

**МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ**

**Федеральное государственное бюджетное образовательное учреждение
высшего образования**

**«Луганский государственный университет имени Владимира Даля»
(ФГБОУ ВО «ЛГУ им. В. Даля»)**

**Северодонецкий технологический институт
Кафедра информационных технологий, приборостроения и электротехники**

УТВЕРЖДАЮ:
Врио. директора СТИ (филиал)
ФГБОУ ВО «ЛГУ им. В. Даля»
Ю.В. Бородач
(подпись)
« 26 » 2024 года



РАБОЧАЯ ПРОГРАММА УЧЕБНОЙ ДИСЦИПЛИНЫ

«Машинное обучение»

По направлению подготовки: 09.04.01 Информатика и вычислительная техника

Магистерская программа «Цифровые технологии в экономике»

Лист согласования РПУД

Рабочая программа учебной дисциплины «Машинное обучение» по направлению подготовки 09.04.01 Информатика и вычислительная техника (магистерская программа «Цифровые технологии в экономике») – 22 с.

Рабочая программа учебной дисциплины «Машинное обучение» разработана в соответствии с Федеральным государственным образовательным стандартом высшего образования – магистратура по направлению подготовки 09.04.01 Информатика и вычислительная техника, утвержденным приказом Министерства образования и науки Российской Федерации от 19 сентября 2017 г. № 918 (с изменениями и дополнениями в соответствии с приказами Министерства образования и науки Российской Федерации № 1456 от 26.11.2020 г., № 82 от 08.02.2021 г.).

СОСТАВИТЕЛЬ:

ст. преподаватель Е.В. Кузнецова

Рабочая программа дисциплины утверждена на заседании кафедры информационных технологий, приборостроения и электротехники « 05 » 09 2024 г., протокол № 1.

Заведующий кафедрой ИТПЭ  В.Г. Чебан

Переутверждена: « ____ » _____ 202 ____ г., протокол № ____.

Рекомендована на заседании учебно-методической комиссии Северодонецкого технологического института (филиал) федерального государственного бюджетного образовательного учреждения высшего образования «Луганский государственный университет имени Владимира Даля» « 16 » 09 2024 г., протокол № 1.

Председатель учебно-методической комиссии
СТИ (филиал) ФГБОУ ВО «ЛГУ им. В. Даля»

 Ю.В. Бородач

© Кузнецова Е. В., 2024 год

© СТИ (филиал) ФГБОУ ВО «ЛГУ им. В. Даля», 2024 г.

1. Цели и задачи освоения дисциплины

Цель изучения дисциплины – формирование компетенций в области использования технологий машинного обучения для анализа данных.

Задачи:

- изучение стадий технологии машинного обучения;
- овладение навыками работы с различными методами построения алгоритмов, способных обучаться;
- получение практических навыков реализации методов машинного обучения;
- изучение теоретических и организационно-методических вопросов построения и функционирования систем, основанных на знаниях.

2. Место дисциплины в структуре ОПОП ВО

Дисциплина «Машинное обучение» входит в обязательную часть дисциплин учебного плана.

Необходимыми условиями для освоения дисциплины являются:

знания: об основных математических подходах и технологиях построения моделей машинного обучения, методах предобработки и анализа данных, способах оценки качества моделей;

умения: использовать современными библиотеками и фреймворками для реализации моделей машинного обучения; выполнять обработку и анализ данных с применением моделей и алгоритмов машинного обучения; применять метрики оценки качества моделей;

навыки: работы с текстовыми данными, изображениями и временными рядами; решения задач регрессии, классификации и кластеризации на основе базовых методик машинного обучения.

Содержание дисциплины является логическим продолжением содержания дисциплины «Информационные технологии в научно-исследовательской и практической деятельности».

Служит основой для изучения следующих дисциплин: «Автоматизированные системы обработки информации и управления», «Современные методы обработки и анализа данных».

3. Компетенции обучающегося, формируемые в результате освоения дисциплины

Код и наименование компетенции	Индикаторы достижений компетенции (по реализуемой дисциплине)	Перечень планируемых результатов
ОПК-1. Способен самостоятельно приобретать, развивать и применять математические, естественнонаучные, социально-экономические и профессиональные знания для решения нестандартных задач, в том числе в новой или незнакомой среде и в междисциплинарном контексте	ОПК-1.1. Знает математические, естественнонаучные и социально-экономические методы для использования в профессиональной деятельности ОПК-1.2. Умеет решать нестандартные профессиональные задачи, в том числе в новой или незнакомой среде и в междисциплинарном контексте, с применением математических, естественнонаучных, социально-экономических и профессиональных знаний ОПК-1.3. Владеет методами теоретического и	Знать: математические, естественнонаучные и социально-экономические методы для использования в профессиональной деятельности Уметь: решать нестандартные профессиональные задачи с применением математических, естественнонаучных, социально-экономических знаний Владеть: навыками теоретического и экспериментального исследования объектов профессиональной деятельности, в том числе в новой или незнакомой среде и в

	экспериментального исследования объектов профессиональной деятельности, в том числе в новой или незнакомой среде и в междисциплинарном контексте	междисциплинарном контексте
ОПК-2. Способен разрабатывать оригинальные алгоритмы и программные средства, в том числе с использованием современных интеллектуальных технологий, для решения профессиональных задач	ОПК-2.1. Знает современные информационно-коммуникационные и интеллектуальные технологии, инструментальные среды, программно-технические платформы для решения профессиональных задач ОПК-2.2. Умеет обосновывать выбор современных информационно-коммуникационных и интеллектуальных технологий, разрабатывать оригинальные программные средства для решения профессиональных задач ОПК-2.3. Владеет методами разработки оригинальных программных средств, в том числе с использованием современных информационно-коммуникационных и интеллектуальных технологий, для решения профессиональных задач	Знать: современные информационно-коммуникационные и интеллектуальные технологии, инструментальные среды, программно-технические платформы для решения профессиональных задач Уметь: выбирать современные информационно-коммуникационные и интеллектуальные технологии, разрабатывать оригинальные программные средства для решения профессиональных задач Владеть: навыками разработки оригинальных программных средств, в том числе с использованием современных информационно-коммуникационных и интеллектуальных технологий, для решения профессиональных задач

4. Структура и содержание дисциплины

4.1. Объем учебной дисциплины и виды учебной работы

Вид учебной работы	Объем часов (зач. ед.)	
	Очная форма	Заочная форма
Общая учебная нагрузка (всего)	144 (4 зач. ед)	144 (4 зач. ед)
Обязательная аудиторная учебная нагрузка (всего) в том числе:	56	24
Лекции	28	12
Семинарские занятия	-	-
Практические занятия	-	-
Лабораторные работы	28	12
Курсовая работа (курсовой проект)	-	-
Другие формы и методы организации образовательного процесса (индивидуальные задания)	25	36
Самостоятельная работа студента (всего)	63	84
Форма аттестации	экзамен	экзамен

4.2. Содержание разделов дисциплины

Тема 1. Основные понятия машинного обучения.

Общие сведения и терминология дисциплины. Машинное обучение как область науки о данных. Проблемы и задачи машинного обучения. Пространство признаков. Общий алгоритм машинного обучения. Сферы применения машинного обучения.

Тема 2. Основы обучения с учителем и оценки его качества.

Типы машинного обучения. Основные понятия регрессии и классификации.

Тема 3. Задачи регрессии.

Линейная регрессия. Проблема мультиколлинеарности признаков. Задачи нелинейной регрессии. Применение линейной регрессии для моделирования кривых второго порядка. Задачи регрессии повышенной сложности.

Тема 4. Задачи классификации.

Линейно разделимые множества. Линейно неразделимые множества и метод комитетов. Разновидности комитетных конструкций.

Тема 5. Прогнозирование временных рядов.

Анализ временных рядов. Оценки стационарности ряда. Классические методы прогнозирования временных рядов. Метрики качества прогнозирования.

Тема 6. Градиентный спуск.

Пакетный метод градиентного спуска. Стохастический градиентный спуск. Алгоритм mini-batch метода. Адаптивный градиентный спуск.

Тема 7. Переобучение модели и методы борьбы с ним.

Суть проблемы переобучения. Причины переобучения модели. Методы борьбы с переобучением: отложенная выборка, кросс-валидация, регуляризация.

Тема 8. Бинарная классификация.

Основные понятия и задачи бинарной классификации. Линейный классификатор. Логистическая регрессия. Метрики качества классификации.

Тема 9. Решающие деревья для обучения с учителем.

Регрессия решающим деревом. Классификация решающим деревом. Обучение решающих деревьев. Решающие деревья и категориальные признаки. Композиция алгоритмов.

Тема 10. Случайный лес.

Алгоритмы обучения с учителем методом «Случайный лес». Bagging. Boosting. Градиентный бустинг. Стохастический градиентный бустинг.

Тема 11. Обучение без учителя.

Основные понятия и области применения. Задачи кластеризации. Задача визуализации многомерных данных на плоскости. Задача поиска аномалий.

Тема 12. Понижение размерности.

Цель понижения размерности. Алгоритмы отбора признаков. Методы обертывания. Методы фильтров. Метод вложения. Понижение размерности способом проекции признаков. Метод главных компонент PCA. Метод SNE. Метод t-SNE.

Тема 13. Кластеризация.

Задача, алгоритм и этапы кластеризации. Меры расстояний между объектами. Метод ближайших соседей (kNN). Метод k средних (k-means). Плотностной алгоритм пространственной кластеризации с присутствием шума. Иерархические методы кластеризации.

4.3. Лекции

№ п/п	Название темы	Объём часов	
		Очная форма	Заочная форма
1.	Основные понятия машинного обучения	2	1
2.	Основы обучения с учителем и оценки его качества	2	
3.	Задачи регрессии	2	1
4.	Задачи классификации	2	
5.	Прогнозирование временных рядов	2	1
6.	Градиентный спуск	2	
7.	Переобучение модели и методы борьбы с ним	2	1
8.	Бинарная классификация	2	
9.	Решающие деревья для обучения с учителем	2	1
10.	Случайный лес	2	
11.	Обучение без учителя	2	1
12.	Понижение размерности	4	
13.	Кластеризация	2	1
Итого:		28	8

4.4. Практические (семинарские) занятия

Практические занятия по дисциплине «Машинное обучение» не предусмотрены учебным планом.

4.5. Лабораторные работы

№ п/п	Название темы	Объём часов	
		Очная форма	Заочная форма
1.	Предобработка данных. Отбор признаков	4	4
2.	Введение в обработку естественного языка	4	
3.	Методы оптимизации в глубоком обучении	4	
4.	Анализ и предсказание временных рядов	4	4
5.	Функции ошибок в машинном обучении	4	
6.	Метрики качества задач классификации	4	
7.	Алгоритмы кластеризации	4	
Итого:		28	8

4.6. Самостоятельная работа студентов

№ п/п	Название темы	Вид СРС	Объем часов	
			Очная форма	Заочная форма
1.	Классификация методов машинного обучения	Изучение лекционного материала. Подготовка к лабораторным работам.	4	6
2.	Математическое программирование и машинное обучение	Изучение лекционного материала. Подготовка к лабораторным работам.	5	7
3.	Задачи регрессии	Изучение лекционного материала. Подготовка к лабораторным работам.	4	6
4.	Задачи классификации	Изучение лекционного материала. Подготовка к лабораторным работам.	5	7
5.	Классические методы классификации	Изучение лекционного материала. Подготовка к лабораторным работам.	5	7
6.	Обучение с учителем	Изучение лекционного материала. Подготовка к лабораторным работам.	5	6
7.	Переобучение модели	Изучение лекционного материала. Подготовка к лабораторным работам.	5	7
8.	Классификация в пространстве бинарных переменных	Изучение лекционного материала. Подготовка к лабораторным работам.	5	6
9.	Обучение без учителя	Изучение лекционного материала. Подготовка к лабораторным работам.	5	7
10.	Кластерный анализ	Изучение лекционного материала. Подготовка к лабораторным работам.	5	6
11.	Методы совместной классификации	Изучение лекционного материала. Подготовка к лабораторным работам.	5	7
12.	Методы обучения с подкреплением	Изучение лекционного материала. Подготовка к лабораторным работам.	5	6
13.	Обучения искусственных нейронных сетей	Изучение лекционного материала. Подготовка к лабораторным работам.	5	7
Итого:			63	92

4.7. Курсовые работы/проекты

Согласно учебному плану, курсовой проект (работа) по данной дисциплине не предусмотрен.

5. Образовательные технологии

Преподавание дисциплины ведётся с применением следующих видов образовательных технологий:

- традиционные объяснительно-иллюстративные технологии, которые обеспечивают доступность учебного материала для большинства студентов, системность, отработанность организационных форм и привычных методов, относительно малые затраты времени;
- технологии проблемного обучения, направленные на развитие познавательной активности, творческой самостоятельности студентов и предполагающие последовательное и целенаправленное выдвижение перед студентом познавательных задач, разрешение которых позволяет студентам активно усваивать знания (используются поисковые методы, постановка познавательных задач);
- технологии развивающего обучения, позволяющие ориентировать учебный процесс на потенциальные возможности студентов, их реализацию и развитие;
- технологии концентрированного обучения, суть которых состоит в создании максимально близкой к естественным психологическим особенностям человеческого восприятия структуры учебного процесса, и которые дают возможность глубокого и системного изучения содержания учебных дисциплин за счёт объединения занятий в тематические блоки;
- технологии модульного обучения, дающие возможность обеспечения гибкости процесса обучения, адаптации его к индивидуальным потребностям и особенностям обучающихся (применяются, как правило, при самостоятельном обучении студентов по индивидуальному учебному плану);
- технологии дифференцированного обучения, обеспечивающие возможность создания оптимальных условий для развития интересов и способностей студентов, в том числе и студентов с особыми образовательными потребностями, что позволяет реализовать в культурно-образовательном пространстве университета идею создания равных возможностей для получения образования;
- технологии активного (контекстного) обучения, с помощью которых осуществляется моделирование предметного, проблемного и социального содержания будущей профессиональной деятельности студентов (используются активные и интерактивные методы обучения) и т.д.

Максимальная эффективность педагогического процесса достигается путём конструирования оптимального комплекса педагогических технологий и (или) их элементов на личностно-ориентированной, деятельностной, диалогической основе и использования необходимых современных средств обучения.

6. Учебно-методическое и информационное обеспечение дисциплины

а) Основная литература:

1. Лимановская, О.В. Основы машинного обучения : учебное пособие / О.В. Лимановская, Т.И. Алферьева ; Мин-во науки и высш. образования РФ. – Екатеринбург : Изд-во Урал. ун-та, 2020. – 88 с
2. Чернавин, П. Ф. Практический курс классического машинного обучения с использованием моделей математического программирования : учебно-методическое пособие / П. Ф. Чернавин, Н. П. Чернавин, Ф. П. Чернавин ; М-во науки и высшего образования РФ. – Екатеринбург : Изд-во Урал. ун-та, 2023. – 124 с.
3. Воронина, В.В. Теория и практика машинного обучения : учебное пособие / В. В. Воронина, А. В. Михеев, Н. Г. Ярушкина, К. В. Святков. – Ульяновск : УлГТУ, 2017. – 290 с.
4. Гладилин П.Е., Боченина К.О., Технологии машинного обучения – СПб: Университет ИТМО, 2020. – 75 с.
5. Мясников, В.В. Распознавание образов и машинное обучение. Основные подходы: учебное пособие / В.В. Мясников. – Самара: Издательство Самарского университета, 2023. – 196 с.

б) Дополнительная литература:

1. Николенко С., Кадури А., Архангельская Е. Глубокое обучение. Погружение в мир нейронных сетей. – Изд-во «Питер». – 2018. – 476 с.
2. К.В. Воронцов. Машинное обучение (курс лекций) – 2009. <http://www.machinelearning.ru/>, <http://www.ccas.ru/voron/teaching.html>.
3. Афанасьев В.Н., Юзбашев М.М. Анализ временных рядов и прогнозирование – М.: Финансы и статистика, 2001. – 228 с.

в) Интернет-ресурсы:

1. Министерство образования и науки Российской Федерации – <http://минобрнауки.рф>
2. Министерства природных ресурсов и экологии Российской Федерации – <http://www.mnr.gov.ru>
3. Федеральная служба по надзору в сфере образования и науки – <http://obrnadzor.gov.ru>
4. Министерство образования и науки Луганской Народной Республики – <https://minobr.su>
5. Министерство природных ресурсов и экологической безопасности ЛНР – <https://www.mprlnr.su>
6. Народный совет Луганской Народной Республики – <https://nslnr.su>
7. Портал Федеральных государственных образовательных стандартов высшего образования – <http://fgosvo.ru>
8. Федеральный портал «Российское образование» – <http://www.edu.ru>
9. Информационная система «Единое окно доступа к образовательным ресурсам» – <http://window.edu.ru>
10. Федеральный центр информационно-образовательных ресурсов – <http://fcior.edu.ru>

Электронные библиотечные системы и ресурсы:

1. Электронно-библиотечная система «Консультант студента» – <http://www.studentlibrary.ru/cgi-bin/mb4x>
2. Электронно-библиотечная система «StudMed.ru» – <https://www.studmed.ru>
3. Научная электронная библиотека eLIBRARY.RU» – <http://elibrary.ru>
4. ЭБС Издательства «ЛАНЬ» – <https://e.lanbook.com>

Информационный ресурс библиотеки образовательной организации

1. Научная библиотека имени А. Н. Коняева – <http://biblio.dahluniver.ru>

7. Материально-техническое обеспечение дисциплины

Освоение дисциплины «Машинное обучение» предполагает использование академических аудиторий, соответствующих действующим санитарным и противопожарным правилам и нормам.

Лекционные занятия: комплект электронных презентаций/слайдов; аудитория, оснащенная презентационной техникой (проектор, экран, компьютер/ноутбук).

Лабораторные работы: учебный компьютерный класс, имеющий рабочие места студентов, оснащённые компьютерами с доступом в Интернет, презентационная техника (проектор, экран, компьютер/ноутбук), бесплатное программное обеспечение.

Прочее: рабочее место преподавателя, оснащённое компьютером с доступом в Интернет. В качестве материально-технического обеспечения дисциплины могут быть использованы мультимедийные средства; наборы слайдов, демонстрационные приборы, при необходимости – средства мониторинга и т.д.

Программное обеспечение:

Функциональное назначение	Бесплатное программное обеспечение	Ссылки
Офисный пакет	Libre Office 6.3.1	https://www.libreoffice.org/ https://ru.wikipedia.org/wiki/LibreOffice
Операционная система	UBUNTU 19.04	https://ubuntu.com/ https://ru.wikipedia.org/wiki/Ubuntu
Браузер	Firefox Mozilla	http://www.mozilla.org/ru/firefox/fx
Браузер	Opera	http://www.opera.com
Почтовый клиент	Mozilla Thunderbird	http://www.mozilla.org/ru/thunderbird
Файл-менеджер	Far Manager	http://www.farmanager.com/download.php
Архиватор	7Zip	http://www.7-zip.org/
Графический редактор	GIMP (GNU Image Manipulation Program)	http://www.gimp.org/ http://gimp.ru/viewpage.php?page_id=8 http://ru.wikipedia.org/wiki/GIMP
Редактор PDF	PDFCreator	http://www.pdfforge.org/pdfcreator
Аудиоплеер	VLC	http://www.videolan.org/vlc/

**8. Фонд оценочных средств для проведения текущего контроля
и промежуточной аттестации по дисциплине**

**Паспорт
оценочных средств по учебной дисциплине
«Машинное обучение»**

Перечень компетенций, формируемых в результате освоения учебной дисциплины

№ п/п	Код компетенции	Формулировка контролируемой компетенции	Индикаторы достижений компетенции (по дисциплине)	Темы учебной дисциплины	Этапы формирован ия (семестр изучения)
1	ОПК-1	Способен самостоятельно приобретать, развивать и применять математические, естественнонаучные, социально-экономические и профессиональные знания для решения нестандартных задач, в том числе в новой или незнакомой среде и в междисциплинарном контексте	ОПК-1.1. Знает математические, естественнонаучные и социально-экономические методы для использования в профессиональной деятельности ОПК-1.2. Умеет решать нестандартные профессиональные задачи, в том числе в новой или незнакомой среде и в междисциплинарном контексте, с применением математических, естественнонаучных, социально-экономических и профессиональных знаний ОПК-1.3. Владеет методами теоретического и экспериментального исследования объектов профессиональной деятельности, в том числе в новой или незнакомой среде и в междисциплинарном контексте	Тема 1. Основные понятия машинного обучения	2
				Тема 2. Основы обучения с учителем и оценки его качества	2
				Тема 3. Задачи регрессии	2
				Тема 4. Задачи классификации	2
				Тема 5. Прогнозирование временных рядов	2
				Тема 6. Градиентный спуск	2
				Тема 7. Переобучение модели и методы борьбы с ним	2
				Тема 8. Бинарная классификация	2
				Тема 9. Решающие деревья для обучения с учителем	2
				Тема 10. Случайный лес	2
				Тема 11. Обучение без учителя	2
				Тема 12. Понижение размерности	4
				Тема 13. Кластеризация	2

2	ОПК-2	Способен разрабатывать оригинальные алгоритмы и программные средства, в том числе с использованием современных интеллектуальных технологий, для решения профессиональных задач	ОПК-2.1. Знает современные информационно-коммуникационные и интеллектуальные технологии, инструментальные среды, программно-технические платформы для решения профессиональных задач ОПК-2.2. Умеет обосновывать выбор современных информационно-коммуникационных и интеллектуальных технологий, разрабатывать оригинальные программные средства для решения профессиональных задач ОПК-2.3. Владеет методами разработки оригинальных программных средств, в том числе с использованием современных информационно-коммуникационных и интеллектуальных технологий, для решения профессиональных задач	Тема 1. Основные понятия машинного обучения	2
				Тема 2. Основы обучения с учителем и оценки его качества	2
				Тема 3. Задачи регрессии	2
				Тема 4. Задачи классификации	2
				Тема 5. Прогнозирование временных рядов	2
				Тема 6. Градиентный спуск	2
				Тема 7. Переобучение модели и методы борьбы с ним	2
				Тема 8. Бинарная классификация	2
				Тема 9. Решающие деревья для обучения с учителем	2
				Тема 10. Случайный лес	2
				Тема 11. Обучение без учителя	2
				Тема 12. Понижение размерности	4
				Тема 13. Кластеризация	2

Показатели и критерии оценивания компетенций, описание шкал оценивания

№ п/п	Код компетенции	Индикаторы достижений компетенции	Планируемые результаты обучения по дисциплине	Контролируемые темы учебной дисциплины	Наименование оценочного средства
1	ОПК-1	ОПК-1.1. Знает математические, естественнонаучные и социально-экономические методы для использования в профессиональной деятельности ОПК-1.2. Умеет решать нестандартные профессиональные задачи, в том числе в новой или незнакомой среде и в междисциплинарном контексте, с применением математических, естественнонаучных, социально-экономических и профессиональных знаний ОПК-1.3. Владеет методами теоретического и экспериментального исследования объектов профессиональной деятельности, в том числе в новой или незнакомой среде и в междисциплинарном контексте	Знать: математические, естественнонаучные и социально-экономические методы для использования в профессиональной деятельности Уметь: решать нестандартные профессиональные задачи с применением математических, естественнонаучных, социально-экономических знаний Владеть: навыками теоретического и экспериментального исследования объектов профессиональной деятельности, в том числе в новой или незнакомой среде и в междисциплинарном контексте	Тема 1. Тема 2. Тема 3. Тема 4. Тема 5. Тема 6. Тема 7. Тема 8. Тема 9. Тема 10. Тема 11. Тема 12. Тема 13.	Вопросы для контроля усвоения теоретического материала, тестовые задания, выполнение задания на лабораторных работах

2	ОПК-2	ОПК-2.1. Знает современные информационно-коммуникационные и интеллектуальные технологии, инструментальные среды, программно-технические платформы для решения профессиональных задач ОПК-2.2. Умеет обосновывать выбор современных информационно-коммуникационных и интеллектуальных технологий, разрабатывать оригинальные программные средства для решения профессиональных задач ОПК-2.3. Владеет методами разработки оригинальных программных средств, в том числе с использованием современных информационно-коммуникационных и интеллектуальных технологий, для решения профессиональных задач	Знать: современные информационно-коммуникационные и интеллектуальные технологии, инструментальные среды, программно-технические платформы для решения профессиональных задач Уметь: выбирать современные информационно-коммуникационные и интеллектуальные технологии, разрабатывать оригинальные программные средства для решения профессиональных задач Владеть: навыками разработки оригинальных программных средств, в том числе с использованием современных информационно-коммуникационных и интеллектуальных технологий, для решения профессиональных задач	Тема 1. Тема 2. Тема 3. Тема 4. Тема 5. Тема 6. Тема 7. Тема 8. Тема 9. Тема 10. Тема 11. Тема 12. Тема 13.	Вопросы для контроля усвоения теоретического материала, тестовые задания, выполнение задания на лабораторных работах
---	-------	---	--	--	---

8.1. Тестовые задания

(низкий уровень)

1. К какому методу познания относится машинное обучение?

- а) дедукция;
- б) абстрагирование;
- в) индукция;
- г) анализ;
- д) синтез.

2. Выберите верные утверждения

- а) одна из задач машинного обучения – научиться делать прогнозы для признаков

- б) объекты описываются с помощью признаков
 - в) одна из задач машинного обучения – научиться делать прогнозы для объектов
 - д) признаки описываются с помощью объектов
3. Какие из приведенных ниже задач относятся к задачам обучения с учителем?
- а) классификация;
 - б) кластеризация;
 - в) фильтрация;
 - г) регрессия;
 - д) сокращение размерности.
4. Какие из этих задач являются задачами классификации?
- а) прогноз температуры на следующий день
 - б) разделение книг, хранящихся в электронной библиотеке, на научные и художественные
 - в) поиск групп похожих пользователей интернет-магазина
 - г) прогноз оценки студента по пятибалльной шкале на экзамене по машинному обучению в следующей сессии
5. Модель называют переобученной, если:
- а) в ней большое количество наблюдений;
 - б) в ней большое количество признаков;
 - в) средняя ошибка алгоритма на наблюдениях контрольной выборки значительно выше, чем на обучающей;
 - г) размер тестовой выборки значительно выше обучающей.
6. Какие признаки называют мультиколлинеарными?
- а) имеющие одинаковое значение;
 - б) линейно зависимые;
 - в) сильно коррелирующие друг с другом;
 - г) имеющие низкую дисперсию.
7. Зависимость между входными и выходными признаками имеет полиномиальный характер. Как данную задачу свести к задаче линейной регрессии?
- а) попытаться аппроксимировать полином касательными;
 - б) создать новые признаки на основе существующих;
 - в) попытаться аппроксимировать полином отрезками, соединяющими точки полинома.
7. Какие множества являются линейно не разделимыми?
- а) выпуклые;
 - б) невыпуклые;
 - в) не разделяемые гиперплоскостью;
 - г) не разделяемые любой гиперповерхностью.
8. Какой из комитетов является универсальным?
- а) единогласия;
 - б) большинства;
 - в) старшинства;
 - г) единодушия.
9. Возможно или нет в методе комитетов использовать нелинейные функции?
- а) да;
 - б) нет.
10. В методе опорных векторов разделяющая гиперплоскость:
- а) опирается на одно из множеств;
 - б) находится на одинаковом расстоянии от ближайших точек обоих множеств;
 - в) находится на одинаковом расстоянии от всех точек обоих множеств.
11. Какие из перечисленных слов не имеют никакого отношения к названиям научных школ машинного обучения?
- а) символизм;

- б) абстракционизм;
- в) коннекционизм;
- г) эволюционизм;
- д) экзистенциализм;
- е) байесианизм;
- ж) аналогизм;
- з) композиционизм.

12. Какие из перечисленных методов являются ансамблевыми?

- а) логистическая регрессия;
- б) метод комитетов;
- в) дерево решений;
- г) наивный байесовский;
- д) нейросети;
- е) ближайших соседей;
- ж) случайный лес;
- з) потенциальных функций.

13. В каких моделях регрессии совмещено обучение модели и отбор признаков?

- а) с критерием MAE;
- б) с критерием MSE;
- в) гребневая;
- г) лассо

14. Какие из методов можно использовать для решения задач регрессии?

- а) ближайших соседей;
- б) случайный лес;
- в) нейросети.

15. Сеть какого типа лучше использовать для прогнозирования временных рядов?

- а) сверточную
- б) ART MAP
- в) импульсную
- г) MLP
- д) рекуррентную
- е) когнитрон

Критерии и шкала оценивания по оценочному средству «Тестовые задания»

Шкала оценивания	Критерий оценивания
5 (отлично)	85 – 100% правильных ответов
4 (хорошо)	71 – 85% правильных ответов
3 (удовлетворительно)	61 – 70% правильных ответов
2 (неудовлетворительно)	60% правильных ответов и ниже

8.2. Вопросы для контроля усвоения теоретического материала (средний уровень)

1. Что является основной целью науки о данных?
2. Приведите примеры задач, которые можно решать с помощью науки о данных.
3. Дайте определение классификации.
4. Дайте определение кластеризации.
5. Поясните разницу между классификацией и кластеризацией.
6. Определите область применения кластеризации.
7. Запишите выражение для алгоритма линейного регрессора.
8. В чем отличие линейной и логистической регрессий?
9. В чем отличие линейной от нелинейной регрессии?

10. В чем отличие линейной регрессии от полиномиальной?
11. Что позволяет делать LASSO?
12. В чем заключаются особенности Ridge регрессии?
13. Что рассчитывается при обучении модели?
14. В чем смысл коэффициента детерминации?
15. Когда следует применять метрику MAE для оценки качества предсказания временного ряда?
16. Каким образом можно оценить количество регрессионных компонент в модели SARIMA?
17. Как вычисляется изменение весов на каждой итерации градиентного спуска?
18. Что является условием остановки расчета при градиентном спуске?
19. Как считается ошибка на каждой итерации пакетного градиентного спуска?
20. Сколько объектов из выборки используется в методе стохастического градиентного спуска?
21. Почему стохастический градиентный спуск работает быстрее, чем классический градиентный спуск?
22. Какие существуют методы обхода локальных минимумов?
23. Чем должен определяться размер mini-batch для реализации градиентного спуска в конкретном случае?
24. Какие факторы влияют на переобучение?
25. Какие есть способы оценки переобучения?
26. Какие есть способы борьбы с переобучением?
27. В чем смысл процесса регуляризации?
28. Приведите формулу L1-регуляризатора.
29. Почему регуляризатор L1 (Lasso) обнуляет веса при линейно зависимых признаках?
30. Приведите формулу L2-регуляризатора.
31. Опишите задачу бинарной классификации.
32. Какая стандартная функция используется в работе линейного классификатора?
33. Что оценивает логистический регрессор?
34. Что является ответом в методе случайного леса?
35. Как обучается каждый последующий алгоритм в бустинге?
36. Что используется для обучения каждой последующей модели в алгоритме градиентного бустинга?
37. Какую функцию потерь нужно применять в задаче регрессии, чтобы модель больше штрафовала за выбросы в данных?
38. Как функция кросс-энтропии связана с дивергенцией Кульбака-Лейблера?
39. Есть ли ответы на выборке в методе обучения без учителя?
40. По какому принципу объединяются объекты выборки в кластеры?
41. Зачем нужна визуализация данных?
42. По какому критерию выбирается первая компонента, на которую проецируются объекты из многомерного пространства?
43. Какая метрика используется для оценки расстояния между объектом и компонентой в методе главных компонент?
44. Одинаковые ли метрики используются для оценки близости объектов в многомерном пространстве и в маломерном в методе SNE?
45. Одинаковые ли метрики используются для оценки близости объектов в многомерном пространстве и в маломерном в методе t-SNE?
46. На мелких или на глубоких деревьях лучше строить случайный лес с точки зрения качества алгоритма классификации?
47. Нужно ли знать количество кластеров в методе k-means?
48. Задается ли количество кластеров в методе DBSCAN?

49. Каким образом следует выбирать порог разделения для иерархической кластеризации?

50. Как можно улучшить инициализацию центров кластеров в алгоритме k-means с целью нахождения наиболее оптимальных начальных приближений?

Лектор или преподаватель, ведущий практические занятия по дисциплине производит устный опрос по пройденным теоретическим материалам и выставляет оценку в журнале с текущей успеваемостью.

Критерии и шкала оценивания по оценочному средству
«Вопросы для контроля усвоения теоретического материала»

Шкала оценивания	Критерий оценивания
5 (отлично)	Обучающийся глубоко и в полном объеме владеет программным материалом. Грамотно, исчерпывающе и логично его излагает в устной или письменной форме. При этом знает рекомендованную литературу, проявляет творческий подход в ответах на вопросы и правильно обосновывает принятые решения, хорошо владеет умениями и навыками при выполнении практических задач.
4 (хорошо)	Обучающийся знает программный материал, грамотно и по сути излагает его в устной или письменной форме, допуская незначительные неточности в утверждениях, трактовках, определениях и категориях или незначительное количество ошибок. При этом владеет необходимыми умениями и навыками при выполнении практических задач.
3 (удовлетворительно)	Обучающийся знает только основной программный материал, допускает неточности, недостаточно четкие формулировки, непоследовательность в ответах, излагаемых в устной или письменной форме. При этом недостаточно владеет умениями и навыками при выполнении практических задач. Допускает до 30% ошибок в излагаемых ответах.
2 (неудовлетворительно)	Обучающийся не знает значительной части программного материала. При этом допускает принципиальные ошибки в доказательствах, в трактовке понятий и категорий, проявляет низкую культуру знаний, не владеет основными умениями и навыками при выполнении практических задач. Обучающийся отказывается от ответов на дополнительные вопросы.

8.3 Практическое (прикладное) задание (высокий уровень)

Задания, выполняемые на лабораторных работах:

Задание 1. Выполнить предобработку данных, необходимых для качественной настройки моделей машинного обучения.

Задание 2. Реализация базовых методов обработки естественного языка, включая предобработку текста, формирование «мешка слов» («bag-of-words»), выделение стоп-слов и наиболее важных слов в документе, создание тематических моделей.

Задание 3. Реализация современных алгоритмов оптимизации и модификаций алгоритма градиентного спуска для обучения глубоких нейронных сетей.

Задание 4. Анализ временных рядов, проверка стационарности ряда и применение классических методов прогнозирования временных рядов.

Задание 5. Реализовать модель логистической регрессии с функциями потерь.

Задание 6. Рассчитать метрики качества в задаче машинного обучения с учителем и выполнить тонкую настройку алгоритмов бинарной классификации.

Задание 7. Реализация классических алгоритмов кластеризации k-means и иерархической кластеризации.

Критерии и шкала оценивания по оценочному средству «Практическое задание»

Шкала оценивания	Характеристика знания предмета и ответов
5 (отлично)	Обучающийся полностью и правильно выполнил задание. Показал отличные знания, умения и владения навыками, применения их при решении задач в рамках усвоенного учебного материала.
4 (хорошо)	Обучающийся выполнил задание с небольшими неточностями. Показал хорошие знания, умения и владения навыками, применения их при решении задач в рамках усвоенного учебного материала.
3 (удовлетворительно)	Обучающийся выполнил задание с существенными неточностями. Показал удовлетворительные знания, умения и владения навыками, применения их при решении задач.
2 (неудовлетворительно)	Обучающийся выполнил задание неправильно. При выполнении обучающийся продемонстрировал недостаточный уровень знаний, умений и владения ими при решении задач в рамках усвоенного учебного материала.

8.4 Оценочные средства для промежуточной аттестации (экзамен)

Вопросы к экзамену:

1. Машинное обучение как область науки о данных.
2. Общий алгоритм машинного обучения.
3. Пространство признаков.
4. Типы машинного обучения.
5. Линейная регрессия.
6. Мультиколлинеарность признаков.
7. Классические методы прогнозирования временных рядов.
8. Метрики качества прогнозирования.
9. Градиентный спуск.
10. Стохастический градиентный спуск.
11. Переобучение модели.
12. Методы борьбы с переобучением.
13. Бинарная классификация.
14. Логистическая регрессия.
15. Регрессия решающим деревом.
16. Классификация решающим деревом.
17. Кластерный анализ.
18. Методы понижения размерности.
19. Задача, алгоритм и этапы кластеризации.
20. Метод ближайших соседей.
21. Метод k средних.
22. Иерархические методы кластеризации.

Критерии и шкала оценивания к промежуточной аттестации «экзамен»

Шкала оценивания	Критерий оценивания
5 (отлично)	Обучающийся глубоко и в полном объёме владеет программным материалом. Грамотно, исчерпывающе и логично его излагает в устной или письменной форме. При этом знает рекомендованную литературу, проявляет творческий подход в ответах на вопросы и правильно обосновывает принятые решения, хорошо владеет умениями и навыками при выполнении практических задач.
4 (хорошо)	Обучающийся знает программный материал, грамотно и по сути излагает его в устной или письменной форме, допуская незначительные неточности в утверждениях, трактовках, определениях и категориях или незначительное количество ошибок. При этом владеет необходимыми умениями и навыками при выполнении практических задач.
3 (удовлетворительно)	Обучающийся знает только основной программный материал, допускает неточности, недостаточно чёткие формулировки, непоследовательность в ответах, излагаемых в устной или письменной форме. При этом недостаточно владеет умениями и навыками при выполнении практических задач. Допускает до 30% ошибок в излагаемых ответах.
2 (неудовлетворительно)	Обучающийся не знает значительной части программного материала. При этом допускает принципиальные ошибки в доказательствах, в трактовке понятий и категорий, проявляет низкую культуру знаний, не владеет основными умениями и навыками при выполнении практических задач. Обучающийся отказывается от ответов на дополнительные вопросы.

9. Особенности организации обучения для лиц с ограниченными возможностями здоровья и инвалидов

При необходимости рабочая программа учебной дисциплины может быть адаптирована для обеспечения образовательного процесса инвалидов и лиц с ограниченными возможностями здоровья, в том числе с применением электронного обучения и дистанционных образовательных технологий.

Для этого требуется заявление студента (его законного представителя) и заключение психолого-медико-педагогической комиссии (ПМПК). В случае необходимости обучающимся из числа лиц с ограниченными возможностями здоровья (по заявлению обучающегося), а для инвалидов также в соответствии с индивидуальной программой реабилитации инвалида могут предлагаться следующие варианты восприятия учебной информации с учетом их индивидуальных психофизических особенностей:

- создание текстовой версии любого нетекстового контента для его возможного преобразования в альтернативные формы, удобные для различных пользователей;
- создание контента, который можно представить в различных видах без потери данных или структуры, предусмотреть возможность масштабирования текста и изображений без потери качества, предусмотреть доступность управления контентом с клавиатуры;
- создание возможностей для обучающихся воспринимать одну и ту же информацию из разных источников, например, так, чтобы лица с нарушениями слуха получали информацию визуально, с нарушениями зрения – аудиально;
- применение программных средств, обеспечивающих возможность освоения навыков и умений, формируемых дисциплиной (модулем), за счёт альтернативных способов, в том числе виртуальных лабораторий и симуляционных технологий;
- применение электронного обучения, дистанционных образовательных технологий для передачи информации, организации различных форм интерактивной контактной работы обучающегося с преподавателем, в том числе вебинаров, которые могут быть использованы для проведения виртуальных лекций с возможностью взаимодействия всех участников дистанционного обучения, проведения семинаров, выступления с докладами и защиты выполненных работ, проведения тренингов, организации коллективной работы;
- применение электронного обучения, дистанционных образовательных технологий для организации форм текущего и промежуточного контроля;
- увеличение продолжительности сдачи обучающимся инвалидом или лицом с ограниченными возможностями здоровья форм промежуточной аттестации по отношению к установленной продолжительности их сдачи:
 - продолжительность сдачи зачёта или экзамена, проводимого в письменной форме, – не более чем на 90 минут;
 - продолжительность подготовки обучающегося к ответу на зачёте или экзамене, проводимом в устной форме, – не более чем на 20 минут;
 - продолжительность выступления обучающегося при защите курсовой работы – не более чем на 15 минут.

Лист изменений и дополнений

№ п/п	Виды дополнений и изменений с указанием страниц	Дата и номер протокола заседания кафедры (кафедр), на котором были рассмотрены и одобрены изменения и дополнения	Подпись (с расшифровкой) заведующего кафедрой (заведующих кафедрами)
1.			
2.			
3.			
4.			