

# Основные понятия курса

Марк Блуменау

31 марта 2025 г.

## Предисловие

Знание ответов на данные вопросы и возможность ответить на 4 из них за 3 минуты гарантирует студенту оценку 4 'удовл.' за экзамен. Ответы приведены такие, которые точно достаточны, но, естественно, принимается любая верная формулировка и формула. В случае опечаток незамедлительно сообщите @markblumenu.

## Вопросы

### 1. Виды градиентного спуска и функций потерь

1. **Что такое mini-batch стохастический градиентный спуск?**

Ответ: Обновление параметров модели на основе mini-batch, а не полного набора данных.

2. **Какова идея лежит в основе градиентного спуска?**

Ответ: Итеративное изменение параметров для минимизации функции потерь.

3. **Что такое MSE?**

Ответ:  $MSE = \frac{1}{n} \sum (y_{\text{pred}} - y_{\text{true}})^2$ .

4. **Какую функцию потерь можно использовать для обучения в задаче классификации?**

Ответ: Кросс-энтропия. (Любые другие корректные ответы принимаются)

5. **Что такое оптимизатор Adam?**

Ответ: Комбинирует методы AdaGrad и RMSProp для адаптивного шага обучения.

6. **Какую проблему пытается решить моментум в градиентном спуске?**

Ответ: Проблему застревания в неглубоких локальных минимумах.

7. **Каким методом вычисляются градиенты в нейросетях?**

Ответ: Методом обратного распространения ошибки.

8. **Как в PyTorch называется параметр для регуляризации весов?**

Ответ: Weight decay.

## 2. Linear, Dropout + Batch Normalization

1. **Что такое линейный слой (Fully Connected Layer)?**

Ответ: Это слой, который описывается формулой:  $y = xA^T + b$ .

2. **Что такое Dropout (кратко)?**

Ответ: Случайное отключение нейронов для предотвращения переобучения.

3. **Что такое Batch Normalization (кратко)?**

Ответ: Масштабирование батча на входе в слой так, что матожидание 0, а дисперсия 1.

4. **Какие параметры используются в Batch Normalization?**

Ответ: Параметры  $\gamma$  (масштаб) и  $\beta$  (смещение).

5. **Какова общая цель использования Dropout и BatchNorm?**

Ответ: Улучшение обобщения и стабильности обучения модели.

6. **Какое количество параметров у Linear слоя?**

Ответ: (Количество входов + 1) \* количество выходов.

7. **Как работает Dropout с точки зрения формул?**

Ответ:  $y_{\text{dropout}} = (r \odot y)$ , где  $r \sim \text{Bernoulli}(p)$ .

8. **Запишите формулы для вычисления Batch Normalization.**

Ответ:

$$\mu_B = \frac{1}{m} \sum_i x_i, \quad \sigma_B^2 = \frac{1}{m} \sum_i (x_i - \mu_B)^2, \quad \hat{x}_i = \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \varepsilon}}, \quad y_i = \gamma \hat{x}_i + \beta.$$

9. **Какую роль играет функция активации ReLU?**

Ответ: Вводит нелинейность:  $f(x) = \max(0, x)$ .

10. **Почему нужна нелинейность между слоями?**

Ответ: Использование двух линейных слоев подряд эквивалентно использованию одного линейного слоя.

11. **Что такое Универсальная теорема аппроксимации? (Цыбенко)**

Ответ: Линейная нейросеть с одним скрытым слоем может аппроксимировать любую непрерывную функцию многих переменных с любой точностью.

### 3. Операции в CNN

1. **Что такое свёрточный слой в CNN?**

Ответ: Выполняет операцию свёртки изображения с фильтрами для извлечения признаков.

2. **Сколько параметров в Conv2d слое?**

Ответ:  $K \cdot K \cdot C_{in} \cdot C_{out} + C_{out}$ , где  $K$  - размер ядра,  $C$  - количество каналов.

3. **Сколько умножений в Conv2d слое?**

Ответ:  $H \cdot W \cdot C_{out} \cdot (K \cdot K \cdot C_{in})$ , где  $K$  - размер ядра,  $C$  - количество каналов,  $H$  и  $W$  - высота и ширина выходного тензора.

4. **Какова роль ядра свёртки?**

Ответ: Выделяет локальные паттерны через взвешенное суммирование пикселей.

5. **Что такое stride (шаг) в свёрточном слое?**

Ответ: Количество пикселей, на которое сдвигается фильтр при свёртке.

6. **Что такое padding в свёрточном слое?**

Ответ: Добавление дополнительных пикселей для сохранения размерности.

7. **Что делает операция pooling?**

Ответ: Уменьшает размерность, беря заданную функцию в заданном окне.

8. Как вычисляется размер выходного тензора свёрточного слоя?

Ответ:  $O = \frac{W-F+2P}{S} + 1$ , где  $W$  — размер входа,  $F$  — размер ядра,  $P$  — паддинг,  $S$  — шаг.

9. Запишите формулу свёрточной операции для изображения.

Ответ:

$$\text{Im}^{\text{out}}(x, y, t) = \sum_{i=-d}^d \sum_{j=-d}^d \sum_{c=1}^c (K_t(i, j, c) \text{Im}^{\text{in}}(x + i, y + j, c) + b_t)$$

10. Как изменяется размерность при max pooling?

Ответ: Размер =  $\frac{W-F}{S} + 1$ , где  $F$  — размер окна,  $S$  — шаг pooling.

11. Что такое zero-padding и для чего он применяется?

Ответ: Добавление нулевых элементов по границам для сохранения размерности после свёртки.

## 4. Архитектуры CNN

1. Чем характеризуется архитектура VGG?

Ответ: Глубокая сеть с последовательными свёрточными слоями и небольшими фильтрами.

2. Запишите формулу остаточного блока (residual block) в ResNet.

Ответ:  $y = F(x, \{W_i\}) + x$ .

3. Что такое ResNet?

Ответ: Сеть с остаточными (residual) связями, позволяющими строить более глубокие модели.

## 5. Задачи компьютерного зрения (CV tasks)

1. Что такое задача классификации изображений?

Ответ: Определение категории объекта на изображении.

2. Что такое задача детекции объектов?

Ответ: Выделение и локализация объектов с указанием их координат и категории.

3. **Что такое сегментация изображений?**

Ответ: Разделение изображения на области, соответствующие различным объектам.

4. **Запишите формулу для вычисления IoU в детекции объектов.**

Ответ:  $\text{IoU} = \frac{\text{area}(\text{intersection})}{\text{area}(\text{union})}$ .

5. **Как применяется Non-Maximum Suppression (NMS) в обнаружении объектов?**

Ответ: Выбираются боксы с наибольшей уверенностью, удаляя сильно перекрывающиеся с ними.

6. **Как формализуется задача сегментации изображений?**

Ответ: Минимизируется функция  $L(\text{mask}_{\text{pred}}, \text{mask}_{\text{true}})$ , например, dice loss.

7. **Как определяется метрика mAP для оценки детекции?**

Ответ: Усреднением AP по классам.  $\text{mAP} = \frac{1}{N} \sum \text{AP}_i$ , где AP — площадь под кривой precision-recall.

## 6. Эмбединги

1. **Что такое эмбединги?**

Ответ: Представление дискретных данных в виде плотных векторных пространств.

2. **Что такое cosine similarity для эмбедингов?**

Ответ:  $\cos(\theta) = \frac{A \cdot B}{\|A\| \cdot \|B\|}$ .

3. **Как выглядит softmax-функция для нормализации эмбедингов?**

Ответ:  $\text{softmax}(z_i) = \frac{\exp(z_i)}{\sum_j \exp(z_j)}$ .

## 7. RNN, LSTM, GRU

1. **Что такое RNN?**

Ответ: Рекуррентная нейронная сеть, обрабатывающая последовательные данные с сохранением состояния (в так называемом скрытом состоянии).

2. **Что такое LSTM?**

Ответ: RNN с ячейками памяти и гейтами для контроля потока информации, предотвращающих затухание градиентов.

3. **Что такое GRU?**

Ответ: Упрощённая версия LSTM с двумя гейтами (обновления и сброса).

4. **Запишите формулу обновления скрытого состояния в стандартном RNN.**

Ответ:

$$h_t = \tanh(x_t W_{ih}^T + b_{ih} + h_{t-1} W_{hh}^T + b_{hh})$$