<ul style="list-style-type: none"> - Формулировать задачи и подзадачи для разработки моделей машинного обучения. <p>Владеть:</p> <ul style="list-style-type: none"> - Навыками структурирования и декомпозиции задач. - Методами определения ключевых задач и их приоритизации.
		<p>УК-2.2 Способен выбирать оптимальные способы решения задач, исходя из правовых норм, ресурсов и ограничений</p>	<p>Знать:</p> <ul style="list-style-type: none"> - Основные методы и алгоритмы машинного обучения. - Правовые нормы и ограничения, связанные с обработкой данных и разработкой моделей машинного обучения. <p>Уметь:</p> <ul style="list-style-type: none"> - Выбирать методы и алгоритмы для решения конкретных задач машинного обучения. - Оценивать ресурсы и ограничения, влияющие на выбор решений. <p>Владеть:</p> <ul style="list-style-type: none"> - Навыками оценки и выбора оптимальных методов и алгоритмов. - Методами обоснования выбора решений с учетом правовых норм и ресурсов.
ПК-1	Разработка и отладка программного кода	<p>ПК-1.1 Способен разрабатывать программный код</p>	<p>Знать:</p> <ul style="list-style-type: none"> - Основные языки программирования, используемые в машинном обучении (Python, R). - Структуру и синтаксис языков программирования. <p>Уметь:</p> <ul style="list-style-type: none"> - Писать программный код для реализации алгоритмов машинного обучения. - Использовать библиотеки и фреймворки для машинного обучения (scikit-learn, TensorFlow, PyTorch). <p>Владеть:</p> <ul style="list-style-type: none"> - Навыками разработки кода для решения задач машинного обучения. - Методами отладки и оптимизации программного кода.
		<p>ПК-1.2 Способен отлаживать программный код</p>	<p>Знать:</p> <ul style="list-style-type: none"> - Основные методы и инструменты отладки программного кода. - Принципы тестирования и проверки корректности программного обеспечения. <p>Уметь:</p> <ul style="list-style-type: none"> - Отлаживать код для выявления и устранения ошибок. - Использовать инструменты для автоматизированного тестирования. <p>Владеть:</p> <ul style="list-style-type: none"> - Навыками эффективной отладки кода. - Методами тестирования и обеспечения качества программного обеспечения.

ПК-2	Проверка работоспособности и рефакторинг кода программного обеспечения	ПК-2.1 Способен проверять работоспособность программного кода	<p>Знать:</p> <ul style="list-style-type: none"> - Методы и инструменты тестирования программного обеспечения. - Принципы обеспечения надежности и стабильности кода. <p>Уметь:</p> <ul style="list-style-type: none"> - Проводить тестирование и верификацию программного кода. - Оценивать работоспособность и производительность кода. <p>Владеть:</p> <ul style="list-style-type: none"> - Навыками тестирования и верификации программного обеспечения. - Методами анализа и улучшения производительности кода.
		ПК-2.2 Способен рефакторить программный код	<p>Знать:</p> <ul style="list-style-type: none"> - Принципы рефакторинга и оптимизации кода. - Стандарты и лучшие практики написания чистого и поддерживаемого кода. <p>Уметь:</p> <ul style="list-style-type: none"> - Проводить рефакторинг для улучшения структуры и читаемости кода. - Оптимизировать код для повышения его производительности. <p>Владеть:</p> <ul style="list-style-type: none"> - Навыками рефакторинга и оптимизации кода. - Методами обеспечения читаемости и поддерживаемости программного кода.
ПК-5	Анализ экономических данных и применение методов экономического моделирования для обоснования управленческих решений	ПК-5.1 Способен проводить анализ экономических данных	<p>Знать:</p> <ul style="list-style-type: none"> - Методы сбора и анализа экономических данных. - Основные экономические показатели и их интерпретацию. <p>Уметь:</p> <ul style="list-style-type: none"> - Проводить сбор и предварительную обработку экономических данных. - Анализировать экономические данные с использованием методов машинного обучения. <p>Владеть:</p> <ul style="list-style-type: none"> - Навыками анализа и интерпретации экономических данных. - Методами применения машинного обучения для анализа экономических данных.
		ПК-5.2 Способен применять методы экономического моделирования	<p>Знать:</p> <ul style="list-style-type: none"> - Основные методы экономического моделирования. - Принципы построения экономических моделей с использованием машинного обучения. <p>Уметь:</p> <ul style="list-style-type: none"> - Разрабатывать и оценивать экономические модели. - Применять методы машинного обучения для решения задач экономического моделирования. <p>Владеть:</p>

			<ul style="list-style-type: none"> - Навыками разработки и применения экономических моделей. - Методами использования машинного обучения для экономического моделирования.
ПК-6	Применение математических и статистических методов для анализа данных и принятия решений в условиях неопределенности.	ПК-6.1 Способен применять математические методы анализа данных	<p>Знать:</p> <ul style="list-style-type: none"> - Основные математические методы, используемые в анализе данных. - Принципы и методы машинного обучения для анализа данных. <p>Уметь:</p> <ul style="list-style-type: none"> - Применять математические методы для анализа данных. - Использовать алгоритмы машинного обучения для анализа данных. <p>Владеть:</p> <ul style="list-style-type: none"> - Навыками применения математических методов и алгоритмов машинного обучения для анализа данных. - Методами интерпретации результатов анализа данных.
		ПК-6.2 Способен применять статистические методы для анализа данных	<p>Знать:</p> <ul style="list-style-type: none"> - Основные статистические методы, используемые в анализе данных. - Принципы статистического анализа и интерпретации данных. <p>Уметь:</p> <ul style="list-style-type: none"> - Применять статистические методы для анализа данных. - Использовать статистические методы в машинном обучении для построения моделей. <p>Владеть:</p> <ul style="list-style-type: none"> - Навыками применения статистических методов для анализа данных. - Методами использования статистических методов в контексте машинного обучения.

КРИТЕРИИ ОЦЕНИВАНИЯ КОМПЕТЕНЦИЙ

(признак, на основании которого, проводится оценка по выбранному показателю)

<i>Показатель оценивания компетенций</i>	<i>Результат обучения</i>	<i>Критерии оценивания компетенций</i>
Высокий уровень (отлично)	Знать	Обучающийся продемонстрировал: глубокие исчерпывающие знания и понимание учебного материала; содержательные, полные, правильные и конкретные ответы на все вопросы, включая дополнительные; свободное владение основной и дополнительной литературой, рекомендованной учебной программой дисциплины.
	Уметь	Обучающийся продемонстрировал: понимание учебного материала; умение свободно решать практические задания (ситуационные задачи), которые следует выполнить или описание результата, который нужно получить и др.; логически последовательные, содержательные, полные, правильные и конкретные ответы (решения) на все поставленные задания (вопросы), включая дополнительные; свободное владение основной и дополнительной литературой, рекомендованной учебной программой дисциплины.
	Владеть	Обучающийся продемонстрировал: понимание учебного материала; умение свободно решать комплексные практические

		задания (решения задач по нестандартным ситуациям); логически последовательные, полные, правильные и конкретные ответы в ходе защиты задания, включая дополнительные уточняющие вопросы (задания); свободное владение основной и дополнительной литературой, рекомендованной учебной программой дисциплины.
Средний уровень (хорошо)	Знать	Обучающийся продемонстрировал: твердые и достаточно полные знания учебного материала; правильное понимание сущности и взаимосвязи рассматриваемых процессов и явлений; последовательные, правильные, конкретные ответы на поставленные вопросы при свободном устранении замечаний по отдельным вопросам; достаточное владение литературой, рекомендованной учебной программой дисциплины
	Уметь	Обучающийся продемонстрировал: понимание учебного материала; логически последовательные, правильные и конкретные ответы (решения) на основные задания (вопросы), включая дополнительные; устранение замечаний по отдельным элементам задания (вопроса); владение основной и дополнительной литературой, рекомендованной учебной программой дисциплины
	Владеть	Обучающийся продемонстрировал: понимание учебного материала; продемонстрировал логически последовательные, достаточно полные, правильные ответы, включая дополнительные; самостоятельно устранил замечания по отдельным элементам задания (вопроса); владение основной и дополнительной литературой, рекомендованной учебной программой дисциплины
Достаточный уровень (удовлетворительно)	Знать	Обучающийся продемонстрировал: твердые знания и понимание основного учебного материала; правильные, без грубых ошибок, ответы на поставленные вопросы при устранении неточностей и несущественных ошибок в освещении отдельных положений при наводящих вопросах преподавателя; недостаточно полное владение литературой, рекомендованной учебной программой дисциплины
	Уметь	Обучающийся продемонстрировал: понимание основного учебного материала; правильные, без грубых ошибок, ответы (решения) на основные задания (вопросы), включая дополнительные, устранение, при наводящих вопросах преподавателя, замечаний по отдельным элементам задания (вопроса); недостаточно полное владение литературой, рекомендованной учебной программой дисциплины
	Владеть	Обучающийся понимание основного учебного материала; без грубых ошибок дал ответы на поставленные вопросы при устранении неточностей и ошибок в решениях в ходе защиты задания (проекта, портфолио) при наводящих вопросах преподавателя; недостаточно полное владение литературой, рекомендованной учебной программой дисциплины

ОПИСАНИЕ ШКАЛ ОЦЕНИВАНИЯ

При проведении промежуточной аттестации в ЧУВО «ВШП» используются традиционные формы аттестации:

Форма промежуточной аттестации	Шкала оценивания
ЗАЧЕТ	«зачтено», «незачтено»
ЭКЗАМЕН	«отлично», «хорошо», «удовлетворительно», «неудовлетворительно»

КРИТЕРИИ И ПРОЦЕДУРЫ ОЦЕНИВАНИЯ КОМПЕТЕНЦИЙ НА РАЗЛИЧНЫХ ЭТАПАХ ИХ ФОРМИРОВАНИЯ

Для оценивания результатов обучения в виде **ЗНАНИЙ** используются следующие процедуры и технологии:

- тестирование.

Для оценивания результатов обучения в виде **УМЕНИЙ и ВЛАДЕНИЙ** используются следующие процедуры и технологии:

- устный или письменный ответ на вопрос.
- практические задания, включающие одну или несколько задач (вопросов) в виде краткой формулировки действий (комплекса действий), которые следует выполнить, или описать результат, который нужно получить.

Критерии оценивания результата обучения по дисциплине (модулю)

Результат обучения по дисциплине (модулю)	ШКАЛА ОЦЕНИВАНИЯ				Процедуры оценивания
	«отлично»	«хорошо»	«удовлетворительно»	«неудовлетворительно»	
<u>УК-1,</u> <u>УК-2,</u> <u>ПК-1,</u> <u>ПК-2,</u> <u>ПК-5,</u> <u>ПК-6</u> Знать:	Обучаемый продемонстрировал: глубокие исчерпывающие знания и понимание учебного материала; содержательные, полные, правильные и конкретные ответы на все вопросы, включая дополнительные; свободное владение основной и дополнительной литературой, рекомендованной учебной программой дисциплины.	Обучаемый продемонстрировал: твердые и достаточно полные знания учебного материала; правильное понимание сущности и взаимосвязи рассматриваемых процессов и явлений; последовательные, правильные, конкретные ответы на поставленные вопросы при свободном устранении замечаний по отдельным вопросам; достаточное владение литературой.	Обучаемый продемонстрировал: твердые знания и понимание основного учебного материала; правильные, без грубых ошибок, ответы на поставленные вопросы при устранении неточностей и несущественных ошибок в освещении отдельных положений при наводящих вопросах преподавателя; недостаточно полное владение литературой, рекомендованной учебной программой дисциплины.	Обучаемый продемонстрировал: неправильные ответы на основные вопросы; грубые ошибки в ответах; непонимание сущности излагаемых вопросов; неуверенные и неточные ответы на дополнительные вопросы; не владеет основной литературой, рекомендованной учебной программой дисциплины.	Тестовые задания
<u>УК-1,</u> <u>УК-2,</u> <u>ПК-1,</u> <u>ПК-2,</u> <u>ПК-5,</u> <u>ПК-6</u> Уметь:	Обучаемый продемонстрировал: понимание учебного материала, содержательные, полные, правильные и конкретные ответы на все поставленные вопросы, включая дополнительные; свободное владение	Обучаемый продемонстрировал: понимание учебного материала; логически последовательные, правильные и конкретные ответы на основные задания/вопросы, включая дополнительные; устранение	Обучаемый продемонстрировал: понимание основного учебного материала; правильные, без грубых ошибок, ответы на основные вопросы, включая дополнительные, при устранении, при наводящих вопросах преподавателя, замечаний по отдельным	Обучаемый продемонстрировал: непонимание основного учебного материала; не дал правильные ответы на основные вопросы, включая дополнительные; не устранил, при наводящих вопросах преподавателя, замечания и грубые ошибки по вопросу; не владеет основной	Вопросы Практические задания

	основной и дополнительной литературой, рекомендованной учебной программой дисциплины	замечаний по отдельным элементам задания; владение основной и дополнительной литературой, рекомендованной учебной программой дисциплины	элементам задания; недостаточное полное владение литературой, рекомендованной учебной программой дисциплины	учебной литературой, рекомендованной учебной программой дисциплины	
<u>УК-1</u> , <u>УК-2</u> , <u>ПК-1</u> , <u>ПК-2</u> , <u>ПК-5</u> , <u>ПК-6</u> <u>Владеть:</u>	Обучаемый продемонстрировал: понимание учебного материала; правильные и конкретные ответы, включая дополнительные уточняющие вопросы; свободное владение основной и дополнительной литературой, рекомендованной учебной программой дисциплины	Обучаемый продемонстрировал: понимание учебного материала; продемонстрировал логически последовательные, достаточно полные, верные ответы; самостоятельно устранил замечания по отдельным элементам; владение основной и дополнительной литературой, рекомендованной учебной программой дисциплины	Обучаемый продемонстрировал: понимание основного учебного материала; без грубых ошибок дал ответы на поставленные вопросы, в том числе при наводящих вопросах преподавателя; недостаточно полное владение литературой, рекомендованной учебной программой дисциплины	Обучаемый продемонстрировал: непонимание основного учебного материала; дал неправильные ответы на поставленные вопросы; не владеет основной учебной литературой, рекомендованной учебной программой дисциплины	Вопросы Практические задания

1. Оценочные материалы для самостоятельной работы обучающихся (студентов)

1.1 Доклад

Доклад позволит студентам углубиться в основные концепции и методы машинного обучения, а также понять их важность и применение в различных отраслях. Доклад покрывает компетенции УК-1.1, УК-1.2, УК-2.1, УК-2.2, ПК-1.1, ПК-1.2, ПК-2.1, ПК-2.2, ПК-5.1, ПК-5.2, ПК-6.1, ПК-6.2.

Примерная тематика докладов:

1. **Введение в машинное обучение:** Основные понятия и принципы машинного обучения, классификация методов машинного обучения: обучение с учителем, обучение без учителя, обучение с подкреплением. Примеры применения машинного обучения в различных отраслях.
2. **Основы статистики и вероятности для машинного обучения:** Основные понятия статистики и теории вероятностей, дискретные и непрерывные случайные величины, распределения вероятностей, проверка гипотез и доверительные интервалы. Роль статистики в машинном обучении.
3. **Предобработка данных:** Методы очистки, нормализации и трансформации данных, работа с пропущенными значениями, принципы выбора признаков и их влияние на результаты моделей. Использование библиотек для предобработки данных (pandas, sklearn).
4. **Линейные модели в машинном обучении:** Линейная регрессия и классификация, математическое описание и интерпретация линейных моделей, градиентный спуск и его вариации, применение линейных моделей к реальным данным.
5. **Деревья решений и ансамблевые методы:** Основы деревьев решений, преимущества и недостатки деревьев решений, ансамблевые методы: бэггинг, бустинг и случайные леса, применение ансамблевых методов для улучшения точности моделей.
6. **Методы снижения размерности:** Принципы и методы снижения размерности данных, главный компонентный анализ (PCA), линейный дискриминантный анализ (LDA), применение методов снижения размерности в задачах визуализации и предварительного анализа данных.
7. **Кластеризация и методы обучения без учителя:** Основные алгоритмы кластеризации: k-means, иерархическая кластеризация, DBSCAN. Методы оценки качества кластеризации, примеры применения кластеризации в различных отраслях.
8. **Нейронные сети и глубокое обучение:** Основы нейронных сетей, архитектуры и виды нейронных сетей: многослойные перцептроны, сверточные нейронные сети, рекуррентные нейронные сети. Основы обучения глубоких моделей, применение глубокого обучения в реальных задачах.
9. **Оценка моделей и перекрестная проверка:** Методы оценки качества моделей: точность, полнота, F-мера, ROC-кривая. Принципы перекрестной проверки, методы борьбы с переобучением, применение методов оценки моделей на практике.
10. **Применение машинного обучения в экономике:** Примеры использования машинного обучения для анализа экономических данных, прогнозирование временных рядов, модели для оценки кредитного риска, анализ рыночных данных и инвестиционных стратегий с помощью машинного обучения.
11. **Практическое использование библиотек и инструментов для машинного обучения:** Обзор популярных библиотек и инструментов для машинного обучения (scikit-learn, TensorFlow, Keras, PyTorch). Примеры использования библиотек для решения типовых задач машинного обучения, практические задания по разработке и оценке моделей.
12. **Этика и правовые аспекты машинного обучения:** Основные этические вопросы в области машинного обучения, влияние машинного обучения на общество и экономику,

правовые аспекты обработки данных и использования моделей машинного обучения.
Принципы ответственного использования технологий машинного обучения.

Цель написания докладов: Углубить понимание и критическое осмысление роли машинного обучения в современных информационных системах, развивая аналитические и научные навыки студентов.

Структура доклада:

1. **Введение**
 - Краткое описание темы и целей доклада.
 - Актуальность темы.
2. **Основная часть**
 - Теоретические основы темы.
 - История и эволюция (если применимо).
 - Применение в современной экономике/бизнесе.
 - Примеры и кейсы.
 - Проблемы и вызовы.
 - Перспективы и тенденции развития.
3. **Заключение**
 - Выводы по результатам исследования.
 - Значение машинного обучения для современных информационных систем.
4. **Список использованных источников**
 - Перечень использованной литературы и интернет-ресурсов.

Критерии оценивания:

1. **Структура и логика изложения (20%)**
 - Четкая структура работы (введение, основная часть, заключение).
 - Логичность и последовательность изложения материала.
2. **Содержание (40%)**
 - Полнота раскрытия темы.
 - Описание основных методов и принципов машинного обучения.
 - Анализ современных тенденций.
 - Примеры применения машинного обучения в реальных проектах.
3. **Аналитическая часть (20%)**
 - Глубина анализа роли машинного обучения в процессе разработки ПО.
 - Наличие собственных выводов и оценок.
4. **Оформление (10%)**
 - Соответствие требованиям к оформлению докладов (шрифт, отступы, заголовки и т.д.).
 - Корректное оформление ссылок и списка литературы.
5. **Язык и стиль (10%)**
 - Грамотность и точность изложения.
 - Научный стиль текста.

Требования к объему:

Объем доклада должен составлять 8-12 страниц печатного текста (шрифт Times New Roman, размер 12, интервал 1.5, поля 2 см со всех сторон).

2. Оценочные материалы для оценки текущей аттестации обучающихся (студентов)

2.1 Тестовые задания для текущего контроля успеваемости в виде ЗНАНИЙ

В тестовом задании вопросы, которые имеют закрытый характер.

Правильные ответы выделены знаком +.

1. Что такое машинное обучение? (УК-1.1)

- Метод программирования роботов.
- Процесс обучения машины выполнять задачи без явного программирования. +
- Использование машин для выполнения физических задач.
- Метод разработки аппаратного обеспечения.

2. Какой алгоритм используется для регрессии? (ПК-6.1)

- K-means.
- Линейная регрессия. +
- DBSCAN.
- Apriori.

3. Что такое переобучение в машинном обучении? (ПК-2.1)

- Модель плохо обучается на тренировочных данных.
- Модель хорошо обобщает данные на тестовых данных.
- Модель слишком хорошо обучается на тренировочных данных, но плохо обобщает на новых данных. +
- Модель использует все доступные данные для обучения.

4. Какой метод используется для уменьшения размерности данных? (ПК-6.2)

- K-means.
- Линейная регрессия.
- PCA (Анализ главных компонент). +
- Random Forest.

5. Какой алгоритм используется для кластеризации? (УК-2.1)

- Линейная регрессия.
- K-means. +
- Логистическая регрессия.
- Градиентный бустинг.

6. Что такое перекрестная проверка (cross-validation)? (УК-2.2)

- Метод проверки гипотез.
- Метод предобработки данных.
- Метод оценки модели путем разделения данных на обучающую и тестовую выборки несколько раз. +
- Метод визуализации данных.

7. Какой из следующих методов относится к ансамблевым методам? (ПК-5.2)

- Линейная регрессия.
- Деревья решений.
- Бэггинг. +
- K-means.

8. Какую библиотеку часто используют для глубокого обучения? (ПК-1.1)

- pandas.
- scikit-learn.
- TensorFlow. +
- Matplotlib.

9. В чем заключается метод нормализации данных? (ПК-6.1)

- В увеличении размера данных.
- В уменьшении размерности данных.
- В преобразовании данных в диапазон от 0 до 1 или -1 до 1. +

- В добавлении шума к данным.
- 10. Какой метод используется для классификации? (УК-1.2)**
- PCA.
 - Линейная регрессия.
 - Логистическая регрессия. +
 - K-means.
- 11. Что такое ROC-кривая? (ПК-2.1)**
- График зависимости между истинными положительными и ложными отрицательными значениями.
 - График зависимости между истинными положительными и ложными положительными значениями. +
 - График зависимости между тренировочными и тестовыми ошибками.
 - График зависимости между размером выборки и точностью модели.
- 12. Какой метод используется для прогнозирования временных рядов? (ПК-5.1)**
- K-means.
 - ARIMA. +
 - PCA.
 - Random Forest.
- 13. В чем заключается метод бэггинга? (ПК-1.2)**
- В использовании одной модели для всех данных.
 - В создании нескольких моделей на различных подвыборках данных и усреднении их результатов. +
 - В использовании регуляризации для предотвращения переобучения.
 - В преобразовании данных перед обучением.
- 14. Какие принципы этики важны в машинном обучении? (УК-2.2)**
- Прозрачность и справедливость. +
 - Скорость и точность.
 - Объем данных и их доступность.
 - Количество моделей и их сложность.
- 15. Какие задачи решает метод главных компонент (PCA)? (ПК-6.2)**
- Увеличение размерности данных.
 - Снижение размерности данных. +
 - Кластеризация данных.
 - Регрессия.
- 16. Какой метод машинного обучения используется для анализа текстов? (УК-1.1)**
- Линейная регрессия.
 - Нейронные сети. +
 - K-means.
 - PCA.
- 17. Какой показатель используется для оценки точности классификационной модели? (ПК-2.1)**
- Среднеквадратическая ошибка.
 - Точность (accuracy). +
 - Средняя абсолютная ошибка.
 - Коэффициент детерминации.
- 18. Какой из следующих методов используется для борьбы с переобучением? (ПК-2.2)**
- Увеличение размерности данных.
 - Уменьшение объема данных.
 - Регуляризация. +
 - Использование всех доступных данных для обучения.
- 19. Что такое метод K-ближайших соседей (KNN)? (УК-2.1)**
- Алгоритм кластеризации.

- Алгоритм регрессии.
- Алгоритм классификации и регрессии. +
- Метод снижения размерности.

20. Какой метод машинного обучения применяется для анализа временных рядов? (ПК-5.1)

- DBSCAN.
- ARIMA. +
- Random Forest.
- K-means.

21. Каковы основные этапы предобработки данных? (УК-1.2)

- Сбор данных, построение модели.
- Очистка данных, нормализация, трансформация. +
- Анализ данных, интерпретация.
- Визуализация данных, моделирование.

22. В чем заключается метод случайных лесов? (ПК-5.2)

- В использовании одного дерева решений.
- В создании множества деревьев решений и усреднении их результатов. +
- В применении линейной регрессии.
- В использовании K-ближайших соседей.

23. Какой метод используется для прогнозирования временных рядов? (ПК-5.1)

- K-means.
- ARIMA. +
- PCA.
- Random Forest.

24. Какой метод используется для оценки качества кластеризации? (ПК-6.2)

- Коэффициент Силуэт. +
- Среднеквадратическая ошибка.
- Коэффициент детерминации.
- Логарифмическая вероятность.

Критерии оценки результатов теста

1. "Неудовлетворительно" (0-39%)

- Студент ответил правильно на менее 40% вопросов.
- Значительные пробелы в знаниях по большинству тем.
- Неправильное понимание ключевых понятий и принципов.
- Неспособность применить теоретические знания на практике.

2. "Удовлетворительно" (40-59%)

- Студент ответил правильно на 40-59% вопросов.
- Основные понятия и принципы поняты частично, есть ошибки в ответах.
- Знания по большинству тем на базовом уровне, недостаточная глубина понимания.
- Частичная способность применять теоретические знания на практике, нужны дополнительные разъяснения.

3. "Хорошо" (60-79%)

- Студент ответил правильно на 60-79% вопросов.
- Хорошее понимание ключевых понятий и принципов, незначительные ошибки.
- Знания по всем темам на достаточном уровне, однако есть некоторые пробелы.
- Способность применять теоретические знания на практике, но требуется улучшение точности и уверенности.

4. "Отлично" (80-100%)

- Студент ответил правильно на 80-100% вопросов.
- Полное и правильное понимание всех ключевых понятий и принципов.

- Глубокие знания по всем темам, минимальные или отсутствующие ошибки.
- Высокий уровень способности применять теоретические знания на практике, демонстрация уверенности и точности в ответах.

2.2 Вопросы для текущего контроля успеваемости в виде УМЕНИЙ

1. Что такое машинное обучение и какие его основные типы? (УК-1.1)

- Правильный ответ: Машинное обучение — это область искусственного интеллекта, которая занимается разработкой алгоритмов, позволяющих компьютерам обучаться на данных и делать прогнозы или принимать решения без явного программирования. Основные типы машинного обучения включают обучение с учителем (supervised learning), обучение без учителя (unsupervised learning) и обучение с подкреплением (reinforcement learning).

2. В чем разница между обучением с учителем и обучением без учителя? (УК-1.2)

- Правильный ответ: Обучение с учителем предполагает использование размеченных данных, где каждая запись имеет входные данные и соответствующие выходные значения. Цель состоит в том, чтобы обучить модель предсказывать выходные значения для новых данных. Обучение без учителя не использует размеченные данные, и цель состоит в обнаружении скрытых структур или закономерностей в данных.

3. Что такое переобучение и как его избежать? (ПК-2.1)

- Правильный ответ: Переобучение (overfitting) происходит, когда модель слишком хорошо подстраивается под тренировочные данные, но плохо обобщает на новые данные. Методы для предотвращения переобучения включают регуляризацию, использование большего количества данных для обучения, кросс-валидацию и раннюю остановку.

4. Какой алгоритм используется для кластеризации данных? (УК-2.1)

- Правильный ответ: Для кластеризации данных часто используется алгоритм K-means, который делит данные на K групп, минимизируя внутрикластерные расстояния.

5. Что такое линейная регрессия и для чего она используется? (ПК-6.1)

- Правильный ответ: Линейная регрессия — это метод статистического анализа, который используется для моделирования линейных отношений между зависимой переменной и одной или несколькими независимыми переменными. Она применяется для прогнозирования и анализа влияния факторов.

6. Какие метрики используются для оценки классификационных моделей? (ПК-2.1)

- Правильный ответ: Основные метрики для оценки классификационных моделей включают точность (accuracy), полноту (recall), точность (precision) и F-меру. Точность показывает долю правильных предсказаний, полнота отражает способность модели находить все положительные примеры, точность измеряет долю истинных положительных результатов среди всех положительных предсказаний, а F-мера является гармоническим средним точности и полноты.

7. Что такое главные компоненты анализа (PCA) и для чего он используется? (ПК-6.2)

- Правильный ответ: PCA (Principal Component Analysis) — это метод снижения размерности данных, который преобразует исходные переменные в новый набор переменных, главных компонент, упорядоченных по убыванию важности. PCA используется для уменьшения числа переменных в данных, сохранения важной информации и улучшения визуализации.

8. В чем заключается метод случайных лесов (Random Forest)? (ПК-5.2)

- Правильный ответ: Случайный лес (Random Forest) — это ансамблевый метод машинного обучения, который использует множество деревьев решений, обученных на различных подвыборках данных. Итоговое предсказание определяется голосованием или усреднением предсказаний всех деревьев. Этот метод улучшает точность моделей и уменьшает риск переобучения.

9. Какие библиотеки используются для машинного обучения на Python? (ПК-1.1)

- Правильный ответ: Основные библиотеки для машинного обучения на Python включают scikit-learn, TensorFlow, Keras и PyTorch. Scikit-learn используется для базовых алгоритмов и предобработки данных, TensorFlow и Keras — для разработки и обучения нейронных сетей, а PyTorch — для гибкого и динамичного построения нейронных сетей.
- 10. Какой метод используется для оценки модели на основе ROC-кривой? (ПК-2.1)**
- Правильный ответ: ROC-кривая (Receiver Operating Characteristic curve) используется для оценки качества бинарных классификаторов. Она показывает зависимость между долей истинно положительных результатов (True Positive Rate) и долей ложно положительных результатов (False Positive Rate) при различных порогах принятия решения. Чем ближе кривая к верхнему левому углу, тем лучше модель.
- 11. Что такое ARIMA и для чего она используется? (ПК-5.1)**
- Правильный ответ: ARIMA (Autoregressive Integrated Moving Average) — это модель прогнозирования временных рядов, которая сочетает авторегрессию, интегрирование и скользящее среднее. ARIMA используется для анализа и прогнозирования данных, которые зависят от времени, таких как экономические показатели, спрос на продукцию и температурные данные.
- 12. В чем разница между регуляризацией L1 и L2? (ПК-2.2)**
- Правильный ответ: Регуляризация L1 (Lasso) добавляет к функции потерь абсолютное значение весов, что приводит к разреженным моделям и отбору признаков. Регуляризация L2 (Ridge) добавляет к функции потерь квадрат весов, что предотвращает переобучение, уменьшая величину весов, но не приводя к разреженности.
- 13. Что такое метод К-ближайших соседей (KNN) и для чего он используется? (УК-2.1)**
- Правильный ответ: Метод К-ближайших соседей (KNN) — это алгоритм классификации и регрессии, который использует близость к k ближайшим точкам в тренировочном наборе для предсказания метки или значения для новой точки. Он используется для задач классификации, где метка класса определяется голосованием ближайших соседей, и для задач регрессии, где значение определяется усреднением значений ближайших соседей.
- 14. В чем заключается метод бэггинга (bagging) и его преимущества? (ПК-1.2)**
- Правильный ответ: Метод бэггинга (bagging, bootstrap aggregating) заключается в создании нескольких моделей на различных подвыборках данных и усреднении их результатов. Преимущества включают снижение вариации и уменьшение переобучения, что улучшает точность и стабильность моделей.
- 15. Что такое регуляризация и как она помогает предотвратить переобучение? (ПК-2.2)**
- Правильный ответ: Регуляризация — это метод, который добавляет штраф за сложность модели к функции потерь. Это помогает предотвратить переобучение, уменьшая величину весов параметров модели и тем самым делая модель более устойчивой к изменениям в данных. Основные типы регуляризации включают L1 (Lasso) и L2 (Ridge).
- 16. Какой метод используется для прогнозирования временных рядов и как он работает? (ПК-5.1)**
- Правильный ответ: Модель ARIMA (Autoregressive Integrated Moving Average) используется для прогнозирования временных рядов. Она сочетает авторегрессию (AR), интегрирование (I) и скользящее среднее (MA) для моделирования временных зависимостей в данных.
- 17. Какие основные метрики используются для оценки качества регрессионных моделей? (ПК-6.1)**
- Правильный ответ: Основные метрики для оценки качества регрессионных моделей включают среднеквадратическую ошибку (MSE), среднюю абсолютную ошибку (MAE) и коэффициент детерминации (R^2). MSE измеряет средний квадрат разностей между предсказанными и фактическими значениями, MAE измеряет среднюю абсолютную разность, а R^2 показывает долю дисперсии, объясненную моделью.
- 18. Какой алгоритм машинного обучения используется для анализа текстов и как он работает? (УК-1.1)**

- Правильный ответ: Нейронные сети, особенно рекуррентные нейронные сети (RNN) и трансформеры, используются для анализа текстов. Они работают, учитывая последовательности слов и контекст, что позволяет моделям понимать и генерировать текст на естественном языке.

19. В чем заключается метод K-means и как он работает? (ПК-6.2)

- Правильный ответ: Метод K-means — это алгоритм кластеризации, который делит данные на K кластеров. Он работает, итеративно назначая точки кластерам, минимизируя внутрикластерные расстояния, и обновляя центроиды кластеров до тех пор, пока они не стабилизируются.

20. Какова роль предобработки данных в машинном обучении и какие основные этапы она включает? (УК-1.2)

- Правильный ответ: Предобработка данных играет ключевую роль в машинном обучении, так как улучшает качество данных и, соответственно, модели. Основные этапы включают очистку данных, нормализацию, трансформацию и отбор признаков.

21. Какие библиотеки часто используются для машинного обучения на Python и какие их основные функции? (ПК-1.1)

- Правильный ответ: Основные библиотеки для машинного обучения на Python включают scikit-learn (алгоритмы машинного обучения, предобработка данных), TensorFlow и Keras (разработка и обучение нейронных сетей), PyTorch (гибкое и динамичное построение нейронных сетей) и pandas (обработка и анализ данных).

22. Какой метод используется для оценки качества кластеризации? (ПК-6.2)

- Правильный ответ: Коэффициент Силуэт используется для оценки качества кластеризации. Он измеряет, насколько хорошо каждая точка данных лежит в своем кластере по сравнению с другими кластерами, что помогает определить оптимальное число кластеров.

23. В чем заключаются этические вопросы в области машинного обучения и как они могут быть решены? (УК-2.2)

- Правильный ответ: Этические вопросы в области машинного обучения включают предвзятость алгоритмов, защиту данных и приватность, а также прозрачность моделей. Решения включают обеспечение справедливости и прозрачности моделей, соблюдение законодательства о защите данных и внедрение этических принципов при разработке и использовании моделей.

24. Что такое перекрестная проверка и почему она важна? (ПК-2.1)

- Правильный ответ: Перекрестная проверка (cross-validation) — это метод оценки модели путем разделения данных на несколько частей, обучения модели на одной части и тестирования на другой. Это помогает оценить стабильность и обобщающую способность модели, предотвращая переобучение и обеспечивая более надежные результаты.

Критерии оценки ответов на вопросы

- **"Отлично" (5 баллов)**
 - Критерии:
 - Полное и точное объяснение вопроса.
 - Ответ включает все ключевые аспекты и детали.
 - Примеры, если требуются, приведены и правильно объяснены.
 - Ответ демонстрирует глубокое понимание темы.
- **"Хорошо" (4 балла)**
 - Критерии:
 - Корректное объяснение вопроса.
 - Ответ охватывает основные аспекты, но может отсутствовать незначительная деталь или пример.
 - Демонстрируется хорошее, но не полное понимание темы.

- **"Удовлетворительно" (3 балла)**
 - **Критерии:**
 - Общее представление о вопросе.
 - Ответ включает основные аспекты, но содержит неточности или пропуски.
 - Примеры, если требуются, могут отсутствовать или быть неверно объяснены.
 - Демонстрируется базовое понимание темы.
- **"Неудовлетворительно" (2 балла)**
 - **Критерии:**
 - Некорректное или неполное объяснение вопроса.
 - Отсутствие ключевых аспектов и деталей.
 - Примеры, если требуются, отсутствуют или приведены неверные.
 - Ответ демонстрирует недостаточное понимание темы.

2.3 Задачи на соответствие понятий для текущего контроля успеваемости в виде ВЛАДЕНИЙ

Правильные ответы расположены в таблицах друг напротив друга, во время тестирования предполагается что порядок данных в рамках каждого столбца будет случайным.

Задача 1: Соотнесите методы машинного обучения с их описаниями (УК-1.1, ПК-1.1)

Метод	Описание
А - Линейная регрессия	1 - Метод предсказания непрерывных значений на основе линейных зависимостей.
В - K-means	2 - Алгоритм кластеризации, который делит данные на K групп, минимизируя внутрикластерные расстояния.
С - Логистическая регрессия	3 - Метод классификации, который предсказывает вероятность принадлежности к одному из классов.
D - PCA	4 - Метод снижения размерности данных путем преобразования переменных в главные компоненты.

Правильный ответ: А-1, В-2, С-3, D-4

Задача 2: Соотнесите типы данных с их примерами (УК-2.1, ПК-6.1)

Тип данных	Пример
А - Непрерывные данные	1 - Температура в градусах Цельсия.
В - Категориальные данные	2 - Цвета (красный, синий, зеленый).
С - Временные ряды	3 - Стоимость акций по дням.
D - Дискретные данные	4 - Количество учеников в классе.

Правильный ответ: А-1, В-2, С-3, D-4

Задача 3: Соотнесите методы предобработки данных с их функциями (ПК-1.2, ПК-6.2)

Метод	Функция
-------	---------

А - Нормализация	1 - Преобразование данных в диапазон от 0 до 1.
В - Заполнение пропущенных значений	2 - Замена пропущенных данных средним значением.
С - Кодирование категориальных признаков	3 - Преобразование категориальных данных в числовые.
Д - Удаление выбросов	4 - Исключение экстремальных значений из набора данных.

Правильный ответ: А-1, В-2, С-3, Д-4

Задача 4: Соотнесите типы кластеризации с их характеристиками (УК-2.2, ПК-5.2)

Тип кластеризации	Характеристика
А - К-means	1 - Деление данных на К групп с минимизацией внутрикластерных расстояний.
В - Иерархическая кластеризация	2 - Построение дерева кластеров с использованием метода агломерации или деления.
С - DBSCAN	3 - Обнаружение кластеров произвольной формы и исключение выбросов.
Д - К-медианы	4 - Деление данных на К групп с минимизацией медианного расстояния.

Правильный ответ: А-1, В-2, С-3, Д-4

Задача 5: Соотнесите метрики с их применением в машинном обучении (ПК-2.1, ПК-6.1)

Метрика	Применение
А - Точность (ассигасу)	1 - Доля правильных предсказаний в классификационной модели.
В - Среднеквадратическая ошибка (MSE)	2 - Средний квадрат разностей между предсказанными и фактическими значениями в регрессионной модели.
С - Коэффициент детерминации (R^2)	3 - Доля объясненной дисперсии в регрессионной модели.
Д - F-мера	4 - Гармоническое среднее точности и полноты в классификационной модели.

Правильный ответ: А-1, В-2, С-3, Д-4

Задача 6: Соотнесите виды регуляризации с их характеристиками (ПК-2.2, УК-2.2)

Вид регуляризации	Характеристика
A - L1 (Lasso)	1 - Добавление абсолютного значения весов к функции потерь, приводящее к разреженным моделям.
B - L2 (Ridge)	2 - Добавление квадрата весов к функции потерь, уменьшая величину весов и предотвращая переобучение.
C - Elastic Net	3 - Комбинация L1 и L2 регуляризаций.
D - Dropout	4 - Исключение случайных нейронов во время обучения нейронной сети для предотвращения переобучения.

Правильный ответ: A-1, B-2, C-3, D-4

Критерии оценки выполнения задач на соответствие понятий

- **Правильность соответствий:**
 - **Отлично (5):** Все соответствия выполнены правильно.
 - **Хорошо (4):** 1 ошибка в соответствиях.
 - **Удовлетворительно (3):** 2 ошибки в соответствиях.
 - **Неудовлетворительно (2):** 3 и более ошибок в соответствиях.

3. Оценочные материалы для проведения промежуточной аттестации обучающихся (студентов)

3.1 Вопросы для проведения промежуточной аттестации в форме ЭКЗАМЕНА

1. Что такое машинное обучение и какие его основные типы? (УК-1.1)

- **Правильный ответ:** Машинное обучение — это область искусственного интеллекта, которая занимается разработкой алгоритмов, позволяющих компьютерам обучаться на данных и делать прогнозы или принимать решения без явного программирования. Основные типы машинного обучения включают обучение с учителем (supervised learning), обучение без учителя (unsupervised learning) и обучение с подкреплением (reinforcement learning). Обучение с учителем использует размеченные данные для тренировки моделей, обучение без учителя работает с неразмеченными данными для выявления скрытых структур, а обучение с подкреплением оптимизирует действия агента через награды и наказания.

2. В чем разница между обучением с учителем и обучением без учителя? (УК-1.2)

- **Правильный ответ:** Обучение с учителем предполагает использование размеченных данных, где каждая запись имеет входные данные и соответствующие выходные значения. Цель состоит в том, чтобы обучить модель предсказывать выходные значения для новых данных. В обучении без учителя используются неразмеченные данные, и цель заключается в выявлении скрытых структур или закономерностей в данных, таких как кластеры или ассоциации.

3. Что такое переобучение и как его избежать? (ПК-2.1)

- **Правильный ответ:** Переобучение (overfitting) происходит, когда модель слишком хорошо подстраивается под тренировочные данные, но плохо обобщает на новые данные. Это может произойти из-за чрезмерной сложности модели. Методы для предотвращения

переобучения включают регуляризацию (L1 и L2), использование большего количества данных для обучения, кросс-валидацию, снижение сложности модели и раннюю остановку обучения.

4. Какие методы и алгоритмы используются для кластеризации данных? (УК-2.1)

- Правильный ответ: Основные алгоритмы кластеризации включают K-means, иерархическую кластеризацию и DBSCAN. K-means делит данные на K кластеров с минимизацией внутрикластерных расстояний. Иерархическая кластеризация строит дерево кластеров, объединяя или разделяя группы данных. DBSCAN обнаруживает кластеры произвольной формы и исключает выбросы, основываясь на плотности данных.

5. Что такое линейная регрессия и как она применяется в машинном обучении? (ПК-6.1)

- Правильный ответ: Линейная регрессия — это метод статистического анализа, который используется для моделирования линейных отношений между зависимой переменной и одной или несколькими независимыми переменными. В машинном обучении линейная регрессия применяется для прогнозирования непрерывных значений, таких как цены на жилье, заработная плата или спрос на продукцию. Модель строится путем минимизации суммы квадратов разностей между предсказанными и фактическими значениями.

6. Какие метрики используются для оценки классификационных моделей и как их интерпретировать? (ПК-2.1)

- Правильный ответ: Основные метрики для оценки классификационных моделей включают точность (accuracy), полноту (recall), точность (precision) и F-меру. Точность показывает долю правильных предсказаний среди всех предсказаний. Полнота измеряет способность модели находить все положительные примеры. Точность отражает долю истинных положительных результатов среди всех положительных предсказаний. F-мера является гармоническим средним точности и полноты, показывая баланс между ними.

7. Что такое анализ главных компонент (PCA) и как он используется? (ПК-6.2)

- Правильный ответ: Анализ главных компонент (PCA) — это метод снижения размерности данных, который преобразует исходные переменные в новый набор переменных, главных компонент, упорядоченных по убыванию важности. PCA используется для уменьшения числа переменных в данных, сохраняя важную информацию и улучшая визуализацию. Это помогает ускорить обучение моделей и снизить вероятность переобучения.

8. Как работает метод случайных лесов (Random Forest) и какие его преимущества? (ПК-5.2)

- Правильный ответ: Метод случайных лесов (Random Forest) использует множество деревьев решений, обученных на различных подвыборках данных, для создания ансамблевой модели. Итоговое предсказание определяется голосованием или усреднением предсказаний всех деревьев. Преимущества случайных лесов включают высокую точность, устойчивость к переобучению и способность обрабатывать большое количество признаков и наблюдений.

9. Какие библиотеки используются для машинного обучения на Python и какие их основные функции? (ПК-1.1)

- Правильный ответ: Основные библиотеки для машинного обучения на Python включают scikit-learn, TensorFlow, Keras и PyTorch. Scikit-learn предоставляет инструменты для базовых алгоритмов и предобработки данных. TensorFlow и Keras используются для разработки и обучения нейронных сетей. PyTorch предлагает гибкие и динамичные возможности для построения нейронных сетей. Эти библиотеки помогают решать широкий спектр задач машинного обучения, включая классификацию, регрессию и кластеризацию.

10. Что такое ROC-кривая и как она используется для оценки моделей? (ПК-2.1)

- Правильный ответ: ROC-кривая (Receiver Operating Characteristic curve) — это график, который используется для оценки качества бинарных классификаторов. Она показывает

зависимость между долей истинно положительных результатов (True Positive Rate) и долей ложно положительных результатов (False Positive Rate) при различных порогах принятия решения. Чем ближе кривая к верхнему левому углу, тем лучше модель. ROC-кривая помогает выбрать оптимальный порог и сравнить различные модели.

11. Что такое ARIMA и как она используется для прогнозирования временных рядов? (ПК-5.1)

- Правильный ответ: ARIMA (Autoregressive Integrated Moving Average) — это модель прогнозирования временных рядов, которая сочетает авторегрессию, интегрирование и скользящее среднее. ARIMA используется для анализа и прогнозирования данных, которые зависят от времени, таких как экономические показатели, спрос на продукцию и температурные данные. Модель учитывает как тренды, так и сезонные колебания в данных.

12. В чем разница между регуляризацией L1 и L2 и как они помогают предотвратить переобучение? (ПК-2.2)

- Правильный ответ: Регуляризация L1 (Lasso) добавляет к функции потерь абсолютное значение весов, что приводит к разреженным моделям и отбору признаков. Регуляризация L2 (Ridge) добавляет к функции потерь квадрат весов, что предотвращает переобучение, уменьшая величину весов. Оба метода снижают сложность модели, делая ее более устойчивой к изменениям в данных.

13. Как работает метод K-ближайших соседей (KNN) и в каких задачах он применяется? (УК-2.1)

- Правильный ответ: Метод K-ближайших соседей (KNN) — это алгоритм классификации и регрессии, который использует близость к k ближайшим точкам в тренировочном наборе для предсказания метки или значения для новой точки. Для классификации метка класса определяется голосованием ближайших соседей, для регрессии — усреднением значений ближайших соседей. KNN используется для задач классификации, где требуется разделить данные на категории, и для регрессии, где необходимо предсказать непрерывное значение.

14. Что такое регуляризация и как она помогает предотвратить переобучение? (ПК-2.2)

- Правильный ответ: Регуляризация — это метод, который добавляет штраф за сложность модели к функции потерь. Это помогает предотвратить переобучение, уменьшая величину весов параметров модели и тем самым делая модель более устойчивой к изменениям в данных. Основные типы регуляризации включают L1 (Lasso) и L2 (Ridge). L1 регуляризация приводит к разреженности модели, в то время как L2 регуляризация уменьшает величину всех весов, не делая их равными нулю.

15. Как используются нейронные сети для анализа текстов? (УК-1.1)

- Правильный ответ: Нейронные сети, особенно рекуррентные нейронные сети (RNN) и трансформеры, используются для анализа текстов. RNN учитывают последовательности слов и контекст, что позволяет моделям понимать и генерировать текст на естественном языке. Трансформеры, такие как BERT и GPT, обеспечивают более глубокое понимание контекста благодаря механизмам внимания, которые фокусируются на важных частях текста.

16. Что такое метод главных компонент (PCA) и для чего он используется? (ПК-6.2)

- Правильный ответ: Метод главных компонент (PCA) — это метод снижения размерности данных, который преобразует исходные переменные в новый набор переменных, главных компонент, упорядоченных по убыванию важности. PCA используется для уменьшения числа переменных, сохраняя важную информацию, и улучшения визуализации. Это помогает ускорить обучение моделей и снизить вероятность переобучения.

17. Как используется метод случайных лесов (Random Forest) и какие его преимущества? (ПК-5.2)

- Правильный ответ: Метод случайных лесов (Random Forest) использует множество деревьев решений, обученных на различных подвыборках данных, для создания ансамблевой модели. Итоговое предсказание определяется голосованием или усреднением предсказаний всех деревьев. Преимущества случайных лесов включают высокую точность, устойчивость к переобучению и способность обрабатывать большое количество признаков и наблюдений.

18. Какие библиотеки используются для машинного обучения на Python и какие их основные функции? (ПК-1.2)

- Правильный ответ: Основные библиотеки для машинного обучения на Python включают scikit-learn, TensorFlow, Keras и PyTorch. Scikit-learn предоставляет инструменты для базовых алгоритмов и предобработки данных. TensorFlow и Keras используются для разработки и обучения нейронных сетей. PyTorch предлагает гибкие и динамичные возможности для построения нейронных сетей. Эти библиотеки помогают решать широкий спектр задач машинного обучения, включая классификацию, регрессию и кластеризацию.

19. Что такое ROC-кривая и как она используется для оценки моделей? (ПК-2.1)

- Правильный ответ: ROC-кривая (Receiver Operating Characteristic curve) — это график, который используется для оценки качества бинарных классификаторов. Она показывает зависимость между долей истинно положительных результатов (True Positive Rate) и долей ложно положительных результатов (False Positive Rate) при различных порогах принятия решения. Чем ближе кривая к верхнему левому углу, тем лучше модель. ROC-кривая помогает выбрать оптимальный порог и сравнить различные модели.

20. Как используется модель ARIMA для прогнозирования временных рядов? (ПК-5.1)

- Правильный ответ: Модель ARIMA (Autoregressive Integrated Moving Average) используется для анализа и прогнозирования временных рядов. Она сочетает авторегрессию (AR), интегрирование (I) и скользящее среднее (MA) для моделирования временных зависимостей в данных. ARIMA учитывает как тренды, так и сезонные колебания в данных, что делает её эффективной для прогнозирования экономических показателей, спроса на продукцию и других временных данных.

21. Каковы основные этапы предобработки данных и зачем они необходимы? (УК-1.2)

- Правильный ответ: Основные этапы предобработки данных включают очистку данных, нормализацию, трансформацию и отбор признаков. Очистка данных устраняет ошибки и пропуски, нормализация приводит данные к единому масштабу, трансформация изменяет данные для улучшения их свойств, а отбор признаков удаляет нерелевантные или избыточные признаки. Эти этапы необходимы для улучшения качества данных и, соответственно, точности моделей.

22. Что такое коэффициент Силуэт и как он используется для оценки качества кластеризации? (ПК-6.2)

- Правильный ответ: Коэффициент Силуэт используется для оценки качества кластеризации. Он измеряет, насколько хорошо каждая точка данных лежит в своем кластере по сравнению с другими кластерами. Значение коэффициента Силуэт варьируется от -1 до 1, где высокие значения указывают на хорошо определенные кластеры. Этот метод помогает определить оптимальное число кластеров и оценить эффективность алгоритма кластеризации.

23. Какие этические вопросы возникают в области машинного обучения и как они могут быть решены? (УК-2.2)

- Правильный ответ: Этические вопросы в области машинного обучения включают предвзятость алгоритмов, защиту данных и приватность, а также прозрачность моделей. Решения включают обеспечение справедливости и прозрачности моделей, соблюдение законодательства о защите данных и внедрение этических принципов при разработке и использовании моделей. Это помогает избежать дискриминации и других негативных последствий использования машинного обучения.

24. Что такое перекрестная проверка (cross-validation) и почему она важна? (ПК-2.1)

- Правильный ответ: Перекрестная проверка (cross-validation) — это метод оценки модели путем разделения данных на несколько частей, обучения модели на одной части и тестирования на другой. Этот процесс повторяется несколько раз с разными частями данных. Перекрестная проверка важна, так как она помогает оценить стабильность и обобщающую способность модели, предотвращая переобучение и обеспечивая более надежные результаты.

Критерии оценки ответов на экзамене

- **"Отлично" (5 баллов)**
 - Критерии:
 - Полное и точное объяснение вопроса.
 - Ответ включает все ключевые аспекты и детали.
 - Примеры, если требуются, приведены и правильно объяснены.
 - Ответ демонстрирует глубокое понимание темы.
- **"Хорошо" (4 балла)**
 - Критерии:
 - Корректное объяснение вопроса.
 - Ответ охватывает основные аспекты, но может отсутствовать незначительная деталь или пример.
 - Демонстрируется хорошее, но не полное понимание темы.
- **"Удовлетворительно" (3 балла)**
 - Критерии:
 - Общее представление о вопросе.
 - Ответ включает основные аспекты, но содержит неточности или пропуски.
 - Примеры, если требуются, могут отсутствовать или быть неверно объяснены.
 - Демонстрируется базовое понимание темы.
- **"Неудовлетворительно" (2 балла)**
 - Критерии:
 - Некорректное или неполное объяснение вопроса.
 - Отсутствие ключевых аспектов и деталей.
 - Примеры, если требуются, отсутствуют или приведены неверные.
 - Ответ демонстрирует недостаточное понимание темы.