

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ОБРАЗОВАНИЯ
«СЕВЕРО-КАВКАЗСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»

МЕТОДИЧЕСКИЕ УКАЗАНИЯ

по выполнению лабораторных работ
по дисциплине «Искусственный интеллект»
для студентов направления подготовки
04.03.01 Химия

Химия окружающей среды, химическая экспертиза и экологическая
безопасность

СОДЕРЖАНИЕ

	стр
Введение	3
Лабораторная работа 1. Структурная и объектно-ориентированная технология программирования в Python	5
Лабораторная работа 2. Библиотеки NumPy и Matplotlib	12
Лабораторная работа 3. Библиотеки TensorFlow и Keras	20
Лабораторная работа 4. Подготовка входных данных для нейросети и интерпретация результата	26
Лабораторная работа 5. Одиночные нейроны и многослойный персептрон	36
Лабораторная работа 6. Энкодеры и генеративно-состязательные сети	44
Лабораторная работа 7. Сверточные нейросети	52
Лабораторная работа 8. Рекуррентные нейросети	60
Лабораторная работа 9. Нейронные сети трансформеры	69
Список литературы	76
Приложение 1. Оборудование, материалы и программное обеспечение, используемые для организации и проведения лабораторных работ	78
Приложение 2. Правила поведения в компьютерном классе (аудитории, лаборатории)	79

Введение

Цель освоения дисциплины: формирование у студентов универсальных и профессиональных компетенций, основанных на формировании цифровой культуры профессиональной деятельности.

Задачи освоения дисциплины:

1. знакомство с основами создания систем искусственного интеллекта, базовыми принципами обработки данных, возможностями и ограничениями современных нейросетей для решения задач профессиональной деятельности;
2. выработка умений анализировать задачу обработки данных, подбирать тип искусственной нейронной сети для ее решения, использовать научные и образовательные ресурсы сети Интернет в профессиональной деятельности;
3. выработка навыков владения инструментами и форматами представления информации, умения решать основные задачи обработки данных при помощи искусственных нейронных сетей на языке программирования Python, представлять данные в графическом формате, работать с поисковыми и справочными системами.

В результате освоения дисциплины «Искусственный интеллект» формируются следующие компетенции:

Код	Формулировка:
УК-1	Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач (УК- 1)
ОПК-3	Способен применять расчетно-теоретические методы для изучения свойств веществ и процессов с их участием с использованием современной вычислительной техники

Выполняя лабораторную работу, студенту следует вначале повторить все приведенные в пособии примеры, а затем выполнить задания.

Основная цель, которая ставится студенту при выполнении задания – практическое освоение всех этапов работы с цифровым инструментом разработки и отладки систем искусственного интеллекта. Лабораторные работы выполняются индивидуально в соответствии с вариантом задания. Выполнив лабораторную работу, студент оформляет отчет.

При защите лабораторной работы необходимо отвечать на вопросы и уметь пояснять работу.

Общие требования к оформлению отчета

Отчет должен быть выполнен печатным способом на одной стороне листа белой бумаги формата А4 через 1 интервал.

Цвет шрифта должен быть черным, высота букв, цифр и других знаков — не менее 1,8 мм (кегель не менее 12).

Шрифт Times New Roman.

Выравнивание – по ширине, абзац 0,75.

Текст отчета следует печатать, соблюдая следующие размеры полей: правое – 15 мм, верхнее и нижнее – 20 мм, левое – 30 мм.

Нумерация страниц – сквозная на весь отчет, в центре нижней части листа, арабской цифрой, без точки.

Титульный лист включают в общую нумерацию страниц отчета.

Номер страницы на титульном листе не проставляют

Разделы, подразделы, пункты и подпункты следует нумеровать арабскими цифрами и записывать с абзацного отступа, без точки.

Разделы, подразделы должны иметь заголовки. Пункты, как правило, заголовков не имеют. Заголовки должны четко и кратко отражать содержание разделов, подразделов.

Заголовки разделов, подразделов и пунктов следует печатать с абзацного отступа с прописной буквы без точки в конце, не подчеркивая.

Нумерация иллюстраций, таблиц, формул – сквозная или по разделам, арабскими цифрами.

Лабораторная работа 1

Тема: «Структурная и объектно-ориентированная технология программирования в Python»

Цель: «Рассмотреть основные возможности языка Python в рамках изучаемой дисциплины»

Знания, умения и владения, приобретаемые обучающимся в результате освоения темы

Знать: онлайн-среды и локальные IDE для программирования на языке Python, Каркас Python-программы. Операторы ввода-вывода, основные типы данных языка программирования Python (числа, строки, списки, множества, кортежи, словари), форматированный вывод. Индексирование и срезы списков и строк. Функции split и join. Агрегирующие функции (сумма, минимум, максимум). Управление потоком исполнения операторов (условия и циклы). Генераторы списков. Стандартные функции Python, создание подпрограмм в Python, лямбды, подключение внешних библиотек. Сохранение и загрузка файлов. Классы и объекты. Примеры объектно-ориентированных иерархий. Наследование, инкапсуляция, полиморфизм. Конструкторы. Реализация ООП в языке Python. Установка библиотеки PyQt. Каркас PyQt программы, основные виджеты (QWidget, QPushButton, QLabel, QLineEdit, QListWidget). Подключение обработчиков событий к виджетам. Компьютерная графика. Создание и подключение GUI-интерфейса в программе QtDesigner.

Уметь: создавать вычислительные программы с консольным и графическим пользовательским интерфейсом на языке программирования Python. Создавать классы и объекты в Python.

Владеть: навыками создания вычислительные программы с консольным и графическим пользовательским интерфейсом на языке программирования Python.

Теоретическая часть:

см. список рекомендованной литературы

Оборудование и материалы: см. приложение 1 методических рекомендаций по выполнению ЛР.

Техника безопасности: см. приложение 2 методических рекомендаций по выполнению ЛР.

Задания

1. Ознакомиться с теоретической частью лабораторной работы (найти ответы на контрольные вопросы).
2. Сформировать и сохранить пункты 1-4 отчета.
3. Выполнить практическую часть лабораторной работы.
 - зарегистрироваться в google colab
 - создать и переименовать новый Notebook

- ввести и отладить пробные примеры
- сохранить результат работы пробных примеров
- 4. Сформировать и сохранить пункты 5-7 отчета.
- 5. По желанию сформировать и сохранить пункты 8-9 отчета
- 6. Защитить (выслать на проверку) отчет преподавателю

Содержимое отчета по лабораторной работе

Лабораторная работа считается выполненной студентом и засчитывается преподавателем (с зачислением баллов пропорционально выполненным пунктам отчета, т. е. необязательно, чтобы отчет содержал все пункты) при предоставлении в электронном виде к защите (бумажная распечатка не требуется!!!) аккуратно оформленного файла-отчета в формате .pdf (презентация), .odt (OpenOffice, LibreOffice) или .doc (Microsoft Office), именованный по шаблону (при несоответствии работа не допускается к защите):

AI_LR1_(фамилия).pdf

Файл-отчет должен иметь следующие элементы-слайды презентации или пункты текстового отчета-документа в указанном порядке (с сохранением их номера в случае пропуска некоторых пунктов):

1. Титульный лист, содержащий название дисциплины, тему и номер лабораторной работы, ФИО студента, курс, группу.

Форму и дизайн титульного листа студент определяет самостоятельно, ориентируясь на удобство восприятия информации на титульном листе.

2. Отметка об ознакомлении с техникой безопасности (правилами поведения в компьютерном классе (аудитории, лаборатории)).

В данном пункте необходимо наличие записи «С правилами техники безопасности при выполнении лабораторной работы и поведения в компьютерном классе (аудитории, лаборатории) ознакомлен, обязуюсь выполнять _____ (ФИО), _____ дата».

3. Постановка задачи (полное условие в соответствии с выбранным вариантом).

4. Реферативный обзор информационных (интернет) источников по заданию.

В данном пункте указать по шагам:

4.1. Назначение языка программирования Python (кратко).

4.2. Основные offline среды ввода и отладки компьютерных программ на языке Python.

4.3. Основные online среды ввода и отладки компьютерных программ на языке Python.

4.4. Как установить язык программирования Python на локальную ЭВМ?

4.5. Онлайн-среда google colab для ввода и отладки программ на Python.

4.6. Регистрация и начало работы в среде google colab.

4.7. Сохранение проекта из google colab на локальную ЭВМ.

4.8. Загрузка проекта из локальной ЭВМ в google colab.

4.9. Каркас Python-программы.

- 4.10. Операторы ввода-вывода,
- 4.11. Основные типы данных языка программирования Python (числа, строки, списки, множества, кортежи, словари),
- 4.12. Форматированный вывод.
- 4.13. Индексирование и срезы списков и строк.
- 4.14. Функции split и join.
- 4.15. Агрегирующие функции (сумма, минимум, максимум).
- 4.16. Управление потоком исполнения операторов (условия и циклы).
- 4.17. Генераторы списков.
- 4.18. Стандартные функции Python,
- 4.19. Создание подпрограмм в Python,
- 4.20. Лямбды,
- 4.21. Подключение внешних библиотек.
- 4.22. Сохранение и загрузка файлов.
- 4.23. Реализация ООП в языке Python.
- 4.24. Классы и объекты.
- 4.25. Примеры объектно-ориентированных иерархий.
- 4.26. Наследование,
- 4.27. Инкапсуляция,
- 4.28. Полиморфизм.
- 4.29. Конструкторы.
- 4.30. Основные возможности библиотеки PyQt
- 4.31. Установка библиотеки PyQt
- 4.32. Каркас PyQt программы,
- 4.33. Основные виджеты PyQt (QWidget, QPushButton, QLabel, QLineEdit, QListWidget).
- 4.34. Подключение обработчиков событий к виджетам.
- 4.35. Создание и подключение GUI-интерфейса в программе QtDesigner.
- 4.36. Компьютерная графика PyQt.
- 4.37. Компьютерная графика библиотеки Turtle.

Необходимо указать информационные источники и их содержимое по заданию. Допустимо использовать текст, гиперссылки, таблицы, рисунки и т. д. любую информацию, облегчающую просмотр и защиту презентации-отчета. Разрешено использование только открытых публичных источников информации, не требующих передачи авторских прав или каких-либо иных отчислений. Рекомендуемый объем реферативного обзора — 3-10 страниц (5-10 слайдов).

5. Зарегистрироваться в системе google colab. В данном пункте сделать экранные снимки пошаговой регистрации с подписями, расшифровывающими ФИО автора (подписи реальных ФИО можно добавить на скриншот в графическом редакторе уже после процесса регистрации). Рекомендуемый объем — 2-3 страницы (2-5 слайдов)

6. Создать новый Notebook в системе google colab, и переименовать его по шаблону PNN_LR1_(фамилия).ipynb, (пример PNN_LR1_Ionisyan.ipynb).

Сделать скриншоты. Вставить скриншоты в отчет. Рекомендуемый объем — 1-2 страницы (2-5 слайдов)

7. Ввести и отладить в системе google colab программу решения основной задачи лабораторной работы, выбранной по варианту:

1. Напишите функцию `atomic_mass(symbol)`, которая по символу химического элемента (например, "Na") возвращает его атомную массу (используйте словарь внутри функции для 10 элементов: H, He, Li, Be, B, C, N, O, F, Ne). В основной программе запросите у пользователя символ элемента и выведите его атомную массу, используя эту функцию.

2. Создайте функцию `molar_mass_binary(formula)`, которая принимает строку с формулой бинарного соединения без индексов (например, "NaCl" или "CO"), и возвращает молярную массу. Считайте, что соединение состоит ровно из двух элементов. Атомные массы заархивируйте внутри функции для 6 пар: NaCl, KCl, CaO, MgO, CO, CO₂ (учтите, что CO₂ имеет индекс 2, но это частный случай — обработайте его отдельно через проверку строки). Основная программа: ввод формулы → вывод молярной массы.

3. Напишите функцию `mass_fraction(element_symbol, compound_formula)`, где `compound_formula` — это строка вида "H₂O" или "CO₂" (индексы только однозначные). Функция возвращает массовую долю указанного элемента (число от 0 до 1). Внутри функции реализуйте парсинг формулы и расчет. Атомные массы установите константами для H, C, N, O, Na, Cl, Ca. Основная программа: запрос элемента и формулы → вывод массовой доли.

4. Реализуйте функцию `compare_molar_mass(formula1, formula2)`, которая принимает две строки с формулами простых веществ или бинарных соединений (без индексов, кроме CO₂ и H₂ O как исключения). Функция возвращает строку: "Первое тяжелее", "Второе тяжелее" или "Одинаковы". Значения атомных массы (5-7 элементов) установите константами внутри функции. Основная программа: ввод двух формул → вывод результата сравнения.

5. Создайте функцию `extract_elements(formula)`, которая принимает строку с химической формулой, содержащей заглавные и строчные буквы (например, "KMnO₄"), и возвращает список символов элементов в том порядке, как они встречаются, без повторений и без учета индексов. Основная программа: ввод формулы → вызов функции → вывод списка элементов.

6. Напишите функцию `moles_from_mass(mass, molar_mass)`, которая возвращает количество вещества (в молях). Основная программа: запросить у пользователя массу вещества (г) и его молярную массу (г/моль), вызвать функцию и вывести количество молей.

7. Создайте функцию `reaction_moles(coeff_A, coeff_B, moles_A)`, которая для реакции $\text{coeff_A} * \text{A} \rightarrow \text{coeff_B} * \text{B}$ возвращает количество молей

вещества В. Основная программа: ввод коэффициентов А и В и количества молей А → вывод молей В.

8. Реализуйте функцию `limiting_reactant(moles_A, moles_B, coeff_A, coeff_B)`, которая возвращает строку "А", "В" или "оба в стехиометрии" в зависимости от того, какой реагент лимитирующий. Основная программа: ввод молей и коэффициентов → вывод результата.

9. Напишите функцию `gas_volume_STP(moles)`, которая возвращает объем газа (в литрах) при нормальных условиях (22.4 л/моль). Основная программа: запросить количество молей → вывести объем.

10. Создайте функцию `dilution_volume(C1, V1, C2)`, которая возвращает объем растворителя (в литрах), который нужно добавить к V1 литрам раствора концентрации C1 моль/л, чтобы получить концентрацию C2 моль/л. Основная программа: ввод C1, V1, C2 → вывод объема растворителя.

11. Напишите функцию `oxide_class(formula)`, которая по формуле оксида (например, "CO2", "CaO", "Al2O3") возвращает "кислотный", "основной" или "амфотерный". Используйте внутри словарь с 8 примерами оксидов. Основная программа: ввод формулы → вывод класса.

12. Создайте функцию `acid_or_base(formula)`, которая для формулы (например, "HCl", "H2SO4", "NaOH", "Ca(OH)2") возвращает "кислота" или "основание". Критерий: если формула начинается с 'H' и не содержит "(OH)" — кислота; если содержит "(OH)" — основание. Для простоты обработайте только эти 4 примера. Основная программа: ввод формулы → вывод типа.

13. Реализуйте функцию `neutralization_salt(acid, base)`, которая принимает строки с формулами кислоты и основания (из ограниченного набора: "HCl", "HNO3", "NaOH", "KOH") и возвращает формулу образующейся соли. Используйте внутри словарь соответствий. Основная программа: ввод кислоты и основания → вывод соли.

14. Напишите функцию `salt_formula(metal, residue)`, где metal — строка с русским названием металла ("натрий", "кальций", "алюминий"), а residue — строка с названием кислотного остатка ("хлорид", "сульфат", "нитрат"). Функция возвращает формулу соли (например, "NaCl"). Используйте внутри словари. Основная программа: ввод названий → вывод формулы.

15. Создайте функцию `classify_inorganic(formula)`, которая для строки вида "CO2", "H2SO4", "NaOH", "NaCl" возвращает класс вещества: "оксид", "кислота", "основание" или "соль". Правила: если ровно два элемента и есть кислород — оксид; если начинается с H и не содержит OH — кислота; если содержит OH — основание; иначе — соль. Основная программа: ввод формулы → вывод класса.

16. Напишите функцию `period_by_number(Z)`, которая по атомному номеру (целое число от 1 до 36) возвращает номер периода (1–4). Используйте список границ периодов: 2, 10, 18, 36. Основная программа: ввод Z → вывод периода.

17. Создайте функцию `group_by_symbol(symbol)`, которая для символов первых 20 элементов (H–Ca) возвращает номер группы (1,2,13–18). Используйте внутри словарь. Основная программа: ввод символа → вывод группы.

18. Реализуйте функцию `is_halogen(symbol)`, которая возвращает `True`, если элемент является галогеном (F, Cl, Br, I, At), иначе `False`.

Основная программа: ввод символа → вывод "Да" или "Нет".

19. Напишите функцию `next_noble_gas(Z)`, которая принимает атомный номер и возвращает символ следующего инертного газа с бóльшим номером. Используйте список пар: 2→He, 10→Ne, 18→Ar, 36→Kr. Если Z уже больше 36 — вернуть "Нет". Основная программа: ввод Z → вывод символа.

20. Создайте функцию `electron_shells(Z)`, которая для атомного номера от 1 до 20 возвращает строку с количеством электронов на оболочках через запятую (например, для 11 → "2,8,1"). Установите константами распределение для всех 20 элементов внутри функции. Основная программа: ввод Z → вывод электронной конфигурации.

21. Напишите функцию `mass_for_molarity(C, V, molar_mass)`, возвращающую массу вещества (в граммах), необходимую для приготовления V литров раствора с молярной концентрацией C (моль/л). Основная программа: ввод C , V и молярной массы → вывод массы.

22. Создайте функцию `molarity_from_mass(mass, molar_mass, V)`, возвращающую молярность раствора, полученного растворением $mass$ граммов вещества молярной массы $molar_mass$ в объеме V литров. Основная программа: ввод массы, молярной массы, объема → вывод молярности.

23. Реализуйте функцию `mix_concentration(C1, V1, C2, V2)`, которая возвращает концентрацию раствора после смешения $V1$ литров концентрации $C1$ с $V2$ литрами концентрации $C2$ (вещество одно и то же). Основная программа: ввод $C1$, $V1$, $C2$, $V2$ → вывод итоговой концентрации.

24. Напишите функцию `ppm(mg_solute, kg_solution)`, которая возвращает концентрацию в миллионных долях (ppm) для раствора, содержащего mg_solute миллиграммов растворенного вещества в $kg_solution$ килограммах раствора. Основная программа: ввод mg и kg → вывод ppm.

25. Создайте функцию `normality(molarity, formula)`, где `formula` — строка с формулой кислоты или основания ("HCl", "H2SO4", "NaOH", "Ca(OH)2"). Функция возвращает нормальность, умножая молярность на фактор эквивалентности (1 для HCl и NaOH, 2 для H2SO4 и Ca(OH)2). Основная программа: ввод молярности и формулы → вывод нормальности.

26. Напишите функцию `heat_water(mass_kg, delta_T)`, которая возвращает количество теплоты (в джоулях) для нагрева воды массой $mass_kg$ килограмм на $delta_T$ градусов Цельсия. Удельная теплоемкость воды 4200 Дж/(кг·°C). Основная программа: ввод массы и ΔT → вывод теплоты.

27. Создайте функцию `enthalpy_sign(delta_H)`, которая принимает число — изменение энтальпии (в кДж), и возвращает строку "Эндотермическая", если $delta_H > 0$, или "Экзотермическая", если $delta_H < 0$. Основная программа: ввод ΔH → вывод типа реакции.

28. Реализуйте функцию `final_temp_mix(m1, T1, m2, T2)`, которая возвращает конечную температуру (в °C) после смешения двух порций воды с массами `m1`, `m2` (в кг) и температурами `T1`, `T2` (в °C). Теплоемкость воды считайте одинаковой и сократите. Основная программа: ввод масс и температур → вывод конечной температуры.

29. Напишите функцию `joules_to_calories(J)`, которая переводит джоули в калории (1 кал = 4.184 Дж). Результат округлите до 2 знаков. Основная программа: ввод энергии в джоулях → вывод в калориях.

30. Создайте функцию `reaction_heat(moles, delta_H_molar)`, которая возвращает общее количество теплоты (в кДж) для реакции, где участвует `moles` молей вещества, а молярная теплота реакции равна `delta_H_molar` (кДж/моль). Основная программа: ввод молей и молярной теплоты → вывод общей теплоты (с указанием знака).

Сделать скриншоты результата работы примера и сохранить их в отчете. Рекомендуемый объем — 1-2 страницы (2-5 слайдов)

8. Необязательный пункт отчета для получения повышенных баллов. Дополнительные фотографии процесса выполнения лабораторной работы (ознакомление с методичкой, литературой, чтение сайтов из интернета, написание отчета — 5-10 фотографий (5-10 слайдов) с пояснениями кто на фото и что именно делает студент.

9. Необязательный пункт отчета для получения повышенных баллов. Публичная защита отчета по лабораторной работе. В данном пункте привести ссылку на Rutube-видео или иной публичный видео-ресурс, на котором записана защита-выступление отчета или презентации по лабораторной работе лично студентом. Рекомендуемое время видеоролика — 5-10 минут.

Критерии оценивания отчета:

1. Выполнены пункты 1-4 отчета - оценка «удовлетворительно».
2. Выполнены пункты 1-6 отчета - оценка «хорошо».
3. Выполнены пункты 1-7 отчета - оценка «отлично».
4. Отсутствие пункта 4 отчета понижает оценку на 1 балл. (Можно заменить по желанию студента пункт 4 устными ответами на все контрольные вопросы без понижения оценки).
5. Отсутствие нумерации пунктов отчета и их названий понижает оценку на 1 балл. (Можно не детализировать в названии пункта, что именно должно быть в данном пункте отчета, но само название должно быть!)
6. «Перемешивание» (идут не по порядку) пунктов отчета понижает оценку на 1 балл.
7. Наличие любого из пунктов 8, 9 или их вместе повышает оценку на 1 балл (но не выше «отлично»).

Лабораторная работа 2

Тема: «Библиотеки NumPy и Matplotlib»

Цель: «Рассмотреть основные возможности библиотек NumPy и Matplotlib языка Python в рамках изучаемой дисциплины»

Знания, умения и владения, приобретаемые обучающимся в результате освоения темы

Знать: Назначение библиотеки NumPy, методы установка библиотеки NumPy. Массивы и многомерные массивы, типы данных для элементов массива. Создание массивов (из списка, из нулей, из единиц, из случайных чисел). Создание обычных списков из массивов методом toList. Атрибуты массивов (размерность, размер, форма, оси). Изменение формы массива. Сортировка массива. Индексирование и срезы массивов. Высокопроизводительные базовые операции над массивами (математические операции с скалярами, операции с другими массивами - сложение, вычитание, перемножение массивов, обработка стандартными функциями, битовые маски), агрегирующие функции (сумма, произведение, среднее, стандартное отклонение, минимум, максимум, индексы минимума и максимума для элементов массива). Булева индексация массивов. Сохранение и загрузка массивов. Конкатенация массивов. Объединение массивов по вертикали и по горизонтали. Назначение библиотеки Matplotlib, Методы установки библиотеки Matplotlib, загрузка и сохранение изображений, доступ к отдельным пикселям изображения, рисование графиков одномерных функций, рисование множеств точек (кластеров), создание легенды графика и подписывание изображения, совмещение нескольких графиков на одном полотне, совмещение нескольких изображений на одном полотне, рисование двумерных графиков.

Уметь: создавать вычислительные программы с отображением результатов вычислительных экспериментов в виде графиков и изображений с использованием библиотек NumPy и Matplotlib на языке программирования Python..

Владеть: навыками создания вычислительных программ на языке программирования Python, моделирующих работу искусственных нейронных сетей.

Теоретическая часть:

см. список рекомендованной литературы

Оборудование и материалы: см. приложение 1 методических рекомендаций по выполнению ЛР.

Техника безопасности: см. приложение 2 методических рекомендаций по выполнению ЛР.

Задания

1. Ознакомиться с теоретической частью лабораторной работы (найти

ответы на контрольные вопросы).

2. Сформировать и сохранить пункты 1-4 отчета.
3. Выполнить практическую часть лабораторной работы.
 - войти в созданную в ЛР1 личную учетную запись google colab
 - создать и переименовать новый Notebook
 - ввести и отладить пробные примеры
 - сохранить результат работы пробных примеров
4. Сформировать и сохранить пункты 5-7 отчета.
5. По желанию сформировать и сохранить пункты 8-9 отчета
6. Защитить (выслать на проверку) отчет преподавателю

Содержимое отчета по лабораторной работе

Лабораторная работа считается выполненной студентом и засчитывается преподавателем (с зачислением баллов пропорционально выполненным пунктам отчета, т. е. необязательно, чтобы отчет содержал все пункты) при предоставлении в электронном виде к защите (бумажная распечатка не требуется!!!) аккуратно оформленного файла-отчета в формате .pdf (презентация), .odt (OpenOffice, LibreOffice) или .doc (Microsoft Office), именованный по шаблону (при несоответствии работа не допускается к защите):

AI_LR2_(фамилия).pdf

Файл-отчет должен иметь следующие элементы-слайды презентации или пункты текстового отчета-документа в указанном порядке (с сохранением их номера в случае пропуска некоторых пунктов):

1. Титульный лист, содержащий название дисциплины, тему и номер лабораторной работы, ФИО студента, курс, группу.

Форму и дизайн титульного листа студент определяет самостоятельно, ориентируясь на удобство восприятия информации на титульном листе.

2. Отметка об ознакомлении с техникой безопасности (правилами поведения в компьютерном классе (аудитории, лаборатории)).

В данном пункте необходимо наличие записи «С правилами техники безопасности при выполнении лабораторной работы и поведения в компьютерном классе (аудитории, лаборатории) ознакомлен, обязуюсь выполнять _____ (ФИО), _____ дата».

3. Постановка задачи (полное условие в соответствии с выбранным вариантом).

4. Реферативный обзор информационных (интернет) источников по заданию.

4.1. Назначение библиотеки NumPy,

4.2. методы установка библиотеки NumPy.

4.3. Массивы и многомерные массивы,

4.4. типы данных для элементов массива.

4.5. Создание массивов (из списка, из нулей, из единиц, из случайных чисел).

4.6. Создание обычных списков из массивов методом toList.

4.7. Атрибуты массивов (размерность, размер, форма, оси).

- 4.8. Изменение формы массива.
- 4.9. Сортировка массива.
- 4.10. Индексирование и срезы массивов.
- 4.11. Высокопроизводительные базовые операции над массивами (математические операции с скалярами, операции с другими массивами - сложение, вычитание, перемножение массивов, обработка стандартными функциями, битовые маски),
- 4.12. агрегирующие функции (сумма, произведение, среднее, стандартное отклонение, минимум, максимум, индексы минимума и максимума для элементов массива).
- 4.13. Булева индексация массивов.
- 4.14. Сохранение и загрузка массивов.
- 4.15. Конкатенация массивов.
- 4.16. Объединение массивов по вертикали и по горизонтали.
- 4.17. Назначение библиотеки Matplotlib,
- 4.18. Методы установки библиотеки Matplotlib,
- 4.19. загрузка и сохранение изображений,
- 4.20. доступ к отдельным пикселям изображения,
- 4.21. рисование графиков одномерных функций,
- 4.22. рисование множеств точек (кластеров),
- 4.23. создание легенды графика и подписывание изображения, совмещение
- 4.24. нескольких графиков на одном полотне,
- 4.25. совмещение нескольких изображений на одном полотне,
- 4.26. рисование двумерных графиков.

Необходимо указать по порядку информационные источники и их содержимое по заданию. Допустимо использовать текст, гиперссылки, таблицы, рисунки и т. д. любую информацию, облегчающую просмотр и защиту презентации-отчета. Разрешено использование только открытых публичных источников информации, не требующих передачи авторских прав или каких-либо иных отчислений. Если информация по какому-либо из разделов реально отсутствует в публичном доступе или засекречена, то напечатать «информация засекречена» с указанием ведомства или организации, засекретившей данную информацию. Рекомендуемый объем реферативного обзора — 3-10 страниц (5-10 слайдов).

5. Войти в личную учетную запись в системе google colab, созданную в ЛР1. В данном пункте сделать экранные снимки процесса успешного входа с подписями, расшифровывающими ФИО автора (подписи реальных ФИО можно добавить на скриншот в графическом редакторе уже после процесса регистрации). Рекомендуемый объем — 1-2 страницы (1-2 слайда)

6. Создать новый Notebook в системе google colab, и переименовать его по шаблону: AI_LR2_(фамилия).ipynb

Сделать скриншоты. Вставить скриншоты в отчет. Рекомендуемый объем — 1-2 страниц (1-2 слайда)

7. Ввести и отладить в системе google colab программу решения основной задачи лабораторной работы, выбранной по варианту:

1. Закон Бугера-Ламберта-Бера (массивы)

Создайте массив концентраций $C = \text{np.array}([0.01, 0.02, 0.05, 0.10, 0.20])$ моль/л. Задайте коэффициент экстинкции $\varepsilon = 5000$ л/(моль·см) и длину кюветы $l = 1$ см. Вычислите оптическую плотность $A = \varepsilon \cdot l \cdot C$ для каждой концентрации. Выведите на экран пары (C, A) .

2. Построение калибровочного графика

Используйте данные из задачи 1 (C и A). Постройте точечный график (`marker='o'`) зависимости $A(C)$. Подпишите оси: "Концентрация, моль/л" и "Оптическая плотность". Добавьте заголовок "Калибровочный график".

3. Линейная регрессия (вручную через numpy)

Для данных из задачи 1.1 вычислите коэффициенты линейной регрессии $A = k \cdot C$ с помощью `np.polyfit(C, A, 1)`. Выведите значение k (должно быть близко к $\varepsilon \cdot l$). Постройте на том же графике исходные точки и линию регрессии.

4. Работа с зашумленными данными

Создайте массив $C = \text{np.linspace}(0, 0.2, 10)$ и вычислите идеальную $A = 5000 \cdot C$. Добавьте шум: $A_{\text{noisy}} = A + \text{np.random.normal}(0, 0.02, \text{len}(C))$. Постройте оба графика: идеальный (сплошная линия) и экспериментальный (точки).

5. Вычисление неизвестной концентрации по калибровке

Постройте калибровку по точкам из задачи 1.1. Пусть оптическая плотность неизвестного образца $A_x = 0.35$. Используйте коэффициенты регрессии, чтобы вычислить концентрацию C_x по формуле $C_x = (A_x - \text{intercept}) / \text{slope}$. Выведите значение.

6. Моделирование простого спектра поглощения

Создайте массив длин волн $\text{wavelength} = \text{np.linspace}(400, 700, 200)$ нм. Смоделируйте спектр поглощения в виде гауссова пика:

$$A = \text{np.exp}-((\text{wavelength} - 550)**2) / (2 * 30**2))$$

Постройте график зависимости $A(\lambda)$. Подпишите оси: "Длина волны, нм" и "Поглощение, отн. ед."

7. Два перекрывающихся пика

Создайте два гауссовых пика с центрами 500 нм и 600 нм, каждый с амплитудой 1. Постройте их сумму. Подпишите график "Спектр смеси двух компонентов".

8. Поиск максимума спектра (argmax)

Для данных из задачи 6 найдите индекс максимума с помощью `np.argmax(A)`, затем выведите длину волны максимума поглощения λ_{max} . Подпишите максимум стрелкой на графике с помощью `plt.annotate()`.

9. Нормировка спектра

Создайте спектр из задачи 6. Нормируйте его так, чтобы максимум стал равен 1. Постройте исходный и нормированный спектр на одном графике разными цветами.

10. Интегрирование пика (площадь под кривой)

Для спектра из задачи 6 вычислите площадь под кривой с помощью `np.trapz(A, wavelength)`. Выведите значение. Объясните, какой физический смысл может иметь эта площадь (например, общее поглощение).

11. Построение кривой охлаждения

Создайте массив времени `t = np.linspace(0, 60, 100)` минут. Задайте закон охлаждения Ньютона: $T = 20 + 80 * \text{np.exp}(-0.05 * t)$, где T — температура в °C. Постройте график $T(t)$. Подпишите оси.

12. Кинетика реакции первого порядка

Для реакции $A \rightarrow B$ с константой скорости $k = 0.1 \text{ мин}^{-1}$ постройте график зависимости концентрации A от времени: $C_A = C0 * \text{np.exp}(-k * t)$, где $C0 = 1 \text{ моль/л}$, $t = \text{np.linspace}(0, 50, 100)$. На том же графике постройте концентрацию продукта B : $C_B = C0 * (1 - \text{np.exp}(-k * t))$.

13. Время полупревращения (визуально)

Для данных из задачи 12 найдите время, когда C_A снизится до $C0/2$. Сделайте это с помощью `np.where()` или визуально добавьте горизонтальную и вертикальную линии (`plt.axhline`, `plt.axvline`).

14. Аррениусовская зависимость

Создайте массив температур `T_K = np.array([273, 293, 313, 333, 353])` К. Задайте константу скорости по Аррениусу: $k = A * \text{np.exp}(-Ea/(R*T))$, где $A = 1e10$, $Ea = 50000 \text{ Дж/моль}$, $R = 8.314$. Постройте график $k(T)$. На втором графике (`subplot`) постройте зависимость $\ln(k)$ от $1/T$ — она должна быть линейной.

15. Сравнение двух кинетических кривых

Постройте на одном графике две кривые распада для $k1 = 0.05$ и $k2 = 0.15 \text{ мин}^{-1}$. Добавьте легенду, разные цвета и стили линий (например, сплошная и пунктирная). Подпишите оси: "Время, мин" и "Концентрация A, моль/л".

16. Плавление и кипение (кусочно-линейная функция)

Смоделируйте нагревание 1 кг воды от -10°C до 110°C через 0°C и 100°C с плато. Создайте массив температур и вычислите подведенное тепло Q (проще всего через `np.select` или цикл). Постройте график $Q(T)$.

17. Построение диаграммы растворимости

Задайте зависимость растворимости соли от температуры: $S = 30 + 0.5 * T + 0.01 * T^2$, где T от 0 до 100°C . Постройте график. Закрасьте область ниже кривой (`plt.fill_between`).

18. Титрование (кривая pH)

Смоделируйте кривую титрования сильной кислоты сильным основанием: создайте массив объема титранта V от 0 до 20 мл, вычислите pH (упрощенно: до точки эквивалентности $\text{pH} = -\log(C_{\text{кислоты}})$, после — $\text{pH} = 14 + \log(C_{\text{основания}})$). Постройте график $\text{pH}(V)$. Отметьте точку эквивалентности.

19. Сравнение спектров двух веществ

Создайте массив длин волн $\text{lam} = \text{np.linspace}(400, 700, 100)$ нм. Задайте два гауссовых пика: вещество А с центром 520 нм, ширина 25 нм, вещество В с центром 620 нм, ширина 30 нм. Постройте оба спектра на одном графике разными цветами. Добавьте легенду. Подпишите оси.

20. Влияние концентрации на спектр

Для вещества с пиком при 550 нм (ширина 40 нм) постройте три спектра для концентраций 0.5, 1.0 и 2.0 отн. ед. (просто умножьте исходный спектр на концентрацию). Используйте разные стили линий: сплошная, пунктирная, штрих-пунктирная. Добавьте легенду с указанием концентраций.

21. Температурная зависимость растворимости (два соединения)

Задайте две функции растворимости (г/100г воды) от температуры T (°C) от 0 до 100:

- Соль 1: $S1 = 20 + 0.3 \cdot T$

- Соль 2: $S2 = 50 * (1 - \text{np.exp}(-0.05 \cdot T))$

Постройте оба графика. Найдите температуру, где растворимости равны (визуально или с помощью $\text{np.argmin}(\text{np.abs}(S1-S2))$). Отметьте эту точку маркером.

22. Множественные кинетические кривые (разные k)

Для реакции первого порядка $C = C0 * \text{exp}(-k \cdot t)$ постройте семейство кривых для $k = 0.02, 0.05, 0.10, 0.20 \text{ мин}^{-1}$ на одном графике. Используйте циклическую смену цветов (можно задать список вручную). Добавьте легенду.

23. Subplot: спектр и его производная

Для спектра из двух перекрывающихся пиков (центры 500 и 600 нм, амплитуды 0.8 и 1.0) создайте figure с двумя subplot (вертикально):

- Верхний: сам спектр $A(\lambda)$

- Нижний: производная спектра $dA = \text{np.gradient}(A, \text{wavelength})$

Подпишите оба графика. Объясните, почему производная помогает находить скрытые пики.

24. Сглаживание шумного спектра (скользящее среднее)

Создайте «чистый» спектр: гауссов пик с центром 550 нм, ширина 20 нм. Добавьте шум: $\text{noise} = \text{np.random.normal}(0, 0.05, \text{len}(\text{wavelength}))$. Реализуйте функцию $\text{moving_average}(\text{data}, \text{window_size}=5)$, которая усредняет по соседним точкам. Постройте три графика: чистый, зашумленный, сглаженный.

25. Вычитание фона (базовой линии)

Создайте спектр с линейным растущим фоном: $A = \text{np.exp}(-((\text{lam}-550)/30)**2) + 0.002 \cdot (\text{lam}-400)$. Постройте исходный спектр. Вычтите линейный фон, оценив его с помощью $\text{np.polyfit}(\text{lam}, A, 1)$. Постройте исправленный спектр на том же графике.

26. Поиск пиков выше порога

Создайте спектр с тремя гауссовыми пиками (центры 450, 550, 650 нм) на небольшом фоне. Найдите все точки, где $A > 0.5$ (порог). Выведите длины

волн этих точек. Отметьте их на графике вертикальными пунктирными линиями (plt.axvline).

27. Преобразование шкалы: нм → эВ

Для массива длин волн $\text{lam_nm} = \text{np.linspace}(200, 800, 100)$ пересчитайте энергию фотона в электронвольтах: $E_eV = 1239.84 / \text{lam_nm}$. Постройте спектр поглощения (гауссов пик с центром 500 нм) в координатах E_eV . Подпишите ось X как "Энергия, эВ". Обратите внимание, что шкала сжимается и переворачивается.

28. Нормализация и центрирование данных

Создайте массив измерений pH в ходе титрования (смоделируйте S-образную кривую). Нормализуйте данные так, чтобы минимум стал 0, максимум стал 1. Центрируйте (вычтите среднее). Постройте исходные и преобразованные данные рядом (subplot 1x2).

29. Радиоактивный распад двух изотопов одновременно

Пусть изотоп A распадается в изотоп B с $k_1 = 0.1 \text{ мин}^{-1}$, а B распадается в стабильный C с $k_2 = 0.05 \text{ мин}^{-1}$. Смоделируйте численно (или аналитически, если знаете решение) концентрации $A(t)$, $B(t)$, $C(t)$ для t от 0 до 60 мин. Постройте все три кривые на одном графике. Упрощенно: можно решить систему через экспоненты или просто показать затухание A и рост/спад B.

30. Смещение двух растворов разной температуры (тепловой баланс)

Создайте сетку объемов $V_1 = \text{np.linspace}(0, 1, 50)$ и $V_2 = 1 - V_1$. Пусть $T_1 = 80^\circ\text{C}$, $T_2 = 20^\circ\text{C}$. Вычислите конечную температуру $T_final = (V_1 * T_1 + V_2 * T_2) / (V_1 + V_2)$. Постройте график $T_final(V_1)$. Найдите, при каком объеме V_1 температура будет 50°C (с помощью np.interp или np.where).

31. Зависимость скорости реакции от pH (колоколообразная кривая)

Многие ферменты (и катализаторы) работают в узком диапазоне pH. Смоделируйте: $v = v_max * \text{np.exp}(-((\text{pH} - \text{pH_opt})^2) / (2 * \text{sigma}^2))$, где $\text{pH_opt} = 7.0$, $\text{sigma} = 1.5$. Постройте график $v(\text{pH})$ для pH от 3 до 11. Найдите, при каких pH скорость падает вдвое (визуально).

32. Построение изотермы адсорбции Ленгмюра

Изотерма: $\theta = (K * C) / (1 + K * C)$, где θ — степень заполнения поверхности, C — концентрация (моль/л), $K = 2 \text{ л/моль}$. Создайте массив $C = \text{np.logspace}(-2, 1, 50)$ (от 0.01 до 10). Постройте график $\theta(C)$ в полулогарифмической шкале: plt.semilogx . Подпишите оси. Отметьте точку, где $\theta = 0.5$ (найдите C , решив уравнение: $C = 1/K$).

33. Визуализация распределения Больцмана по уровням энергии

Для трех уровней энергии $E = [0, 1000, 2000] \text{ Дж/моль}$ (или см^{-1}) при $T = 300 \text{ К}$ вычислите заселенности: $n_i = g_i * \text{np.exp}(-E_i / (R * T))$, где все $g_i = 1$, $R = 8.314$. Нормируйте на сумму. Постройте столбчатую диаграмму (plt.bar) с подписями уровней. Покажите, как меняются заселенности при $T = 1000 \text{ К}$ (два графика рядом).

Сделать скриншоты результата работы примера и сохранить их в отчете. Рекомендуемый объем — 2-3 страницы (2-5 слайдов)

8. Необязательный пункт отчета для получения повышенных баллов. Дополнительные фотографии процесса выполнения лабораторной работы (ознакомление с методичкой, литературой, чтение сайтов из интернета, написание отчета — 5-10 фотографий (5-10 слайдов) с пояснениями кто на фото и что именно делает студент.

9. Необязательный пункт отчета для получения повышенных баллов. Публичная защита отчета по лабораторной работе.

В данном пункте привести ссылку на Rutube-видео или иной публичный видео-ресурс, на котором записана защита-выступление отчета или презентации по лабораторной работе лично студентом. Рекомендуемое время видеоролика — 5-10 минут.

Критерии оценивания отчета:

1. Выполнены пункты 1-4 отчета - оценка «удовлетворительно».
2. Выполнены пункты 1-6 отчета - оценка «хорошо».
3. Выполнены пункты 1-7 отчета - оценка «отлично».
4. Отсутствие пункта 4 отчета понижает оценку на 1 балл. (Можно заменить по желанию студента пункт 4 устными ответами на все контрольные вопросы без понижения оценки).
5. Отсутствие нумерации пунктов отчета и их названий понижает оценку на 1 балл. (Можно не детализировать в названии пункта, что именно должно быть в данном пункте отчета, но само название должно быть!)
6. «Перемешивание» (идут не по порядку) пунктов отчета понижает оценку на 1 балл.
7. Наличие любого из пунктов 8, 9 или их вместе повышает оценку на 1 балл (но не выше «отлично»).

Лабораторная работа 3

Тема: «Библиотеки TensorFlow и Keras»

Цель: «Рассмотреть основные возможности библиотек TensorFlow и Keras в рамках изучаемой дисциплины»

Знания, умения и владения, приобретаемые обучающимся в результате освоения темы

Знать: Основные типы современных искусственных нейронных сетей. Обучение ИНС. Датасеты. Установка библиотек TensorFlow и Keras. Возможности библиотеки TensorFlow и Keras для создания и обучения искусственных нейронных сетей, создание одиночных нейронов, многослойных персептронов, энкодеров, рекуррентных нейронных сетей, сверточных нейронных сетей. Обучение искусственных нейронных сетей библиотекой Keras. Оценка качества обучения. Визуализация процесса обучения и работы искусственных нейронных сетей.

Уметь: .

Владеть: навыками создания вычислительных программ на языке программирования Python, моделирующих работу искусственных нейронных сетей.

Теоретическая часть:

см. список рекомендованной литературы

Оборудование и материалы: см. приложение 1 методических рекомендаций по выполнению ЛР.

Техника безопасности: см. приложение 2 методических рекомендаций по выполнению ЛР.

Задания

1. Ознакомиться с теоретической частью лабораторной работы (найти ответы на контрольные вопросы).
2. Сформировать и сохранить пункты 1-4 отчета.
3. Выполнить практическую часть лабораторной работы.
 - войти в созданную в ЛР1 личную учетную запись google colab
 - создать и переименовать новый Notebook
 - ввести и отладить пробные примеры
 - сохранить результат работы пробных примеров
4. Сформировать и сохранить пункты 5-7 отчета.
5. По желанию сформировать и сохранить пункты 8-9 отчета
6. Защитить (выслать на проверку) отчет преподавателю

Содержимое отчета по лабораторной работе

Лабораторная работа считается выполненной студентом и засчитывается преподавателем (с зачислением баллов пропорционально

выполненным пунктам отчета, т. е. необязательно, чтобы отчет содержал все пункты) при предоставлении в электронном виде к защите (бумажная распечатка не требуется!!!) аккуратно оформленного файла-отчета в формате .pdf (презентация), .odt (OpenOffice, LibreOffice) или .doc (Microsoft Office), именованный по шаблону (при несоответствии работа не допускается к защите):

AI_LR3_(фамилия).pdf

Файл-отчет должен иметь следующие элементы-слайды презентации или пункты текстового отчета-документа в указанном порядке (с сохранением их номера в случае пропуска некоторых пунктов):

1. Титульный лист, содержащий название дисциплины, тему и номер лабораторной работы, ФИО студента, курс, группу.

Форму и дизайн титульного листа студент определяет самостоятельно, ориентируясь на удобство восприятия информации на титульном листе.

2. Отметка об ознакомлении с техникой безопасности (правилами поведения в компьютерном классе (аудитории, лаборатории)).

В данном пункте необходимо наличие записи «С правилами техники безопасности при выполнении лабораторной работы и поведения в компьютерном классе (аудитории, лаборатории) ознакомлен, обязуюсь выполнять _____ (ФИО), _____ дата».

3. Постановка задачи (полное условие в соответствии с выбранным вариантом).

4. Реферативный обзор информационных (интернет) источников по заданию.

4.1. Основные типы современных искусственных нейронных сетей.

4.2. Обучение ИНС.

4.3. Датасеты.

4.4. Установка библиотек TensorFlow и Keras.

4.5. Возможности библиотеки TensorFlow и Keras для создания и обучения искусственных нейронных сетей,

4.6. создание одиночных нейронов (кратко),

4.7. создание многослойных персептронов (кратко),

4.8. создание энкодеров (кратко),

4.9. создание рекуррентных нейронных сетей (кратко),

4.10. создание сверточных нейронных сетей (кратко).

4.11. функция активации

4.12. функция потерь

4.13. обучение с подкреплением

4.14. обратное распространение ошибки

4.15. автоматическое дифференцирование

4.16. оптимизаторы

4.17. регуляризация

4.18. переобучение

4.19. Обучение искусственных нейронных сетей библиотекой Keras.

4.20. Оценка качества обучения.

4.21. метрики качества обучения

4.22. Визуализация процесса обучения и работы искусственных нейронных сетей (кратко).

Необходимо указать по порядку информационные источники и их содержимое по заданию. Допустимо использовать текст, гиперссылки, таблицы, рисунки и т. д. любую информацию, облегчающую просмотр и защиту презентации-отчета. Разрешено использование только открытых публичных источников информации, не требующих передачи авторских прав или каких-либо иных отчислений. Если информация по какому-либо из разделов реально отсутствует в публичном доступе или засекречена, то напечатать «информация засекречена» с указанием ведомства или организации, засекретившей данную информацию. Рекомендуемый объем реферативного обзора — 3-10 страниц (5-10 слайдов).

5. Войти в личную учетную запись в системе google colab, созданную в ЛР1. В данном пункте сделать экранные снимки процесса успешного входа с подписями, расшифровывающими ФИО автора (подписи реальных ФИО можно добавить на скриншот в графическом редакторе уже после процесса регистрации). Рекомендуемый объем — 1-2 страницы (1-2 слайда)

6. Создать новый Notebook в системе google colab, и переименовать его по шаблону: AI_LR3_(фамилия).ipynb
Сделать скриншоты. Вставить скриншоты в отчет. Рекомендуемый объем — 1-2 страниц (1-2 слайда)

7. Ввести и отладить в системе google colab программу решения основной задачи лабораторной работы, выбранной по варианту:

1. Создайте датасет из 100 точек: концентрация вещества C (0–1 моль/л) → оптическая плотность $A = 0.5 \cdot C + \text{шум}$. Обучите одиночный нейрон без функции активации (linear) предсказывать A по C . Выведите итоговые веса (они должны быть близки к 0.5).

2. Сгенерируйте бинарные данные: pH раствора (4–10) → кислая/щелочная среда (0 или 1). Обучите одиночный нейрон с сигмоидной активацией (бинарная классификация). Выведите точность на тестовой выборке.

3. Создайте простой датасет: температура T (0–100°C) → растворимость соли (20–80 г/100г). Обучите линейную модель (один нейрон, линейная активация). Постройте график предсказаний vs реальные значения.

4. Сгенерируйте данные для реакции первого порядка: время t (0–10 мин) → концентрация реагента $C = \exp(-k \cdot t) + \text{шум}$. Обучите одиночный нейрон с функцией активации 'exponential' (или используйте экспоненту как слой). Сравните предсказанный k с истинным.

5. Возьмите 50 образцов с двумя признаками: [молярная масса, заряд иона] → целевой признак "растворим/нерастворим" (0/1). Обучите логистическую регрессию (один нейрон с сигмоидой). Выведите матрицу ошибок (confusion matrix) с помощью sklearn.metrics.

6. Создайте датасет: входы [температура, давление] → выход "выход газа" (м³/ч). Постройте MLP с одним скрытым слоем (4 нейрона, ReLU). Обучите на 200 точках, оцените MSE на тесте.

7. Сгенерируйте нелинейную зависимость: pH (3–11) → скорость ферментативной реакции (колокол, $v = \max * \exp(-(pH-7)^2/2)$). Обучите MLP с двумя скрытыми слоями (8 и 4 нейрона) предсказывать скорость. Сравните с истинной кривой на графике.

8. Возьмите 4 химических дескриптора (например, logP, MW, количество Н-доноров, количество Н-акцепторов) для 100 молекул. Цель — температура кипения (°C). Обучите MLP (один скрытый слой, 10 нейронов) предсказывать температуру кипения. Выведите R² на тесте.

9. Создайте искусственный датасет: концентрации трех компонентов [A, B, C] → скорость реакции $r = k * A^{0.5} * B^{1.5}$ (степенная зависимость). Обучите MLP с двумя скрытыми слоями. Проверьте, может ли сеть аппроксимировать нецелые степени.

10. Сгенерируйте данные: время t (0–20 с) и начальная концентрация C₀ → текущая концентрация $C(t) = C_0 * \exp(-0.2*t)$. Обучите MLP с одним скрытым слоем (6 нейронов). Сравните предсказания для C₀=1.0 и t=5,10,15 с формулой.

11. Создайте датасет из трех классов оксидов: кислотные, основные, амфотерные. Входные признаки: электроотрицательность элемента и степень окисления. Обучите MLP с одним скрытым слоем (8 нейронов, softmax на выходе). Выведите ассигасу.

12. Сгенерируйте 200 образцов по правилу: если молярная масса > 100 и заряд > 0 → класс "катион", иначе "анион" (с шумом 10%). Обучите классификатор (MLP, скрытый слой 5 нейронов). Постройте ROC-кривую (разрешите импорт sklearn.metrics).

13. Возьмите данные по 5 типам кристаллических решеток (кубическая, тетрагональная и т.д.) с 3 числовыми признаками (плотность, модуль упругости, температура плавления). Обучите MLP с двумя скрытыми слоями (16 → 8 → 5) классифицировать тип решетки.

14. Создайте бинарный датасет: "токсично/нетоксично" по двум признакам (LD50 и logP). Обучите MLP со слоями: вход (2) → скрытый (4, ReLU) → выход (1, сигмоида). Оцените precision и recall.

15. Сгенерируйте данные для трех состояний вещества (твердое/жидкое/газ) при разных температурах и давлениях. Обучите MLP с одним скрытым слоем (10 нейронов). Выведите classification report (можно через sklearn).

16. Создайте датасет из 1000 спектров (100 точек каждый) с одним гауссовым пиком со случайным центром (450–650 нм). Постройте простой

автоэнкодер: вход (100) → скрытый слой (10) → выход (100). Обучите восстанавливать спектры. Выведите MSE восстановления.

17. Возьмите 500 образцов с 10 химическими дескрипторами. Обучите автоэнкодер сжатию до 2 нейронов в середине. Визуализируйте полученные 2-мерные представления (точки на плоскости), окрасьте по классу вещества.

18. Сгенерируйте 200 кривых титрования (100 точек каждая) с разными точками эквивалентности. Обучите автоэнкодер (вход 100 → сжатие до 3 → выход 100). Проверьте, можно ли по сжатому представлению предсказать точку эквивалентности (дополните линейной регрессией после энкодера).

19. Создайте датасет из шумных спектров (добавьте случайный шум). Обучите автоэнкодер очищать шум: вход — зашумленный спектр, выход — чистый. Сравните MSE до и после.

20. Возьмите 50 молекул, представленных векторами fingerprints длиной 256 бит. Обучите автоэнкодер сжатию до 16 бит, а затем декодированию обратно. Посчитайте долю восстановленных битов.

21. Реализуйте простую среду: химический реактор, который может нагревать (действие 0) или охлаждать (действие 1). Цель — достичь 80°C за минимальное число шагов. Обучите агента с одним скрытым слоем (4 нейрона) через Q-learning на Keras (имитируйте среду вручную). Выведите суммарное вознаграждение.

22. Среда: выбор растворителя для реакции (3 варианта). Вознаграждение: +1, если реакция идет быстро, -1 если медленно. Обучите простую policy network (один скрытый слой, 3 выхода softmax). После 100 эпизодов выведите предпочтение агента.

23. Симуляция подбора pH: агент может добавить кислоту (0) или щелочь (1) шагами по 0.1. Целевой pH = 7.0. Обучите DQN агента с двумя скрытыми слоями. Постройте график среднего вознаграждения за эпизод.

24. Среда: выбор дозы катализатора (0.1, 0.5, 1.0 мг). Вознаграждение = конверсия (0–1) минус штраф за перерасход катализатора. Обучите агента находить оптимальную дозу.

25. Обучите агента выбирать один из двух модов реакции (A→B или A→C). Вознаграждение зависит от температуры, которая меняется случайно. Агент должен адаптироваться. Используйте простейший REINFORCE (policy gradient) с одним скрытым слоем.

26. Создайте маленький датасет (50 точек) с сильным шумом. Обучите большой MLP (скрытые слои 64→64) без регуляризации, затем добавьте Dropout(0.5). Сравните accuracy на тесте и покажите, что переобучение уменьшилось.

27. Для задачи регрессии (предсказание теплоты образования по 5 признакам) добавьте L2-регуляризацию (kernel_regularizer=l2(0.01)). Сравните веса модели с регуляризацией и без нее (они должны быть меньше).

28. Обучите MLP классифицировать 2 класса (растворим/не растворим) на 500 примерах. Используйте EarlyStopping с patience=5, мониторинг validation_loss. Покажите, на какой эпохе обучение остановилось.

29. Разделите датасет (300 образцов) на train/val/test в соотношении 60/20/20. Обучите модель и постройте графики train_loss и val_loss на одной оси. Определите, есть ли переобучение.

30. Возьмите задачу регрессии (10 признаков → активность ингибитора). Добавьте BatchNormalization после каждого скрытого слоя. Сравните скорость сходимости (loss по эпохам) с моделью без нормализации.

Сделать скриншоты результата работы примера и сохранить их в отчете. Рекомендуемый объем — 2-3 страницы (2-5 слайдов)

8. Необязательный пункт отчета для получения повышенных баллов. Дополнительные фотографии процесса выполнения лабораторной работы (ознакомление с методичкой, литературой, чтение сайтов из интернета, написание отчета — 5-10 фотографий (5-10 слайдов) с пояснениями кто на фото и что именно делает студент.

9. Необязательный пункт отчета для получения повышенных баллов. Публичная защита отчета по лабораторной работе.

В данном пункте привести ссылку на Rutube-видео или иной публичный видео-ресурс, на котором записана защита-выступление отчета или презентации по лабораторной работе лично студентом. Рекомендуемое время видеоролика — 5-10 минут.

Критерии оценивания отчета:

1. Выполнены пункты 1-4 отчета - оценка «удовлетворительно».
2. Выполнены пункты 1-6 отчета - оценка «хорошо».
3. Выполнены пункты 1-7 отчета - оценка «отлично».
4. Отсутствие пункта 4 отчета понижает оценку на 1 балл. (Можно заменить по желанию студента пункт 4 устными ответами на все контрольные вопросы без понижения оценки).
5. Отсутствие нумерации пунктов отчета и их названий понижает оценку на 1 балл. (Можно не детализировать в названии пункта, что именно должно быть в данном пункте отчета, но само название должно быть!)
6. «Перемешивание» (идут не по порядку) пунктов отчета понижает оценку на 1 балл.
7. Наличие любого из пунктов 8, 9 или их вместе повышает оценку на 1 балл (но не выше «отлично»).

Лабораторная работа 4

Тема: «Подготовка входных данных для нейросети и интерпретация результата»

Цель: «Рассмотреть основные методы подготовки входных данных для нейросетей и интерпретации результатов их работы в рамках изучаемой дисциплины»

Знания, умения и владения, приобретаемые обучающимся в результате освоения темы

Знать: методы подготовки входных данных для нейросети и интерпретация результата работы нейросетей.

Уметь: при помощи библиотек numpy и keras подготавливать входные данные, при помощи библиотеки matplotlib визуализировать результат работы нейросети.

Владеть: навыками создания вычислительных программ на языке программирования Python, моделирующих работу искусственных нейронных сетей.

Теоретическая часть:

см. список рекомендованной литературы

Оборудование и материалы: см. приложение 1 методических рекомендаций по выполнению ЛР.

Техника безопасности: см. приложение 2 методических рекомендаций по выполнению ЛР.

Задания

1. Ознакомиться с теоретической частью лабораторной работы (найти ответы на контрольные вопросы).
2. Сформировать и сохранить пункты 1-4 отчета.
3. Выполнить практическую часть лабораторной работы.
 - войти в созданную в ЛР1 личную учетную запись google colab
 - создать и переименовать новый Notebook
 - ввести и отладить пробные примеры
 - сохранить результат работы пробных примеров
4. Сформировать и сохранить пункты 5-7 отчета.
5. По желанию сформировать и сохранить пункты 8-9 отчета
6. Защитить (выслать на проверку) отчет преподавателю

Содержимое отчета по лабораторной работе

Лабораторная работа считается выполненной студентом и засчитывается преподавателем (с зачислением баллов пропорционально выполненным пунктам отчета, т. е. необязательно, чтобы отчет содержал все пункты) при предоставлении в электронном виде к защите (бумажная распечатка не требуется!!!) аккуратно оформленного файла-отчета в

формате .pdf (презентация), .odt (OpenOffice, LibreOffice) или .doc (Microsoft Office), именованный по шаблону (при несоответствии работа не допускается к защите):

AI_LR4_(фамилия).pdf

Файл-отчет должен иметь следующие элементы-слайды презентации или пункты текстового отчета-документа в указанном порядке (с сохранением их номера в случае пропуска некоторых пунктов):

1. Титульный лист, содержащий название дисциплины, тему и номер лабораторной работы, ФИО студента, курс, группу.

Форму и дизайн титульного листа студент определяет самостоятельно, ориентируясь на удобство восприятия информации на титульном листе.

2. Отметка об ознакомлении с техникой безопасности (правилами поведения в компьютерном классе (аудитории, лаборатории)).

В данном пункте необходимо наличие записи «С правилами техники безопасности при выполнении лабораторной работы и поведения в компьютерном классе (аудитории, лаборатории) ознакомлен, обязуюсь выполнять _____ (ФИО), _____ дата».

3. Постановка задачи (полное условие в соответствии с выбранным вариантом).

4. Реферативный обзор информационных (интернет) источников по заданию.

4.1. Форматы входных и выходных данных современных нейросетей при моделировании на языке Python

4.2. нормализация данных

4.3. категориальные метки

4.4. аугментация данных

4.5. матрица ошибок

4.6. перемешивание (shuffle) входных данных

4.7. разделение датасета на обучающую, валидационную и тестовую выборки

4.8. эмбединги

4.9. методы получения эмбедингов

4.10. слои эмбединга

4.11. меры близости при работе с эмбедингами

4.12. предобученные эмбединги

4.13. визуализация процесса обучения нейросети

4.14. визуализация результата работы нейросети

Необходимо указать по порядку информационные источники и их содержимое по заданию. Допустимо использовать текст, гиперссылки, таблицы, рисунки и т. д. любую информацию, облегчающую просмотр и защиту презентации-отчета. Разрешено использование только открытых публичных источников информации, не требующих передачи авторских прав или каких-либо иных отчислений. Если информация по какому-либо из разделов реально отсутствует в публичном доступе или засекречена, то напечатать «информация засекречена» с указанием ведомства или

организации, засекретившей данную информацию. Рекомендуемый объем реферативного обзора — 3-10 страниц (5-10 слайдов).

5. Войти в личную учетную запись в системе google colab, созданную в ЛР1. В данном пункте сделать экранные снимки процесса успешного входа с подписями, расшифровывающими ФИО автора (подписи реальных ФИО можно добавить на скриншот в графическом редакторе уже после процесса регистрации). Рекомендуемый объем — 1-2 страницы (1-2 слайда)

6. Создать новый Notebook в системе google colab, и переименовать его по шаблону: AI_LR4_(фамилия).ipynb
Сделать скриншоты. Вставить скриншоты в отчет. Рекомендуемый объем — 1-2 страниц (1-2 слайда)

7. Ввести и отладить в системе google colab программу решения основной задачи лабораторной работы, выбранной по варианту:

1. Датасет «Синтез аммиака»

У вас есть CSV-файл с 200 записями экспериментов по синтезу аммиака (реакция Габера-Боша). Колонки: температура (300–500 °C), давление (100–300 атм), концентрация катализатора (0.5–2.5 %), время реакции (1–6 ч). Целевая переменная — выход аммиака (0–25 %). Прочитайте данные в pandas, отделите признаки X (первые 4 колонки) от целевой переменной y. Примените MinMaxScaler к признакам. Обучите трехслойный перцептрон (входной слой 4 нейрона, скрытый 8 нейронов с ReLU, выходной 1 нейрон с линейной активацией) предсказывать выход аммиака. Сравните MSE на тестовой выборке с нормализацией и без нее.

2. Датасет «Растворители в реакции»

У вас есть данные о 100 реакциях, где использовались разные растворители: «вода», «этанол», «ацетон», «толуол», «ДМСО», «гексан», «ТГФ», «дихлорметан», «этилацетат», «пиридин» (10 категорий). Другие признаки: температура (25–150 °C), время реакции (0.5–24 ч). Целевая переменная — выход продукта (0–100 %). Преобразуйте текстовые названия растворителей в числовые метки с помощью LabelEncoder. Обучите MLP (вход: 2 числовых признака + 1 категориальный как целое число; после входного слоя добавьте Dense без эмбединга). Постройте график зависимости потерь от эпохи. Оцените, насколько хорошо сеть учитывает тип растворителя.

3. Датасет «Ионы металлов»

У вас есть 50 образцов водных растворов с разными катионами: Li^+ , Na^+ , K^+ , Mg^{2+} , Ca^{2+} , Al^{3+} , Fe^{3+} , Cu^{2+} , Zn^{2+} , Ag^+ (10 категорий). Измерены признаки: электропроводность (0.1–10 мСм/см), pH (2–8), температура (10–40 °C). Цель — предсказать, является ли ион «тяжелым металлом» (бинарная метка: 1 — тяжелый, 0 — легкий). Создайте one-hot кодирование катиона с

помощью `pd.get_dummies()`. Обучите MLP с одним скрытым слоем (6 нейронов, ReLU) и выходом с сигмойдой. Сравните ассигасу модели с one-hot и модели, где катион подан как целое число (без кодирования).

4. Датасет «Титрометрический анализ»

У вас есть 300 результатов титрования сильных кислот (HCl , HNO_3 , H_2SO_4 , HClO_4) и оснований (NaOH , KOH , Ca(OH)_2). Признаки: объем титранта (10–30 мл), начальная концентрация кислоты (0.01–0.1 М), тип кислоты (4 категории), тип основания (3 категории). Цель — pH в точке эквивалентности (число от 1 до 13). Преобразуйте типы кислот и оснований в one-hot векторы, объедините с числовыми признаками. Нормализуйте все числовые признаки `StandardScaler`. Обучите MLP (вход: $4+3+2=9$ признаков, скрытые слои: 12 и 6 нейронов, ReLU, выход линейный). Выведите R^2 на тестовой выборке.

5. Датасет «Кристаллогидраты»

У вас есть 120 формул кристаллогидратов вида $\text{X} \cdot n\text{H}_2\text{O}$, где X — неорганическая соль (NaCl , CaCl_2 , MgSO_4 , CuSO_4 , AlCl_3 и 5 других — всего 10 солей), а n — число молекул воды (1–12). Признаки, которые вы вычисляете из формулы: молярная масса безводной соли (г/моль), заряд катиона, радиус катиона (пм). Цель — температура дегидратации ($^{\circ}\text{C}$, 50–300). Примените `StandardScaler` к трем числовым признакам. Обучите MLP. Постройте гистограмму предсказанных значений температуры vs реальных. Вычислите MAE и RMSE.

6. Датасет «Спектры поглощения красителей»

У вас есть 80 реальных спектров поглощения (400–700 нм, 61 точка на спектр) для 4 разных красителей. Каждый спектр измерен один раз. Вы хотите увеличить датасет. Для каждого спектра создайте 4 аугментированные копии: добавьте гауссов шум с $\sigma = 0.01$, сдвиньте по длине волны на ± 2 нм, умножьте на случайный коэффициент 0.95–1.05. Итоговый датасет: $80 \times 5 = 400$ спектров. Обучите MLP (вход 61 нейрон, скрытые слои 32 и 16, выход 4 нейрона с softmax) классифицировать краситель. Сравните ассигасу до и после аугментации.

7. Датасет «Растворимость соли в воде»

У вас есть 50 экспериментальных точек: температура (0–80 $^{\circ}\text{C}$) → растворимость соли (г/100г воды) для нитрата калия KNO_3 . Зависимость нелинейная, сильно растет с температурой. Примените аугментацию: между каждыми двумя соседними точками добавьте 3 интерполированные точки (линейная интерполяция). Также добавьте небольшой случайный шум (± 0.5 г). Получите ~200 точек. Обучите MLP (вход 1, скрытый 10 нейронов, ReLU, выход линейный) предсказывать растворимость. Постройте график исходных данных, аугментированных точек и предсказаний модели.

8. Датасет «Кинетика гидролиза»

У вас есть 30 кинетических кривых гидролиза сложного эфира: время (0–100 мин) → концентрация продукта (0–1 М). Каждая кривая имеет свою константу скорости k (0.02–0.1 мин^{-1}). Аугментируйте датасет: для каждой

кривой создайте 5 вариантов, сдвигая время на ± 5 мин и масштабируя амплитуду на 0.9–1.1. Кроме того, добавьте шум в измерение концентрации ($\sigma = 0.02$). Обучите MLP (вход — весь временной ряд, 101 точка) предсказывать константу скорости k (регрессия). Сравните MSE до и после аугментации.

9. Датасет «Смешивание растворов»

У вас есть 20 реальных смесей трех кислот (HCl , H_2SO_4 , HNO_3) с известными объемными долями (сумма = 1). Измерено одно свойство — pH смеси. Вы хотите предсказывать состав смеси по pH. Используйте аугментацию: создайте 500 синтетических смесей, генерируя случайные объемные доли (Dirichlet распределение), и вычислите pH смеси по простой формуле (средневзвешенный $\text{pH} = -\log_{10}(\sum c_i \cdot 10^{-\text{pH}_i})$). Обучите MLP (вход 1 (pH) $\rightarrow$ выход 3 (доли кислот), softmax) восстанавливать состав. Оцените среднюю абсолютную ошибку предсказанных долей.

10. Датасет «Токсичность ионов» (несбалансированный)

У вас есть датасет из 100 ионов с 5 признаками (заряд, радиус, электроотрицательность, поляризуемость, масса). Целевая метка: «токсичный» (1) — 100 образцов, «нетоксичный» (0) — 900 образцов. Примените аугментацию к меньшему классу с помощью SMOTE (синтез новых образцов, разрешите imbalanced-learn). Обучите MLP (вход 5, скрытый слой 8 нейронов, выход 1 с сигмой) на исходных данных и на сбалансированных. Сравните recall и F1-score для класса «токсичный». Объясните, почему accuracy выросла, но качество по токсичности упало без балансировки.

11. Датасет «Химические дескрипторы молекулы»

У вас есть 100 молекул с 10 дескрипторами (логарифм коэффициента распределения $\log P$, молекулярная масса, число доноров водородных связей, число акцепторов, поляризуемость, дипольный момент и др.). Цель — активность ингибитора фермента (0–100 % ингибирования). Разделите данные на обучающую (70 %), валидационную (15 %) и тестовую (15 %) выборки с помощью двух последовательных вызовов `train_test_split`. Обучите MLP с EarlyStopping по валидационной потере. Оцените финальную R^2 на тестовой выборке. Убедитесь, что выборки не пересекаются.

12. Датасет «Мониторинг реактора во времени»

У вас есть данные с химического реактора, записанные каждую минуту с 8:00 до 20:00 (720 точек). Признаки: температура, давление, расход реагента. Цель — концентрация продукта. Обучите MLP, НЕ перемешивая данные: первые 80 % точек (8:00–17:36) — обучение, последние 20 % (17:36–20:00) — тест. Затем обучите ту же архитектуру, предварительно перемешав все строки случайным образом. Сравните ошибки на тесте. Объясните, почему для временных рядов перемешивание дает завышенное качество и почему это опасно.

13. Датасет «Партии катализатора»

У вас есть 50 образцов, изготовленных из 5 различных партий катализатора (по 10 образцов от партии). Внутри партии образцы похожи. Разделите данные так, чтобы все образцы из одной партии попадали только в одну выборку (GroupShuffleSplit, группа = номер партии). Обучите MLP предсказывать конверсию реагента. Сравните с обычным случайным разделением (без группировки). Покажите, что случайное разделение завышает качество, потому что в тест попадают образцы из тех же партий, что и в обучение.

14. Датасет «Кинетические эксперименты» (кросс-валидация)

У вас есть 150 кинетических кривых (время → концентрация) для разных начальных условий. Используйте 3-кратную кросс-валидацию (KFold с `n_splits=3`). Для каждого фолда обучите MLP (вход: время, начальная концентрация; выход: текущая концентрация). Запишите `test_MSE` для каждого фолда. Выведите среднюю MSE и стандартное отклонение. Сравните результат с одним случайным разделением на train/test (80/20). Объясните, почему кросс-валидация дает более надежную оценку для маленького датасета.

15. Датасет «Спектры ЯМР»

У вас есть 60 образцов органических молекул. Для каждого образца записан спектр ЯМР ^1H (200 точек). Цель — определить тип функциональной группы (5 классов: спирт, альдегид, карбоновая кислота, амин, ароматическое кольцо). Разделите выборку на train/val/test (40/10/10). Перед каждым обучением меняйте seed перемешивания (например, 42, 123, 999). Обучите MLP 5 раз, запишите `test_accuracy`. Вычислите среднее и стандартное отклонение. Покажите, что без перемешивания (если образцы идут в порядке съемки) результат может сильно упасть.

16. Датасет «Типы химической связи»

У вас есть 50 записей о химических связях в неорганических соединениях. Категория связи: «одинарная», «двойная», «тройная», «ароматическая» (4 категории). Также есть признак: разница электроотрицательностей атомов (0–3.5). Цель — энергия связи (кДж/моль, 100–1000). Создайте MLP, где первый слой — Embedding(input_dim=4, output_dim=2) для типа связи, затем конкатенация с разностью электроотрицательностей, затем скрытый слой 8 нейронов и выход. Обучите сеть. Визуализируйте 2D-эмбединги четырех типов связи (нарисуйте точки на плоскости). Проверьте, близко ли расположены «двойная» и «тройная» (более похожие по энергии) по сравнению с «одинарной».

17. Датасет «Неорганические кислоты»

У вас есть 50 образцов 10 различных неорганических кислот: HCl, HBr, HI, HF, HNO_3 , H_2SO_4 , H_3PO_4 , HClO_4 , H_2CO_3 , H_2S . Признак — только название кислоты (категория). Цель — pK_a (значение от -10 до 10). Создайте MLP с Embedding(input_dim=10, output_dim=3) для названия кислоты и двумя полносвязными слоями (6 и 4 нейрона). Обучите предсказывать pK_a . После обучения извлеките эмбединги всех 10 кислот (весовая матрица слоя Embedding). Постройте scatter plot первых двух

компонент эмбедингов (или примените PCA к трем размерностям). Подпишите точки. Объясните, почему HI и HClO₄ (сильные кислоты) оказались рядом.

18. Датасет «Катионы в ряду активности»

У вас есть 8 катионов: Li⁺, Na⁺, K⁺, Mg²⁺, Ca²⁺, Al³⁺, Zn²⁺, Cu²⁺. Для каждого известен стандартный электродный потенциал (В) — целевая переменная. Создайте MLP с Embedding(8,2) для катиона и выходным Dense слоем (без дополнительных признаков). Обучите сеть предсказывать потенциал. Визуализируйте 2D-эмбединги катионов. Вычислите косинусную близость между эмбедингами Li⁺ и K⁺ (щелочные металлы) и между Li⁺ и Cu²⁺. Сравните с близостью по реальному химическому сходству (заряд, радиус). Сделайте вывод.

19. Датасет «Органические растворители»

У вас есть 20 разных органических растворителей (метанол, этанол, ацетон, ТГФ, дихлорметан, хлороформ, бензол, толуол, гексан, циклогексан, этилацетат, ДМСО, ДМФА, ацетонитрил, диоксан, пиридин, уксусная кислота, трет-бутанол, изопропанол, глицерин). Цель — температура кипения (°C). Обучите две модели: (1) one-hot кодирование (20 входов) + Dense слои; (2) Embedding(20,4) + Dense слои. Сравните количество параметров в обеих моделях. Сравните MAE на тесте. Объясните, почему эмбединги дают меньше параметров при том же или лучшем качестве.

20. Датасет «Периодическая таблица (первые 20 элементов)»

У вас есть 100 соединений, каждое описано двумя элементами (символы из H, He, Li, Be, B, C, N, O, F, Ne, Na, Mg, Al, Si, P, S, Cl, Ar, K, Ca). Для каждой пары известна энергия кристаллической решетки (кДж/моль). Создайте два отдельных Embedding слоя (по одному на каждый элемент, оба с output_dim=3), затем сложите (или конкатенируйте) эмбединги, затем Dense(12) и Dense(1). Обучите MLP. Визуализируйте эмбединги всех 20 элементов на 2D-плоскости (примените PCA к 3-мерным эмбедингам). Посмотрите, образуют ли щелочные металлы (Li, Na, K) кластер, галогены (F, Cl) — другой кластер.

21. Датасет «Активные vs неактивные молекулы»

У вас есть 50 молекул с 5 дескрипторами. Цель — бинарная: активна (1) или неактивна (0) в тесте на антибактериальную активность. Обучите MLP (бинарная классификация, сигмоида на выходе). После предсказания на тестовой выборке (100 образцов) постройте confusion matrix (через sklearn.metrics.confusion_matrix). Вручную вычислите accuracy, precision, recall, F1-score по элементам матрицы. Объясните, что означает каждый элемент матрицы (TP, TN, FP, FN) в химическом контексте (ложноположительный — предсказали активность, а молекула неактивна).

22. Датасет «Классификация оксидов»

У вас есть 300 образцов оксидов с тремя классами: кислотные (класс 0), основные (класс 1), амфотерные (класс 2). Признаки: степень окисления

элемента, электроотрицательность, радиус иона. Обучите MLP с тремя выходами (softmax). Постройте нормализованную confusion matrix в виде heatmap (matplotlib/seaborn). Выясните, какие классы чаще путаются. Предположите химическую причину (например, амфотерные оксиды путаются с кислотными при высокой степени окисления).

23. Датасет «Теплота образования веществ» (регрессия)

У вас есть 100 неорганических веществ с 8 признаками (атомные номера элементов, стехиометрия и др.). Цель — стандартная энтальпия образования ΔH° (кДж/моль, отрицательные значения). Обучите MLP. Постройте scatter plot: реальные ΔH° по оси X, предсказанные по оси Y. Добавьте линию $y=x$. Вычислите R^2 и MAE. Отметьте точки с наибольшей ошибкой (высокий остаток). Посмотрите, что это за вещества (например, пероксиды или азиды, которые плохо описываются простыми признаками).

24. Датасет «Конверсия катализатора» (визуализация обучения)

У вас есть 100 записей о работе катализатора (признаки: температура, давление, концентрация). Цель — конверсия (0–100 %). Обучите MLP с валидационной выборкой (20 %). В процессе обучения сохраняйте history. Постройте два графика: (1) train_loss и val_loss от эпохи; (2) train_accuracy (или MSE) и val_accuracy от эпохи. Определите эпоху, где val_loss минимальна. Укажите, наступает ли переобучение и когда его лучше остановить.

25. Датасет «Классификация типов реакций» (ROC-кривые)

У вас есть 100 образцов химических реакций, отнесенных к 4 типам: окисление, восстановление, замещение, комплексообразование (4 класса). Обучите MLP с 4 выходами (softmax). Для каждого класса постройте ROC-кривую и вычислите AUC (используйте roc_curve и roc_auc_score из sklearn, режим OneVsRest). Постройте все 4 ROC-кривые на одном графике. Какой класс предсказывается лучше всего? Почему?

26. Датасет «logP молекул» (гистограмма ошибок)

У вас есть 300 молекул с 6 дескрипторами. Цель — коэффициент распределения logP (от -2 до 6). Обучите MLP. Вычислите ошибки предсказаний на тесте (реальный logP минус предсказанный). Постройте гистограмму ошибок (20 бинов). Добавьте вертикальную линию на нуле. Нормально ли распределены ошибки? Рассчитайте среднюю ошибку (bias) и стандартное отклонение. Есть ли систематическое смещение (все ошибки положительные или отрицательные)?

27. Датасет «Типы кристаллических решеток» (classification_report)

У вас есть 150 образцов неорганических соединений с 7 признаками (плотность, твердость, температура плавления, модуль упругости, теплопроводность, коэффициент теплового расширения, электропроводность). Цель — тип решетки: кубическая (0), тетрагональная (1), гексагональная (2), ромбическая (3) (4 класса). Обучите MLP. На тестовой выборке (50 образцов) выведите classification_report из sklearn. Объясните, что означают precision, recall, f1-score и support для каждого

класса. Какой класс предсказывается хуже всего? С чем это может быть связано?

28. Датасет «Химическая кинетика» (визуализация весов слоев)

У вас есть 100 кинетических экспериментов. Входы: начальная концентрация, температура, pH, время реакции. Выход: текущая концентрация реагента. Обучите MLP с одним скрытым слоем из 10 нейронов (вход 4, скрытый 10, выход 1). Сохраните веса первого полносвязного слоя после 1-й, 5-й и последней (20-й) эпохи с помощью callback (можно просто переобучить три раза). Визуализируйте веса как тепловую карту (heatmap, оси: входные признаки $\times$ нейроны). Покажите, как веса меняются и становится ли картина более «осмысленной» (некоторые нейроны специализируются на определенных признаках).

29. Датасет «Растворимость газа» (анализ остатков)

У вас есть 100 экспериментальных точек: температура (0–100 °C) и давление (1–50 атм) $\rightarrow$ растворимость газа в воде (мг/л). Обучите MLP. На тестовой выборке постройте график: предсказанная растворимость по оси X, остаток (реальная минус предсказанная) по оси Y. Если остатки систематически растут с ростом предсказанного значения — модель гетероскедастична (ошибка зависит от величины). Сделайте вывод для вашего датасета. Предложите, как преобразовать целевую переменную (например, логарифм) для улучшения гомоскедастичности.

30. Датасет «Редкие реакции» (сильный дисбаланс классов)

У вас есть датасет из 100 химических реакций: 95 относятся к типу « $A + B \rightarrow C$ » (обычный синтез), 30 — к типу « $A \rightarrow B + C$ » (разложение), 20 — к типу «комплексобразование». Обучите MLP классифицировать три типа реакций. Выведите ассигасу (она будет высокой — около 0.95). Затем постройте confusion matrix. Объясните, почему ассигасу вводит в заблуждение, и какие метрики лучше использовать (macro/micro F1, recall для редких классов). Увеличьте вес минорных классов в функции потерь (class_weight) и снова выведите метрики.

Сделать скриншоты результата работы примера и сохранить их в отчете. Рекомендуемый объем — 2-3 страницы (2-5 слайдов)

8. Необязательный пункт отчета для получения повышенных баллов. Дополнительные фотографии процесса выполнения лабораторной работы (ознакомление с методичкой, литературой, чтение сайтов из интернета, написание отчета — 5-10 фотографий (5-10 слайдов) с пояснениями кто на фото и что именно делает студент.

9. Необязательный пункт отчета для получения повышенных баллов. Публичная защита отчета по лабораторной работе.

В данном пункте привести ссылку на Rutube-видео или иной публичный видео-ресурс, на котором записана защита-выступление отчета

или презентации по лабораторной работе лично студентом. Рекомендуемое время видеоролика — 5-10 минут.

Критерии оценивания отчета:

1. Выполнены пункты 1-4 отчета - оценка «удовлетворительно».
2. Выполнены пункты 1-6 отчета - оценка «хорошо».
3. Выполнены пункты 1-7 отчета - оценка «отлично».
4. Отсутствие пункта 4 отчета понижает оценку на 1 балл. (Можно заменить по желанию студента пункт 4 устными ответами на все контрольные вопросы без понижения оценки).
5. Отсутствие нумерации пунктов отчета и их названий понижает оценку на 1 балл. (Можно не детализировать в названии пункта, что именно должно быть в данном пункте отчета, но само название должно быть!)
6. «Перемешивание» (идут не по порядку) пунктов отчета понижает оценку на 1 балл.
7. Наличие любого из пунктов 8, 9 или их вместе повышает оценку на 1 балл (но не выше «отлично»).

Лабораторная работа 5

Тема: «Одиночные нейроны и многослойный персептрон»

Цель: «Рассмотреть методы моделирования одиночные нейронов и многослойных персептронов в рамках изучаемой дисциплины»

Знания, умения и владения, приобретаемые обучающимся в результате освоения темы

Знать: методы создания и обучения одиночных нейронов и многослойных персептронов при помощи библиотеки Keras и TensorFlow.

Уметь: создавать и обучать одиночные нейроны и многослойные персептроны.

Владеть: навыками создания вычислительных программ на языке программирования Python, моделирующих работу искусственных нейронных сетей.

Теоретическая часть:

см. список рекомендованной литературы

Оборудование и материалы: см. приложение 1 методических рекомендаций по выполнению ЛР.

Техника безопасности: см. приложение 2 методических рекомендаций по выполнению ЛР.

Задания

1. Ознакомиться с теоретической частью лабораторной работы (найти ответы на контрольные вопросы).
2. Сформировать и сохранить пункты 1-4 отчета.
3. Выполнить практическую часть лабораторной работы.
 - войти в созданную в ЛР1 личную учетную запись google colab
 - создать и переименовать новый Notebook
 - ввести и отладить пробные примеры
 - сохранить результат работы пробных примеров
4. Сформировать и сохранить пункты 5-7 отчета.
5. По желанию сформировать и сохранить пункты 8-9 отчета
6. Защитить (выслать на проверку) отчет преподавателю

Содержимое отчета по лабораторной работе

Лабораторная работа считается выполненной студентом и засчитывается преподавателем (с зачислением баллов пропорционально выполненным пунктам отчета, т. е. необязательно, чтобы отчет содержал все пункты) при предоставлении в электронном виде к защите (бумажная распечатка не требуется!!!) аккуратно оформленного файла-отчета в формате .pdf (презентация), .odt (OpenOffice, LibreOffice) или .doc (Microsoft Office), именованный по шаблону (при несоответствии работа не допускается к защите):

Файл-отчет должен иметь следующие элементы-слайды презентации или пункты текстового отчета-документа в указанном порядке (с сохранением их номера в случае пропуска некоторых пунктов):

1. Титульный лист, содержащий название дисциплины, тему и номер лабораторной работы, ФИО студента, курс, группу.

Форму и дизайн титульного листа студент определяет самостоятельно, ориентируясь на удобство восприятия информации на титульном листе.

2. Отметка об ознакомлении с техникой безопасности (правилами поведения в компьютерном классе (аудитории, лаборатории)).

В данном пункте необходимо наличие записи «С правилами техники безопасности при выполнении лабораторной работы и поведения в компьютерном классе (аудитории, лаборатории) ознакомлен, обязуюсь выполнять _____ (ФИО), _____ дата».

3. Постановка задачи (полное условие в соответствии с выбранным вариантом).

4. Реферативный обзор информационных (интернет) источников по заданию.

4.1. Одиночный нейрон

4.2. многослойный персептрон

4.3. функция активации

4.4. функция потерь

4.5. обратное распространение ошибки

4.6. автоматическое дифференцирование

4.7. оптимизаторы

4.8. регуляризация

4.9. переобучение

4.10. методы создания одиночных нейронов при помощи библиотеки Keras

4.11. методы обучения одиночных нейронов при помощи библиотеки Keras

4.12. методы создания многослойных персептронов при помощи библиотеки Keras.

4.13. методы обучения многослойных персептронов при помощи библиотеки Keras.

Необходимо указать по порядку информационные источники и их содержимое по заданию. Допустимо использовать текст, гиперссылки, таблицы, рисунки и т. д. любую информацию, облегчающую просмотр и защиту презентации-отчета. Разрешено использование только открытых публичных источников информации, не требующих передачи авторских прав или каких-либо иных отчислений. Если информация по какому-либо из разделов реально отсутствует в публичном доступе или засекречена, то напечатать «информация засекречена» с указанием ведомства или организации, засекретившей данную информацию. Рекомендуемый объем реферативного обзора — 3-10 страниц (5-10 слайдов).

5. Войти в личную учетную запись в системе google colab, созданную в ЛР1. В данном пункте сделать экранные снимки процесса успешного входа с подписями, расшифровывающими ФИО автора (подписи реальных ФИО можно добавить на скриншот в графическом редакторе уже после процесса регистрации). Рекомендуемый объем — 1-2 страницы (1-2 слайда)

6. Создать новый Notebook в системе google colab, и переименовать его по шаблону: AI_LR5_(фамилия).ipynb
Сделать скриншоты. Вставить скриншоты в отчет. Рекомендуемый объем — 1-2 страниц (1-2 слайда)

7. Ввести и отладить в системе google colab программу решения основной задачи лабораторной работы, выбранной по варианту:

1. Закон Бугера-Ламберта-Бера

Создайте синтетический датасет из 100 точек: концентрация вещества C равномерно распределена от 0 до 0.1 моль/л. Оптическая плотность A вычисляется по формуле $A = 5000 \cdot C + \text{случайный шум}$ (нормальное распределение с $\sigma=0.02$). Постройте и обучите одиночный нейрон без функции активации (linear) предсказывать A по C . Выведите полученные веса модели и сравните их с теоретическим коэффициентом 5000.

2. Зависимость температуры плавления от молярной массы

Сгенерируйте 80 точек для гомологического ряда алканов: молярная масса от 50 до 250 г/моль, температура плавления линейно растет от 0 до 1000°C с добавлением случайного шума ($\sigma=20^\circ\text{C}$). Обучите одиночный нейрон с линейной активацией предсказывать температуру плавления по молярной массе. Постройте график, показывающий исходные точки и предсказания модели.

3. Классификация металлов и неметаллов по электроотрицательности

Создайте датасет из 60 химических элементов. Признак — электроотрицательность (значения от 0.7 у франция до 4.0 у фтора). Целевая метка: 1 для металлов, 0 для неметаллов. Обучите одиночный нейрон с сигмоидной функцией активации (бинарная классификация). Выведите точность модели на тестовой выборке и проверьте предсказание для натрия (электроотрицательность 0.93) и хлора (3.16).

4. Определение кислотности среды по pH

Сгенерируйте 200 точек: значение pH равномерно распределено от 3 до 11. Класс 0 (кислая среда) присваивается при $\text{pH} < 7$, класс 1 (щелочная среда) при $\text{pH} > 7$. Точки с $\text{pH}=7$ исключите. Обучите одиночный нейрон с сигмоидной активацией классифицировать среду по одному признаку pH. Проверьте предсказания для $\text{pH}=5$ и $\text{pH}=9$.

5. Три состояния воды по температуре

Создайте 150 точек: температура воды равномерно распределена от -20 до 120°C. Классы: 0 — твердое состояние (температура $< 0^\circ\text{C}$), 1 — жидкое (0–

100°C), 2 — газообразное (>100°C). Обучите одиночный нейрон с функцией активации softmax (3 выходных нейрона) предсказывать фазовое состояние по температуре. Выведите ассигнатуру и проверьте предсказания для -10°C, 50°C и 110°C.

6. Растворимость нитрата калия от температуры

Сгенерируйте 200 точек: температура T равномерно от 0 до 80°C. Растворимость KNO_3 в г/100г воды вычисляется по эмпирической формуле $S = 13 + 0.7 \cdot T + 0.005 \cdot T^2 + \text{шум}$ ($\sigma=3$). Постройте MLP со структурой: вход 1 нейрон, скрытый слой 8 нейронов с ReLU, выходной линейный слой. Обучите сеть предсказывать растворимость. Постройте график сравнения истинных и предсказанных значений.

7. Конверсия катализатора от трех параметров

Создайте 300 точек синтетических данных. Три признака: температура (300–500°C), давление (1–10 атм), время реакции (1–5 ч). Конверсия (в процентах) вычисляется как $30 + 0.15 \cdot T + 2.5 \cdot P + 8 \cdot t + \text{нелинейные члены} + \text{шум}$ ($\sigma=5$). Обучите MLP с двумя скрытыми слоями по 16 и 8 нейронов (ReLU) и линейным выходом. Вычислите MSE на тестовой выборке и R^2 .

8. Вязкость водного раствора от концентрации и температуры

Сгенерируйте 250 точек: концентрация сахарозы C от 0 до 30%, температура T от 20 до 80°C. Вязкость μ (в cP) вычисляется по формуле $\mu = 1 + 0.05 \cdot C + 0.01 \cdot (T-20) + 0.002 \cdot C \cdot (T-20) + \text{шум}$. Обучите MLP с входом 2 признака, одним скрытым слоем из 10 нейронов (ReLU) и линейным выходом. Сравните качество этой модели с одиночным линейным нейроном.

9. Энергия активации по молекулярным дескрипторам

Создайте 400 точек. Четыре дескриптора: поляризуемость (0.5–5.0), дипольный момент (0–4.0), молекулярный объем (50–200 Å³), электроотрицательность (2.0–4.0). Энергия активации $E_a = 50 + 10 \cdot \text{поляризуемость} + 5 \cdot \text{диполь} + 0.3 \cdot \text{объем} + 15 \cdot \text{электроотрицательность} + \text{квадратичные члены} + \text{шум}$ (кДж/моль). Обучите MLP: вход 4, скрытые слои 12 и 6 нейронов, выход линейный. Выведите R^2 и MAE.

10. Температура кипения от молекулярной массы и давления

Сгенерируйте 200 точек для органических жидкостей. Молекулярная масса M от 50 до 300 г/моль, давление P от 0.5 до 2.0 атм. Температура кипения $T_{\text{кип}} = 50 + 0.8 \cdot M + 40 \cdot (P-1) + 0.1 \cdot M \cdot (P-1) + \text{шум}$ ($\sigma=8$). Обучите MLP с одним скрытым слоем из 8 нейронов (ReLU) и линейным выходом. Постройте 3D-график (можно цветовую карту) предсказанной температуры кипения от M и P .

11. Тип оксида по электроотрицательности и степени окисления

Создайте 150 образцов оксидов. Два признака: электроотрицательность элемента (0.7–4.0) и степень окисления (1–7). Три класса: 0 — кислотный оксид, 1 — основной оксид, 2 — амфотерный оксид. Обучите MLP: вход 2, скрытый слой 6 нейронов (ReLU), выход 3 нейрона с softmax. Постройте confusion matrix на тестовой выборке и определите, какие классы чаще путаются.

12. Классификация характера раствора по pH, температуре и ионной силе

Сгенерируйте 300 точек. Признаки: pH (3–11), температура (0–100°C), ионная сила (0.01–1.0 моль/л). Класс 0 — кислый раствор, класс 1 — щелочной раствор. Обучите MLP с одним скрытым слоем из 4 нейронов (ReLU) и выходным нейроном с сигмоидой. Выведите precision, recall и F1-score.

13. Кристаллические решетки по физическим свойствам

Создайте 200 образцов неорганических соединений. Три признака: плотность (1–8 г/см³), твердость по Моосу (1–10), температура плавления (100–2000°C). Три класса: 0 — кубическая решетка, 1 — гексагональная, 2 — тетрагональная. Обучите MLP со структурой 3→8→3 (softmax). Постройте ROC-кривые для каждого класса (OneVsRest).

14. Прогнозирование растворимости в воде

Сгенерируйте 500 точек для органических молекул. Три дескриптора: logP (от -2 до 5), молекулярная масса (50–500 г/моль), число доноров водородных связей (0–4). Бинарная метка: 1 — растворимо (>10 г/л), 0 — нерастворимо. Обучите MLP с двумя скрытыми слоями (10 и 5 нейронов, ReLU) и Dropout(0.2) перед выходным сигмоидным слоем. Сравните accuracy с моделью без Dropout.

15. Токсичность ионов по заряду, радиусу и электроотрицательности

Создайте 300 образцов различных ионов. Признаки: заряд (1–3), радиус (30–200 пм), электроотрицательность (0.8–2.5). Целевая метка: 1 — токсичный ион, 0 — нетоксичный. Обучите MLP: вход 3, скрытый слой 6 нейронов (ReLU), выход 1 нейрон с сигмоидой. Выведите F1-score и матрицу ошибок.

16. Сравнение ReLU и tanh для аппроксимации синусоидальной зависимости

Сгенерируйте датасет из 200 точек: x равномерно от 0 до 10, $y = \sin(x) + \text{нормальный шум} (\sigma=0.05)$. Постройте две одинаковые MLP с одним скрытым слоем из 10 нейронов: первую с активацией ReLU, вторую с tanh. Обучите обе и сравните MSE на тестовой выборке. Постройте графики предсказаний обеих моделей.

17. Влияние выбора оптимизатора на сходимость

Для задачи предсказания растворимости соли от температуры (как в задаче 2.1) обучите MLP с оптимизатором SGD (скорость обучения 0.01) и с оптимизатором Adam. Постройте на одном графике кривые loss от эпохи для обоих оптимизаторов. Объясните, почему Adam сходится быстрее и стабильнее.

18. L2-регуляризация для борьбы с переобучением

Создайте маленький датасет (50 точек) со сложной нелинейной зависимостью и сильным шумом. Постройте большую MLP с одним скрытым слоем из 64 нейронов. Обучите две версии: без регуляризации и с L2-регуляризацией (kernel_regularizer=l2(0.01)). Сравните разрыв между train_loss и val_loss на графиках и сделайте вывод о переобучении.

19. Dropout как метод регуляризации

Для задачи классификации кислотных и щелочных растворов (3.2) с 1000 точек обучите MLP с двумя скрытыми слоями (16 и 8 нейронов). Обучите две модели: без Dropout и с Dropout(0.5) после каждого скрытого слоя. Сравните ассигурацию на тестовой выборке и постройте графики `train_loss` и `val_loss` для обеих моделей.

20. Разные функции потерь для регрессии с выбросами

Сгенерируйте 300 точек для зависимости: x от 0 до 10, $y = 2 \cdot x + \text{шум}$. Добавьте 10% выбросов (аномально большие значения y). Обучите две MLP-регрессии: одну с функцией потерь MSE, вторую с MAE. Какая модель оказалась более устойчивой к выбросам? Сравните предсказания на тесте, содержащем выбросы.

21. Визуализация уменьшения функции потерь в процессе обучения

Для задачи линейной регрессии (закон Бера, как в задаче 1) обучите одиночный нейрон с оптимизатором SGD в течение 100 эпох. После каждой эпохи сохраняйте значение `loss`. Постройте график `loss` от эпохи. Объясните, как алгоритм обратного распространения ошибки постепенно уменьшает значение `loss`, подстраивая веса.

22. Отслеживание нормы градиента в процессе обучения

Обучите MLP для задачи предсказания конверсии катализатора (задача 7). Используйте `callback` для записи средней нормы градиента первого слоя после каждой эпохи. Постройте график нормы градиента от эпохи. Объясните, почему градиенты обычно уменьшаются по мере приближения к минимуму.

23. Влияние размера пакета (batch_size) на обучение

Для задачи классификации кристаллических решеток (задача 13) обучите три одинаковые MLP с разными `batch_size`: 1, 16 и 128. Сравните время обучения (секунды) и гладкость кривой `loss`. Постройте три графика `loss` от эпохи на одном рисунке. Объясните, почему маленький `batch` дает шумную кривую, а большой — сглаженную.

24. Проблема затухания градиента в глубоких сетях

Постройте глубокую MLP для задачи предсказания температуры кипения (задача 10) с архитектурой: 6 скрытых слоев по 100 нейронов (активация `sigmoid`). Обучите сеть (всего на несколько эпох). Вычислите среднюю норму градиента для первого и последнего скрытого слоя. Объясните, почему градиенты первого слоя значительно меньше — это явление называется затуханием градиента.

25. Sequential API: от концентрации ионов водорода к pH

Создайте датасет: 200 точек с концентрацией ионов H^+ от 10^{-10} до 10^{-1} моль/л (логарифмическая шкала). Вычислите $pH = -\log_{10} [H^+]$. Обучите одиночный нейрон с активацией 'exponential' предсказывать pH по H^+ . Подберите веса так, чтобы модель приблизилась формуле $pH = -\log_{10} (x)$. Выведите полученные веса.

26. Functional API для модели с двумя независимыми входами

Создайте датасет для химического реактора: первый вход — температура (300–500°C), второй вход — концентрация реагента (0.1–2.0 моль/л). Выход — скорость реакции (0–50 ммоль/л·мин). Постройте MLP с двумя отдельными входными слоями, объедините их через Concatenate, затем добавьте скрытый слой из 10 нейронов и выходной слой. Обучите модель.

27. EarlyStopping для предотвращения переобучения

Для задачи классификации растворимости (задача 19) обучите MLP с валидационной выборкой 20%. Добавьте callback EarlyStopping с параметрами `patience=10`, `monitor='val_loss'`, `restore_best_weights=True`. Выведите эпоху, на которой обучение остановилось. Покажите, что `val_loss` перестала улучшаться после этой эпохи.

28. ModelCheckpoint: сохранение лучших весов в файл

Обучите MLP для предсказания конверсии катализатора (задача 7). Используйте callback ModelCheckpoint для сохранения весов модели при наименьшей `val_loss`. Обучите модель 50 эпох, затем загрузите сохраненные веса и оцените их на тестовой выборке. Сравните с последними весами после 50 эпох.

29. model.summary() и plot_model для визуализации архитектуры

Постройте MLP для предсказания вязкости раствора (задача 8) со следующей архитектурой: вход 2 нейрона, скрытые слои 16, 8, 4 нейрона (ReLU), выход 1 нейрон. Выведите текстовое описание `model.summary()`. Если установлен `pydot`, нарисуйте диаграмму архитектуры с помощью `plot_model`. Назовите каждый слой вручную через параметр `name`.

30. Ранняя остановка как форма регуляризации

Создайте датасет с малым количеством данных (100 точек) и сильной нелинейной зависимостью. Постройте большую MLP (скрытый слой 64 нейрона). Обучите две модели: одну с EarlyStopping (`patience=5`), другую — без остановки (100 эпох). Сравните `test_loss` обеих моделей. Объясните, почему ранняя остановка предотвращает переобучение.

31. Ручной подбор гиперпараметров: количество нейронов

Для задачи классификации оксидов (задача 11) переберите три варианта архитектуры: скрытый слой из 4 нейронов, из 8 нейронов, и два слоя 12→6 нейронов. Для каждой архитектуры проведите обучение и вычислите `validation accuracy`. Выберите лучшую архитектуру и объясните свой выбор.

32. Влияние доли валидационной выборки

Для задачи растворимости (задача 6) обучите MLP с `validation_split = 0.1` и с `validation_split = 0.4`. Постройте кривые `train_loss` и `val_loss` для обоих случаев на одном графике. Объясните, почему при малой валидационной выборке (10%) кривая `val_loss` более шумная и менее надежная.

33. Различные форматы целевых меток: sparse vs categorical

Для задачи классификации фазовых состояний (задача 5) обучите две MLP: первую с метками как целыми числами (0,1,2) и `loss='sparse_categorical_crossentropy'`, вторую с one-hot векторами и

`loss='categorical_crossentropy'`. Покажите, как преобразовать метки из одного формата в другой с помощью `keras.utils.to_categorical`.

34. Получение вероятностей классов вместо жестких меток

Обучите MLP для бинарной классификации металлов/неметаллов (задача 3). На тестовой выборке получите не только итоговые классы (0 или 1) с помощью `model.predict`, округленного до целых, но и вероятности принадлежности к классу 1 с помощью `predict_proba` (или просто `predict` без округления). Постройте ROC-кривую, используя эти вероятности.

Сделать скриншоты результата работы примера и сохранить их в отчете. Рекомендуемый объем — 2-3 страницы (2-5 слайдов)

8. Необязательный пункт отчета для получения повышенных баллов. Дополнительные фотографии процесса выполнения лабораторной работы (ознакомление с методичкой, литературой, чтение сайтов из интернета, написание отчета — 5-10 фотографий (5-10 слайдов) с пояснениями кто на фото и что именно делает студент.

9. Необязательный пункт отчета для получения повышенных баллов. Публичная защита отчета по лабораторной работе.

В данном пункте привести ссылку на Rutube-видео или иной публичный видео-ресурс, на котором записана защита-выступление отчета или презентации по лабораторной работе лично студентом. Рекомендуемое время видеоролика — 5-10 минут.

Критерии оценивания отчета:

1. Выполнены пункты 1-4 отчета - оценка «удовлетворительно».
2. Выполнены пункты 1-6 отчета - оценка «хорошо».
3. Выполнены пункты 1-7 отчета - оценка «отлично».
4. Отсутствие пункта 4 отчета понижает оценку на 1 балл. (Можно заменить по желанию студента пункт 4 устными ответами на все контрольные вопросы без понижения оценки).
5. Отсутствие нумерации пунктов отчета и их названий понижает оценку на 1 балл. (Можно не детализировать в названии пункта, что именно должно быть в данном пункте отчета, но само название должно быть!)
6. «Перемешивание» (идут не по порядку) пунктов отчета понижает оценку на 1 балл.
7. Наличие любого из пунктов 8, 9 или их вместе повышает оценку на 1 балл (но не выше «отлично»).

Лабораторная работа 6

Тема: «Энкодеры и генеративно-состязательные сети»

Цель: «Рассмотреть основные возможности энкодеров и генеративно-состязательных сетей в рамках изучаемой дисциплины»

Знания, умения и владения, приобретаемые обучающимся в результате освоения темы

Знать: методы создания и обучения нейронных сетей встречного распространения (энкодеры), принципы работы генеративно-состязательных сетей.

Уметь: создавать и обучать энкодеры и генеративно-состязательные нейронные сети.

Владеть: навыками создания вычислительных программ на языке программирования Python, моделирующих работу искусственных нейронных сетей.

Теоретическая часть:

см. список рекомендованной литературы

Оборудование и материалы: см. приложение 1 методических рекомендаций по выполнению ЛР.

Техника безопасности: см. приложение 2 методических рекомендаций по выполнению ЛР.

Задания

1. Ознакомиться с теоретической частью лабораторной работы (найти ответы на контрольные вопросы).
2. Сформировать и сохранить пункты 1-4 отчета.
3. Выполнить практическую часть лабораторной работы.
 - войти в созданную в ЛР1 личную учетную запись google colab
 - создать и переименовать новый Notebook
 - ввести и отладить пробные примеры
 - сохранить результат работы пробных примеров
4. Сформировать и сохранить пункты 5-7 отчета.
5. По желанию сформировать и сохранить пункты 8-9 отчета
6. Защитить (выслать на проверку) отчет преподавателю

Содержимое отчета по лабораторной работе

Лабораторная работа считается выполненной студентом и засчитывается преподавателем (с зачислением баллов пропорционально выполненным пунктам отчета, т. е. необязательно, чтобы отчет содержал все пункты) при предоставлении в электронном виде к защите (бумажная распечатка не требуется!!!) аккуратно оформленного файла-отчета в формате .pdf (презентация), .odt (OpenOffice, LibreOffice) или .doc (Microsoft

Office), именованный по шаблону (при несоответствии работа не допускается к защите):

AI_LR6_(фамилия).pdf

Файл-отчет должен иметь следующие элементы-слайды презентации или пункты текстового отчета-документа в указанном порядке (с сохранением их номера в случае пропуска некоторых пунктов):

1. Титульный лист, содержащий название дисциплины, тему и номер лабораторной работы, ФИО студента, курс, группу.

Форму и дизайн титульного листа студент определяет самостоятельно, ориентируясь на удобство восприятия информации на титульном листе.

2. Отметка об ознакомлении с техникой безопасности (правилами поведения в компьютерном классе (аудитории, лаборатории)).

В данном пункте необходимо наличие записи «С правилами техники безопасности при выполнении лабораторной работы и поведения в компьютерном классе (аудитории, лаборатории) ознакомлен, обязуюсь выполнять _____ (ФИО), _____ дата».

3. Постановка задачи (полное условие в соответствии с выбранным вариантом).

4. Реферативный обзор информационных (интернет) источников по заданию.

4.1. нейронные сети встречного распространения

4.2. слой Кохонена

4.3. слой Гроссберга

4.4. энкодеры

4.5. автоэнкодеры

4.6. вариативные автоэнкодеры

4.7. генеративно-состязательные нейронные сети

4.8. архитектура генеративно-состязательной нейронной сети

4.9. методы создания энкодеров при помощи библиотеки Keras

4.10. методы обучения энкодеров при помощи библиотеки Keras

4.11. методы создания генеративно-состязательных сетей при помощи библиотеки Keras

4.12. Методы обучения генеративно-состязательных сетей при помощи библиотеки Keras

Необходимо указать по порядку информационные источники и их содержимое по заданию. Допустимо использовать текст, гиперссылки, таблицы, рисунки и т. д. любую информацию, облегчающую просмотр и защиту презентации-отчета. Разрешено использование только открытых публичных источников информации, не требующих передачи авторских прав или каких-либо иных отчислений. Если информация по какому-либо из разделов реально отсутствует в публичном доступе или засекречена, то напечатать «информация засекречена» с указанием ведомства или организации, засекретившей данную информацию. Рекомендуемый объем реферативного обзора — 3-10 страниц (5-10 слайдов).

5. Войти в личную учетную запись в системе google colab, созданную в ЛР1. В данном пункте сделать экранные снимки процесса успешного входа с подписями, расшифровывающими ФИО автора (подписи реальных ФИО можно добавить на скриншот в графическом редакторе уже после процесса регистрации). Рекомендуемый объем — 1-2 страницы (1-2 слайда)

6. Создать новый Notebook в системе google colab, и переименовать его по шаблону: AI_LR6_(фамилия).ipynb
Сделать скриншоты. Вставить скриншоты в отчет. Рекомендуемый объем — 1-2 страниц (1-2 слайда)

7. Ввести и отладить в системе google colab программу решения основной задачи лабораторной работы, выбранной по варианту:

1. Сжатие спектров поглощения (5 точек → 2 нейрона)

Создайте датасет из 500 спектров: каждый спектр имеет 5 длин волн (400, 450, 500, 550, 600 нм). Значения поглощения генерируются как сумма двух гауссовых пиков со случайными параметрами. Постройте энкодер (вход 5 → скрытый слой 4 → сжатие 2 нейрона). Обучите сеть восстанавливать вход (автоэнкодер). Визуализируйте 2-мерные сжатые представления спектров (точки на плоскости).

2. Сжатие кинетических кривых (10 временных точек → 3 нейрона)

Сгенерируйте 400 кинетических кривых: время от 0 до 45 мин (10 точек). Каждая кривая описывается экспоненциальным законом $C = C_0 \cdot \exp(-k \cdot t)$ со случайными C_0 (0.5–2.0) и k (0.05–0.30). Постройте энкодер (вход 10 → скрытый слой 6 → выход 3 нейрона). Обучите автоэнкодер. Проверьте, можно ли по 3-мерному сжатому представлению восстановить исходную кривую с $MSE < 0.01$.

3. Энкодер для температурных зависимостей растворимости

Создайте 300 кривых растворимости: температура от 0 до 100°C (8 точек). Растворимость $S = a + b \cdot T + c \cdot T^2$ со случайными a , b , c . Постройте энкодер (вход 8 → сжатие до 2 нейронов) и декодер (2 → 8). Обучите автоэнкодер. Визуализируйте сжатые представления и раскрасьте их по параметру b (температурному коэффициенту). Сделайте вывод о кластеризации.

4. Понижение размерности данных о катализаторах (15 признаков → 4)

Создайте 600 образцов катализаторов с 15 признаками (температура обжига, концентрация активного компонента, удельная поверхность, пористость и др.). Эти признаки коррелированы между собой. Постройте энкодер (15 → 10 → 4). Обучите автоэнкодер восстанавливать исходные 15 признаков. После обучения используйте только энкодер для сжатия данных. Сравните качество кластеризации на исходных и сжатых данных.

5. Шумоподавление спектров ЯМР с помощью энкодера

Создайте 500 чистых спектров ЯМР (100 точек каждый) с 3 пиками. Добавьте к каждому спектру гауссов шум ($\sigma=0.1$). Постройте автоэнкодер (вход 100 → скрытые слои 50 → 20 → 50 → выход 100, активация ReLU). Обучите сеть на парах (зашумленный спектр → чистый спектр). Проверьте на тестовом зашумленном спектре и визуализируйте результат шумоподавления.

6. Генерация новых спектров поглощения через автоэнкодер

Обучите автоэнкодер на 800 спектрах поглощения (60 точек) так, как в задаче 1. После обучения возьмите случайные точки в сжатом латентном пространстве (2 нейрона) из нормального распределения. Пропустите их через декодер. Получите и визуализируйте 5 новых «синтетических» спектров. Похожи ли они на реальные?

7. Аугментация малого датасета кинетических кривых

У вас есть только 50 экспериментальных кинетических кривых (20 временных точек). Обучите автоэнкодер на этих данных (вход 20 → сжатие 4 → выход 20). После обучения сгенерируйте 200 новых кривых, проходя случайные точки в латентном пространстве через декодер. Объедините с исходными. Обучите простой MLP для предсказания константы скорости k по кривой и сравните качество до и после аугментации.

8. Интерполяция между двумя спектрами через латентное пространство

Возьмите два реальных спектра (например, при концентрациях 0.1 и 0.5 моль/л). Обучите автоэнкодер на множестве спектров разных концентраций. Найдите их латентные векторы z_1 и z_2 . Создайте 5 промежуточных векторов: $z = z_1 + \lambda \cdot (z_2 - z_1)$ для $\lambda=0.2, 0.4, 0.6, 0.8$. Пропустите через декодер. Постройте полученные спектры. Должны ли они соответствовать промежуточным концентрациям?

9. Генерация новых кривых титрования

Создайте датасет из 200 кривых титрования (pH от объема титранта, 50 точек). Обучите автоэнкодер (вход 50 → сжатие до 3 → выход 50). Случайно сэмплируйте 10 точек в трехмерном латентном пространстве из нормального распределения $N(0,1)$. Сгенерируйте новые кривые. Визуализируйте их. Отличаются ли они от обучающих? Есть ли среди них физически некорректные (например, pH не монотонный)?

10. Восстановление пропущенных участков спектра

Создайте датасет полных спектров (400–700 нм, 61 точка). В каждом спектре искусственно удалите участок (500–550 нм). Обучите автоэнкодер (вход 61 → сжатие 10 → выход 61) на полных спектрах. Теперь подайте на вход спектры с пропуском (пропуск заполните нулями или средним). Сможет ли автоэнкодер восстановить пропущенный участок? Оцените MSE на восстановленном участке.

11. VAE для генерации спектров с контролируемой формой пика

Обучите VAE (вход 50 → латентное пространство 2 → выход 50) на 1000 спектрах с разной шириной пика (от 10 до 50 нм). После обучения сгенерируйте новые спектры, перемещаясь по латентному пространству.

Определите, какая компонента латентного пространства отвечает за ширину пика. Постройте график зависимости ширины пика от значения латентной переменной.

12. VAE для генерации новых кинетических кривых

Обучите VAE на 600 кинетических кривых (время $\rightarrow$ концентрация). Латинское пространство — 3 нейрона. Сгенерируйте 20 новых кривых. Проверьте, все ли они монотонно убывают (физическое ограничение). Если нет, измените функцию активации выходного слоя или добавьте регуляризацию. Сравните качество генерации с обычным автоэнкодером из задачи 6.

13. VAE с регуляризацией KL-дивергенцией

Для датасета из 500 спектров (40 точек) обучите два автоэнкодера: обычный и VAE (с добавлением KL-дивергенции в функцию потерь). Визуализируйте латентное пространство (2 нейрона) для обоих случаев. В обычном автоэнкодере точки будут разбросаны неравномерно, в VAE — сгруппированы и близки к $N(0,1)$. Объясните, почему VAE лучше подходит для генерации новых данных.

14. Интерполяция между двумя химическими структурами (через дескрипторы)

Создайте датасет из 300 молекул, каждая представлена 10 дескрипторами ($\log P$, MW, Tкип и др.). Обучите VAE (10 $\rightarrow$ 2 $\rightarrow$ 10). Выберите две молекулы — метанол и этанол. Найдите их латентные векторы z_1 и z_2 . Сгенерируйте промежуточные точки и декодируйте их в дескрипторы. Получится ли молекула с дескрипторами, промежуточными между метанолом и этанолом? (Например, пропанол?)

15. Генерация новых температурных зависимостей вязкости

У вас есть 400 кривых вязкости от температуры (0–100°C, 10 точек) для разных жидкостей. Обучите VAE (вход 10 $\rightarrow$ латентное пространство 2 $\rightarrow$ выход 10). Сгенерируйте 100 новых кривых. Постройте их на графике вместе с реальными. Есть ли среди сгенерированных физически невозможные (например, отрицательная вязкость или немонотонная зависимость)? Отфильтруйте их.

16. GAN для генерации одномерных спектров (простой MLP GAN)

Постройте GAN для генерации спектров поглощения (61 точка). Генератор: вход — шум 10 размерности $\rightarrow$ скрытый слой 30 $\rightarrow$ выход 61 (tanh). Дискриминатор: вход 61 $\rightarrow$ скрытый слой 30 $\rightarrow$ выход 1 (сигмоида). Обучите GAN на 2000 реальных спектров. Визуализируйте сгенерированные спектры после 10, 50 и 200 эпох. Как меняется качество?

17. GAN для генерации кинетических кривых (экспоненциальное затухание)

Обучите GAN на 1500 кинетических кривых (20 точек). Генератор: 8 шумов $\rightarrow$ скрытый слой 20 $\rightarrow$ выход 20. Дискриминатор: вход 20 $\rightarrow$ скрытый слой 20 $\rightarrow$ выход 1. Сгенерируйте 50 кривых и проверьте, насколько они похожи на экспоненту (визуально и по MSE относительно экспоненциальной аппроксимации). Сравните с VAE из задачи 12.

18. Условный GAN (сGAN) для генерации спектров с заданной концентрацией

Модифицируйте GAN из задачи 16: добавьте в генератор дополнительный вход — желаемую концентрацию (число от 0.1 до 1.0). В дискриминатор тоже подавайте концентрацию (конкатенация со спектром). Обучите сGAN генерировать спектр, соответствующий заданной концентрации. Сгенерируйте спектры для $C=0.2, 0.5, 0.8$ и проверьте, растет ли оптическая плотность с концентрацией.

19. GAN для генерации новых кривых титрования

Создайте датасет кривых титрования (pH от объема, 50 точек). Обучите GAN (генератор: 10 шумов $\rightarrow$ скрытый слой 40 $\rightarrow$ выход 50; дискриминатор: 50 $\rightarrow$ 25 $\rightarrow$ 1). Сгенерируйте 20 новых кривых. Проверьте, все ли они начинаются с низкого pH (кислота) и заканчиваются высоким (щелочь). Если нет, добавьте штраф в функцию потерь генератора.

20. Генерация новых составов тройных смесей (3 компонента)

У вас есть 1000 точек, представляющих составы тройных смесей (доли A, B, C, сумма = 1). Обучите GAN генерировать новые допустимые составы. Генератор должен выдавать 3 числа, которые после softmax будут давать доли (сумма 1). Дискриминатор различает реальные и сгенерированные смеси. Оцените, охватывает ли генератор все пространство составов или только некоторые области.

21. Слой Кохонена для классификации типов оксидов

Создайте датасет из 200 образцов оксидов с 3 признаками (электроотрицательность, степень окисления, радиус иона). Реализуйте в Keras слой Кохонена (можно через пользовательский слой или используйте MiniBatchKMeans из sklearn как аналог). Обучите его кластеризовать данные на 3 кластера (соответствуют кислотным, основным, амфотерным оксидам). Сравьте с true-метками.

22. Сеть встречного распространения для сжатия + классификации

Постройте сеть встречного распространения: входной слой (10 признаков молекулы), слой Кохонена (5 нейронов), слой Гроссберга (выходной, 3 класса). Обучите сеть на 500 образцах. Сравните точность классификации с обычным MLP. Объясните, в чем преимущество сети встречного распространения (одновременное сжатие и классификация).

23. Кохоненовская карта для визуализации химического пространства

Возьмите 300 молекул, каждая описана 8 дескрипторами. Обучите самоорганизующуюся карту Кохонена (SOM) с сеткой 5×5 нейронов (библиотека minisom или самодельная в Keras). Визуализируйте карту, окрасив нейроны по среднему logP молекул, попавших в нейрон. Объясните, какие области карты соответствуют гидрофобным молекулам.

24. Сеть встречного распространения для предсказания активности

Постройте сеть: вход 6 дескрипторов катализатора $\rightarrow$ слой Кохонена (4 нейрона) $\rightarrow$ слой Гроссберга (1 нейрон, линейный) для регрессии конверсии.

Обучите на 400 образцах. Сравните MSE с MLP из 2 скрытых слоев (6→12→6→1). Какая сеть обучилась быстрее? Какая дала меньшую ошибку?

25. Анализ чувствительности: влияние признаков на нейроны Кохонена

Обучите слой Кохонена (4×4 нейрона) на 600 спектрах (10 признаков — интенсивности в 10 длинах волн). После обучения для каждого нейрона найдите, какой признак (длина волны) имеет максимальный вес. Постройте карту 4×4, где в каждой ячейке указана длина волны, на которую этот нейрон больше всего реагирует. Сделайте вывод о специализации нейронов.

26. VAE для генерации новых спектров с указанием номера пика

Обучите VAE на 800 спектрах, имеющих 1, 2 или 3 пика (метка — количество пиков). В латентное пространство добавьте условную переменную «число пиков» (один горячий вектор). После обучения сгенерируйте спектры с заданным числом пиков: 1, 2, 3. Проверьте, у всех ли сгенерированных спектров нужно число пиков.

27. Генерация новых RGB-изображений кристаллов (10×10 пикселей) через GAN

Создайте датасет из 50 маленьких изображений (10×10, черно-белые) кристаллов под микроскопом (синтетические: круги, шестиугольники с шумом). Обучите простой GAN (генератор: 20 шумов → скрытый слой 50 → выход 100; дискриминатор: 100 → 50 → 1). Сгенерируйте новые изображения. Похожи ли они на реальные кристаллические структуры?

28. VAE для поиска аномалий в данных хроматографа

Обучите VAE на 1000 нормальных хроматограммах (50 точек). Вычислите ошибку восстановления для каждой хроматограммы (MSE). Для новых данных (200 образцов, из которых 20 — аномальные, например, с выбросом) вычислите ошибку восстановления. Все ли аномалии имеют ошибку выше порога (например, 95-й перцентиль нормальных)? Используйте VAE как детектор аномалий.

29. GAN с регуляризацией градиентного штрафа (WGAN-GP)

Реализуйте WGAN-GP для генерации спектров (61 точка). Генератор: шум 10 → 40 → 61; дискриминатор (критик): 61 → 40 → 1 (без сигмоиды). Используйте градиентный штраф для дискриминатора. Обучите на 2000 спектрах. Сравните качество сгенерированных спектров с обычным GAN из задачи 4.1 (по визуальной оценке и по разнообразию).

30. Циклический VAE для перевода спектра из одного состояния в другое

Создайте парные данные: 300 спектров раствора при 25°C и те же образцы при 50°C. Обучите два VAE (общий энкодер и декодер, но с условным входом температуры). Сделайте так, чтобы можно было перевести спектр из 25°C в 50°C: закодировать спектр при 25°C в латентное пространство, а затем декодировать с условием 50°C. Проверьте визуально, похож ли полученный спектр на реальный спектр при 50°C.

Сделать скриншоты результата работы примера и сохранить их в отчете. Рекомендуемый объем — 2-3 страницы (2-5 слайдов)

8. Необязательный пункт отчета для получения повышенных баллов. Дополнительные фотографии процесса выполнения лабораторной работы (ознакомление с методичкой, литературой, чтение сайтов из интернета, написание отчета — 5-10 фотографий (5-10 слайдов) с пояснениями кто на фото и что именно делает студент.

9. Необязательный пункт отчета для получения повышенных баллов. Публичная защита отчета по лабораторной работе.

В данном пункте привести ссылку на Rutube-видео или иной публичный видео-ресурс, на котором записана защита-выступление отчета или презентации по лабораторной работе лично студентом. Рекомендуемое время видеоролика — 5-10 минут.

Критерии оценивания отчета:

1. Выполнены пункты 1-4 отчета - оценка «удовлетворительно».
2. Выполнены пункты 1-6 отчета - оценка «хорошо».
3. Выполнены пункты 1-7 отчета - оценка «отлично».
4. Отсутствие пункта 4 отчета понижает оценку на 1 балл. (Можно заменить по желанию студента пункт 4 устными ответами на все контрольные вопросы без понижения оценки).
5. Отсутствие нумерации пунктов отчета и их названий понижает оценку на 1 балл. (Можно не детализировать в названии пункта, что именно должно быть в данном пункте отчета, но само название должно быть!)
6. «Перемешивание» (идут не по порядку) пунктов отчета понижает оценку на 1 балл.
7. Наличие любого из пунктов 8, 9 или их вместе повышает оценку на 1 балл (но не выше «отлично»).

Лабораторная работа 7

Тема: «Сверточные нейросети»

Цель: «Рассмотреть основные возможности сверточных нейросетей в рамках изучаемой дисциплины»

Знания, умения и владения, приобретаемые обучающимся в результате освоения темы

Знать: методы создания и обучения сверточных нейронных сетей.

Уметь: создавать программы моделирования сверточных нейронных сетей (распознавания графических образов) на языке программирования Python.

Владеть: навыками создания вычислительных программ на языке программирования Python, моделирующих работу искусственных нейронных сетей.

Теоретическая часть:

см. список рекомендованной литературы

Оборудование и материалы: см. приложение 1 методических рекомендаций по выполнению ЛР.

Техника безопасности: см. приложение 2 методических рекомендаций по выполнению ЛР.

Задания

1. Ознакомиться с теоретической частью лабораторной работы (найти ответы на контрольные вопросы).
2. Сформировать и сохранить пункты 1-4 отчета.
3. Выполнить практическую часть лабораторной работы.
 - войти в созданную в ЛР1 личную учетную запись google colab
 - создать и переименовать новый Notebook
 - ввести и отладить пробные примеры
 - сохранить результат работы пробных примеров
4. Сформировать и сохранить пункты 5-7 отчета.
5. По желанию сформировать и сохранить пункты 8-9 отчета
6. Защитить (выслать на проверку) отчет преподавателю

Содержимое отчета по лабораторной работе

Лабораторная работа считается выполненной студентом и засчитывается преподавателем (с зачислением баллов пропорционально выполненным пунктам отчета, т. е. необязательно, чтобы отчет содержал все пункты) при предоставлении в электронном виде к защите (бумажная распечатка не требуется!!!) аккуратно оформленного файла-отчета в формате .pdf (презентация), .odt (OpenOffice, LibreOffice) или .doc (Microsoft Office), именованный по шаблону (при несоответствии работа не допускается к защите):

Файл-отчет должен иметь следующие элементы-слайды презентации или пункты текстового отчета-документа в указанном порядке (с сохранением их номера в случае пропуска некоторых пунктов):

1. Титульный лист, содержащий название дисциплины, тему и номер лабораторной работы, ФИО студента, курс, группу.

Форму и дизайн титульного листа студент определяет самостоятельно, ориентируясь на удобство восприятия информации на титульном листе.

2. Отметка об ознакомлении с техникой безопасности (правилами поведения в компьютерном классе (аудитории, лаборатории)).

В данном пункте необходимо наличие записи «С правилами техники безопасности при выполнении лабораторной работы и поведения в компьютерном классе (аудитории, лаборатории) ознакомлен, обязуюсь выполнять _____ (ФИО), _____ дата».

3. Постановка задачи (полное условие в соответствии с выбранным вариантом).

4. Реферативный обзор информационных (интернет) источников по заданию.

4.1. Операция свертки

4.2. Двумерная свертка

4.3. фильтрация графических изображений

4.4. субдискретизация графических изображений

4.5. архитектура сверточной нейронной сети

4.6. создание сверточной нейронной сети при помощи библиотеки Keras

4.7. обучение сверточной нейронной сети при помощи библиотеки Keras

Необходимо указать по порядку информационные источники и их содержимое по заданию. Допустимо использовать текст, гиперссылки, таблицы, рисунки и т. д. любую информацию, облегчающую просмотр и защиту презентации-отчета. Разрешено использование только открытых публичных источников информации, не требующих передачи авторских прав или каких-либо иных отчислений. Если информация по какому-либо из разделов реально отсутствует в публичном доступе или засекречена, то напечатать «информация засекречена» с указанием ведомства или организации, засекретившей данную информацию. Рекомендуемый объем реферативного обзора — 3-10 страниц (5-10 слайдов).

5. Войти в личную учетную запись в системе google colab, созданную в ЛР1. В данном пункте сделать экранные снимки процесса успешного входа с подписями, расшифровывающими ФИО автора (подписи реальных ФИО можно добавить на скриншот в графическом редакторе уже после процесса регистрации). Рекомендуемый объем — 1-2 страницы (1-2 слайда)

6. Создать новый Notebook в системе google colab, и переименовать его по шаблону: AI_LR7_(фамилия).ipynb

Сделать скриншоты. Вставить скриншоты в отчет. Рекомендуемый объем — 1-2 страниц (1-2 слайда)

7. Ввести и отладить в системе google colab программу решения основной задачи лабораторной работы, выбранной по варианту:

1. Обнаружение пика поглощения в спектре

Создайте 500 одномерных спектров (длина 100 точек). В половине спектров есть гауссов пик с центром 50 и шириной 10, в половине — только шум. Постройте CNN: вход (100,1) → Conv1D(8, kernel_size=5, activation='relu') → GlobalMaxPooling1D → Dense(1, sigmoid). Обучите сеть обнаруживать наличие пика. Выведите точность.

2. Классификация типов хроматограмм

Сгенерируйте 600 хроматограмм (80 точек) трех типов: один пик, два разделенных пика, два перекрывающихся пика. Постройте CNN: вход (80,1) → Conv1D(16, 7) → Conv1D(8, 5) → GlobalAveragePooling1D → Dense(3, softmax). Обучите классифицировать тип хроматограммы. Постройте confusion matrix.

3. Определение числа компонентов в смеси по спектру

Создайте 800 спектров (120 точек), каждый из которых является суммой 1, 2 или 3 гауссовых пиков со случайными центрами. Постройте CNN: Conv1D(16, 5) → MaxPooling1D(2) → Conv1D(32, 5) → GlobalMaxPooling1D → Dense(3, softmax). Обучите сеть предсказывать количество компонентов. Выведите accuracy.

4. Поиск положения максимума пика (регрессия)

Сгенерируйте 1000 спектров с одним гауссовым пиком, центр которого равномерно распределен от 20 до 80 (при длине спектра 100 точек). Постройте CNN: вход (100,1) → Conv1D(16, 7) → Flatten → Dense(32) → Dense(1, linear). Обучите сеть предсказывать позицию максимума. Выведите MAE в точках.

5. Сглаживание зашумленного спектра (сверточный автоэнкодер)

Создайте 1000 чистых спектров (60 точек) с одним пиком. Добавьте шум ($\sigma=0.1$). Постройте сверточный автоэнкодер: вход (60,1) → Conv1D(8, 3) → Conv1D(4, 3) → Conv1DTranspose(8, 3) → Conv1DTranspose(1, 3). Обучите восстанавливать чистый спектр из зашумленного. Визуализируйте результат для тестового примера.

6. Классификация кинетических кривых (порядок реакции)

Сгенерируйте 500 кинетических кривых (50 точек): нулевой порядок (линейное убывание), первый порядок (экспонента), второй порядок (гипербола). Постройте CNN: Conv1D(8, 5) → MaxPooling1D(2) → Conv1D(16, 5) → Flatten → Dense(3, softmax). Обучите определять порядок реакции. Выведите F1-score.

7. Обнаружение аномалии в хроматограмме

Создайте 800 нормальных хроматограмм (три стандартных пика) и 200 аномальных (дополнительный пик или сдвиг). Постройте CNN: Conv1D(8, 5)

→ GlobalAveragePooling1D → Dense(1, sigmoid). Обучите бинарной классификации (норма/аномалия). Выведите precision и recall.

8. Предсказание концентрации по спектру (одномерная регрессия)

Сгенерируйте 600 спектров (80 точек) с одним пиком, высота которого пропорциональна концентрации (0.1–1.0). Постройте CNN: Conv1D(8, 5) → MaxPooling1D(2) → Conv1D(16, 5) → Flatten → Dense(1, linear). Обучите предсказывать концентрацию. Выведите R^2 на тесте.

9. Разделение перекрывающихся пиков (задача деконволюции)

Создайте 500 спектров, каждый из которых является суммой двух близких гауссовых пиков с известными параметрами. Постройте CNN: вход (100,1) → Conv1D(16, 7) → Conv1D(32, 7) → Dense(100) (выход — восстановленные два пика). Обучите сеть восстанавливать отдельные пики. Визуализируйте результат.

10. Классификация типов ЯМР-спектров (^1H , ^{13}C , DEPT)

Создайте по 300 синтетических спектров (200 точек) для каждого типа ЯМР (^1H — много пиков, ^{13}C — мало пиков, DEPT — с положительными и отрицательными пиками). Постройте CNN: Conv1D(8, 9) → Conv1D(16, 9) → GlobalMaxPooling1D → Dense(3, softmax). Обучите классифицировать тип спектра. Выведите accuracy.

11. Распознавание формы кристаллов (круг, квадрат, треугольник)

Создайте 90 черно-белых изображений 32×32 пикселя: круги, квадраты, треугольники (по 30 каждого). Постройте CNN: Conv2D(8, 3×3) → MaxPooling2D(2×2) → Conv2D(16, 3×3) → Flatten → Dense(3, softmax). Обучите классифицировать форму кристалла. Выведите confusion matrix.

12. Классификация типов кристаллических решеток по микрофотографиям

Создайте 80 изображений 28×28 (синтетические текстуры) для кубической, гексагональной и тетрагональной решеток. Постройте CNN: Conv2D(8, 3) → Conv2D(16, 3) → MaxPooling2D(2) → Flatten → Dense(3, softmax). Обучите классифицировать тип решетки. Сравните точность с MLP на тех же данных (без свертки).

13. Обнаружение дефектов на поверхности кристалла

Создайте 60 изображений 40×40: 40 без дефектов, 20 с трещинами или включениями. Постройте CNN: Conv2D(8, 5) → MaxPooling2D(2) → Conv2D(16, 5) → GlobalAveragePooling2D → Dense(1, sigmoid). Обучите бинарной классификации. Выведите ROC-кривую и AUC.

14. Распознавание химических формул на изображениях (цифры и буквы)

Создайте 100 маленьких изображений 28×28 с символами: H, O, C, N, Na, Cl, цифры 1–9 (14 классов). Постройте CNN: Conv2D(16, 3) → Conv2D(32, 3) → MaxPooling2D(2) → Flatten → Dense(14, softmax). Обучите распознавать символы. Выведите accuracy по каждому классу.

15. Оценка размера кристаллитов по SEM-изображению

Создайте 50 изображений 32×32 с кругами разного радиуса (3–15 пикселей). Постройте CNN-регрессию: Conv2D(8, 3) → Conv2D(16, 3) →

Flatten $\rightarrow$ Dense(32) $\rightarrow$ Dense(1, linear). Обучите предсказывать радиус кристаллита. Выведите MAE в пикселях.

16. Классификация типов частиц (сферические, игольчатые, пластинчатые)

Сгенерируйте 90 изображений 64×64 частиц трех форм. Постройте CNN: Conv2D(16, 5) $\rightarrow$ MaxPooling2D(2) $\rightarrow$ Conv2D(32, 5) $\rightarrow$ MaxPooling2D(2) $\rightarrow$ Flatten $\rightarrow$ Dense(3, softmax). Обучите классифицировать форму частиц. Выведите F1-score.

17. Определение цвета раствора по изображению (RGB $\rightarrow$ класс)

Создайте 60 цветных изображений 16×16 , раскрашенных в один из 4 цветов: красный (pH низкий), желтый (средний), зеленый (нейтральный), синий (щелочной). Постройте CNN: Conv2D(8, 3) $\rightarrow$ Flatten $\rightarrow$ Dense(4, softmax). Обучите классифицировать цвет (pH). Выведите confusion matrix.

18. Поиск границ кристаллов на изображении (задача сегментации)

Создайте 40 изображений 32×32 с кругами и их маски (граница). Постройте сверточную сеть: Conv2D(8, 3) $\rightarrow$ Conv2D(16, 3) $\rightarrow$ Conv2D(1, 3, activation='sigmoid') на выходе (пиксельная классификация). Обучите предсказывать границы кристаллов. Визуализируйте предсказанную маску для тестового изображения.

19. Классификация типов спектроскопических карт (2D COSY, TOCSY, HSQC)

Создайте по 20 синтетических 2D-карт 40×40 для трех типов ЯМР-экспериментов. Постройте CNN: Conv2D(8, 5) $\rightarrow$ Conv2D(16, 5) $\rightarrow$ GlobalMaxPooling2D $\rightarrow$ Dense(3, softmax). Обучите классифицировать тип эксперимента. Выведите accuracy.

20. Детектирование положения пятна на ТСХ-пластинке

Создайте 50 изображений 50×50 с одним круглым пятном в случайной позиции. Постройте CNN-регрессию: Conv2D(8, 3) $\rightarrow$ MaxPooling2D(2) $\rightarrow$ Flatten $\rightarrow$ Dense(64) $\rightarrow$ Dense(2, linear) (координаты x,y). Обучите предсказывать координаты пятна. Выведите среднюю евклидову ошибку в пикселях.

21. Сравнение MaxPooling и AveragePooling для спектров

Для задачи классификации хроматограмм (задача 2) постройте две одинаковые CNN: одну с MaxPooling1D, другую с AveragePooling1D. Сравните accuracy и время обучения. Объясните, в каком случае какой пулинг лучше.

22. Влияние размера пулинга на качество классификации

Для задачи распознавания формы кристаллов (задача 11) постройте CNN с MaxPooling2D(2×2), с MaxPooling2D(4×4) и без пулинга. Сравните accuracy и количество параметров. Сделайте вывод о влиянии субдискретизации.

23. Global pooling вместо Flatten

Для задачи классификации кристаллических решеток (задача 12) сравните две модели: с Flatten + Dense и с GlobalAveragePooling2D + Dense. Сравните количество параметров и склонность к переобучению. Какая модель лучше работает на маленьком датасете (200 изображений)?

24. Свертка с субдискретизацией (stride=2) вместо пулинга

Для задачи обнаружения дефектов (задача 13) постройте две CNN: первая с MaxPooling2D после свертки, вторая с Conv2D(strides=2) без отдельного пулинга. Сравните ассурасу и скорость обучения. Объясните разницу.

25. Многомасштабная обработка через разные размеры пулинга

Создайте CNN с двумя параллельными путями: один с MaxPooling2D(2), другой с MaxPooling2D(4), затем объедините их через Concatenate и добавьте Dense слои. Для задачи классификации частиц (задача 11) сравните эту модель с обычной CNN. Улучшилось ли качество?

26. Простая CNN с одним сверточным слоем

Для задачи классификации спектров (задача 1) постройте минимальную CNN: Conv1D(4, kernel_size=3) → GlobalMaxPooling1D → Dense(1, sigmoid). Обучите и оцените ассурасу. Достаточно ли одного сверточного слоя для обнаружения пика?

27. Два сверточных слоя с разным размером ядра

Для задачи определения числа компонентов (задача 3) постройте CNN: Conv1D(8, 3) → Conv1D(16, 7) → GlobalMaxPooling1D → Dense(3, softmax). Сравните с сетью, где оба ядра имеют размер 3. Какая архитектура лучше справляется с разномасштабными пиками?

28. CNN с dropout для регуляризации

Для задачи классификации ЯМР-спектров (задача 10) постройте CNN с Dropout(0.3) после каждого сверточного слоя и перед Dense. Обучите на 100 спектров. Сравните ассурасу на тесте с моделью без dropout. Уменьшилось ли переобучение?

29. CNN с batch normalization

Для задачи распознавания формул (задача 14) добавьте BatchNormalization после каждого Conv2D слоя. Обучите и постройте график loss. Сравните скорость сходимости с моделью без нормализации.

30. Остаточный блок (один skip connection) в пределах 5 слоев

Постройте CNN для задачи 15 (оценка радиуса) с архитектурой: Conv2D(8,3) → Conv2D(8,3) → Add(входной сигнал) → GlobalAvgPooling2D → Dense(1). Обучите и сравните с аналогичной сетью без skip connection. Улучшилась ли точность?

31. CNN с увеличенным количеством фильтров

Для задачи 11 (форма частиц) постройте CNN с количеством фильтров: 8, 16, 32. Обучите и сравните ассурасу. Выясните, есть ли переобучение при 32 фильтрах. Сделайте вывод о выборе количества фильтров.

32. CNN + LSTM для последовательности спектров (временная динамика)

Создайте 400 последовательностей из 5 спектров (каждый 40 точек) — реакция во времени. Постройте CNN-LSTM: TimeDistributed(Conv1D(8,3) + MaxPooling1D) → LSTM(16) → Dense(1) для предсказания константы скорости. Обучите и сравните с MLP на тех же данных.

33. CNN для регрессии с несколькими выходами

Для задачи 4 (поиск максимума пика) добавьте второй выход — предсказание ширины пика. Постройте CNN с двумя головами (Dense(1) для центра, Dense(1) для ширины) и общей сверточной базой. Обучите с двумя функциями потерь (MSE + MSE). Выведите ошибки для обоих параметров.

34. Использование предобученных весов (имитация) на химических изображениях

Для задачи 11 (кристаллы) сначала обучите CNN на 100 синтетических изображениях геометрических фигур (перенос стиля), затем заморозьте сверточные слои и дообучите классификатор на 30 кристаллах. Сравните ассигасу с обучением с нуля.

35. Визуализация сверточных фильтров

Обучите CNN для задачи 12 (кристаллические решетки). После обучения визуализируйте веса первых 8 фильтров первого сверточного слоя (как изображения 3×3). Есть ли фильтры, похожие на детекторы краев, пятен или текстур? Сделайте вывод.

36. Анализ 2D-спектроскопических карт (интенсивность → признак)

Создайте 50 2D-карт 30×30 с одним пиком разной интенсивности. Постройте CNN-регрессию: Conv2D(8, 3) → MaxPooling2D(2) → Conv2D(16, 3) → Flatten → Dense(1, linear). Обучите предсказывать интенсивность пика. Выведите MSE.

37. Классификация 2D-карт распределения элементов (EDS)

Сгенерируйте 60 2D-карт 20×20, показывающих распределение 3 элементов (каждый — свое гауссово пятно). Постройте CNN: Conv2D(8, 3) → Conv2D(16, 3) → GlobalMaxPooling2D → Dense(3, softmax) для классификации, какой элемент преобладает в центре. Выведите ассигасу.

38. Обнаружение трещин на 2D-изображениях катализатора

Создайте 50 изображений 64×64 катализатора (40 — целые, 10 — с трещинами). Постройте CNN: Conv2D(16, 3) → MaxPooling2D(2) → Conv2D(32, 3) → GlobalAveragePooling2D → Dense(1, sigmoid). Обучите детектировать трещины. Выведите ROC-кривую.

39. Определение пористости по 2D-микрофотографии

Создайте 40 изображений 32×32 бинарных (поры — белые, матрица — черные) с известной пористостью (0.1–0.6). Постройте CNN-регрессию: Conv2D(8, 5) → Conv2D(16, 5) → Flatten → Dense(1, sigmoid) (выход нормирован от 0 до 1). Обучите предсказывать пористость. Выведите MAE.

40. Перенос стиля с микрофотографий одного кристалла на другой

Возьмите 20 изображений кристаллов типа А (квадратные) и 20 типа В (гексагональные). Обучите простую CNN для переноса стиля (имитация CycleGAN в упрощенном виде — два автоэнкодера). Сгенерируйте квадратный кристалл в стиле гексагонального. Визуализируйте результат.

Сделать скриншоты результата работы примера и сохранить их в отчете. Рекомендуемый объем — 2-3 страницы (2-5 слайдов)

8. Необязательный пункт отчета для получения повышенных баллов. Дополнительные фотографии процесса выполнения лабораторной работы (ознакомление с методичкой, литературой, чтение сайтов из интернета, написание отчета — 5-10 фотографий (5-10 слайдов) с пояснениями кто на фото и что именно делает студент.

9. Необязательный пункт отчета для получения повышенных баллов. Публичная защита отчета по лабораторной работе.

В данном пункте привести ссылку на Rutube-видео или иной публичный видео-ресурс, на котором записана защита-выступление отчета или презентации по лабораторной работе лично студентом. Рекомендуемое время видеоролика — 5-10 минут.

Критерии оценивания отчета:

1. Выполнены пункты 1-4 отчета - оценка «удовлетворительно».
2. Выполнены пункты 1-6 отчета - оценка «хорошо».
3. Выполнены пункты 1-7 отчета - оценка «отлично».
4. Отсутствие пункта 4 отчета понижает оценку на 1 балл. (Можно заменить по желанию студента пункт 4 устными ответами на все контрольные вопросы без понижения оценки).
5. Отсутствие нумерации пунктов отчета и их названий понижает оценку на 1 балл. (Можно не детализировать в названии пункта, что именно должно быть в данном пункте отчета, но само название должно быть!)
6. «Перемешивание» (идут не по порядку) пунктов отчета понижает оценку на 1 балл.
7. Наличие любого из пунктов 8, 9 или их вместе повышает оценку на 1 балл (но не выше «отлично»).

Лабораторная работа 8

Тема: «Рекуррентные нейросети»

Цель: «Рассмотреть основные возможности рекуррентных нейросетей в рамках изучаемой дисциплины»

Знания, умения и владения, приобретаемые обучающимся в результате освоения темы

Знать: методы создания и обучения искусственных рекуррентных нейронных сетей.

Уметь: создавать программы моделирования рекуррентных нейронных сетей на языке программирования Python.

Владеть: навыками создания вычислительных программ на языке программирования Python, моделирующих работу искусственных нейронных сетей.

Теоретическая часть:

см. список рекомендованной литературы

Оборудование и материалы: см. приложение 1 методических рекомендаций по выполнению ЛР.

Техника безопасности: см. приложение 2 методических рекомендаций по выполнению ЛР.

Задания

1. Ознакомиться с теоретической частью лабораторной работы (найти ответы на контрольные вопросы).
2. Сформировать и сохранить пункты 1-4 отчета.
3. Выполнить практическую часть лабораторной работы.
 - войти в созданную в ЛР1 личную учетную запись google colab
 - создать и переименовать новый Notebook
 - ввести и отладить пробные примеры
 - сохранить результат работы пробных примеров
4. Сформировать и сохранить пункты 5-7 отчета.
5. По желанию сформировать и сохранить пункты 8-9 отчета
6. Защитить (выслать на проверку) отчет преподавателю

Содержимое отчета по лабораторной работе

Лабораторная работа считается выполненной студентом и засчитывается преподавателем (с зачислением баллов пропорционально выполненным пунктам отчета, т. е. необязательно, чтобы отчет содержал все пункты) при предоставлении в электронном виде к защите (бумажная распечатка не требуется!!!) аккуратно оформленного файла-отчета в формате .pdf (презентация), .odt (OpenOffice, LibreOffice) или .doc (Microsoft Office), именованный по шаблону (при несоответствии работа не допускается к защите):

Файл-отчет должен иметь следующие элементы-слайды презентации или пункты текстового отчета-документа в указанном порядке (с сохранением их номера в случае пропуска некоторых пунктов):

1. Титульный лист, содержащий название дисциплины, тему и номер лабораторной работы, ФИО студента, курс, группу.

Форму и дизайн титульного листа студент определяет самостоятельно, ориентируясь на удобство восприятия информации на титульном листе.

2. Отметка об ознакомлении с техникой безопасности (правилами поведения в компьютерном классе (аудитории, лаборатории)).

В данном пункте необходимо наличие записи «С правилами техники безопасности при выполнении лабораторной работы и поведения в компьютерном классе (аудитории, лаборатории) ознакомлен, обязуюсь выполнять _____ (ФИО), _____ дата».

3. Постановка задачи (полное условие в соответствии с выбранным вариантом).

4. Реферативный обзор информационных (интернет) источников по заданию.

4.1. рекуррентные нейросети

4.2. нейросеть Хопфилда

4.3. нейросеть Хэминга

4.4. нейросеть Джордана

4.5. нейросеть Элмана

4.6. LSTM-ячейка

4.7. GRU-ячейка

4.8. создание рекуррентной нейросети при помощи библиотеки Keras

4.9. обучение рекуррентной нейросети при помощи библиотеки Keras

Необходимо указать по порядку информационные источники и их содержимое по заданию. Допустимо использовать текст, гиперссылки, таблицы, рисунки и т. д. любую информацию, облегчающую просмотр и защиту презентации-отчета. Разрешено использование только открытых публичных источников информации, не требующих передачи авторских прав или каких-либо иных отчислений. Если информация по какому-либо из разделов реально отсутствует в публичном доступе или засекречена, то напечатать «информация засекречена» с указанием ведомства или организации, засекретившей данную информацию. Рекомендуемый объем реферативного обзора — 3-10 страниц (5-10 слайдов).

5. Войти в личную учетную запись в системе google colab, созданную в ЛР1. В данном пункте сделать экранные снимки процесса успешного входа с подписями, расшифровывающими ФИО автора (подписи реальных ФИО можно добавить на скриншот в графическом редакторе уже после процесса регистрации). Рекомендуемый объем — 1-2 страницы (1-2 слайда)

6. Создать новый Notebook в системе google colab, и переименовать его по шаблону: AI_LR8_(фамилия).ipynb
Сделать скриншоты. Вставить скриншоты в отчет. Рекомендуемый объем — 1-2 страниц (1-2 слайда)

7. Ввести и отладить в системе google colab программу решения основной задачи лабораторной работы, выбранной по варианту:

1. Предсказание следующего значения pH в процессе титрования

У вас есть 500 последовательностей титрования: каждые 2 секунды записывается pH (всего 50 шагов). Обучите SimpleRNN предсказывать следующее значение pH по предыдущим 10 значениям. Используйте архитектуру: SimpleRNN(16, return_sequences=False) → Dense(1). Выведите MSE на тесте.

2. Классификация кинетических кривых по порядку реакции

Сгенерируйте 600 последовательностей (40 шагов): концентрация реагента во времени для нулевого, первого и второго порядков. Постройте RNN: SimpleRNN(8) → Dense(3, softmax). Обучите классифицировать порядок реакции. Выведите accuracy и confusion matrix.

3. Восстановление пропущенных значений в хроматограмме

Создайте 400 хроматограмм (60 точек), в каждой случайно удалите 5 подряд идущих значений. Постройте двунаправленную RNN (Bidirectional(SimpleRNN(8))) → Dense(1) для предсказания пропущенного участка. Оцените MSE восстановления.

4. Предсказание концентрации продукта в реакторе по временному ряду

Создайте 300 последовательностей (30 шагов) с признаками: температура, давление, расход реагента. Цель — концентрация продукта на следующем шаге. Обучите SimpleRNN(16) → Dense(1). Сравните с MLP, который использует только текущие значения.

5. Раннее обнаружение точки эквивалентности в титровании

У вас есть 500 кривых титрования (40 шагов). Обучите RNN: SimpleRNN(12) → Dense(1, sigmoid) предсказывать, достигнута ли точка эквивалентности (1) или нет (0) на каждом шаге. Выведите precision и recall.

6. Нейросеть Элмана для моделирования колебательной реакции Белоусова-Жаботинского

Сгенерируйте 200 временных рядов концентрации катализатора (80 шагов) с колебательным поведением. Обучите сеть Элмана (SimpleRNN с обратной связью от выхода к входу) предсказывать следующий шаг. Визуализируйте предсказанный и реальный ряды.

7. Классификация типов хроматограмм (нормальная, с хвостом, с размытием)

Создайте по 200 хроматограмм (50 точек) трех типов. Обучите RNN: SimpleRNN(10) → Dense(3, softmax). Сравните accuracy с CNN из

предыдущего урока. Какая сеть лучше работает с временными последовательностями?

8. Предсказание выхода реакции по последовательности температур

Создайте 400 последовательностей: 20 шагов, на каждом шаге температура и текущий выход. Обучите RNN (SimpleRNN(8) → Dense(1)) предсказывать выход на последнем шаге. Выведите R^2 .

9. Детектирование аномалий в работе хроматографа по сигналу

У вас есть 800 временных рядов сигнала детектора (60 шагов), из которых 10% содержат аномалии (резкий выброс). Обучите автоэнкодер на основе RNN: SimpleRNN(6, return_sequences=True) → TimeDistributed(Dense(1)). Ошибка восстановления выше для аномалий. Проверьте на тесте.

10. Нейросеть Хопфилда для восстановления зашумленной спектральной кривой

Создайте 100 спектров (30 точек) и обучите сеть Хопфилда (реализуйте самостоятельно через веса Хебба) запоминать эталонные спектры. Подайте зашумленный спектр (50% битов инвертированы). Восстановите исходный. Оцените долю правильно восстановленных точек.

11. Предсказание константы скорости по длинной кинетической кривой

Сгенерируйте 600 кинетических кривых (100 шагов) с экспоненциальным затуханием, k от 0.02 до 0.2. LSTM(16) → Dense(1). Обучите предсказывать k . Сравните с SimpleRNN — какая сеть лучше запоминает зависимость на длинном горизонте?

12. Прогнозирование температуры в экзотермической реакции

У вас есть 500 временных рядов температуры реактора (80 шагов), где реакция начинается случайно между 20 и 40 шагами. LSTM(20) → Dense(1) предсказывает температуру на следующем шаге. Вычислите MAE.

13. Классификация спектральных временных рядов (ЯМР-релаксация)

Создайте 400 последовательностей (50 шагов) T1-релаксации для трех типов тканей (аналог химических сред). LSTM(8) → Dense(3, softmax). Выведите F1-score.

14. Предсказание момента выхода пика в хроматографии

У вас есть 700 хроматограмм (50 шагов). LSTM(12) → Dense(1) предсказывает, на каком шаге (от 20 до 40) произойдет максимум пика. Выведите MAE в шагах.

15. Восстановление пропущенных участков в длинных кинетических кривых

Создайте 500 кривых (100 шагов), удалите непрерывный участок длиной 15 шагов. Двунаправленный LSTM(16) → Dense(1) предсказывает пропущенные значения. Сравните MSE с SimpleRNN.

16. Классификация типов химических осцилляторов (BZ, Briggs-Rauscher)

Сгенерируйте 300 временных рядов для каждого типа осциллятора (80 шагов). LSTM(20) → Dense(2, softmax). Выведите ROC-кривую и AUC.

17. Предсказание остаточной концентрации реагента по истории процесса

Создайте 600 последовательностей: 40 шагов с признаками [концентрация A, концентрация B, температура]. LSTM(10, return_sequences=False) → Dense(1) предсказывает конечную концентрацию. Выведите R^2 .

18. LSTM-автоэнкодер для обнаружения неисправностей в реакторе

Обучите LSTM-автоэнкодер (LSTM(8, return_sequences=True) → TimeDistributed(Dense(1))) на 800 нормальных временных рядах (60 шагов). Подайте 200 рядов (100 нормальных, 100 с неисправностью). Ошибка восстановления выше для неисправных. Выведите ассигуру детектирования.

19. Предсказание времени полупревращения по кинетической кривой

Сгенерируйте 500 кривых (70 шагов) первого порядка с разным $T_1 / 2$ (10–40 шагов). LSTM(16) → Dense(1) предсказывает $T_1 / 2$. Выведите MAE в шагах.

20. Обработка последовательностей разной длины (padding)

Создайте 400 кинетических кривых разной длины (40–80 шагов). Добавьте паддинг (заполните нулями до 80). Обучите LSTM(12, mask_zero=True) → Dense(1) предсказывать начальную концентрацию. Объясните, зачем нужна маскировка нулей.

21. Сравнение LSTM и GRU для предсказания pH в титровании

Для задачи 1 обучите две одинаковые сети: LSTM(16) и GRU(16). Сравните время обучения и MSE. Какая сеть быстрее и точнее для этой задачи?

22. GRU для классификации хроматограмм

Для задачи 7 замените SimpleRNN на GRU(10) → Dense(3, softmax). Сравните ассигуру. GRU должна быть лучше на длинных последовательностях.

23. Двухнаправленная GRU для предсказания концентрации

Создайте 500 последовательностей (40 шагов) с признаками [t, C_A, C_B, T]. Постройте Bidirectional(GRU(12)) → Dense(1) для предсказания C_A на следующем шаге. Сравните с однонаправленной GRU.

24. GRU с возвратом последовательностей (return_sequences=True)

Для задачи 15 (восстановление пропусков) используйте GRU(16, return_sequences=True) → TimeDistributed(Dense(1)). Обучите предсказывать всю последовательность целиком. Выведите MSE.

25. Стэкинг двух GRU-слоев

Для задачи 11 (предсказание k) постройте GRU(16, return_sequences=True) → GRU(8) → Dense(1). Сравните с однослойной GRU. Улучшилось ли качество? Не появилось ли переобучение?

26. GRU для реальных временных рядов (имитация)

Импортируйте встроенный датасет `air_passengers` (не химический, как пример). Примените GRU(10) для предсказания следующего значения. Затем замените данные на химические (температура реактора). Оцените MAE.

27. Регуляризация GRU через Dropout и RecurrentDropout

Для задачи 17 добавьте параметры `dropout=0.2`, `recurrent_dropout=0.2` в GRU-слой. Обучите на 2000 последовательностях. Сравните переобучение (разрыв между train и val loss) с моделью без регуляризации.

28. GRU для многомерных выходов (предсказание всей кривой)

Создайте 400 последовательностей: вход — первые 20 шагов (концентрация), выход — следующие 30 шагов. GRU(20, `return_sequences=True`) → TimeDistributed(Dense(1)). Обучите предсказывать продолжение кинетической кривой. Визуализируйте результат.

29. Визуализация состояния GRU-ячеек

Обучите GRU(4) для задачи 2. Извлеките состояния ячеек (`outputs`, `state`) после каждого шага. Визуализируйте 4 канала состояния на одном графике. Видно ли, как ячейки «запоминают» форму кинетической кривой?

30. GRU с attention-механизмом (упрощенным)

Для задачи 14 добавьте простой механизм внимания: вычислите среднее взвешенное выходов GRU по времени (через Dense(1, softmax) по временной оси). Сравните с обычной GRU.

31. Нейросеть Хопфилда для запоминания эталонных спектров

Создайте 5 эталонных спектров (20 точек, бинарные: 0 и 1). Обучите сеть Хопфилда по правилу Хебба. Подайте искаженный спектр (20% битов перевернуты). Восстановите ближайший эталон. Оцените процент успешного восстановления.

32. Сравнение Хопфилда и LSTM для восстановления зашумленных данных

Создайте 100 кинетических кривых (30 точек, бинаризованные). Сравните точность восстановления при 10%, 30%, 50% шума для сети Хопфилда и LSTM-автоэнкодера. Какая сеть лучше работает при высоком шуме?

33. Нейросеть Хэмминга для классификации химических сигнатур

Создайте 4 эталонных вектора (15 бит), описывающих разные типы функциональных групп. Реализуйте сеть Хэмминга (слой Хопфилда + слой MAXNET). Подайте зашумленные сигнатуры. Классифицируйте, к какому эталону они ближе всего.

34. Ограничение емкости сети Хопфилда

Обучите сеть Хопфилда на 3, 5, 7, 10 эталонных спектрах (20 точек). Для каждого случая проверьте, все ли эталоны являются стабильными состояниями (восстанавливаются без ошибок). Найдите максимальную емкость (правило: $\sim 0.14N$).

35. Биологическая интерпретация: Хопфилд как модель химической памяти

Создайте 4 паттерна химических соединений (векторы из 25 бит, где каждый бит — наличие функциональной группы). Обучите сеть Хопфилда.

Покажите, как сеть «вспоминает» полную структуру по частичной (например, по 50% битов).

36. Хопфилд с непрерывными значениями для спектров

Модифицируйте сеть Хопфилда для работы с непрерывными значениями (функция активации $\tanh$). Запомните 3 спектра (50 точек, значения от -1 до 1). Подайте искаженную версию. Сравните с дискретной версией.

37. MAXNET для выбора ближайшего эталона в сети Хэмминга

Реализуйте MAXNET (слой с латеральным торможением) для выбора победителя среди выходов сети Хэмминга. Проверьте на 4 эталонах, что MAXNET всегда оставляет активным только один нейрон (с максимальным входом).

38. Сеть Хэмминга для поиска ближайшей формулы

Создайте 6 эталонных химических формул (закодированных как 20-битные векторы). Подайте новую формулу. Используйте сеть Хэмминга для нахождения ближайшей эталонной формулы (минимальное расстояние Хэмминга). Выведите результат.

39. Влияние ортогональности эталонов на качество Хопфилда

Создайте 4 попарно ортогональных эталонных вектора (20 бит) и 4 неортогональных (с перекрытием). Сравните вероятность ошибки восстановления при 20% шума. Ортогональные эталоны запоминаются лучше.

40. Применение Хопфилда для исправления ошибок в ЯМР-спектрах

Создайте 8 эталонных спектров (40 бит). Подайте спектр с тремя случайно перевернутыми битами (ошибки измерения). Запустите сеть Хопфилда до сходимости. Сравните исправленный спектр с исходным эталоном. Оцените долю исправленных ошибок.

41. Sequential API для SimpleRNN

Для задачи 1 создайте модель `Sequential: Input(shape=(10,1)) → SimpleRNN(8) → Dense(1)`. Выведите `model.summary()`. Обучите и сохраните веса.

42. Functional API для RNN с двумя входами

Создайте модель для предсказания выхода реакции: один вход — временной ряд температуры (последовательность), другой — скаляр начальной концентрации. Объедините через `Concatenate` после RNN. Обучите.

43. Stacked RNN (два слоя SimpleRNN)

Для задачи 11 постройте модель: `SimpleRNN(16, return_sequences=True) → SimpleRNN(8) → Dense(1)`. Сравните с однослойной RNN. Какая модель лучше на длинных последовательностях (100 шагов)?

44. Callback EarlyStopping для RNN

Для задачи 12 используйте `EarlyStopping(monitor='val_loss', patience=5)` при обучении LSTM. Постройте график `loss`. На какой эпохе остановилось обучение?

45. Сохранение и загрузка обученной RNN

Обучите LSTM для задачи 13. Сохраните модель в файл (`model.save`). Загрузите ее заново и проверьте на тестовой выборке. Результаты совпадают?

46. RNN с masking для последовательностей разной длины

Создайте 400 последовательностей с длинами от 30 до 70. Примените padding и Masking() слой. Обучите LSTM(8) → Dense(3, softmax). Объясните, зачем нужен Masking.

47. Stateful RNN (состояние сохраняется между батчами)

Для задачи 4 обучите Stateful=True LSTM(8, stateful=True). Вручную сбрасывайте состояния после каждой эпохи. Сравните качество с обычной (stateless) LSTM.

48. RNN для прогнозирования нескольких шагов вперед

Создайте модель: LSTM(20, return_sequences=True) → TimeDistributed(Dense(1)) для предсказания следующих 10 шагов кинетической кривой по предыдущим 20 шагам. Обучите. Визуализируйте прогноз на 10 шагов.

49. Использование Embedding перед RNN для категориальных данных

Создайте 500 последовательностей из 20 символов (каждый символ — тип реагента: A,B,C,D). LSTM после Embedding(4,4) → Dense(1) предсказывает выход реакции. Обучите.

50. Визуализация градиентов в RNN (проблема затухания)

Обучите SimpleRNN(16) для задачи с длинной последовательностью (200 шагов). С помощью callback вычислите норму градиента на первых и последних временных шагах. Сравните с LSTM(16). У LSTM градиенты затухают меньше.

Сделать скриншоты результата работы примера и сохранить их в отчете. Рекомендуемый объем — 2-3 страницы (2-5 слайдов)

8. Необязательный пункт отчета для получения повышенных баллов. Дополнительные фотографии процесса выполнения лабораторной работы (ознакомление с методичкой, литературой, чтение сайтов из интернета, написание отчета — 5-10 фотографий (5-10 слайдов) с пояснениями кто на фото и что именно делает студент.

9. Необязательный пункт отчета для получения повышенных баллов. Публичная защита отчета по лабораторной работе.

В данном пункте привести ссылку на Rutube-видео или иной публичный видео-ресурс, на котором записана защита-выступление отчета или презентации по лабораторной работе лично студентом. Рекомендуемое время видеоролика — 5-10 минут.

Критерии оценивания отчета:

1. Выполнены пункты 1-4 отчета - оценка «удовлетворительно».
2. Выполнены пункты 1-6 отчета - оценка «хорошо».
3. Выполнены пункты 1-7 отчета - оценка «отлично».

4. Отсутствие пункта 4 отчета понижает оценку на 1 балл. (Можно заменить по желанию студента пункт 4 устными ответами на все контрольные вопросы без понижения оценки).
5. Отсутствие нумерации пунктов отчета и их названий понижает оценку на 1 балл. (Можно не детализировать в названии пункта, что именно должно быть в данном пункте отчета, но само название должно быть!)
6. «Перемешивание» (идут не по порядку) пунктов отчета понижает оценку на 1 балл.
7. Наличие любого из пунктов 8, 9 или их вместе повышает оценку на 1 балл (но не выше «отлично»).

Лабораторная работа 9

Тема: «Нейронные сети трансформеры»

Цель: «Рассмотреть основные возможности нейронных сетей трансформеров в рамках изучаемой дисциплины»

Знания, умения и владения, приобретаемые обучающимся в результате освоения темы

Знать: основные методы создания и обучения нейросетей-трансформеров.

Уметь: создавать простые программы моделирования нейросетей-трансформеров на языке программирования Python.

Владеть: навыками создания вычислительных программ на языке программирования Python, моделирующих работу искусственных нейронных сетей.

Теоретическая часть:

см. список рекомендованной литературы

Оборудование и материалы: см. приложение 1 методических рекомендаций по выполнению ЛР.

Техника безопасности: см. приложение 2 методических рекомендаций по выполнению ЛР.

Задания

1. Ознакомиться с теоретической частью лабораторной работы (найти ответы на контрольные вопросы).
2. Сформировать и сохранить пункты 1-4 отчета.
3. Выполнить практическую часть лабораторной работы.
 - войти в созданную в ЛР1 личную учетную запись google colab
 - создать и переименовать новый Notebook
 - ввести и отладить пробные примеры
 - сохранить результат работы пробных примеров
4. Сформировать и сохранить пункты 5-7 отчета.
5. По желанию сформировать и сохранить пункты 8-9 отчета
6. Защитить (выслать на проверку) отчет преподавателю

Содержимое отчета по лабораторной работе

Лабораторная работа считается выполненной студентом и засчитывается преподавателем (с зачислением баллов пропорционально выполненным пунктам отчета, т. е. необязательно, чтобы отчет содержал все пункты) при предоставлении в электронном виде к защите (бумажная распечатка не требуется!!!) аккуратно оформленного файла-отчета в формате .pdf (презентация), .odt (OpenOffice, LibreOffice) или .doc (Microsoft Office), именованный по шаблону (при несоответствии работа не допускается к защите):

AI_LR9_(фамилия).pdf

Файл-отчет должен иметь следующие элементы-слайды презентации или пункты текстового отчета-документа в указанном порядке (с сохранением их номера в случае пропуска некоторых пунктов):

1. Титульный лист, содержащий название дисциплины, тему и номер лабораторной работы, ФИО студента, курс, группу.

Форму и дизайн титульного листа студент определяет самостоятельно, ориентируясь на удобство восприятия информации на титульном листе.

2. Отметка об ознакомлении с техникой безопасности (правилами поведения в компьютерном классе (аудитории, лаборатории)).

В данном пункте необходимо наличие записи «С правилами техники безопасности при выполнении лабораторной работы и поведения в компьютерном классе (аудитории, лаборатории) ознакомлен, обязуюсь выполнять _____ (ФИО), _____ дата».

3. Постановка задачи (полное условие в соответствии с выбранным вариантом).

4. Реферативный обзор информационных (интернет) источников по заданию.

4.1. общая архитектура нейросети-трансформера

4.2. блок кодер в нейросети трансформере

4.3. блок декодер в нейросети трансформере

4.4. позиционное кодирование

4.5. механизм самовнимания

4.6. создание нейросети-трансформер при помощи библиотек numpy и keras

4.7. обучение нейросети-трансформер при помощи библиотек numpy и keras

Необходимо указать по порядку информационные источники и их содержимое по заданию. Допустимо использовать текст, гиперссылки, таблицы, рисунки и т. д. любую информацию, облегчающую просмотр и защиту презентации-отчета. Разрешено использование только открытых публичных источников информации, не требующих передачи авторских прав или каких-либо иных отчислений. Если информация по какому-либо из разделов реально отсутствует в публичном доступе или засекречена, то напечатать «информация засекречена» с указанием ведомства или организации, засекретившей данную информацию. Рекомендуемый объем реферативного обзора — 3-10 страниц (5-10 слайдов).

5. Войти в личную учетную запись в системе google colab, созданную в ЛР1. В данном пункте сделать экранные снимки процесса успешного входа с подписями, расшифровывающими ФИО автора (подписи реальных ФИО можно добавить на скриншот в графическом редакторе уже после процесса регистрации). Рекомендуемый объем — 1-2 страницы (1-2 слайда)

6. Создать новый Notebook в системе google colab, и переименовать его по шаблону: AI_LR9_(фамилия).ipynb

Сделать скриншоты. Вставить скриншоты в отчет. Рекомендуемый объем — 1-2 страниц (1-2 слайда)

7. Ввести и отладить в системе google colab программу решения основной задачи лабораторной работы, выбранной по варианту:

1. Визуализация внимания в спектре из 8 точек

Создайте 200 спектров длины 8. Каждый спектр — два гауссовых пика с центрами на позициях 2 и 6. Возьмите первый спектр, вычислите матрицу внимания 8×8 через `MultiHeadAttention(num_heads=1, key_dim=4)`. Визуализируйте heatmap. Определите, у какой пары позиций наибольший вес внимания.

2. Классификация кинетических кривых (8 шагов, 3 класса)

Сгенерируйте 300 последовательностей длины 8: класс 0 — линейное убывание, класс 1 — экспонента, класс 2 — гипербола. Добавьте шум. Модель: `Input(shape=(8,1)) → MultiHeadAttention(1,4) → GlobalAvgPool1D → Dense(4, ReLU) → Dense(3, softmax)`. Выведите accuracy на тесте (60 образцов).

3. Обнаружение корреляции между позициями 3 и 6 в хроматограмме (10 точек)

Сгенерируйте 400 хроматограмм длины 10. В 50% образцов пик на позиции 3 коррелирует с пиком на позиции 6 (оба >0.5 или оба <0.5). Обучите только слой внимания (без классификации), извлеките веса. Вычислите средний вес внимания между позициями 3 и 6 для коррелированных и некоррелированных образцов.

4. Предсказание 9-й точки спектра по первым 8 (регрессия)

Сгенерируйте 500 спектров длины 9 (один гауссов пик с центром 5). Обучите: вход $(8,1) \rightarrow \text{MultiHeadAttention}(1,4) \rightarrow \text{Dense}(1)$. Вычислите MSE на тесте (100 образцов). Сравните с `Dense(8) → Dense(1)` без внимания.

5. Сглаживание зашумленного спектра (12 точек)

Создайте 300 чистых спектров длины 12 (один гауссов пик с центром 6). Добавьте шум $N(0,0.1)$. Модель: вход $(12,1) \rightarrow \text{MultiHeadAttention}(1,4) \rightarrow \text{Dense}(12)$. Обучите предсказывать чистый спектр. Вычислите MSE. Сравните со скользящим средним (окно 3).

6. Детектирование аномального скачка pH (10 шагов)

Сгенерируйте 400 последовательностей pH длины 10: нормальные — плавное изменение от 4 до 7, аномальные — скачок на шаге 5. Автоэнкодер: `MultiHeadAttention(1,4) → Dense(10)`. На тесте (80 образцов, 10% аномалий) измерьте ошибку восстановления. Порог = 95-й перцентиль ошибки на нормальных. Выведите precision и recall.

7. Синусоидальное позиционное кодирование (длина 10, размерность 4)

Создайте матрицу PE размера 10×4 по формуле: $PE(pos, 2i) = \sin(pos/10000^{(2i/4)})$. Визуализируйте heatmap. Объясните, почему у разных позиций разные кодировки.

8. Сравнение позиционного кодирования и Embedding (8 точек, 3 класса)

Сгенерируйте 300 спектров длины 8 с одним пиком. Позиция пика (2, 4 или 6) определяет класс. Обучите две модели: (1) вход + синусоидальное PE, (2) вход + Embedding(8,4). Сравните ассурасу.

9. Влияние PE на внимание (визуализация, длина 6)

Возьмите вектор длины 6: [0,0,1,2,1,0]. Пропустите через MultiHeadAttention(1,4) без PE и с PE. Визуализируйте матрицы внимания 6×6. Покажите, что без PE внимание симметрично, с PE — учитывает порядок.

10. Обучаемое PE для предсказания температуры (8 шагов)

Сгенерируйте 400 последовательностей длины 8 (температура реактора 50–100°C). Цель — температура на следующем шаге. Добавьте Embedding(8,4). Модель: MultiHeadAttention(1,4) → Dense(1). Сравните MSE с моделью без PE.

11. Инвариантность к перестановкам без PE (длина 4)

Создайте 200 последовательностей длины 4: [a,b,c] → сумма a+b+c. Переставьте входы случайно. Обучите трансформер без PE. Покажите, что предсказание не зависит от порядка (проверьте на [1,2,3] и [3,1,2]).

12. Один слой энкодера для классификации (вход 10×2, 3 класса)

Сгенерируйте 500 последовательностей длины 10 с 2 признаками (температура и давление). Класс: 0 — если средняя температура >60°C, 1 — если 40–60°C, 2 — если <40°C. Энкодер: MultiHeadAttention(1,4) → Add → LayerNorm → Dense(8, ReLU) → Dense(4) → Add → LayerNorm → GlobalAvgPool1D → Dense(3, softmax). Выведите ассурасу.

13. Влияние количества голов (1 или 2) на точность

Для задачи 12 сравните модели с 1 и 2 головами (key_dim=4). Какая дает лучшую val_ассурасу? Какая быстрее обучается (эпох до сходимости)?

14. Энкодер для восстановления пропущенных точек (вход 10, пропуск 2 и 3)

Сгенерируйте 400 спектров длины 10 (один пик с центром 5). В каждом обнулите позиции 4 и 5. Обучите энкодер: MultiHeadAttention(1,4) → LayerNorm → Dense(8) → Dense(4) → Dense(10). Вычислите MSE на восстановленных позициях 4 и 5.

15. Dropout для борьбы с переобучением (200 образцов, 6 классов)

Сгенерируйте 200 последовательностей длины 8 для классификации (6 классов). Обучите энкодер с Dropout(0.3) и без. Сравните разрыв между train и val accuracy.

16. LayerNorm до или после остаточной связи (Pre-LN vs Post-LN)

Для задачи 3.1 реализуйте два варианта. Какой сходится быстрее (меньше эпох до val_loss<0.5)?

17. Декодер для предсказания 9-го шага по первым 8

Сгенерируйте 500 кинетических кривых длины 9 (экспонента $C = \exp(-0.1 \cdot t) + \text{шум}$). Обучите декодер: вход (8,1) $\rightarrow$ MultiHeadAttention(1,4) с causal маской $\rightarrow$ Dense(1). Вычислите MSE на предсказании шага 9.

18. Авторегрессивная генерация 4 шагов вперед (всего 12 шагов)

Обучите декодер на 400 кривых длины 12. На тесте подайте первые 8 шагов, сгенерируйте шаги 9–12. Сравните MSE с учительским форсингом (когда на вход подаются реальные значения).

19. Генерация спектра по первым 5 точкам (всего 10 точек)

Сгенерируйте 400 спектров длины 10 (сумма двух пиков с центрами 3 и 7). Обучите декодер: вход (5,1) $\rightarrow$ выход (5,1) (шаги 6–10). На тесте сгенерируйте недостающие 5 точек. Визуализируйте реальный и сгенерированный спектр.

20. Causal маска для длины 5 (визуализация)

Создайте causal маску 5×5: 1 - `tf.linalg.band_part(tf.ones((5,5)), -1, 0)`. Визуализируйте. Подайте случайную последовательность в MultiHeadAttention(attention_mask=mask). Убедитесь, что элемент i видит только $j \leq i$.

21. Температура Softmax при генерации (0.5, 1.0, 2.0)

Для обученного декодера из задачи 4.3 генерируйте спектры при температурах 0.5, 1.0, 2.0. Для каждой температуры вычислите MSE (качество) и дисперсию между 5 генерациями (разнообразие). Какая температура оптимальна?

22. Трансформер: ИК (8 точек) $\rightarrow$ Раман (8 точек)

Сгенерируйте 1000 пар длины 8. ИК — сумма двух пиков (центры 2 и 6), Раман — те же пики с другими амплитудами. Энкодер: 1 слой (1 голова, `d_model=4`), декодер: 1 слой. Обучите. Вычислите MSE на тесте (200 пар).

23. Трансформер: реагенты (4 символа) $\rightarrow$ продукт (4 символа)

Словарь из 6 символов (A,B,C,D,E,F). Сгенерируйте 1000 пар: реагенты (4 символа), продукт — сдвиг каждого символа на +1 (A $\rightarrow$ B, B $\rightarrow$ C, ...). Эмбединги: Embedding(6,4). Обучите. Выведите посимвольную ассигасу.

24. Визуализация кросс-внимания (4×4 матрица)

Для обученной модели из задачи 5.2 возьмите один пример. Извлеките веса кросс-внимания. Визуализируйте матрицу 4×4. Какие символы реагентов имеют наибольшие веса для каждого символа продукта?

25. Трансформер: формула (6 символов) $\rightarrow$ спектр (10 точек)

Словарь из 8 символов. Сгенерируйте 800 пар: 6 символов $\rightarrow$ 10 точек спектра (случайная сумма 1–2 гауссовых пиков). Обучите. Вычислите MSE на тесте (160 пар). Визуализируйте реальный и предсказанный спектр для одного примера.

26. Сравнение трансформера и LSTM для перевода спектров (8 $\rightarrow$ 8 точек)

Для задачи 22 обучите LSTM: энкодер LSTM(8), декодер LSTM(8) $\rightarrow$ TimeDistributed(Dense(1)). Сравните MSE и время обучения (эпохи). Что лучше для маленьких данных?

27. Классификация типа реакции (10 шагов, 3 типа)

Сгенерируйте 600 последовательностей длины 10 (концентрации A,B). Три типа реакций заданы разными формулами. Трансформер (только энкодер, 1 слой, 1 голова) → GlobalAvgPool1D → Dense(3, softmax). Выведите confusion matrix.

28. Трансформер с warmup learning rate

Для задачи 22 реализуйте warmup: lr растет от 0 до 0.001 за первые 500 шагов. Обучите 10 эпох. Постройте график lr и loss от шага. Сравните с постоянным lr=0.001.

29. Заполнение пропусков в спектре (10 точек, пропуск позиций 4 и 7)

Сгенерируйте 500 спектров длины 10 (сумма двух пиков). В каждом обнулите позиции 4 и 7. Обучите трансформер предсказывать эти точки. Вычислите MSE только на пропущенных позициях.

30. Генерация кинетической кривой по первым 5 шагам (всего 12)

Сгенерируйте 400 кинетических кривых (экспонента с $k=0.08, 0.12, 0.16$). Обучите декодер. На тесте подайте первые 5 шагов, сгенерируйте шаги 6–12. Вычислите MSE для финальной концентрации (шаг 12).

31. Минимальный трансформер на CPU (200 образцов, 10 эпох < 1 минуты)

Соберите трансформер: d_model=4, heads=1, энкодер=1 слой, декодер=1 слой, FFN=8. Сгенерируйте 200 пар (6→6 точек). Обучите на CPU. Замерьте время. Достигнуто ли время <1 минуты? Выведите итоговую MSE.

Сделать скриншоты результата работы примера и сохранить их в отчете. Рекомендуемый объем — 2-3 страницы (2-5 слайдов)

8. Необязательный пункт отчета для получения повышенных баллов. Дополнительные фотографии процесса выполнения лабораторной работы (ознакомление с методичкой, литературой, чтение сайтов из интернета, написание отчета — 5-10 фотографий (5-10 слайдов) с пояснениями кто на фото и что именно делает студент.

9. Необязательный пункт отчета для получения повышенных баллов. Публичная защита отчета по лабораторной работе.

В данном пункте привести ссылку на Rutube-видео или иной публичный видео-ресурс, на котором записана защита-выступление отчета или презентации по лабораторной работе лично студентом. Рекомендуемое время видеоролика — 5-10 минут.

Критерии оценивания отчета:

1. Выполнены пункты 1-4 отчета - оценка «удовлетворительно».
2. Выполнены пункты 1-6 отчета - оценка «хорошо».
3. Выполнены пункты 1-7 отчета - оценка «отлично».
4. Отсутствие пункта 4 отчета понижает оценку на 1 балл. (Можно заменить по желанию студента пункт 4 устными ответами на все контрольные вопросы без понижения оценки).

5. Отсутствие нумерации пунктов отчета и их названий понижает оценку на 1 балл. (Можно не детализировать в названии пункта, что именно должно быть в данном пункте отчета, но само название должно быть!)
6. «Перемешивание» (идут не по порядку) пунктов отчета понижает оценку на 1 балл.
7. Наличие любого из пунктов 8, 9 или их вместе повышает оценку на 1 балл (но не выше «отлично»).

Список литературы

Перечень основной литературы:

1. Барский, А.Б. Введение в нейронные сети : учебное пособие / А. Б. Барский. - Введение в нейронные сети, 2021-12-05. - Электрон. дан. (1 файл). - Москва, Саратов : Интернет-Университет Информационных Технологий (ИНТУИТ), Ай Пи Ар Медиа, 2020. - 357 с. - электронный. - Книга находится в премиум-версии ЭБС IPR BOOKS. - ISBN 978-5-4497-0309-5, экземпляров неограничено

2. Павлова, А.И. Искусственные нейронные сети : учебное пособие / А. И. Павлова. - Искусственные нейронные сети, 2031-05-31. - Электрон. дан. (1 файл). - Москва : Ай Пи Ар Медиа, 2021. - 190 с. - электронный. - Книга находится в премиум-версии ЭБС IPR BOOKS. - ISBN 978-5-4497-1165-6, экземпляров неограничено

3. Ростовцев, В.С. Искусственные нейронные сети Электронный ресурс / Ростовцев В. С. : учебник для вузов. - 2-е изд., стер. - Санкт-Петербург : Лань, 2021. - 216 с. - ISBN 978-5-8114-7462-2, экземпляров неограничено

Перечень дополнительной литературы:

1. Галушкин, А.И. Теория нейронных сетей / А. И. Галушкин, Кн. 1. - М. : ИПРЖР, 2000. - 416с.

2. Искусственный интеллект . - В 3-х кн. Кн. 1 Системы общения и экспертные системы: Справочник / под ред. Д.А. Поспелова - М.: Радио и связь, 1990.

3. Искусственный интеллект . - В 3-х кн. Кн. 2 Модели и методы: Справочник / под ред. Д.А. Поспелова - М.: Радио и связь, 1990.

4. Искусственный интеллект . - В 3-х кн. Кн. 3 Программные и аппаратные средства: Справочник / под ред. Д.А. Поспелова - М.: Радио и связь, 1990.

5. Лутц М.. Изучаем Python, том I, 5-е издание. – Пер. с англ. – СПб.: Диалектика, 2019.

6. Лутц М.. Изучаем Python, том II, 5-е издание. – Пер. с англ. – СПб.: Вильямс, 2020.

7. Хайкин С. Нейронные сети: полный курс : науч.-поп. изд. / С. Хайкин ; пер. с англ. Н.Н. Куссуль, А.Ю. Шелестова. - 2-е изд. - М. : Вильямс, 2008. - 1104 с.

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

1. Курс лекций по дисциплине "Искусственный интеллект" (электронный ресурс)

2. Методические рекомендации по выполнению лабораторных работ по дисциплине "Искусственный интеллект" (электронный ресурс)

3. Методические рекомендации по организации самостоятельной работы студентов по дисциплине "Искусственный интеллект" (электронный ресурс)

Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины (модуля)

1. <http://el.ncfu.ru/> – система управления обучением ФГАОУ ВО СКФУ. Дистанционная поддержка дисциплины «Искусственный интеллект».

2. <https://colab.google.com/> - сайт создания и обработки данных машинного обучения и искусственного интеллекта.

3. <https://www.python.org> - сайт языка программирования Python.

**Оборудование, материалы и программное обеспечение, используемые
для организации и проведения лабораторных работ**

Оборудование:

Компьютерный класс (лаборатория, аудитория) с числом исправных ЭВМ по числу студентов группы плюс рабочее место преподавателя и производительностью, достаточной для освоения дисциплины. Обязательно наличие выхода в сеть Интернет и доступ к системе Google colab.

Программное обеспечение:

1. Операционная система (Windows не ниже 10 или Linux Mint не ниже 22)
2. Среда разработки и отладки программ OpenIDE с установленным плагином Python или Visual Studio Code с установленным плагином Python.
3. Текстовый процессор для оформления отчетов (Open/Libre Office.org)
4. Язык программирования Python (версия 3.10) с оболочкой Jupiter, Google colab и библиотеками NumPy, Matplotlib, TensorFlow, Keras.

Правила поведения в компьютерном классе (аудитории, лаборатории)

1. К работе в компьютерном классе допускаются лица, ознакомленные с инструкцией по технике безопасности и охране труда, с правилами поведения и размещения информационных ресурсов.
2. Работа студентов в компьютерном классе разрешается только в присутствии преподавателя (инженера, лаборанта).
3. Во время групповых занятий посторонние лица могут находиться в классе только с разрешения преподавателя.
4. **Перед началом работы необходимо:**
 - убедиться в отсутствии видимых повреждений на рабочем месте;
 - разместить на столе тетради, учебные пособия так, чтобы они не мешали работе на компьютере;
 - принять правильную рабочую позу;
 - если сеанс работы предыдущего пользователя не был завершен, завершить его;
 - ввести регистрационную информацию (при необходимости).
5. **При работе в компьютерном классе категорически запрещается:**
 - находиться в классе в верхней одежде;
 - размещать одежду и сумки на рабочих местах;
 - находиться в классе с едой и напитками;
 - класть книги, тетради и т.п. на клавиатуру;
 - располагаться сбоку или сзади от включенного монитора;
 - присоединять или отсоединять кабели, трогать разъемы, провода и розетки;
 - передвигать компьютеры;
 - открывать системный блок;
 - пытаться самостоятельно устранять неисправности в работе аппаратуры;
 - перекрывать вентиляционные отверстия на системном блоке и мониторе;
 - ударять по клавиатуре, нажимать бесцельно на клавиши;
 - удалять или перемещать чужие файлы;
 - устанавливать и запускать компьютерные игры;
 - использовать Интернет-ресурсы неучебного назначения.
6. **Находясь в компьютерном классе, необходимо:**
 - соблюдать тишину и порядок, выключать мобильные телефоны от громкой связи;
 - выполнять все требования преподавателя, инженера и лаборанта;
 - работать только под своим именем и паролем;
 - соблюдать режим работы (продолжительность непрерывной работы за компьютером не более двух часов с обязательным 10-минутным перерывом и

гимнастикой для глаз; продолжительность интенсивной работы с клавиатурой не более 30 минут с последующей гимнастикой для рук; общая продолжительность работы не более 4 часов в день);

– при появлении рези в глазах, резком ухудшении видимости, невозможности сфокусировать взгляд или навести его на резкость, появлении боли в пальцах и кистях рук, усилении сердцебиения немедленно покинуть рабочее место, сообщить о происшедшем преподавателю и обратиться к врачу;

– после окончания работы завершить все активные программы и корректно завершить сеанс;

– оставить рабочее место чистым.

7. Работая за компьютером, необходимо соблюдать правильную позу:

– расстояние от экрана до глаз 70-80 см (расстояние вытянутой руки);

– вертикально прямая спина;

– плечи опущены и расслаблены;

– ноги на полу и не скрещены;

– локти, запястья и кисти рук на одном уровне;

– локтевые, тазобедренные, коленные, голеностопные суставы под прямым углом.

8. При появлении программных ошибок или сбоях оборудования студент обязан немедленно обратиться к преподавателю (инженеру, лаборанту).

9. В случае порчи или выхода из строя оборудования компьютерного класса по вине пользователя ремонт или замена оборудования производится за счет пользователя.

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РФ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ
ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«СЕВЕРО-КАВКАЗСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»

Методические указания
для обучающихся по организации и проведению самостоятельной работы
по дисциплине
«ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ»

Ставрополь, 2026

СОДЕРЖАНИЕ

	стр
Общие положения	3
Цель и задачи самостоятельной работы	4
<u>Технологическая карта самостоятельной работы студента</u>	5
<u>Порядок выполнения самостоятельной работы студентом</u>	5
<u>Контроль самостоятельной работы студентов</u>	14
<u>Список литературы для выполнения СРС</u>	14

Общие положения

Самостоятельная работа - планируемая учебная, учебно-исследовательская, научно-исследовательская работа студентов, выполняемая во внеаудиторное (аудиторное) время по заданию и при методическом руководстве преподавателя, но без его непосредственного участия (при частичном непосредственном участии преподавателя, оставляющем ведущую роль за работой студентов).

Самостоятельная работа студентов (СРС) в ВУЗе является важным видом учебной и научной деятельности студента. Самостоятельная работа студентов играет значительную роль в рейтинговой технологии обучения.

К основным видам самостоятельной работы студентов относятся:

- формирование и усвоение содержания конспекта лекций на базе рекомендованной лектором учебной литературы, включая информационные образовательные ресурсы (электронные учебники, электронные библиотеки и др.);
- написание докладов;
- подготовка к семинарам, практическим и лабораторным работам, их оформление;
- составление аннотированного списка статей из соответствующих журналов по отраслям знаний (педагогических, психологических, методических и др.);
- выполнение учебно-исследовательских работ, проектная деятельность;
- подготовка практических разработок и рекомендаций по решению проблемной ситуации;
- выполнение домашних заданий в виде решения отдельных задач, проведения типовых расчетов, расчетно-компьютерных и индивидуальных работ по отдельным разделам содержания дисциплин и т.д.;
- компьютерный текущий самоконтроль и контроль успеваемости на базе электронных обучающих и аттестующих тестов;
- выполнение курсовых работ (проектов) в рамках дисциплин;
- выполнение выпускной квалификационной работы и др.

Методика организации самостоятельной работы студентов зависит от структуры, характера и особенностей изучаемой дисциплины, объема часов на ее изучение, вида заданий для самостоятельной работы студентов, индивидуальных качеств студентов и условий учебной деятельности.

Процесс организации самостоятельной работы студентов включает в себя следующие этапы:

1. подготовительный (определение целей, составление программы, подготовка методического обеспечения, подготовка оборудования);
4. основной (реализация программы, использование приемов поиска информации, усвоения, переработки, применения, передачи знаний, фиксирование результатов, самоорганизация процесса работы);

5. заключительный (оценка значимости и анализ результатов, их систематизация, оценка эффективности программы и приемов работы, выводы о направлениях оптимизации труда).

Самостоятельная работа по дисциплине «Искусственный интеллект» направлена на формирование следующих компетенций:

Код	Формулировка:
УК-1	Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач (УК- 1)
ОПК-3	Способен применять расчетно-теоретические методы для изучения свойств веществ и процессов с их участием с использованием современной вычислительной техники
ОПК-5	Способен понимать принципы работы современных информационных технологий и использовать их для решения задач профессиональной деятельности
ОПК-6	Способен представлять результаты своей работы в устной и письменной форме в соответствии с нормами и правилами, принятыми в профессиональном сообществе

Цель и задачи самостоятельной работы

Ведущая цель организации и осуществления СРС совпадает с целью обучения студента – формирование набора общенаучных, профессиональных и специальных компетенций будущего бакалавра по соответствующему направлению подготовки

При организации СРС важным и необходимым условием становятся формирование умения самостоятельной работы для приобретения знаний, навыков и возможности организации учебной и научной деятельности. Целью самостоятельной работы студентов является овладение фундаментальными знаниями, профессиональными умениями и навыками деятельности по профилю, опытом творческой, исследовательской деятельности. Самостоятельная работа студентов способствует развитию самостоятельности, ответственности и организованности, творческого подхода к решению проблем учебного и профессионального уровня.

Задачами СРС являются:

1. систематизация и закрепление полученных теоретических знаний и практических умений студентов;
6. углубление и расширение теоретических знаний;
7. формирование умений использовать нормативную, правовую, справочную документацию и специальную литературу;

8. развитие познавательных способностей и активности студентов: творческой инициативы, самостоятельности, ответственности и организованности;

9. формирование самостоятельности мышления, способностей к саморазвитию, самосовершенствованию и самореализации;

10. развитие исследовательских умений;

11. использование материала, собранного и полученного в ходе самостоятельных занятий на семинарах, на практических и лабораторных занятиях, при написании курсовых и выпускной квалификационной работ, для эффективной подготовки к итоговым зачетам и экзаменам.

Технологическая карта самостоятельной работы студента

Коды реализуемых компетенций	Вид деятельности студентов	Средства и технологии и оценки	Объем часов, в том числе (акад.)		
			СРС	Контактная работа с преподавателем	Всего
3 семестр					
УК-1; ОПК-3; ОПК-5; ОПК-6	Самостоятельное изучение литературы и источников	Собеседование	36	0	36
УК-1; ОПК-3; ОПК-5; ОПК-6	Подготовка к лабораторным занятиям	Защита ЛР	18	54	72
Итого за 3 семестр			54	54	108
Итого			54	54	108

Порядок выполнения самостоятельной работы студентом

4.1. Методические рекомендации по работе с учебной литературой

При работе с книгой необходимо подобрать литературу, научиться правильно ее читать, вести записи. Для подбора литературы в библиотеке используются алфавитный и систематический каталоги.

Важно помнить, что рациональные навыки работы с книгой - это всегда большая экономия времени и сил.

Правильный подбор учебников рекомендуется преподавателем, читающим лекционный курс. Необходимая литература может быть также указана в методических разработках по данному курсу.

Изучая материал по учебнику, следует переходить к следующему вопросу только после правильного уяснения предыдущего, описывая на бумаге все выкладки и вычисления (в том числе те, которые в учебнике опущены или на лекции даны для самостоятельного вывода).

При изучении любой дисциплины большую и важную роль играет самостоятельная индивидуальная работа.

Особое внимание следует обратить на определение основных понятий курса. Студент должен подробно разбирать примеры, которые поясняют такие определения, и уметь строить аналогичные примеры самостоятельно. Нужно добиваться точного представления о том, что изучаешь. Полезно составлять опорные конспекты. При изучении материала по учебнику полезно в тетради (на специально отведенных полях) дополнять конспект лекций. Там же следует отмечать вопросы, выделенные студентом для консультации с преподавателем.

Выводы, полученные в результате изучения, рекомендуется в конспекте выделять, чтобы они при перечитывании записей лучше запоминались.

Опыт показывает, что многим студентам помогает составление листа опорных сигналов, содержащего важнейшие и наиболее часто употребляемые формулы и понятия. Такой лист помогает запомнить формулы, основные положения лекции, а также может служить постоянным справочником для студента.

Чтение научного текста является частью познавательной деятельности. Ее цель – извлечение из текста необходимой информации. От того на сколько осознанна читающим собственная внутренняя установка при обращении к печатному слову (найти нужные сведения, усвоить информацию полностью или частично, критически проанализировать материал и т.п.) во многом зависит эффективность осуществляемого действия.

Выделяют **четыре основные установки в чтении научного текста**:

1. информационно-поисковый (задача – найти, выделить искомую информацию)
2. усваивающая (усилия читателя направлены на то, чтобы как можно полнее осознать и запомнить как сами сведения излагаемые автором, так и всю логику его рассуждений)
3. аналитико-критическая (читатель стремится критически осмыслить материал, проанализировав его, определив свое отношение к нему)
4. творческая (создает у читателя готовность в том или ином виде – как отправной пункт для своих рассуждений, как образ для действия по аналогии и т.п. – использовать суждения автора, ход его мыслей, результат наблюдения, разработанную методику, дополнить их, подвергнуть новой проверке).

Основные виды систематизированной записи прочитанного:

Аннотирование – предельно краткое связное описание просмотренной или прочитанной книги (статьи), ее содержания, источников, характера и назначения;

Планирование – краткая логическая организация текста, раскрывающая содержание и структуру изучаемого материала;

Тезирование – лаконичное воспроизведение основных утверждений автора без привлечения фактического материала;

Цитирование – дословное выписывание из текста выдержек, извлечений, наиболее существенно отражающих ту или иную мысль автора;

Конспектирование – краткое и последовательное изложение содержания прочитанного.

Конспект – сложный способ изложения содержания книги или статьи в логической последовательности. Конспект аккумулирует в себе предыдущие виды записи, позволяет всесторонне охватить содержание книги, статьи. Поэтому умение составлять план, тезисы, делать выписки и другие записи определяет и технологию составления конспекта.

Методические рекомендации по составлению конспекта:

1. Внимательно прочитайте текст. Уточните в справочной литературе непонятные слова. При записи не забудьте вынести справочные данные на поля конспекта;

2. Выделите главное, составьте план;

3. Кратко сформулируйте основные положения текста, отметьте аргументацию автора;

4. Законспектируйте материал, четко следуя пунктам плана. При конспектировании старайтесь выразить мысль своими словами. Записи следует вести четко, ясно.

5. Грамотно записывайте цитаты. Цитируя, учитывайте лаконичность, значимость мысли.

В тексте конспекта желательно приводить не только тезисные положения, но и их доказательства. При оформлении конспекта необходимо стремиться к емкости каждого предложения. Мысли автора книги следует излагать кратко, заботясь о стиле и выразительности написанного. Число дополнительных элементов конспекта должно быть логически обоснованным, записи должны распределяться в определенной последовательности, отвечающей логической структуре произведения. Для уточнения и дополнения необходимо оставлять поля.

Овладение навыками конспектирования требует от студента целеустремленности, повседневной самостоятельной работы.

4.2. Методические рекомендации по подготовке к практическим и лабораторным занятиям

Для того чтобы практические и лабораторные занятия приносили максимальную пользу, необходимо помнить, что упражнение и решение задач проводятся по вычитанному на лекциях материалу и связаны, как правило, с детальным разбором отдельных вопросов лекционного курса. Следует подчеркнуть, что только после усвоения лекционного материала с определенной точки зрения (а именно с той, с которой он излагается на лекциях) он будет закрепляться на практических занятиях как в результате

обсуждения и анализа лекционного материала, так и с помощью решения проблемных ситуаций, задач. При этих условиях студент не только хорошо усвоит материал, но и научится применять его на практике, а также получит дополнительный стимул (и это очень важно) для активной проработки лекции.

При самостоятельном решении задач нужно обосновывать каждый этап решения, исходя из теоретических положений курса. Если студент видит несколько путей решения проблемы (задачи), то нужно сравнить их и выбрать самый рациональный. Полезно до начала вычислений составить краткий план решения проблемы (задачи). Решение проблемных задач или примеров следует излагать подробно, вычисления располагать в строгом порядке, отделяя вспомогательные вычисления от основных. Решения при необходимости нужно сопровождать комментариями, схемами, чертежами и рисунками.

Следует помнить, что решение каждой учебной задачи должно доводиться до окончательного логического ответа, которого требует условие, и по возможности с выводом. Полученный ответ следует проверить способами, вытекающими из существа данной задачи. Полезно также (если возможно) решать несколькими способами и сравнить полученные результаты. Решение задач данного типа нужно продолжать до приобретения твердых навыков в их решении.

4.3. Методические рекомендации по самопроверке знаний

После изучения определенной темы по записям в конспекте и учебнику, а также решения достаточного количества соответствующих задач на практических занятиях и самостоятельно студенту рекомендуется, провести самопроверку усвоенных знаний, ответив на контрольные вопросы по изученной теме.

В случае необходимости нужно еще раз внимательно разобраться в материале.

Иногда недостаточность усвоения того или иного вопроса выясняется только при изучении дальнейшего материала. В этом случае надо вернуться назад и повторить плохо усвоенный материал. Важный критерий усвоения теоретического материала - умение решать задачи или пройти тестирование по пройденному материалу. Однако следует помнить, что правильное решение задачи может получиться в результате применения механически заученных формул без понимания сущности теоретических положений.

4.4. Методические рекомендации по написанию научных текстов (докладов, докладов, эссе, научных статей и т.д.)

Перед тем, как приступить к написанию научного текста, важно разобраться, какова истинная цель вашего научного текста - это поможет вам разумно распределить свои силы и время.

Во-первых, сначала нужно определиться с идеей научного текста, а для этого необходимо научиться либо относиться к разным явлениям и фактам

несколько критически (своя идея – как иная точка зрения), либо научиться увлекаться какими-то известными идеями, которые нуждаются в доработке (идея – как оптимистическая позиция и направленность на дальнейшее совершенствование уже известного). Во-вторых, научиться организовывать свое время, ведь, как известно, свободное (от всяких глупостей) время – важнейшее условие настоящего творчества, для него наконец-то появляется время. Иногда именно на организацию такого времени уходит немалая часть сил и талантов.

Писать следует ясно и понятно, стараясь основные положения формулировать четко и недвусмысленно (чтобы и самому понятно было), а также стремясь структурировать свой текст. Каждый раз надо представлять, что ваш текст будет кто-то читать и ему захочется сориентироваться в нем, быстро находить ответы на интересующие вопросы (заодно представьте себя на месте такого человека). Понятно, что работа, написанная «сплошным текстом» (без заголовков, без выделения крупным шрифтом наиболее важным мест и т. п.), у культурного читателя должна вызывать брезгливость и даже жалость к автору (исключения составляют некоторые древние тексты, когда и жанр был иной и к текстам относились иначе, да и самих текстов было гораздо меньше – не то, что в эпоху «информационного взрыва» и соответствующего «информационного мусора»).

Объем текста и различные оформительские требования во многом зависят от принятых в конкретном учебном заведении порядков.

Доклад - это самостоятельное исследование студентом определенной проблемы, комплекса взаимосвязанных вопросов.

Доклад не должна составляться из фрагментов статей, монографий, пособий. Кроме простого изложения фактов и цитат, в доклад е должно проявляться авторское видение проблемы и ее решения.

Рассмотрим основные этапы подготовки
а студентом.

Выполнение доклада начинается с выбора темы.

Затем студент приходит на первую консультацию к руководителю, которая предусматривает:

- обсуждение цели и задач работы, основных моментов избранной темы;
- консультирование по вопросам подбора литературы;
- составление предварительного плана.

Следующим этапом является работа с литературой. Необходимая литература подбирается студентом самостоятельно.

После подбора литературы целесообразно сделать рабочий вариант плана работы. В нем нужно выделить основные вопросы темы и параграфы, раскрывающие их содержание.

Составленный список литературы и предварительный вариант плана уточняются, согласуются на очередной консультации с руководителем.

Затем начинается следующий этап работы - изучение литературы. Только внимательно читая и конспектируя литературу, можно разобраться в основных вопросах темы и подготовиться к самостоятельному (авторскому)

изложению содержания доклада. Конспектируя первоисточники, необходимо отразить основную идею автора и его позицию по исследуемому вопросу, выявить проблемы и наметить задачи для дальнейшего изучения данных проблем.

Систематизация и анализ изученной литературы по проблеме исследования позволяют студенту написать работу.

Рабочий вариант текста доклада предоставляется руководителю на проверку. На основе рабочего варианта текста руководитель вместе со студентом обсуждает возможности доработки текста, его оформление. После доработки доклад сдается на кафедру для его оценивания руководителем.

Требования к написанию доклада

Написание 1 доклада является обязательным условием выполнения плана СРС по любой дисциплине профессионального цикла.

Тема доклада может быть выбрана студентом из предложенных в рабочей программе или фонде оценочных средств дисциплины, либо определена самостоятельно, исходя из интересов студента (в рамках изучаемой дисциплины). Выбранную тему необходимо согласовать с преподавателем.

Доклад должен быть написан научным языком.

Объем доклада должен составлять 20-25 стр.

Структура доклада:

1) Введение (не более 3-4 страниц). Во введении необходимо обосновать выбор темы, ее актуальность, очертить область исследования, объект исследования, основные цели и задачи исследования.

2) Основная часть состоит из 2-3 разделов. В них раскрывается суть исследуемой проблемы, проводится обзор мировой литературы и источников Интернет по предмету исследования, в котором дается характеристика степени разработанности проблемы и авторская аналитическая оценка основных теоретических подходов к ее решению. Изложение материала не должно ограничиваться лишь описательным подходом к раскрытию выбранной темы. Оно также должно содержать собственное видение рассматриваемой проблемы и изложение собственной точки зрения на возможные пути ее решения.

3) Заключение (1-2 страницы). В заключении кратко излагаются достигнутые при изучении проблемы цели, перспективы развития исследуемого вопроса

4) Список использованной литературы (не меньше 10 источников), в алфавитном порядке, оформленный в соответствии с принятыми правилами. В список использованной литературы рекомендуется включать работы отечественных и зарубежных авторов, в том числе статьи, опубликованные в научных журналах в течение последних 3-х лет и ссылки на ресурсы сети Интернет.

1) Приложение (при необходимости).

Требования к оформлению:

5) текст с одной стороны листа;

6) шрифт Times New Roman;

- 7) кегль шрифта 14;
- 8) межстрочное расстояние 1,5;
- 9) поля: сверху 2,5 см, снизу – 2,5 см, слева - 3 см, справа 1,5 см;
- 10) доклад должен быть представлен в сброшюрованном виде.

Порядок защиты доклада:

Защита доклада проводится на практических занятиях, после окончания работы студента над ним и исправления всех недочетов, выявленных преподавателем в ходе консультаций. На защиту доклада отводится 5-7 минут времени, в ходе которого студент должен показать свободное владение материалом по заявленной теме. При защите доклада приветствуется использование мультимедиа-презентации.

Оценка доклада

Доклад оценивается по следующим критериям:

- соблюдение требований к его оформлению;
- необходимость и достаточность для раскрытия темы приведенной в тексте доклада информации;
- умение студента свободно излагать основные идеи, отраженные в докладе;
- способность студента понять суть задаваемых преподавателем и сокурсниками вопросов и сформулировать точные ответы на них.

Критерии оценки:

Оценка «отлично» выставляется студенту, если в докладе студент исчерпывающе, последовательно, четко и логически стройно излагает материал; свободно справляется с задачами, вопросами и другими видами применения знаний; использует для написания доклада современные научные материалы; анализирует полученную информацию; проявляет самостоятельность при написании доклада.

Оценка «хорошо» выставляется студенту, если качество выполнения доклада достаточно высокое. Студент твердо знает материал, грамотно и по существу излагает его, не допуская существенных неточностей в ответе на вопросы по теме доклада.

Оценка «удовлетворительно» выставляется студенту, если материал доклада излагается частично, но пробелы не носят существенного характера, студент допускает неточности и ошибки при защите доклада, дает недостаточно правильные формулировки, наблюдаются нарушения логической последовательности в изложении материала.

Оценка «неудовлетворительно» выставляется студенту, если он не подготовил доклад или допустил существенные ошибки. Студент неуверенно излагает материал доклада, не отвечает на вопросы преподавателя.

4.5. Методические рекомендации по выполнению исследовательских проектов

Исследовательская проектная работа – это групповая работа, для выполнения которой необходим выбор и приложение научной методики к поставленной задаче, получение собственного теоретического или

экспериментального материала, на основании которого необходимо провести анализ и сделать выводы об исследуемом явлении. Выполнение проекта – это всегда коллективная, творческая практическая работа, предназначенная для получения определенного продукта или научно-технического результата. Такая работа подразумевает четкое, однозначное формирование поставленной задачи, определение сроков выполнения намеченного, определение требований к разрабатываемому объекту.

Выполнение 1 группового проекта является обязательным условием выполнения самостоятельной работы по любой дисциплине цифрового цикла. Тема проектного задания может быть выбрана студентом из предложенных в рабочей программе или фонде оценочных средств дисциплины, либо определена самостоятельно, исходя из интересов студента (в рамках изучаемой дисциплины). Выбранную тему необходимо согласовать с преподавателем.

Требования по выполнению и оформлению проекта

При выполнении проекта приветствуется работа в группе (2-3 человека). Проект – это исследовательская работа, в ходе которой студенты должны продемонстрировать владение навыками научного исследования, умения проводить анализ, обобщать информацию, делать выводы, предлагать свои решения проблемы, рассматриваемой в проекте.

При подготовке материалов проекта студенты должны продемонстрировать владение современными методами компьютерной обработки данных.

Критерии оценки работы участника проекта.

Для каждого из участников проекта оцениваются:

- профессиональные теоретические знания в соответствующей области;
- умение работать со справочной и научной литературой, осуществлять поиск необходимой информации в Интернет;
- умение работать с техническими средствами;
- умение пользоваться соответствующими выполняемому проекту информационными технологиями;
- умение готовить материалы проекта для презентации: составлять и редактировать тексты, формировать презентацию проекта;
- умение работать в команде;
- умение публично представлять результаты собственной деятельности;
- коммуникабельность, инициативность, творческие способности.

Критерии выставления оценки участникам проекта

Оценка	Профессиональные компетенции	Компетенции, связанные с использованием соответствующих их выполняемому проекту технических средств и информационных технологий	Иные универсальные компетенции (коммуникабельность, инициативность, умение работать в «команде», управленческие навыки и т.д.)	Отчетность
«Отлично»	Работа выполнена на высоком профессиональном уровне. Представленный материал в основном фактически верен, допускаются негрубые фактические неточности. Студент свободно отвечает на вопросы, связанные с проектом.	Технические средства и информационные технологии освоены и использованы для реализации проекта полностью	Студент проявил инициативу, творческий подход, способность к выполнению сложных заданий, навыки работы в коллективе, организационные способности.	Проект представлен полностью и в срок.
«Хорошо»	Работа выполнена на достаточно высоком профессиональном уровне. Допущено до 4–5 фактических ошибок. Студент отвечает на вопросы, связанные с проектом, но недостаточно полно.	Обнаруживаются некоторые ошибки в использовании соответствующих их технических средств и информационных технологий	Студент достаточно полно, но без инициативы и творческих находок выполнил возложенные на него задачи.	Проект представлен достаточно полно и в срок, но с некоторыми недоработками.

Оценка	Профессиональные компетенции	Компетенции, связанные с использованием соответствующих их выполняемому проекту технических средств и информационных технологий	Иные универсальные компетенции (коммуникабельность, инициативность, умение работать в «команде», управленческие навыки и т.д.)	Отчетность
«Удовлетворительно»	Уровень недостаточно высок. Допущено до 8 фактических ошибок. Студент может ответить лишь на некоторые из заданных вопросов, связанных с проектом.	Обнаруживает недостаточное владение навыками работы с техническими средствами и соответствующими информационными технологиями	Студент выполнил большую часть возложенной на него работы.	Проект сдан со значительным опозданием (более недели) и не полностью
«Неудовлетворительно»	Работа не выполнена или выполнена на низком уровне. Допущено более 8 фактических ошибок. Ответы на связанные с проектом вопросы обнаруживают непонимание предмета и отсутствие ориентации в материале проекта.	Навыков работы с техническими средствами нет, информационные технологии не освоены	Студент практически не работал, не выполнил свои задачи или выполнил лишь отдельные не существенные поручения в групповом проекте.	Проект не сдан.

4.6. Методические рекомендации по подготовке к экзаменам и зачетам

Общий результат оценивания знаний и умений для выставления дифференцированного зачета может формироваться как интегральная оценка, складывающаяся из текущего контроля (например, выполнение заданий на лабораторно-практических занятиях, защиты индивидуальных творческих работ) и промежуточного контроля (презентация доклада, собеседование, тестирование).

Дифференцированный зачет может проводиться в форме тестирования и просмотра представленных творческих работ. На тестирование обычно отводится ограниченное время (например, 15–20 минут).

Критерии оценивания знаний и умений для выставления дифференцированного зачета:

- Оценка «Отлично»: Студент показал высокий уровень сформированности компетенций, всесторонние, систематизированные, глубокие знания дисциплины и умение уверенно применять их на практике при решении конкретных задач, свободное и правильное обоснование принятых решений.

- «Хорошо» Студент показал уровень сформированности компетенций выше среднего или высокий уровень сформированности большей части компетенций. Он твердо знает материал, грамотно и по существу излагает его, умеет применять полученные знания на практике, но допускает не критичные неточности в ответе или решении задач.

- «Удовлетворительно» Студент показал уровень сформированности компетенций ниже среднего, фрагментарный характер знаний, недостаточно правильные формулировки базовых понятий, нарушения логической последовательности в изложении программного материала. При этом он владеет основными разделами учебной программы, необходимыми для дальнейшего обучения, и может применять полученные знания по образцу в стандартной ситуации.

- «Неудовлетворительно» Студент не усвоил значительную часть программного материала, ответ содержит существенные ошибки.

Контроль самостоятельной работы студентов

Контроль самостоятельной работы проводится преподавателем в аудитории.

Предусмотрены следующие виды контроля: собеседование, оценка доклада, оценка презентации, оценка выполнения проекта и/или индивидуальных творческих работ.

Подробные критерии оценивания компетенций приведены в Фонде оценочных средств для проведения текущей и промежуточной аттестации.

Список литературы для выполнения СРС

Перечень основной литературы:

1. Барский, А.Б. Введение в нейронные сети : учебное пособие / А. Б. Барский. - Введение в нейронные сети, 2021-12-05. - Электрон. дан. (1 файл). - Москва, Саратов : Интернет-Университет Информационных Технологий (ИНТУИТ), Ай Пи Ар Медиа, 2020. - 357 с. - электронный. - Книга находится в премиум-версии ЭБС IPR BOOKS. - ISBN 978-5-4497-0309-5, экземпляров неограничено

2. Павлова, А.И. Искусственные нейронные сети : учебное пособие / А. И. Павлова. - Искусственные нейронные сети, 2031-05-31. - Электрон. дан. (1 файл). - Москва : Ай Пи Ар Медиа, 2021. - 190 с. - электронный. - Книга находится в премиум-версии ЭБС IPR BOOKS. - ISBN 978-5-4497-1165-6, экземпляров неограничено

3. Ростовцев, В.С. Искусственные нейронные сети Электронный ресурс / Ростовцев В. С. : учебник для вузов. - 2-е изд., стер. - Санкт-Петербург : Лань, 2021. - 216 с. - ISBN 978-5-8114-7462-2, экземпляров неограничено

Перечень дополнительной литературы:

1. Галушкин, А.И. Теория нейронных сетей / А. И. Галушкин, Кн. 1. - М. : ИПРЖР, 2000. - 416с.

2. Искусственный интеллект . - В 3-х кн. Кн. 1 Системы общения и экспертные системы: Справочник / под ред. Д.А. Поспелова - М.: Радио и связь, 1990.

3. Искусственный интеллект . - В 3-х кн. Кн. 2 Модели и методы: Справочник / под ред. Д.А. Поспелова - М.: Радио и связь, 1990.

4. Искусственный интеллект . - В 3-х кн. Кн. 3 Программные и аппаратные средства: Справочник / под ред. Д.А. Поспелова - М.: Радио и связь, 1990.

5. Лутц М.. Изучаем Python, том I, 5-е издание. – Пер. с англ. – СПб.: Диалектика, 2019.

6. Лутц М.. Изучаем Python, том II, 5-е издание. – Пер. с англ. – СПб.: Вильямс, 2020.

7. Хайкин С. Нейронные сети: полный курс : науч.-поп. изд. / С. Хайкин ; пер. с англ. Н.Н. Куссуль, А.Ю. Шелестова. - 2-е изд. - М. : Вильямс, 2008. - 1104 с.

Перечень учебно-методического обеспечения СРС

1. Курс лекций по дисциплине "Искусственный интеллект" (электронный ресурс)

2. Методические рекомендации по выполнению лабораторных работ по дисциплине "Искусственный интеллект" (электронный ресурс)

3. Методические рекомендации по организации самостоятельной работы студентов по дисциплине "Искусственный интеллект" (электронный ресурс)

Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины (модуля)

1. <http://el.ncfu.ru/> – система управления обучением ФГАОУ ВО СКФУ. Дистанционная поддержка дисциплины «Искусственный интеллект».

2. <https://colab.google.com/> - сайт создания и обработки данных машинного обучения и искусственного интеллекта.
3. <https://www.python.org> - сайт языка программирования Python.