



УТВЕРЖДАЮ
Заместитель директора Ур СЭИ (филиал)
ОУП ВО «АТиСО» по УВР

_____ О.В. Зубкова

« 17 » февраля 2026 г.

РАБОЧАЯ ПРОГРАММА УЧЕБНОЙ ДИСЦИПЛИНЫ (МОДУЛЯ)

ОПЦ.20 АНАЛИЗ ДАННЫХ И МАШИННОЕ ОБУЧЕНИЕ

**СПЕЦИАЛЬНОСТЬ СРЕДНЕГО ПРОФЕССИОНАЛЬНОГО
ОБРАЗОВАНИЯ**

09.02.11 Разработка и управление программным обеспечением
(код профессии, специальности СПО)

Программист
(наименование квалификации)

Кафедра: Информатики и математики

Разработчик программы:

Зав. кафедрой

Маскайкин Е.П.

17 февраля 2026 г.

Оглавление

1. ОБЛАСТЬ ПРИМЕНЕНИЯ РАБОЧЕЙ ПРОГРАММЫ УЧЕБНОГО ПРЕДМЕТА, КУРСА, ДИСЦИПЛИНЫ (МОДУЛЯ)	3
2. ЦЕЛИ И ЗАДАЧИ УЧЕБНОЙ ДИСЦИПЛИНЫ.....	3
3. МЕСТО УЧЕБНОЙ ДИСЦИПЛИНЫ В СТРУКТУРЕ ПРОГРАММЫ ПОДГОТОВКИ СПЕЦИАЛИСТОВ СРЕДНЕГО ЗВЕНА	3
4. ТРЕБОВАНИЯ К РЕЗУЛЬТАТАМ ОСВОЕНИЯ УЧЕБНОЙ ДИСЦИПЛИНЫ:.....	3
5. СТРУКТУРА И СОДЕРЖАНИЕ УЧЕБНОЙ ДИСЦИПЛИНЫ	4
6. УСЛОВИЯ РЕАЛИЗАЦИИ ПРОГРАММЫ УЧЕБНОЙ ДИСЦИПЛИНЫ	8
7. КОНТРОЛЬ И ОЦЕНКА РЕЗУЛЬТАТОВ ОСВОЕНИЯ УЧЕБНОЙ ДИСЦИПЛИНЫ	9
8. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ	10
9. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ПО УЧЕБНОЙ ДИСЦИПЛИНЕ СО.02.02. ИНФОРМАТИКА	10

1. ОБЛАСТЬ ПРИМЕНЕНИЯ РАБОЧЕЙ ПРОГРАММЫ УЧЕБНОГО ПРЕДМЕТА, КУРСА, ДИСЦИПЛИНЫ (МОДУЛЯ)

Рабочая программа учебной дисциплины «Анализ данных и машинное обучение» является частью программы подготовки специалистов среднего звена (ППССЗ) в соответствии с ФГОС СПО по специальности 09.02.11 Разработка и управление программным обеспечением, квалификация – программист.

Программа может быть использована в дополнительном профессиональном образовании и профессиональной подготовке специалистов в области информационных технологий.

2. ЦЕЛИ И ЗАДАЧИ УЧЕБНОЙ ДИСЦИПЛИНЫ

Цель изучения: обучить студентов современным методам интеллектуального анализа данных и познакомить с основными алгоритмами машинного обучения, особенностями их применения, а также сформировать практические навыки использования этих методов для решения профессиональных задач в области разработки и управления программным обеспечением.

Задачи:

- изучить методы интеллектуального анализа данных (очистка, преобразование, визуализация, корреляционный анализ);
- сформировать базовые понятия машинного обучения (обучение с учителем и без учителя, задачи классификации, регрессии, кластеризации);
- изучить основные алгоритмы машинного обучения: метод ближайших соседей, решающие деревья, линейные методы, логистическая регрессия, метод опорных векторов;
- освоить инструменты анализа данных на языке Python (библиотеки Pandas, NumPy, Scikit-learn, Matplotlib, Seaborn);
- сформировать практические навыки применения алгоритмов машинного обучения для решения бизнес-задач;
- развить умение обосновывать выбор алгоритма для конкретной задачи и оценивать качество моделей..

3. МЕСТО УЧЕБНОЙ ДИСЦИПЛИНЫ В СТРУКТУРЕ ПРОГРАММЫ ПОДГОТОВКИ СПЕЦИАЛИСТОВ СРЕДНЕГО ЗВЕНА

Учебная дисциплина «Анализ данных и машинное обучение» относится к общепрофессиональному циклу (ОПЦ.20) специальности 09.02.11 Разработка и управление программным обеспечением.

Дисциплина базируется на знаниях, умениях и навыках, полученных при изучении информатики, математики, основ алгоритмизации и программирования, дискретной математики, электронных таблиц.

Знания, умения и навыки, полученные при изучении данной дисциплины, являются основой для освоения профессиональных модулей, связанных с разработкой и модификацией бизнес-приложений, а также для практической деятельности программиста в области анализа данных, машинного обучения и искусственного интеллекта.

4. ТРЕБОВАНИЯ К РЕЗУЛЬТАТАМ ОСВОЕНИЯ УЧЕБНОЙ ДИСЦИПЛИНЫ:

В результате освоения дисциплины обучающийся должен овладеть следующими компетенциями:

Код	Формулировка компетенции
ОК 01	Выбирать способы решения задач профессиональной деятельности применительно к различным контекстам.
ОК 02	Использовать современные средства поиска, анализа и интерпретации информации, и информационные технологии для выполнения задач профессиональной деятельности.
ПК 3.2	Разрабатывать бизнес-приложения.
ПК 3.3	Модифицировать бизнес-приложения.
ПК 3.4	Выполнять тестирование и отладку бизнес-приложений.

Таким образом, в результате освоения дисциплины обучающийся должен:
уметь:

- У1 – проводить анализ данных, включая предобработку, визуализацию и корреляционный анализ;
- У2 – применять на практике алгоритмы машинного обучения (метод ближайших соседей, решающие деревья, линейные методы, логистическую регрессию, метод опорных векторов);
- У3 – обосновать применение того или иного алгоритма машинного обучения для решения конкретной задачи;
- У4 – оценивать качество моделей, проводить кросс-валидацию, настраивать гиперпараметры;
- У5 – интерпретировать результаты моделей и внедрять их в бизнес-приложения.

знать:

- 1 – методы интеллектуального анализа данных (включая преобразование и очистку, работу с пропущенными значениями, основные способы визуализации данных, корреляционный анализ, поиск нелинейных ассоциаций);
- 32 – возможности алгоритмов машинного обучения;
- 33 – классы задач, решаемых с помощью алгоритмов машинного обучения (классификация, регрессия, кластеризация);
- 34 – метрики качества моделей, методы оценки обобщающей способности, проблему переобучения;
- 35 – основные библиотеки Python для анализа данных и машинного обучения.

5. СТРУКТУРА И СОДЕРЖАНИЕ УЧЕБНОЙ ДИСЦИПЛИНЫ

5.1 Объем учебной дисциплины и виды учебной работы

Вид учебной работы	7 семестр	Итого
Максимальная учебная нагрузка (всего)	130	130
Обязательная аудиторная учебная нагрузка (всего)	50	50
в том числе: лекции	20	20

Вид учебной работы	7 семестр	Итого
практические занятия	30	30
Самостоятельная работа обучающегося (всего)	80	80
Итоговая аттестация	Дифференцированный зачет (7 семестр)	Дифф. зачет

5.2. Тематический план и содержание учебной дисциплины

Наименование разделов и тем	Содержание учебного материала, практические занятия, самостоятельная работа обучающихся	Объем часов	Уровень освоения	Компетенции
Тема 1. Большие данные и машинное обучение	Лекция: Общая постановка задачи машинного обучения. Обучение с учителем и без учителя, обучение с подкреплением. Задачи классификации, регрессии, предсказания. Модели алгоритмов, признаки, типы признаков. Функционал качества, вероятностная постановка. Оценка обобщающей способности, проблема переобучения. Критерии оценки качества (accuracy, precision, recall, F1, ROC-кривые). Примеры практических задач. Практические занятия: Первичный анализ данных с библиотекой Pandas. Визуальный анализ данных с библиотеками Seaborn и Matplotlib. Самостоятельная работа: изучение теоретического материала, подготовка докладов по темам.	28 (6 лекций + 8 практик + 14 СР)	1-2	ОК 01, ОК 02
Тема 2. Метрические методы классификации	Лекция: Обобщенный метрический классификатор. Виды метрик. Метод ближайшего соседа. Алгоритм k-ближайших соседей (kNN). Взвешенная версия алгоритма. Метод окна Парзена, метод потенциальных функций. Понятие эталона. Отступы и классификация объектов. Практические занятия: Классификация с использованием kNN. Кросс-валидация, подбор параметра k. Самостоятельная работа: решение задач, подготовка докладов.	26 (5 лекций + 6 практик + 15 СР)	2	ОК 01, ОК 02, ПК 3.2, ПК 3.4

Наименование разделов и тем	Содержание учебного материала, практические занятия, самостоятельная работа обучающихся	Объем часов	Уровень освоения	Компетенции
Тема 3. Логические методы классификации	Лекция: Основные принципы построения логических алгоритмов классификации. Критерии информативности (энтропийный, статистический). Алгоритм построения дерева решений ID3, C4.5. Решающие деревья в Scikit-learn. Важность признаков. Практические занятия: Построение решающих деревьев. Визуализация деревьев. Анализ важности признаков. Самостоятельная работа: решение задач, подготовка докладов.	26 (5 лекций + 6 практик + 15 СР)	2	ОК 01, ОК 02, ПК 3.2, ПК 3.3
Тема 4. Линейные методы классификации и регрессии	Лекция: Постановка задач линейной регрессии и линейной классификации. Метод наименьших квадратов. Регуляризация: гребневая регрессия (Ridge), Лассо (Lasso). Линейные классификаторы: перцептрон, логистическая регрессия, метод опорных векторов (SVM). Линейно разделимые выборки, двойственная задача, нелинейные обобщения (ядра). Практические занятия: Реализация линейных моделей в Scikit-learn. Сравнение Ridge, Lasso, LogisticRegression, SVM. Подбор гиперпараметров. Самостоятельная работа: решение задач, подготовка докладов.	50 (4 лекции + 10 практик + 36 СР)	2-3	ОК 01, ОК 02, ПК 3.2, ПК 3.3, ПК 3.4
Всего		130		

6. УСЛОВИЯ РЕАЛИЗАЦИИ ПРОГРАММЫ УЧЕБНОЙ ДИСЦИПЛИНЫ

6.1. Требования к минимальному материально-техническому обеспечению

Реализация программы требует наличия компьютерного класса, оснащённого в соответствии с требованиями ФГОС.

Оборудование:

- рабочие места по количеству обучающихся (компьютеры с доступом в Интернет);
- рабочее место преподавателя;
- мультимедийное оборудование (проектор, экран).

Программное обеспечение:

- ОС Windows / Linux;
- Python (версия 3.x) с библиотеками: NumPy, Pandas, Scikit-learn, Matplotlib, Seaborn, Jupyter Notebook;
- среда разработки (Jupyter Notebook, VS Code, PyCharm);
- браузеры.

Технические средства обучения:

- компьютер, мультимедийный проектор, экран;
- доступ к сети Интернет для использования электронных образовательных ресурсов.

6.2. Информационное обеспечение обучения

Основная литература

1. Машинное обучение : учебник : [16+] / Е. Ю. Бутырский, В. В. Цехановский, Н. А. Жукова [и др.]. – Москва : Директ-Медиа, 2023. – 368 с. : ил., табл., схем., граф. – Режим доступа: по подписке. – URL: <https://biblioclub.ru/index.php?page=book&id=701807>. – Библиогр. в кн. – ISBN 978-5-4499-3778-0. – DOI 10.23681/701807. – Текст : электронный.
2. Платонов, А. В. Машинное обучение : учебное пособие / А. В. Платонов. — 2-е изд. — Москва : Издательство Юрайт, 2026. — 89 с. — (Высшее образование). — ISBN 978-5-534-20732-3. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/589132>.
3. Чертыковцев, В. К. Проектирование интерфейсов пользователя. Человеко-машинное взаимодействие : учебник для среднего профессионального образования / В. К. Чертыковцев. — Москва : Издательство Юрайт, 2026. — 111 с. — (Профессиональное образование). — ISBN 978-5-534-20809-2. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://urait.ru/bcode/590007>

Дополнительная литература

4. Баюк, О. А. Практикум по анализу данных на языках Python и R : учебное пособие для студентов, обучающихся по направлениям 38.03.01 «Экономика», 38.03.05 «Бизнес-информатика» : [16+] / О. А. Баюк, М. Р. Исаева, М. О. Самсонкин ; Финансовый университет при Правительстве Российской Федерации. – Москва : Прометей, 2023. – 100 с. : ил., табл., схем., граф. – Режим доступа: по подписке. – URL: <https://biblioclub.ru/index.php?page=book&id=700938>. – Библиогр. в кн. – ISBN 978-5-00172-356-1. – Текст : электронный.
5. Сахарова, Л. В. Методы оптимизации для машинного обучения : учебное пособие : [16+] / Л. В. Сахарова, Г. В. Лукьянова ; Ростовский государственный экономический университет (РИНХ). – Ростов-на-Дону : Издательско-полиграфический комплекс РГЭУ (РИНХ), 2023. – 87 с. : ил. – Режим доступа: по подписке. – URL: <https://biblioclub.ru/index.php?page=book&id=711220>. – Библиогр. в кн. – ISBN 978-5-7972-3139-4. – Текст : электронный.

6. Татарникова, Т. М. Интеллектуальный анализ данных : учебное пособие : [16+] / Т. М. Татарникова. – Москва ; Вологда : Инфра-Инженерия, 2024. – 172 с. : ил., табл. – Режим доступа: по подписке. – URL: <https://biblioclub.ru/index.php?page=book&id=725643>. – Библиогр. в кн. – ISBN 978-5-9729-1772-3. – Текст : электронный.

Ресурсы информационно-телекоммуникационной сети «Интернет», необходимые для освоения дисциплины

<http://citforum.ru/> – IT-портал
<https://habrahabr.ru/> – ресурс для IT-специалистов
<http://stackoverflow.com/> – вопросы и ответы
<http://www.intuit.ru> – Интернет-университет информационных технологий
<https://scikit-learn.org/> – документация Scikit-learn
<https://pandas.pydata.org/> – документация Pandas

7. КОНТРОЛЬ И ОЦЕНКА РЕЗУЛЬТАТОВ ОСВОЕНИЯ УЧЕБНОЙ ДИСЦИПЛИНЫ

Рабочая программа учебной дисциплины обеспечена оценочными материалами для проведения текущего контроля и промежуточной аттестации.

Оценочные материалы включают в себя тестовые задания, практические задания, вопросы к дифференцированному зачёту.

Текущий контроль осуществляется в форме проверки выполнения практических работ, устных опросов, тестирования, защиты отчётов по самостоятельной работе.

Промежуточная аттестация – дифференцированный зачёт в 7 семестре (включает теоретические вопросы и практическое задание – решение задачи анализа данных с использованием Python).

Примерный перечень вопросов к дифференцированному зачёту:

1. Основные понятия: данные, информация, знания. Классификация задач Data Mining.
2. Машинное обучение: формализация задачи, типы обучения (с учителем, без учителя, с подкреплением).
3. Задачи классификации, регрессии, кластеризации.
4. Признаки и их типы. Масштабирование признаков.
5. Функционалы качества, проблема переобучения и обобщения.
6. Кросс-валидация: смысл, методы (K-fold, leave-one-out).
7. Метрики качества классификации: accuracy, precision, recall, F1, ROC-кривая.
8. Метрики качества регрессии: MSE, MAE, R^2 .
9. Метод k-ближайших соседей (kNN): принцип работы, выбор k.
10. Метод Парзена, метод потенциальных функций.
11. Решающие деревья: принцип построения, критерии информативности (энтропия, Gini).
12. Алгоритмы ID3, C4.5.
13. Линейная регрессия: метод наименьших квадратов, регуляризация (Ridge, Lasso).
14. Логистическая регрессия.
15. Метод опорных векторов (SVM): линейный случай, ядра.
16. Библиотеки Python для анализа данных: Pandas, NumPy, Matplotlib, Seaborn.
17. Библиотека Scikit-learn: основные классы и методы.
18. Оценка важности признаков.
19. Применение машинного обучения в бизнес-приложениях.
20. Этические аспекты использования машинного обучения.

8. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ

По выполнению лекционных занятий:

- вести конспект, фиксируя основные понятия, формулы, алгоритмы;
- после лекции прорабатывать материал по учебнику и документации;
- выделять неясные вопросы для обсуждения с преподавателем.

По выполнению практических занятий:

- выполнять лабораторные работы в Jupyter Notebook, следуя инструкциям;
- экспериментировать с параметрами алгоритмов, анализировать результаты;
- фиксировать выводы, оформлять отчёты.

По самостоятельной работе:

- выполнять домашние задания, решать задачи по анализу данных;
- изучать дополнительную литературу и онлайн-ресурсы;
- готовиться к контрольным работам и дифференцированному зачёту.

По подготовке к дифференцированному зачёту:

- повторить теоретические положения по разделам;
- проработать типовые практические задачи;
- использовать вопросы для самопроверки.

9. ФОНД ОЦЕНОЧНЫХ СРЕДСТВ ПО УЧЕБНОЙ ДИСЦИПЛИНЕ

Фонд оценочных средств включает задания разных типов, проверяющих как знание теории, так и умение применять математические методы.

Тип задания	Инструкция	Пример задания
Задание закрытого типа на установление соответствия (Знать)	Прочитайте текст и установите соответствие между алгоритмами машинного обучения и их характеристиками.	Установите соответствие между алгоритмом и его характеристикой: А) kNN; Б) Дерево решений; В) Логистическая регрессия; Г) SVM. 1) Линейный классификатор, вероятностная интерпретация; 2) Метрический метод, основанный на близости объектов; 3) Максимизация зазора между классами; 4) Логическая модель, использующая разбиение по признакам.
Задание закрытого типа на установление соответствия (Уметь)	Прочитайте текст и установите соответствие между методами предобработки данных и их описанием.	Установите соответствие между методом предобработки и его описанием: А) Стандартизация; Б) Нормализация; В) Заполнение пропусков; Г) Кодирование категорий. 1) Приведение данных к диапазону [0,1]; 2) Преобразование текстовых признаков в числа; 3) Замена NaN на среднее/медиану;

Тип задания	Инструкция	Пример задания
		4) Приведение к нулевому среднему и единичной дисперсии.
Задание закрытого типа на установление последовательности (Знать)	Прочитайте текст и установите правильную последовательность этапов работы над задачей машинного обучения.	Расположите этапы работы над задачей ML в правильном порядке: 1) Обучение модели; 2) Сбор данных; 3) Оценка качества; 4) Предобработка и разведочный анализ.
Задание закрытого типа на установление последовательности (Уметь)	Прочитайте текст и установите правильную последовательность действий при кросс-валидации.	Расположите шаги кросс-валидации в правильном порядке: 1) Обучение на обучающем подмножестве; 2) Разбиение данных на K блоков; 3) Оценка на валидационном подмножестве; 4) Усреднение метрик по всем блокам.
Задание с выбором одного верного ответа (Знать)	Прочитайте текст, выберите правильный ответ и запишите аргументы.	Какая метрика используется для оценки качества регрессии? 1) Accuracy; 2) F1-score; 3) MSE; 4) ROC-AUC.
Задание с выбором одного верного ответа (Уметь)	Прочитайте текст, выберите правильный ответ и запишите аргументы.	Для задачи классификации с несбалансированными классами лучше использовать: 1) Accuracy; 2) F1-score; 3) MSE; 4) R^2 .
Задание с выбором нескольких вариантов (Знать)	Прочитайте текст, выберите все правильные ответы и запишите обоснование.	Какие из перечисленных алгоритмов относятся к методам обучения с учителем? 1) kNN; 2) Решающие деревья; 3) k-means; 4) Логистическая регрессия.
Задание с выбором нескольких вариантов (Уметь)	Прочитайте текст, выберите все правильные ответы и запишите обоснование.	Какие методы можно использовать для борьбы с переобучением? 1) Увеличение обучающей выборки; 2) Регуляризация; 3) Уменьшение сложности модели; 4) Использование всех признаков.
Задание комбинированное с выбором одного верного ответа и	Прочитайте текст, выберите правильный ответ и запишите	Что такое переобучение? 1) Модель хорошо работает на обучающих данных, но плохо на тестовых; 2) Модель плохо работает на

Тип задания	Инструкция	Пример задания
обоснованием (Знать)	развёрнутое обоснование.	всех данных; 3) Модель идеально обобщает; 4) Модель не может обучиться.
Задание комбинированное с выбором одного верного ответа и обоснованием (Уметь)	Прочитайте текст, выберите правильный ответ и запишите развёрнутое обоснование.	Какой параметр в kNN влияет на сложность модели? 1) k – число соседей; 2) Метрика расстояния; 3) Веса; 4) Размер выборки.
Задание комбинированное с выбором нескольких вариантов и обоснованием (Знать)	Прочитайте текст, выберите все правильные ответы и запишите развёрнутое обоснование.	Какие методы регуляризации используются в линейных моделях? 1) L1 (Lasso); 2) L2 (Ridge); 3) ElasticNet; 4) Dropout.
Задание комбинированное с выбором нескольких вариантов и обоснованием (Уметь)	Прочитайте текст, выберите все правильные ответы и запишите развёрнутое обоснование.	Какие библиотеки Python используются для машинного обучения? 1) Scikit-learn; 2) Pandas; 3) NumPy; 4) Matplotlib.
Задание открытого типа с развёрнутым ответом (Знать)	Прочитайте текст и запишите развёрнутый обоснованный ответ.	Объясните разницу между обучением с учителем и без учителя. Приведите примеры задач.
Задание открытого типа с развёрнутым ответом (Уметь)	Прочитайте текст и запишите развёрнутый обоснованный ответ.	Опишите этапы предобработки данных перед применением алгоритмов машинного обучения.
Задание открытого типа с развёрнутым ответом (Знать)	Прочитайте текст и запишите развёрнутый обоснованный ответ.	Что такое ROC-кривая и для чего она используется?
Задание открытого типа с развёрнутым ответом (Уметь)	Прочитайте текст и запишите развёрнутый	Напишите код на Python для загрузки набора данных, разделения на

Тип задания	Инструкция	Пример задания
	обоснованный ответ.	обучающую и тестовую выборки и обучения модели kNN.
Практическое задание (Уметь)	Выполните задание в Python и предоставьте код и результаты.	Загрузите набор данных Iris. Проведите разведочный анализ (describe, info, pairplot).
Практическое задание (Уметь)	Выполните задание в Python и предоставьте код и результаты.	Обучите модель kNN на наборе данных Wine, выберите оптимальное k с помощью кросс-валидации.
Практическое задание (Уметь)	Выполните задание в Python и предоставьте код и результаты.	Постройте дерево решений для набора данных Titanic, визуализируйте дерево, определите важность признаков.
Практическое задание (Уметь)	Выполните задание в Python и предоставьте код и результаты.	Сравните логистическую регрессию и SVM на наборе данных Breast Cancer, оцените качество моделей.