

Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Таскаев Сергей Валерьевич
Должность: Ректор
Дата подписания: 05.09.2025 12:17:41
Уникальный программный ключ:
04c19ed8bb98f3b6cb77a486b9a6788b8522523



МИНОБРАЗОВАНИЯ РОССИИ
Федеральное государственное бюджетное образовательное
учреждение высшего образования
«Челябинский государственный университет» (ФГБОУ ВО «ЧелГУ»)
Физический факультет
Кафедра общей и теоретической физики

Фонд оценочных средств по дисциплине «Системы искусственного интеллекта и машинное обучение»
по направлению подготовки 24.03.03 «Баллистика и гидроаэродинамика» направленности (профилю) Баллистика и
гидроаэродинамика ФГБОУ ВО «ЧелГУ»

Версия документа - 1	стр. 1	Первый экземпляр _____	КОПИЯ № ____
----------------------	--------	------------------------	--------------

**Фонд оценочных средств для промежуточной аттестации
по дисциплине (модулю)
Системы искусственного интеллекта и машинное обучение**

Направление подготовки (специальность)
24.03.03 Баллистика и гидроаэродинамика

Направленность (профиль)
Баллистика и гидроаэродинамика

Присваиваемая квалификация (степень)
Бакалавр

Форма обучения
Очная

Челябинск, 2025 г.



МИНОБРНАУКИ РОССИИ
Федеральное государственное бюджетное образовательное
учреждение высшего образования
«Челябинский государственный университет» (ФГБОУ ВО «ЧелГУ»)
Физический факультет
Кафедра общей и теоретической физики

Фонд оценочных средств по дисциплине «Системы искусственного интеллекта и машинное обучение»
по направлению подготовки 24.03.03 «Баллистика и гидроаэродинамика» направленности (профилю) Баллистика и
гидроаэродинамика ФГБОУ ВО «ЧелГУ»

Версия документа - 1

стр. 2

Первый экземпляр _____

КОПИЯ № _____

Содержание

1. Паспорт фонда оценочных средств
2. Перечень формируемых компетенций
3. Содержание оценочных средств по дисциплине
 - 3.1. Виды оценочных средств
 - 3.2. Содержание оценочных средств
4. Порядок проведения и критерии оценивания промежуточной аттестации
 - 4.1. Порядок проведения промежуточной аттестации
 - 4.2. Результаты промежуточной аттестации и уровни сформированности компетенций



МИНОБРНАУКИ РОССИИ
Федеральное государственное бюджетное образовательное
учреждение высшего образования
«Челябинский государственный университет» (ФГБОУ ВО «ЧелГУ»)
Физический факультет
Кафедра общей и теоретической физики

Фонд оценочных средств по дисциплине «Системы искусственного интеллекта и машинное обучение»
по направлению подготовки 24.03.03 «Баллистика и гидроаэродинамика» направленности (профилю) Баллистика и
гидроаэродинамика ФГБОУ ВО «ЧелГУ»

Версия документа - 1	стр. 3	Первый экземпляр _____	КОПИЯ № _____
----------------------	--------	------------------------	---------------

1. ПАСПОРТ ФОНДА ОЦЕНОЧНЫХ СРЕДСТВ

Направление подготовки: 24.03.03 Баллистика и гидроаэродинамика

Направленность (профиль): Баллистика и гидроаэродинамика

Дисциплина: Системы искусственного интеллекта и машинное обучение

Семестр: 8

Форма промежуточной аттестации: экзамен

Система оценивания экзамена: оценивание результатов осуществляется в рамках 5-балльной системы.

2. ПЕРЕЧЕНЬ ФОРМИРУЕМЫХ КОМПЕТЕНЦИЙ И ЭТАПЫ ИХ ФОРМИРОВАНИЯ

2.1. Компетенции, закреплённые за дисциплиной

Изучение дисциплины «Системы искусственного интеллекта и машинное обучение» направлено на формирование следующих компетенций:

Коды компетенции (по ФГОС)	Содержание компетенций согласно ФГОС	Индикаторы достижения компетенций согласно ОПОП	Перечень планируемых результатов обучения по дисциплине
ОПК-8	Способен разрабатывать алгоритмы и компьютерные программы, пригодные для практического применения	ОПК-8.1. Знать принципах разработки алгоритмов, компьютерных программ для профессиональной деятельности. ОПК-8.2. Уметь использовать основы программирования для профессиональной деятельности. ОПК-8.3. Иметь навыки разработки алгоритмов и программирования для профессиональной деятельности.	Знать: Для достижения ОПК-8.1: терминологию, структуры и этапы построения систем машинного обучения; Уметь: Для достижения ОПК-8.2: применять современные методы и информационные технологии для выполнения отдельных элементов документации и планов проведения отдельных этапов работ; Владеть: Для достижения ОПК-8.3: навыком работы с данными с использованием некоторых компьютерных программ и программных сред, в которых реализованы рассматриваемые методы



МИНОБРНАУКИ РОССИИ
 Федеральное государственное бюджетное образовательное
 учреждение высшего образования
 «Челябинский государственный университет» (ФГБОУ ВО «ЧелГУ»)
 Физический факультет
 Кафедра общей и теоретической физики

Фонд оценочных средств по дисциплине «Системы искусственного интеллекта и машинное обучение»
 по направлению подготовки 24.03.03 «Баллистика и гидроаэродинамика» направленности (профилю) Баллистика и
 гидроаэродинамика ФГБОУ ВО «ЧелГУ»

Версия документа - 1	стр. 4	Первый экземпляр _____	КОПИЯ № _____
----------------------	--------	------------------------	---------------

ПК-3	Способен проводить математическое описание объектов исследования, разрабатывать алгоритмы в области баллистики и гидроаэродинамики	<p>ПК-3.1. Обладает знаниями в своей области научно-исследовательской деятельности о математическом описании объектов исследования, принципах разработки алгоритмов в области баллистики и гидроаэродинамики.</p> <p>ПК-3.2. Демонстрирует умение в своей научно-исследовательской деятельности проводить математическое описание объектов исследования, разрабатывать алгоритмы решения баллистических и гидроаэродинамических задач.</p> <p>ПК-3.3. Имеет практический опыт (навыки) использования в своей научно-исследовательской деятельности математического описания объектов исследования, разработки алгоритмов в области баллистики и гидроаэродинамики.</p>	<p><u>Знать:</u> Для достижения ПК-3.1: основные методы машинного обучения, типы решаемых с их помощью задач, методы подготовки и обработки данных;</p> <p><u>Уметь:</u> Для достижения ПК-3.2: применять системы искусственного интеллекта и методы машинного обучения в научно-исследовательской деятельности;</p> <p><u>Владеть:</u> Для достижения ПК-3.3: общими навыками использования в своей научно-исследовательской деятельности математического описания объектов исследования, разработки алгоритмов</p>
------	--	--	--

3. СОДЕРЖАНИЕ ОЦЕНОЧНЫХ СРЕДСТВ ПО ДИСЦИПЛИНЕ

3.1 Виды оценочных средств

№ п/п	Код компетенции/ планируемые результаты обучения	Контролируемые темы/ разделы	Наименование оценочного средства для текущего контроля	Наименование оценочного средства на промежуточной аттестации/№ задания
-------	--	---------------------------------	--	---



МИНОБРНАУКИ РОССИИ
Федеральное государственное бюджетное образовательное
учреждение высшего образования
«Челябинский государственный университет» (ФГБОУ ВО «ЧелГУ»)
Физический факультет
Кафедра общей и теоретической физики

Фонд оценочных средств по дисциплине «Системы искусственного интеллекта и машинное обучение»
по направлению подготовки 24.03.03 «Баллистика и гидроаэродинамика» направленности (профилю) Баллистика и
гидроаэродинамика ФГБОУ ВО «ЧелГУ»

Версия документа - 1

стр. 5

Первый экземпляр _____

КОПИЯ № _____

1.	ОПК-8 ПК-3	Искусственный интеллект	Отчеты по лабораторным заданиям	вопросы к экзамену
		Предмет машинного обучения	Отчеты по лабораторным заданиям	вопросы к экзамену
		Основные понятия математической статистики используемой в машинном обучении	Отчеты по лабораторным заданиям	вопросы к экзамену
		Байесовская статистика	Отчеты по лабораторным заданиям	вопросы к экзамену
		Методы машинного обучения без учителя	Отчеты по лабораторным заданиям	вопросы к экзамену
		Глубокие нейронные сети прямого распространения	Отчеты по лабораторным заданиям	вопросы к экзамену
		Сверточные нейронные сети	Отчеты по лабораторным заданиям	вопросы к экзамену
		Основные понятия в теории машинного обучения	Отчеты по лабораторным заданиям	вопросы к экзамену

3.2 Содержание оценочных средств

Задачи к лабораторным работам представлены в Приложении 1.

Отчет по лабораторному заданию (ЛЗ)

Отчет по ЛЗ представляется в печатном виде в формате, предусмотренном шаблоном отчета по ЛЗ. Отчет по ЛЗ должен содержать:

- постановку задачи, математическую модель и основные расчетные соотношения используемых методов решения;
- схему расчетной области с характеристиками сетки, краевыми и начальными условиями, реализованными в решаемом варианте;
- графическое представление полученных результатов;
- содержание исследовательского задания, результаты вычислительного моделирования, анализ и выводы по проведенным исследованиям.

Вопросы к экзамену

- Предмет машинного обучения



МИНОБРНАУКИ РОССИИ
Федеральное государственное бюджетное образовательное
учреждение высшего образования
«Челябинский государственный университет» (ФГБОУ ВО «ЧелГУ»)
Физический факультет
Кафедра общей и теоретической физики

Фонд оценочных средств по дисциплине «Системы искусственного интеллекта и машинное обучение»
по направлению подготовки 24.03.03 «Баллистика и гидроаэродинамика» направленности (профилю) Баллистика и
гидроаэродинамика ФГБОУ ВО «ЧелГУ»

Версия документа - 1

стр. 6

Первый экземпляр _____

КОПИЯ № _____

- Задачи, которые решаются методами машинного обучения
- Определение точности моделей машинного обучения
- Основные типы МО: алгоритмы с учителем и без учителя
- Основные понятия в теории машинного обучения
- Емкость модели, переобучение и недообучение
- Регуляризация
- Валидация моделей машинного обучения
- Основные понятия математической статистики, используемой в машинном обучении
- Понятие оценки и ее смещения
- Принцип наилучшего правдоподобия
- Перекрестная энтропия
- Оценка условной вероятности
- Байесовская статистика
- Байесовская линейная регрессия
- Глубокие нейронные сети прямого распространения
- Структура нейронной сети прямого распространения
- Выходные функции активации
- Функции активации в скрытых слоях сети
- Алгоритм обратного распространения
- Сверточные нейронные сети
- Виды операций свертки в нейронных сетях
- Три основные идеи, заложенные в сверточных сетях: редкое взаимодействие, совместное использование параметров, равномерность
- Уровни пула в сверточных сетях
- Метод главных компонент

4. ПОРЯДОК ПРОВЕДЕНИЯ И КРИТЕРИИ ОЦЕНИВАНИЯ ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ

4.1. Порядок проведения промежуточной аттестации

Текущий контроль теоретических знаний и практических навыков производится на лабораторных занятиях.

Студент допускается к сдаче экзамена в конце семестра при успешном выполнении лабораторных заданий.

Зачет ставится на основании устного ответа по билету с вопросами.

Оценка «Зачтено» ставится, если студент знает учебно-программный материал, грамотно и по существу излагает его, не допускает существенных неточностей в ответе на вопрос. Может правильно применить теоретические положения.

Оценка «Не зачтено» ставится, если студент не освоил основной материал.

Экзаменационная оценка ставится на основании письменного и устного ответов по экзаменационному билету. Экзаменационный билет содержит два теоретических вопроса и задачу. Студенты, которые успешно отчитались в течение семестра о решенных задачах по темам лабораторных занятий из предложенного списка задач в методических указаниях к



МИНОБРНАУКИ РОССИИ
Федеральное государственное бюджетное образовательное
учреждение высшего образования
«Челябинский государственный университет» (ФГБОУ ВО «ЧелГУ»)
Физический факультет
Кафедра общей и теоретической физики

Фонд оценочных средств по дисциплине «Системы искусственного интеллекта и машинное обучение»
по направлению подготовки 24.03.03 «Баллистика и гидроаэродинамика» направленности (профилю) Баллистика и
гидроаэродинамика ФГБОУ ВО «ЧелГУ»

Версия документа - 1	стр. 7	Первый экземпляр _____	КОПИЯ № _____
----------------------	--------	------------------------	---------------

курсу, освобождаются от 3-го вопроса в билете (т.е. решения задачи).

На экзамене студент получает оценку «удовлетворительно» в случае успешной сдачи «теоретического минимума», который включает: знание основных понятий, название и физический смысл величин, вид основных распределений и соотношений (без вывода), определяемых вопросом билета.

Оценка «хорошо» – студент твердо знает учебно-программный материал, грамотно и по существу излагает его, но при этом допускаются негрубые ошибки при выводе формул или отсутствие некоторых элементов вывода.

Оценка «отлично» – студент должен продемонстрировать отличное знание материала, как лекционных занятий, так и тем, выносимых на самостоятельное обучение, ответив на оба вопроса билета, воспроизведя соответствующие математические выкладки и логические рассуждения; задача должно быть полностью решена, студент правильно обосновывает принятые решения.

4.2. Результаты промежуточной аттестации и уровни сформированности компетенций

Критерии оценивания экзаменационных вопросов:

Характеристики ответа	Оценка	Уровень освоения проверяемых компетенций
Ответил на оба вопроса билета, воспроизведя соответствующие математические выкладки и логические рассуждения, задача полностью решена, студент правильно обосновывает принятые решения. Возможны незначительные ошибки.	отлично	высокий
Студент твердо знает материал, грамотно и по существу излагает его, но при этом допускаются негрубые ошибки при выводе формул и решении задачи или отсутствие некоторых элементов вывода.	хорошо	средний
Знает «теоретический минимум», т.е. отвечает на вопрос базового уровня и знает основные понятия, соотношения (без вывода), название и физический смысл величин по другим вопросам билета.	удовлетворительно	базовый
Не может ответить на вопрос базового уровня	неудовлетворительно	недостаточный

Особенности проведения процедуры оценивания результатов обучения инвалидов и лиц с ограниченными возможностями здоровья обозначены в рабочей программе дисциплины (модуля).



МИНОБРНАУКИ РОССИИ
Федеральное государственное бюджетное образовательное
учреждение высшего образования
«Челябинский государственный университет» (ФГБОУ ВО «ЧелГУ»)
Физический факультет
Кафедра общей и теоретической физики

Фонд оценочных средств по дисциплине «Системы искусственного интеллекта и машинное обучение»
по направлению подготовки 24.03.03 «Баллистика и гидроаэродинамика» направленности (профилю) Баллистика и
гидроаэродинамика ФГБОУ ВО «ЧелГУ»

Версия документа - 1

стр. 8

Первый экземпляр _____

КОПИЯ № _____

Уровни сформированности компетенций определяется следующим образом:

1. Высокий уровень сформированности компетенций соответствует оценке отлично:
предполагает формирование компетенций на высоком уровне: студент свободно владеет основной терминологией и понятийным аппаратом дисциплины «Системы искусственного интеллекта и машинное обучение», что позволяет формулировать выводы и участвовать в дискуссии по учебным вопросам данной дисциплины; полностью сформировано умение применять полученную теоретическую базу для решения конкретных практических задач и уверенно владеть навыком их решения;
2. Средний уровень соответствует оценке хорошо:
предполагает формирование компетенций на среднем уровне: студент хорошо владеет основной терминологией и понятийным аппаратом дисциплины «Системы искусственного интеллекта и машинное обучение»; сформировано умение применять полученную теоретическую базу для решения конкретных практических задач и владеть навыками решения базовых задач;
3. Базовый уровень соответствует оценке удовлетворительно:
предполагает формирование компетенций на начальном уровне: студент знает «теоретический минимум» и недостаточно владеет методами решения базовых задач;
4. Низкий уровень соответствует оценке неудовлетворительно:
студент не владеет основной терминологией и понятийным аппаратом дисциплины «Системы искусственного интеллекта и машинное обучение»; не владеет навыками решения базовых задач.

Практическое задание 1

Одним из методов обучения без учителя, который может показать особенности в данных, а также снизить их размерность - метод главных компонент (**PCA**)

Использование **PCA** в наборе данных позволяет: 1. Удаляет коррелированные признаки. Поиск коррелированных признаков в данных занимает много времени, особенно если число признаков велико. 2. Улучшает производительность алгоритмов машинного обучения. Поскольку количество признаков уменьшается с помощью **PCA**, время, затрачиваемое на обучение модели, значительно сокращается вследствие более правильной оценки данных функцией стоимости. 3. **PCA** удаляют шум в данных и снижает размерность данных, сводя большое число признаков всего к паре главных компонент.

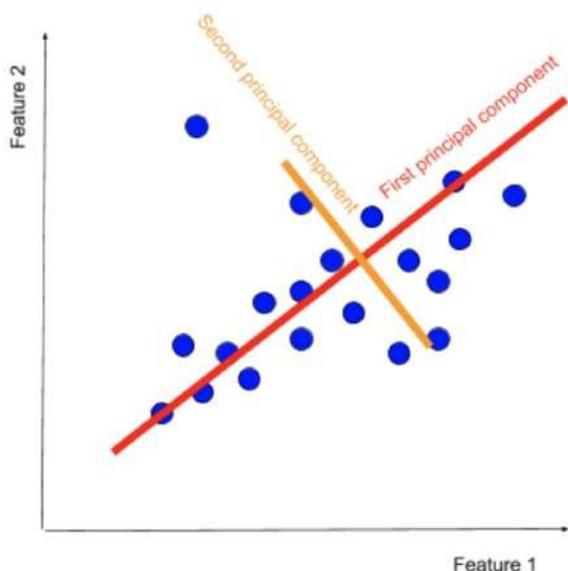


Рис 1. Направление главных компонент в данных.

Для реализации метода **PCA** нужно сделать несколько последовательных шагов.

1. Нормализация данных

На этом этапе нормализуется диапазон значений всех переменных \vec{x}_i до одинакового диапазона, чтобы все они имели равный вклад в анализ.

Основная причина, по которой проводится нормализация, заключается в том, что **PCA** очень чувствителен к дисперсии исходных переменных в наборе данных. Причина в том, что если есть признаки с большими различиями в их начальном диапазоне значений, и признаки с более высоким диапазоном значений будут доминировать в общем анализе, и **PCA** будет более предвзято относиться к этим признакам. Получим нормированные переменных \vec{z}_i через масштабирование по дисперсии и математическому ожиданию:

$$x_i^j = \frac{x_i^j - \mu}{\sigma}$$

где i - индекс по векторам данных, j - индекс по элементам вектора данных, μ - математическое ожидание по вектора \vec{x}_i , σ - стандартное отклонение вектора \vec{x}_i .

2. Вычисление матрицы ковариаций

На данном этапе необходимо понять как переменные набора данных связаны между собой, другими словами вычислить корреляцию между векторами данных. Иногда переменные сильно коррелируют между собой таким образом, что содержат избыточную информацию. Чтобы выявить эти корреляции вычисляются ковариационная матрица. Ковариационная матрица - это симметричная матрица $n \times n$ где n - число векторов данных. в которой в качестве элементов указаны ковариации, связанные со всеми возможными парами исходных переменных:

$$\Sigma = (\sigma_{ij}),$$

$$\sigma_{ij} = \text{cov}(\vec{x}_i, \vec{x}_j) = \mathbb{E}[(\vec{x}_i - \mathbb{E}\vec{x}_i)(\vec{x}_j - \mathbb{E}\vec{x}_j)]$$

где $i, j = 1..n$

3. Вычисление собственных векторов и собственных значений матрицы ковариаций Главные компоненты - это новые вектора данных \vec{z}_i , которые строятся как линейные комбинации исходных данных \vec{x}_i . Линейные комбинации строятся таким образом, чтобы новые данные \vec{z}_i были некоррелированными, а большая часть информации из исходных данных \vec{x}_i была ската в первые компоненты \vec{z}_i . Например, 10-мерные данные дают 10 главных компонент и **PCA** пытается поместить максимум возможной информации в первую компоненту \vec{z}_i , затем максимум оставшейся информации во вторую и так далее.

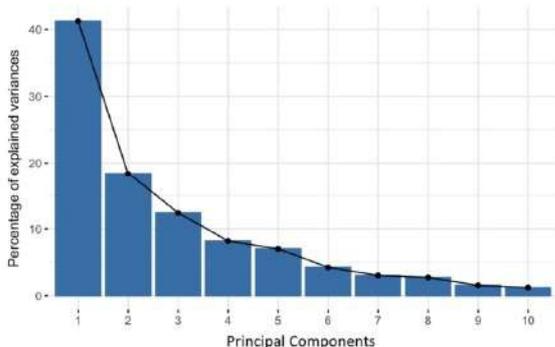


Рис. 2. Распределение дисперсии на каждый главный компонент полученный из набора данных.

Чтобы вычислить новые вектора данных \vec{z}_i как линейную комбинацию, необходимо найти коэффициенты преобразования. Вычисляя собственные вектора и собственные значения матрицы ковариаций, можно найти направления в исходных данных \vec{x}_i дисперсия в которых максимальна и, соответственно, находится основная информация в данных. Собственные вектора и значения матрицы ковариаций можно задать следующим равенством:

$$\Sigma \vec{A}_k = \lambda_k \vec{A}_k$$

где \vec{A}_k, λ_k - собственные вектора и значения матрицы ковариаций. Реализовать данный шаг возможно использованием библиотек линейной алгебры, которые входят, например, в набор библиотек **ScyPy** для **Python**. Полученные собственные значения и вектора \vec{A}_k, λ_k необходимо отсортировать от максимального значения λ_k до минимального (и соответствующие им собственные вектора): наибольшее значение собственного числа определяет статистическую значимость информации в исходных данных заключенную в

собственном векторе.

4. Вычисление главных компонент На заключительном этапе преобразуются данные \vec{x}_i , чтобы использовать вектор новых данных \vec{z}_i . Данные \vec{z}_i сформированы с помощью собственных векторов ковариационной матрицы, для переориентации данных с исходных осей на те, которые представлены главными компонентами (отсюда и название "Метод главных компонент"). Это можно сделать путем матричного умножения сходного набора данных \vec{x}_i , на отсортированные транспонированные собственные вектора матрицы ковариаций \vec{A}_k , при этом можно отбросить часть векторов, соответствующие минимальных собственным значениям:

$$z_i^j = \sum_k x_i^k A_k^j$$

Ссылка на данные

https://www.kaggle.com/competitions/modern-computer-technologies-laboratory/data?select=shok_wave_measurements.csv

Практическое задание 2

Аппроксимация функций ИНС прямого распространения

В данной лабораторной работе рассматриваются **глубокие нейронные сети прямого распространения (ИНС)**, в английской литературе часто можно встретить обозначение **multilayer perceptrons (MPL)**. Такие ИНС являются формой аппроксимации сложных многомерных и нелинейных функциональных зависимостей, с ограничением, что такие зависимости должны быть непрерывными. Пусть есть некоторая функция $\vec{y} = f(\vec{x})$, то нейронная сеть будет представлять собой некоторое приближение к этой функции $\vec{y} = f^*(\vec{x}, \vec{\theta})$, где $\vec{\theta}$ - набор параметров ИНС. Задача машинного обучения в этом случае состоит в том, чтобы подобрать такие параметры $\vec{\theta}$, чтобы получить наилучшее приближение к реальной функции $\vec{y} = f(\vec{x})$. Для того, чтобы подобрать параметры ИНС, или другими словами **обучить** ее, используется набор данных. Другими словами известно, что некоторые данные можно описать функциональной зависимостью, но сам вид функции не знаем и не можем записать его в явном виде, в этом случае и используются ИНС.

В данной лабораторной работе рассматриваются данные численного эксперимента по сжатию чистого алюминия, который проводится методом молекулярной динамики (рис. 1, 2). В эксперименте происходит одноосное сжатие образца, при этом есть однозначная связь между текущей плотностью вещества ρ , температурой T и давлением P , внутренней энергией E . Эта связь фактически выражается через функциональную зависимость $\vec{y} = f(\vec{x})$, где $\vec{y} = (P, E)$, а $\vec{x} = (\rho, T)$. Явный вид функции мы не знаем и задача найти ее аппроксимацию в виде нейронной сети прямого распространения по заданному набору данных \vec{x}_i, \vec{y}_i . Набор данных представляет собой последовательность (ρ_i, T_i) и соответствующие им значения (P_i, E_i) в процессе деформации.

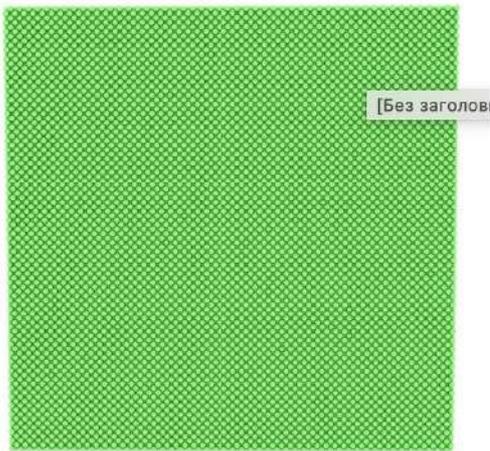


Рис. 1. Ячейка моделирования в численном эксперименте.

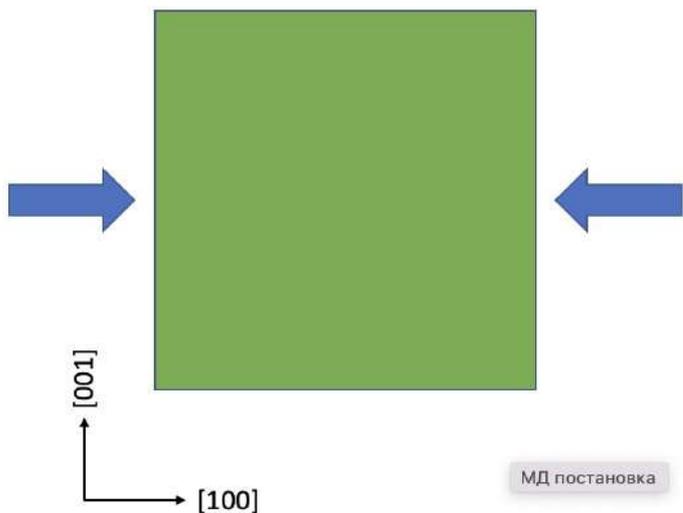


Рис. 2. Схема численного эксперимента.

Перейдем к описанию ИНС и методу ее обучения. ИНС состоит из слоев, в которых располагаются искусственные нейроны в виде функций активации (рис. 3). Можно выделить 3 типа слоев: входной, скрытые и выходной. Входной слой ассоциирован с аргументами функции \vec{x} , в скрытых слоях происходит основное преобразование входных данных и выходной слой ассоциирован с результатом функции \vec{y} . Входной слой содержит только входные данные \vec{x} , а скрытые и выходной слой состоят из последовательно вычисляемых элементарных функций - функций активации. Параметры нейронной сети $\vec{\theta}$ обозначаются как веса W и смещения \vec{b} . На каждом слое ИНС веса W - это матрица ассоциированная с связями в сети, а смещения \vec{b} - это вектор, ассоциированный с нейронами (Рис. 4):

$$z_i^l = \sum_j w_{i,j}^l a_j^{l-1} + b_i^l$$

, где $l = 1, 2, \dots, L$ - индекс по слоям ИНС, L - общее число слоев, i - индекс по нейронам в текущем слое l , j - индекс по всем нейронам в предыдущем слое $l - 1$. Для первого слоя $l = 1$ функция активации равна $a_j^0 = x_k$, где x_k - входной вектор на ИНС. После вычисления входного вектора z_i^l на слой l вычисляется функция активации в этом слое:

$$a_i^l = f(z_i^l)$$

Повторяя предыдущие операции вычисляется функция активации на последнем слое $a_i^L = f(z_i^L)$ которая является приближенным результатом функции, то есть \vec{y} .

Часто используются следующие функции активации:

PRReLU (Parameteric rectified linear unit)

$$f_{pr}(\alpha_{pr}, z) = \begin{cases} \alpha_{pr} z & \text{if } z < 0 \\ z & \text{otherwise} \end{cases}$$

Sigmoid

$$\sigma(z) = \frac{1}{1 + \exp^{-z}}$$

Часто используемые архитектуры:

1. Скрытые слои: сигмоидальная функция; Выходной слой: сигмоидальная функция.
2. Скрытые слои: функция PRReLU; Выходной слой: сигмоидальная функция.

Для начальной конфигурации ИНС матрицы весов W на каждом слое можно задать случайным образом, например, по Гауссовому распределению в диапазоне $[0, 1]$. Значения смещений \vec{b} можно задать как 0.001. Также необходимо нормализовать данные по двум причинам: [1] результат сигмоидальной функции находится в диапазоне $(0, 1)$; [2] Сильное различие в величинах данных может привести к тому, что нейроны не будут "активироваться", то есть их параметры W и \vec{b} будут близки к 0 и они фактически не будут влиять на преобразование данных при прогоне по ИНС. Соответственно, необходимо привести матрицы данных X и Y составленные из строк \vec{x}_i и \vec{y}_i к виду:

$$x_{i,j}^{norm} = \frac{x_{i,j} - x_i^{min}}{x_i^{max} - x_i^{min}}$$

и

$$y_{i,j}^{norm} = \frac{y_{i,j} - y_i^{min}}{y_i^{max} - y_i^{min}}$$

, где i - индекс по столбцам данных, а j - индекс по строкам, x_i^{max} и y_i^{max} - максимальное значение в столбцах, x_i^{min} и y_i^{min} - минимальное значение в столбцах.

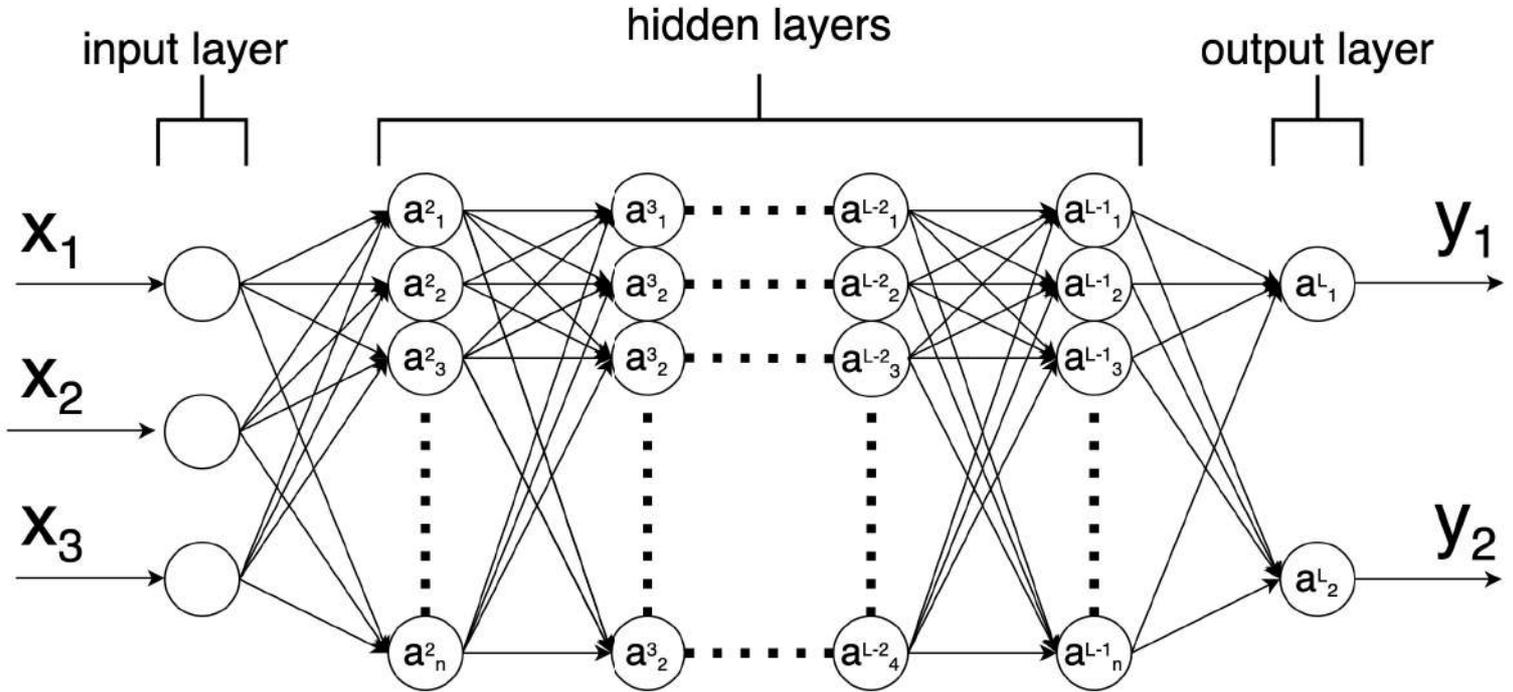
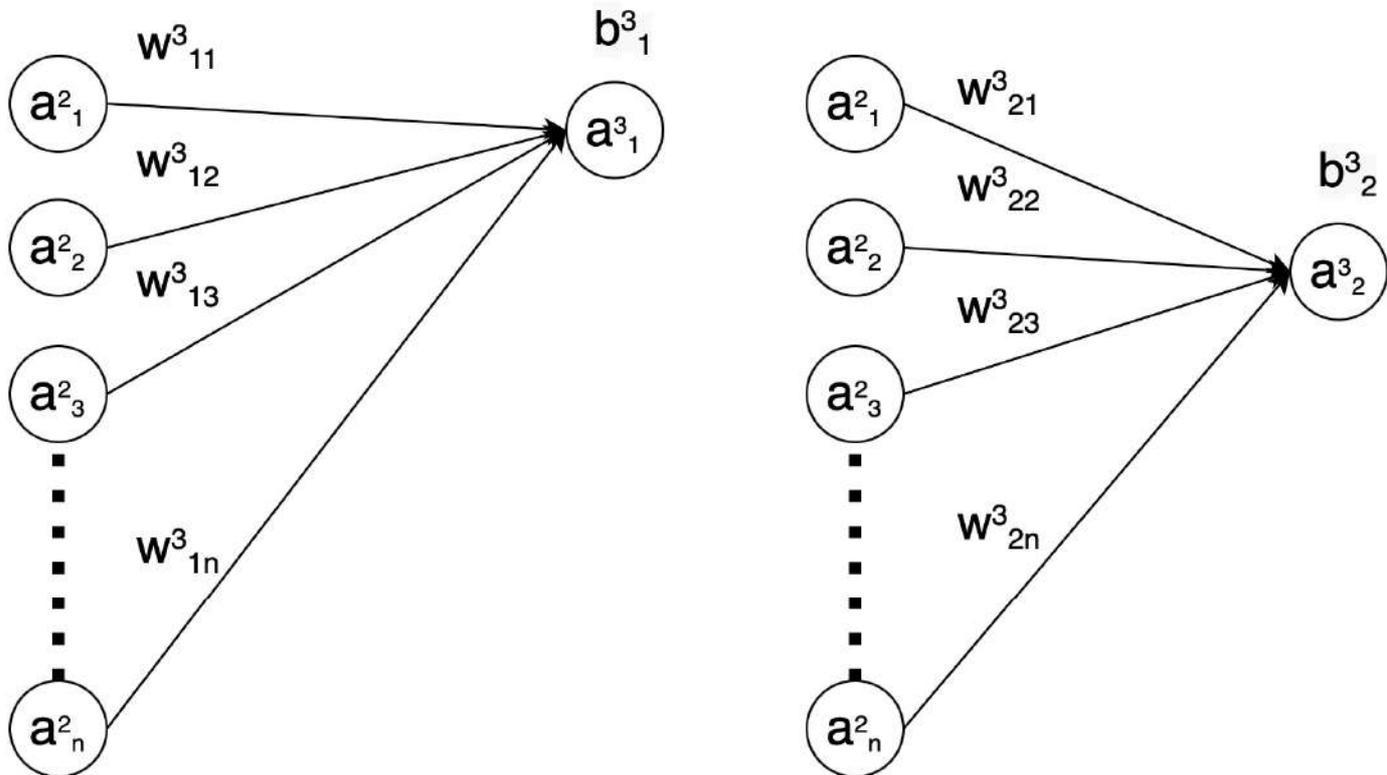


Рис. 3. Схема нейронной сети прямого распространения.



4. Схема связей в скрытом и выходном слое.

Рис.

Обучение нейронной сети В изначальной конфигурации ИНС со случайно инициализированными значениями весов W , выходные значения ИНС a_i^L будут далеки от реальных значений функции y_i , где индекс i пробегает количество переменных в результатах функции и количество выходных нейронов, соответственно. Для того, чтобы ИНС выдавала нужные результаты необходимо подобрать оптимальные значения W и \vec{b} . Это реализуется методом **градиентного спуска** - то есть минимизацией некоторой функции, которая определяет точность ИНС в сравнении с реальными данными. Такая функция называется **функцией стоимости** и обычно задается как среднее квадратичное отклонение:

$$C_j = \frac{1}{n} \sum_i (y_i - a_i^L)^2$$

, где n - количество выходных нейронов в ИНС. Предыдущий формула показывает как вычислить градиент по одному элементу из данных, градиент по всему набору данным соответственно равен

$$C^{full} = \frac{1}{m} \sum_j C_j$$

, где m - количество элементов в наборе данных. Обучение ИНС сводится к обновлению параметров сети путем вычисления градиента функции стоимости и шага значений параметров в сторону минимума функции стоимости с шагом μ

$$w_{i,j}^{l,new} = w_{i,j}^l - \mu \frac{\partial C^{full}}{\partial w_{i,j}^l}$$

$$b_i^{l,new} = b_i^l - \mu \frac{\partial C^{full}}{\partial b_i^l}$$

Вычисление частных производных производится путем применения **алгоритма обратного распространения**

$$\delta_i^L = \nabla_{a_i} C^{full} \odot \sigma'(z_i^L)$$

$$\delta_i^l = (w_{i,j}^{l+1})^T \delta_i^{l+1} \odot a'(z_i^l)$$

$$\frac{\partial C^{full}}{\partial b_i^l} = \delta_i^l$$

$$\frac{\partial C^{full}}{\partial w_{i,j}^l} = a_j^{l-1} \delta_i^l$$

Обучение ИНС состоит из последовательного прямого и обратного проходов с изменением параметров W и \vec{b} до получения необходимой точности.

Ссылка на данные: <https://www.kaggle.com/competitions/modern-computer-technologies-laboratory/data?select=train.csv>

<https://www.kaggle.com/competitions/modern-computer-technologies-laboratory/data?select=test.csv>

Литература: 1. <http://neuralnetworksanddeeplearning.com> 2. <https://www.deeplearningbook.org>

