

# ФОС по курсу Современные архитектуры глубоких ИНС

**1:** Какой порядок вычислений происходит в модели перцептрона?

- (a) Вычисляется функция активации входа, а затем линейная функция  $(Wx + b)$ .
- (b) Вычисляется линейная функция входа  $(Wx + b)$ , к которой затем применяется функция активации.
- (c) Входные данные усредняются, вычисляется линейная функция  $(Wx + b)$ , а затем функция активации.

**2:** К выходному значению искусственного нейрона применяется:

- (a) функция активации
- (b) взятие производной
- (c) нормализация

**3:** Оценка качества ИНС проводится с использованием:

- (a) градиентного спуска
- (b) функции активации
- (c) функции ошибки

**4:** Функция потерь в процессе обучения ИНС выполняет следующую функцию:

- (a) вычисляет ошибку вывода ИНС
- (b) изменяет веса ИНС
- (c) изменяет параметры расписания обучения

**5:** Обратное распространение ошибки это:

- (a) процесс вычисления вывода ИНС на пакете входных примеров
- (b) процесс расчета градиента ошибки и изменение весов
- (c) процесс расчета скорости обучения ИНС

**6:** Скорость обучения это:

- (a) число эпох в цикле обучения ИНС
- (b) время, которое необходимо затратить на обучение ИНС
- (c) параметр шага в алгоритме градиентного спуска

**7:** Одна эпоха цикла обучения ИНС это:

- (a) этап обучения ИНС, на котором обработан один пакет данных
- (b) этап обучения ИНС, на котором обработаны все имеющиеся пакеты данных
- (c) весь цикл обучения ИНС

**8:** В ИНС прямого распространения подряд стоят два слоя из пяти и четырех нейронов соответственно. Сколько весов-связей существует между этими слоями?

- (a) 20
- (b) 9
- (c) 5

**9:** В чем заключается смысл преобразования softmax?

- (a) полученные числа можно интерпретировать как вероятности принадлежности примера к классам в задаче классификации
- (b) данное преобразование необходимо для регуляризации
- (c) данное преобразование необходимо для численной стабилизации выходных данных

**10:** Деактивация части нейронов на каждой итерации обучения ИНС называется:

- (a) дропаут
- (b) регуляризация
- (c) нормализация по мини-батчам

**11:** Если нейронная сеть содержит вес  $w$ , то функция потерь при  $L2$ -регуляризации будет содержать слагаемое

- (a)  $\lambda|w|$
- (b)  $\lambda w^2$
- (c)  $\lambda_1|w| + \lambda_2 w^2$

**12:** Формула  $h(z) = \max(0, z)$  задает функцию активации:

- (a) сигмоида
- (b) ReLU
- (c) Swish

**13:** Какой из следующих приемов решает проблему сдвига распределения выходов слоя ИНС?

- (a) дропаут
- (b) регуляризация
- (c) нормализация по мини-батчам

**14:** Ранняя остановка это:

- (a) предварительное завершение цикла обучения, по причине отсутствия улучшения метрики качества работы ИНС на протяжении нескольких эпох подряд
- (b) отбрасывание пакетов данных, по причине того, что на заданной эпохе уже достигнута нужное значение метрики качества работы ИНС
- (c) прекращение изменения параметров скорости обучения в процессе обучения по расписанию, по причине того, что функция ошибки вышла на плато

**15:** Какой эффект характерен для алгоритма градиентного спуска ADAM?

- (a) сверхбыстрая сходимость к некоторой точке локального максимума
- (b) стабильность обучения и малая вероятность переобучения
- (c) его вычислительная сложность существенно ниже, чем у обычного градиентного спуска

**16:** Сверточные сети рекомендуются использовать для:

- (a) обработки последовательностей
- (b) обработки изображений
- (c) обучения с подкреплением мультиагентной системы

**17:** К изображению размера 25 на 25 применяется свертка с размером окна 5 на 5, параметрами  $\text{stride}=1$ ,  $\text{padding}=0$ . Какой размер получается на выходе?

- (a) 21 на 21
- (b) 20 на 20
- (c) 22 на 22

**18:** Батч-нормализацию применяют:

- (a) перед функцией активации
- (b) после функции активации
- (c) только на выходном слое

**19:** Применяется ли dropout при вычислении метрики качества модели на тестовой выборке?

- (a) да
- (b) нет

**20:** Чему равен результат макспулинга 2x2 для тензора  $[[1, 2, 3, 4], [5, 6, 7, 8]]$ ?

- (a) 48
- (b) 68
- (c) 5678

**21:** Для чего предназначена свертка 1 на 1?

- (a) уменьшения размерности карты признака
- (b) выделение дополнительных признаков и увеличение выразительной способности
- (c) улучшения качества входного изображения?

**22:** Свертки какого размера используются в большинстве сверточных слоях архитектуры VGG?

- (a) 3 на 3
- (b) 1 на 1
- (c) 5 на 5

**23:** В какой из следующих архитектур используются остаточные связи?

- (a) AlexNet
- (b) VGG
- (c) ResNet

**24:** Какую из следующих архитектур можно использовать для решения задачи сегментации изображения?

- (a) UNet
- (b) seq2seq
- (c) VGG

**25:** После обучения GAN может генерировать достоверные изображения яблок. Верно ли, что его дискриминатор можно использовать для задачи классификации: яблоко или не яблоко?

- (a) да
- (b) нет

**26:** Как расшифровывается аббревиатура LSTM?

- (a) Long short-term memory
- (b) Least saturated memory
- (c) Long shifted-time memory

**27:** По ходу обучения стало заметно, что значения весов и активаций принимают значения NaN. Какую проблему вы наблюдаете?

- (a) затухание градиентов
- (b) взрыв градиентов
- (c) недообучение

**28:** Задача улучшения качества медицинских снимков относится к:

- (a) обучению с подкреплением
- (b) сегментации изображений
- (c) генеративным задачам

**29:** Какую технику обычно применяют для решения проблемы взрывающихся градиентов?

- (a) ограничение максимума нормы градиентов
- (b)  $L^2$ -регуляризацию
- (c) дропаут

**30:** Какого типа RNN архитектуры применяют для решения задачи машинного перевода?

- (a) one-to-many
- (b) many-to-many
- (c) many-to-one