



Министерство здравоохранения Российской Федерации

**федеральное государственное бюджетное образовательное учреждение
высшего образования
"Северо-Западный государственный медицинский университет имени И.И. Мечникова"
Министерства здравоохранения Российской Федерации**

(ФГБОУ ВО СЗГМУ им. И.И. Мечникова Минздрава России)

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

«Основы машинного обучения и Data mining»

Специальность: 30.05.03 Медицинская кибернетика

Направленность: Цифровые технологии медицины и здравоохранения

Рабочая программа дисциплины «Основы машинного обучения и Data mining» составлена на основании Федерального государственного образовательного стандарта высшего образования – специалитет по специальности 30.05.03 Медицинская кибернетика, утвержденного приказом Министерства науки и высшего образования Российской Федерации от 13 августа 2020 года № 1006 «Об утверждении федерального государственного образовательного стандарта высшего образования - специалитет по специальности 30.05.03 Медицинская кибернетика»

Составители рабочей программы дисциплины:

Рассказова Альбина Николаевна, кандидат технических наук, доцент, доцент кафедры медицинской информатики и физики; Шматко Алексей Дмитриевич, доктор экономических наук, профессор, заведующий кафедрой медицинской информатики и физики

Рецензент:

Бригаднов Игорь Альбертович – профессор кафедры Информационных систем и вычислительной техники Федерального государственного бюджетного образовательного учреждения высшего образования «Санкт-Петербургский горный университет императрицы Екатерины II», докт. физ.-мат. наук.

Рабочая программа дисциплины обсуждена на заседании кафедр(ы) медицинской информатики и физики

11 ноября 2024 г., Протокол № 11

Заведующий кафедрой  / Шматко А.Д./
(подпись) (Ф.И.О.)

Рассмотрено Методическим советом и рекомендовано для утверждения на Ученом совете 22 ноября 2024 г.

Председатель  /Артюшкин С.А./

Дата обновления:

СОДЕРЖАНИЕ

1. Цель освоения дисциплины.....	4
2. Место дисциплины в структуре образовательной программы	4
3. Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения образовательной программы.....	4
4. Объем дисциплины и виды учебной работы	4
5. Содержание дисциплины, структурированное по разделам (темам) с указанием количества академических часов и видов учебных занятий.....	5
6. Методические указания для обучающихся по освоению дисциплины	7
7. Оценочные материалы	8
8. Перечень учебной литературы и ресурсов сети «Интернет», необходимых для освоения дисциплины	8
9. Перечень информационных технологий, используемых для освоения дисциплины, включая перечень программного обеспечения, профессиональных баз данных и информационных справочных систем.....	8
10. Материально-техническое обеспечение дисциплины	10
Приложение А.....	12

1. Цель освоения дисциплины

Целью освоения дисциплины «Основы машинного обучения и Data mining» является углубление уровня освоения компетенций обучающегося, необходимых при анализе, прогнозировании и моделировании медицинских данных, для разработки персонализированных планов лечения, оптимизации управленческих и клинических процессов в медицинских учреждениях, а также для разработки медицинских приложений.

2. Место дисциплины в структуре образовательной программы

Дисциплина «Основы машинного обучения и Data mining» относится к обязательной части Блока 1 «Дисциплины (модули)» основной профессиональной образовательной программы по специальности 30.05.03 Медицинская кибернетика (уровень образования специалитет), направленность: Цифровые технологии медицины и здравоохранения. Дисциплина является факультативной.

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

Код и наименование компетенции	Код и наименование индикатора достижения компетенции
ПК-4. Способен к организации и проведению научных исследований в области здравоохранения	ИД-1 ПК-4.1. Использует методы естественных наук, статистику и интеллектуальные методы анализа данных для обработки результатов медико-биологических исследований
	ИД-4 ПК-4.4. Выполняет компьютерную обработку и анализ медицинских данных, сигналов и изображений, получаемых от устройств медицинской электроники с помощью информационных технологий и интеллектуальных методов анализа данных

Код индикатора достижения компетенции	Результаты обучения (показатели оценивания)	Оценочные средства
ИД-1 ПК-4.1.	знает основные алгоритмы машинного обучения и методы Data Mining	Тестовые задания Ситуационные задачи Контрольные вопросы
	умеет использовать библиотеки Python для реализации алгоритмов машинного обучения	
ИД-4 ПК-4.4.	знает основные метрики оценки качества моделей машинного обучения	Тестовые задания Ситуационные задачи Контрольные вопросы
	умеет оценивать качество моделей с использованием различных метрик (например, accuracy, precision, recall, F1-score).	

4. Объем дисциплины и виды учебной работы

Вид учебной работы	Трудоемкость	Семестры
		10
Контактная работа обучающихся с преподавателем:	48	48
Лекции	12	12
Практические занятия	34	34

Промежуточная аттестация: зачет, в том числе сдача и групповые консультации	2	2
Самостоятельная работа:	24	24
в период теоретического обучения	20	20
подготовка к сдаче зачета	4	4
Общая трудоемкость:	академических часов	72
	зачетных единиц	2

5. Содержание дисциплины, структурированное по разделам (темам) с указанием количества академических часов и видов учебных занятий

5.1. Содержание разделов дисциплины

№ п/п	Наименование раздела дисциплины	Аннотированное содержание раздела дисциплины	Перечень компетенций, формируемых в процессе освоения раздела
1	Подготовка данных для алгоритмов машинного обучения	Машинное обучение как проект: получение данных, обнаружение и визуализация данных, подготовка данных для алгоритмов машинного обучения, выбор модели и ее обучения, представление решения, запуск, наблюдение и сопровождение системы.	ПК-4
2	Классификация	Обучение двоичного классификатора. Показатели производительности: матрица неточностей, точность и полнота, кривая ROC. Многоклассовая классификация. Анализ ошибок.	ПК-4
3	Обучение моделей	Линейная регрессия. Градиентный спуск. Полиномиальная регрессия. Кривые обучения. Регуляризованные линейные модели: гребневая регрессия, лассо-регрессия. Логистическая регрессия: оценивание вероятностей, обучение и функция издержек, границы решений.	ПК-4
4	Методы опорных векторов	Линейная классификация SVM. Нелинейная классификация: полиномиальное ядро, добавление признаков близости, гауссово ядро RBF. Регрессия SVM.	ПК-4
5	Деревья принятия решений	Обучение и визуализация дерева принятия решений. Выработка прогнозов. Оценивание вероятностей классов. Алгоритм обучения CART. Гиперпараметры регуляризации. Регрессия. Неустойчивость.	ПК-4
6	Ансамблевое обучение и случайные леса	Классификаторы с голосованием: бэггинг и вставка, бэггинг и вставка в Scikit-Learn, оценка на неиспользуемых образцах. Методы случайных участков и случайных подпространств. Случайные леса: значимость признаков. Бустинг: градиентный бустинг. Стекинг.	ПК-4

5.2. Тематический план лекций

№ п/п	Наименование раздела дисциплины	Тематика лекций	Активные формы обучения	Трудоемкость (академических часов)
1	Подготовка данных для алгоритмов машинного обучения	Л.1 Основы машинного обучения и Data mining для анализа данных как проект	ЛБ	2
2	Классификация	Л.2 Применение методов классификации в медицинской кибернетике	ЛБ	2
3	Обучение моделей	Л.3 Методы обучения моделей в медицинской кибернетике	ЛБ	2
4	Методы опорных векторов	Л.4 Методы опорных векторов (SVM) в медицинской кибернетике	ЛБ	2
5	Деревья принятия решений	Л.5 Деревья принятия решений в медицинской кибернетике	ЛБ	2
6	Ансамблевое обучение и случайные леса	Л.6 Ансамблевое обучение и случайные леса в медицинской кибернетике	ЛБ	2
ИТОГО:				12

5.3. Тематический план практических занятий

№ п/п	Наименование раздела дисциплины	Тематика практических занятий	Активные формы обучения	Формы текущего контроля	Трудоемкость (академических часов)
1	Подготовка данных для алгоритмов машинного обучения	ПЗ.1 Очистка и предобработка медицинских данных	АС	Тестирование	4
2		ПЗ.2 Создание и анализ признаков для медицинских данных	АС	Тестирование	4
3	Классификация	ПЗ.3 Реализация и оценка моделей классификации для диагностики заболеваний	АС	Тестирование	4
4	Обучение моделей	ПЗ.4 Обучение модели для предсказания заболеваний	АС	Тестирование	4
5		ПЗ.5 Обучение модели для прогнозирования медицинских расходов	АС	Тестирование	4
6	Методы опорных векторов	ПЗ.6 Применение SVM для классификации медицинских данных	АС	Тестирование	4
7	Деревья принятия решений	ПЗ.7 Применение деревьев решений для диагностики заболеваний	АС	Тестирование	4

8	Ансамблевое обучение и случайные леса	ПЗ.8 Применение ансамблевых методов для улучшения точности классификации медицинских данных	АС	Тестирование	4
9		ПЗ.9 Применение случайных лесов для диагностики заболеваний	АС	Тестирование	4
ИТОГО:					36

5.4. Самостоятельная работа:

№ п/п	Наименование раздела дисциплины	Виды самостоятельной работы	Формы текущего контроля	Трудоемкость (академических часов)
1	Подготовка данных для алгоритмов машинного обучения	Работа с лекционным материалом	Тестирование	3
2	Классификация	Работа с учебной литературой	Тестирование	3
3	Обучение моделей	Работа с учебной литературой	Тестирование	3
4	Методы опорных векторов	Работа с учебной литературой	Тестирование	3
5	Деревья принятия решений	Подготовка к контрольной работе	Тестирование	3
6	Ансамблевое обучение и случайные леса	Подготовка к сдаче зачета	Тестирование	5
Подготовка к сдаче зачета				4
ИТОГО:				24

6. Методические указания для обучающихся по освоению дисциплины

Для успешного освоения учебной дисциплины обучающемуся рекомендуется посещать лекционные и практические занятия в соответствии с расписанием учебных занятий, своевременно и в полном объеме выполнять задания текущего контроля, пройти промежуточную аттестацию.

Подготовка к лекциям

Лекции по дисциплине проводятся в традиционной и интерактивной форме с использованием технических средств обучения. Во время лекций студенту необходимо вести конспект лекции, структура и объем которого определяется самостоятельно. Основой формирования конспекта являются аудио, видеоматериалы, презентации лектора по тематике лекции, а также рекомендованная учебная литература, ресурсы сети «Интернет» и/или нормативные документы.

Подготовка к практическим занятиям и выполнение практических занятий

Практические занятия проводятся с использованием активных форм обучения. При подготовке к практическим занятиям необходимо выполнять задания для самостоятельной работы. В программе дисциплины предусмотрены мероприятия текущего контроля для проверки освоения разделов дисциплины в рамках самостоятельной работы. Контроль выполнения заданий на практических занятиях осуществляется с помощью мероприятий текущего контроля.

Рекомендации по работе с литературой

В программе дисциплины представлен список литературы и ресурсов сети «Интернет», необходимых для освоения дисциплины.

Рекомендации по подготовке к текущему контролю

С целью контроля освоения дисциплины в тематическом плане занятий предусмотрены контрольные мероприятия, которые составляют средства текущего контроля. В рабочей программе дисциплины текущий контроль представлен тестовыми заданиями.

Рекомендации по подготовке к промежуточной аттестации

Промежуточная аттестация по дисциплине проводится в форме зачета. Зачет включает в себя: собеседование по контрольным вопросам, решение ситуационных задач.

7. Оценочные материалы

Оценочные материалы по дисциплине для проведения текущего контроля успеваемости и промежуточной аттестации обучающихся включают в себя примеры оценочных средств (Приложение А к рабочей программе дисциплины), процедуру и критерии оценивания.

8. Перечень учебной литературы и ресурсов сети «Интернет», необходимых для освоения дисциплины

8.1. Учебная литература:

1. Абдулаева, З. И. Информационные компьютерные системы в медицине и здравоохранении : Учеб.-метод. пособие / З. И. Абдулаева, А. Д. Шматко; М-во здравоохранения Рос. Федерации, ФГБОУ ВО Сев.-Зап. гос. мед. ун-т им. И. И. Мечникова, Каф. мед. информатики и физики. - СПб. : Изд-во СЗГМУ им. И. И. Мечникова, 2017. - 43 с. : рис. - (Медицинское образование). - Библиогр.: с. 43 (11 назв.). [https://sdo.szgmu.ru/pluginfile.php/178023/mod_resource/content/1/ Абдулаева З. И., Шматко А. Д. Информационные компьютерные системы в медицине и здравоохранении.pdf](https://sdo.szgmu.ru/pluginfile.php/178023/mod_resource/content/1/Абдулаева_З._И.,_Шматко_А._Д._Информационные_компьютерные_системы_в_медицине_и_здравоохранении.pdf)
2. Абдулаева, З. И. Медицинская информатика [Текст] : учебное пособие : [в 2 ч.] / З. И. Абдулаева, Д. Ф. Курбанбаева. Теоретические основы медицинской информатики. — Санкт-Петербург : Изд-во Политехнического ун-та, 2018-, 2018. — 190 с. : ил., табл.; ISBN 978-5-7422-6240-4. — режим доступа: https://rusneb.ru/catalog/000199_000009_009794428/
3. Медицинская информатика : учебник / Т. В. Зарубина, Б. А. Кобринский, С. С. Белоносков [и др.]; ред. Т. В. Зарубина, Б. А. Кобринский. - М. : ГЭОТАР-Медиа, 2018. - 507 с. : цв. ил., табл. - Терминол. словарь: с. 490-491.- Библиогр.: с. 500-501.- Предм. указ.: с. 502-507. - ISBN 978-5-9704-4573-0.
4. Принципы функционирования интеллектуальной системы динамического контроля факторов риска и формирования рекомендаций по здоровьесбережению / Б. А. Кобринский, А. С. Кадыков, М. Г. Полтавская [и др.] // Профилактическая медицина. - 2019. - Т. 22, № 5. - С. 78-84. - Библиогр.: 26 назв. Ресурсы информационно-телекоммуникационной сети «Интернет»:

9. Перечень информационных технологий, используемых для освоения дисциплины, включая перечень программного обеспечения, профессиональных баз данных и информационных справочных систем

9.1. Перечень программного обеспечения, используемого при осуществлении образовательного процесса (лицензионное и свободно распространяемое программное обеспечение, в том числе отечественного производства):

№ п/п	Наименование программного продукта	Срок действия лицензии	Документы, подтверждающие право использования программных продуктов
лицензионное программное обеспечение			
1.	Dr. Web	1 год	Контракт № 265-2023-ЗК
2.	MS Windows 8 MS Windows 8.1 MS Windows 10 MS Windows Server 2012 Datacenter - 2 Proc MS Windows Server 2012 R2 Datacenter - 2 Proc MS Windows Server 2016 Datacenter Core	Неограниченно	Государственный контракт № 30/2013-О; Государственный контракт № 399/2013-ОА; Государственный контракт № 07/2017-ЭА.
3.	MS Office 2010 MS Office 2013	Неограниченно	Государственный контракт № 30/2013-ОА; Государственный контракт № 399/2013-ОА.
4.	Academic LabVIEW Premium Suite (1 User)	Неограниченно	Государственный контракт № 02/2015
лицензионное программное обеспечение отечественного производства			
1.	Антиплагиат	1 год	Договор № 133/2024-М
2.	«WEBINAR (ВЕБИНАР)» ВЕРСИЯ 3.0	1 год	Контракт № 211/2024-ЭА
3.	«Среда электронного обучения ЗКЛ»	1 год	Контракт № 121/2024-ЗЗЕП
4.	TrueConf Enterprise	1 год	Контракт № 216/2024-ЭА
свободно распространяемое программное обеспечение			
1.	Google Chrome	Неограниченно	Открытое лицензионное соглашение GNU GeneralPublicLicense
2.	NVDA	Неограниченно	Открытое лицензионное соглашение GNU GeneralPublicLicense
свободно распространяемое программное обеспечение отечественного производства			
1.	Moodle	Неограниченно	Открытое лицензионное соглашение GNU GeneralPublicLicense

9.2. Перечень профессиональных баз данных и информационных справочных систем:

№ п/п	Наименование программного продукта	Срок действия лицензии	Документы, подтверждающие право использования программных продуктов	Режим доступа для обучающихся – инвалидов и лиц с ограниченными возможностями здоровья
1.	Консультант Плюс	1 год	Контракт № 1067/2021-ЭА	-
2.	ЭБС «Консультант студента»	1 год	Контракт № 97/2023-ЭА	https://www.studentlibrary.ru/

3.	ЭБС «Айбукс.ру/ibooks.ru»	1 год	Договор № 207/2023-ЗЗЕП	https://ibooks.ru
4.	Цифровой образовательный ресурс IPRsmart	1 год	Договор № 206/2023-ЗЗЕП	http://www.iprbookshop.ru/
5.	Электронно-библиотечная система «Букап»	1 год	Договор № 199/2023-ЗЗЕП	https://www.books-up.ru/
6.	ЭБС «Издательство Лань»	1 год	Договор № 200/2023-ЗЗЕП	https://e.lanbook.com/
7.	Образовательная платформа ЮРАЙТ	1 год	Договор № 155/2023-ПЗ	https://urait.ru/
8.	Электронные издания в составе базы данных НАУЧНАЯ ЭЛЕКТРОННАЯ БИБЛИОТЕКА eLIBRARY.RU	1 год	Лицензионный договор № SU-7139/2024	https://www.elibrary.ru/defaultx.asp
9.	Программное обеспечение «Платформа mb4» в части Справочно-информационной системы «MedBaseGeotar»	1 год	Лицензионный договор № 97/2024-ЗЗЕП	https://mbasegeotar.ru/
10.	Универсальные базы электронных периодических изданий ИВИС	1 год	Лицензионный договор № 116/2023-ЗЗЕП «Журналы России по медицине и здравоохранению» Лицензионный договор № 42/2023-ЗЗЕП «Индивидуальные издания»	https://dlib.eastview.com/
11.	Создание Виртуального читального зала Российской государственной библиотеки (ВЧЗ РГБ) для обслуживания удаленного пользователя	1 год	Лицензионный договор № 120/2024-М14	https://search.rsl.ru/

10. Материально-техническое обеспечение дисциплины

Учебные аудитории для проведения учебных занятий лекционного типа, групповых и индивидуальных консультаций, текущего контроля успеваемости и промежуточной аттестации обучающихся, оснащенные оборудованием и техническими средствами обучения: г. Санкт-Петербург, Пискаревский проспект, д. 47, лит АЛ, лит Р, ФГБОУ ВО СЗГМУ им. И.И. Мечникова Минздрава России;
Оборудование: доска (меловая); стол и стул преподавателя, столы и стулья студенческие

Технические средства обучения: мультимедиа-проектор, ноутбук преподавателя, системный блок, монитор.

Специальные технические средства обучения: Roger Pen (Индивидуальный беспроводной передатчик Roger в форме ручки), Roger MyLink (приемник сигнала системы Roger Pen) (для обучающихся с нарушениями слуха); IntelliKeys (проводная клавиатура с русским шрифтом Брайля с матовым покрытием черного цвета), (г. Санкт-Петербург, Пискаревский проспект, д. 47, лит Р (корп.9), ауд. № 18,19, ФГБОУ ВО СЗГМУ им. И.И. Мечникова Минздрава России).

Учебные аудитории для проведения учебных занятий семинарского типа, групповых и индивидуальных консультаций, текущего контроля успеваемости и промежуточной аттестации обучающихся, оснащенные оборудованием и техническими средствами обучения: г. Санкт-Петербург, Пискаревский проспект, д. 47, лит АЛ, лит Р, ФГБОУ ВО СЗГМУ им. И.И. Мечникова Минздрава России;

Оборудование: столы, стулья, доски маркерные, доска меловая.

Технические средства обучения: 110 автоматизированных рабочих мест (ESET NOD 32, MS Windows 7, MS Office 2010, Moodle, GNU, Academic LabVIEW Premium Suite (1 User), Google Chrome). Специальные технические средства обучения: Roger Pen (Индивидуальный беспроводной передатчик Roger в форме ручки), Roger MyLink (приемник сигнала системы Roger Pen) (для обучающихся с нарушениями слуха); IntelliKeys (проводная клавиатура с русским шрифтом Брайля с матовым покрытием черного цвета), (г. Санкт-Петербург, Пискаревский проспект, д. 47, лит Р (корп.9), ауд. № 18,19, ФГБОУ ВО СЗГМУ им. И.И. Мечникова Минздрава России).

Помещения для самостоятельной работы обучающихся, оснащенные компьютерной техникой с возможностью подключения к сети «Интернет» и обеспеченные доступом к электронной информационно-образовательной среде Университета: г. Санкт-Петербург, Пискаревский проспект, д. 47, лит АЕ (корп.32), ауд. № 1, лит Р (корп.9), ауд. № 18,19 ФГБОУ ВО СЗГМУ им. И.И. Мечникова Минздрава России.

Министерство здравоохранения Российской Федерации
федеральное государственное бюджетное образовательное учреждение
высшего образования
«Северо-Западный государственный медицинский университет
имени И.И. Мечникова»
Министерства здравоохранения Российской Федерации
(ФГБОУ ВО СЗГМУ им. И.И. Мечникова Минздрава России)

ОЦЕНОЧНЫЕ МАТЕРИАЛЫ

(для проведения текущего контроля успеваемости и промежуточной аттестации обучающихся)

Специальность:	30.05.03 Медицинская кибернетика
Направленность:	Цифровые технологии медицины и здравоохранения
Наименование дисциплины:	Основы машинного обучения и Data mining

1. Перечень планируемых результатов обучения по дисциплине

Код индикатора достижения компетенции	Результаты обучения (показатели оценивания)	Оценочные средства
ИД-1 ПК-4.1.	знает основные алгоритмы машинного обучения и методы Data Mining	Тестовые задания Ситуационные задачи Контрольные вопросы
	умеет использовать библиотеки Python для реализации алгоритмов машинного обучения	
ИД-4 ПК-4.4.	знает основные метрики оценки качества моделей машинного обучения	Тестовые задания Ситуационные задачи Контрольные вопросы
	умеет оценивать качество моделей с использованием различных метрик (например, accuracy, precision, recall, F1-score).	

2.1. Примеры входного контроля

Что такое алгоритм?

- 1) Набор данных
- 2) Последовательность действий для решения задачи**
- 3) Язык программирования
- 4) Компьютерная программа

Какой из следующих методов является методом "разделяй и властвуй"?

- 1) Сортировка вставками
- 2) Быстрая сортировка**
- 3) Сортировка выбором
- 4) Линейный поиск

Критерии оценки, шкала оценивания *зачтено/не зачтено*

Оценка	Описание
«зачтено»	Демонстрирует полное понимание проблемы. Все требования, предъявляемые к заданию, выполнены
«не зачтено»	Демонстрирует непонимание проблемы. Многие требования, предъявляемые к заданию, не выполнены. Нет ответа

2.2. Тестовые задания:

ИД-1 ПК-4.1.

Название вопроса: Вопрос № 1.

Какова основная цель этапа подготовки данных (data preprocessing) в проекте машинного обучения для медицинских исследований?

- 5) Немедленное построение финальной модели.
- 6) Создание сложной визуализации для отчета.
- 7) Приведение данных в форму, пригодную для эффективного обучения алгоритмов, и повышение их качества.
- 8) Автоматический выбор лучшего алгоритма.

ИД-4 ПК-4.4.

Название вопроса: Вопрос № 11

При обработке медицинских изображений для последующей классификации CNN, какой этап часто предшествует загрузке в модель и направлен на увеличение размера обучающей выборки?

- 1) Сжатие изображений (compression).
- 2) Аугментация данных (data augmentation) — случайные повороты, отражения, изменения яркости.
- 3) Преобразование в черно-белый формат.
- 4) Удаление всех изображений низкого качества.

3. Процедура проведения текущего контроля

Текущий контроль успеваемости по дисциплине проводится в форме: тестирования.

4. Примеры оценочных средств и критерии оценивания для проведения промежуточной аттестации

4.1. Примерный перечень контрольных вопросов для подготовки к зачету:

ИД-1 ПК-4.1.

1. Что такое машинное обучение и какие основные задачи оно решает?
2. Опишите основные этапы процесса машинного обучения.
3. В чем разница между обучением с учителем и без учителя?
4. Какие методы используются для оценки качества моделей машинного обучения?
5. Что такое переобучение и как его можно избежать?
6. Опишите принцип работы метода опорных векторов (SVM).
7. Какие алгоритмы используются для кластеризации данных?
8. В чем заключается метод главных компонент (РСА) и для чего он используется?
9. Опишите основные методы ансамблевого обучения.
10. Какие преимущества и недостатки имеют деревья решений?
11. Как загрузить и предобработать данные для машинного обучения с использованием библиотеки pandas?
12. Как реализовать модель линейной регрессии на языке Python с использованием библиотеки scikit-learn?

ИД-4 ПК-4.4.

13. Как провести кросс-валидацию модели и настроить гиперпараметры?
14. Как визуализировать результаты работы модели с помощью библиотеки matplotlib или seaborn?
15. Как реализовать и оценить модель случайного леса для задачи классификации?
16. Как использовать алгоритм k-means для кластеризации данных?
17. Как применить метод логистической регрессии для задачи бинарной классификации?
18. Как интерпретировать результаты работы модели и сделать выводы на основе полученных данных?

Критерии оценки, шкала оценивания по контрольным вопросам

Оценка	Описание
«отлично»	Знает весь учебный материал, отлично понимает и прочно усвоил его. На вопросы (в пределах программы) дает правильные, сознательные и уверенные ответы. В устных ответах пользуется литературно правильным языком и не допускает ошибок
«хорошо»	Знает весь требуемый учебный материал, хорошо понимает и прочно усвоил его. На вопросы (в пределах программы) отвечает без затруднений. В устных ответах пользуется литературным

	языком и не делает грубых ошибок
«удовлетворительно»	Знает основной учебный материал. На вопросы (в пределах программы) отвечает с затруднением. В устных ответах допускает ошибки при изложении материала и в построении речи
«неудовлетворительно»	Не знает большей части учебного материала, отвечает, как правило, лишь на наводящие вопросы преподавателя, неуверенно. В устных ответах допускает частые и грубые ошибки

4.2. Примеры ситуационных задач:

ИД-1 ПК-4.1.

Задача 1. Диагностика диабета.

Используя набор данных Pima Indians Diabetes Database, разработайте модель для диагностики диабета. Реализуйте алгоритмы логистической регрессии и случайного леса, сравните их производительность.

ИД-4 ПК-4.4.

Задача 2. Прогнозирование медицинских расходов.

Используя набор данных Medical Cost Personal Datasets, разработайте модель для прогнозирования медицинских расходов. Реализуйте линейную регрессию и градиентный бустинг, оцените качество моделей с использованием метрик MAE и RMSE.

Критерии оценки, шкала оценивания ситуационных задач

Оценка	Описание
«отлично»	Объяснение кода решения ситуационной задачи подробное, последовательное, грамотное, с теоретическими обоснованиями, с необходимым схематическими изображениями и наглядными демонстрациями, с правильным и свободным владением терминологией; ответы на дополнительные вопросы верные, четкие
«хорошо»	Объяснение кода решения ситуационной задачи подробное, но недостаточно логичное, с единичными ошибками в деталях, некоторыми затруднениями в теоретическом обосновании, схематических изображениях и наглядных демонстрациях, ответы на дополнительные вопросы верные, но недостаточно четкие
«удовлетворительно»	Объяснение кода решения ситуационной задачи недостаточно полное, непоследовательное, с ошибками, слабым теоретическим обоснованием, со значительными затруднениями и ошибками в схематических изображениях и наглядных демонстрациях, ответы на дополнительные вопросы недостаточно четкие, с ошибками в деталях
«неудовлетворительно»	Объяснение кода решения ситуационной задачи дано неполное, непоследовательное, с грубыми ошибками, без теоретического обоснования, без умения схематических изображений и наглядных демонстраций или с большим количеством ошибок, ответы на дополнительные вопросы неправильные или отсутствуют

Критерии оценки, шкала итогового оценивания (зачет)

Оценка	Описание
«зачтено»	Демонстрирует полное понимание проблемы. Знает основные понятия в рамках обсуждаемого вопроса, методы изучения и их взаимосвязь между собой, практические проблемы и имеет представление о перспективных направлениях разработки рассматриваемого вопроса
«не зачтено»	Демонстрирует непонимание проблемы. Не знает основные

	понятия, методы изучения, в рамках обсуждаемого вопроса не имеет представления об основных практических проблемах
--	---

5. Процедура проведения промежуточной аттестации

Промежуточная аттестация по дисциплине проводится в форме зачета. Зачет включает в себя: решение ситуационных задач и собеседование по контрольным вопросам.