

Вопросы тестов по курсу «Глубокое обучение»

Александр Дьяконов (dyakonov.org)

версия 12.01.2023 (85 вопросов, каталог [2022_MSUDL]), темы:

- нейронные сети (НС),
- обучение НС,
- PyTorch,
- свёрточные сети (CNN), архитектуры CNN,
- детектирование объектов, семантическая сегментация,
- рекуррентные сети (RNN),
- обработка текстов (NLP).

Ниже представлены вопросы, которые использовались для проверки знаний по курсу «Глубокое обучение».

- это не вопросы проверки знаний темы «Deep Learning», все вопросы контекстные и имеют отношение к конкретному курсу,
- курс был довольно сокращённый (как по перечню тем, так и по глубине каждой темы),
- не все вопросы представлены,
- вопросы типа «из какой статьи рисунок» являются проверкой не знания статей, а понимания, что изображено на рисунках,
- некоторые формулировки сокращены, например в вопросах «Что при обратном распространении градиента передаётся» имеется в виду «какой градиент передаётся»,
- в вопросах «в какой сети впервые...» выбор ограничен только теми архитектурами, которые рассматривались на лекции,
- в каждом вопросе может быть несколько верных ответов, верные ответы не помечены.

♦ Какую функцию активации используют в последнем слое НС при решении задачи бинарной классификации?

- ☐ сигмоида
- ☐ тождественная
- ☐ ReLU
- ☐ softmax
- ☐ ELU
- ☐ swish

♦ Что можно реализовать в виде нейрона НС?

- ☐ линейная регрессия
- ☐ линейный классификатор
- ☐ логистическая регрессия
- ☐ решающее дерево глубины 1

♦ Какую функции активации используют в последнем слое в (общей) задаче регрессии?

- ☐ сигмоида
- ☐ тождественная
- ☐ ReLU
- ☐ softmax
- ☐ ELU
- ☐ swish

♦ Как реализована сигмоида в pytorch?

- ☐ в виде функции `torch.nn.functional.sigmoid`
- ☐ в виде модуля `torch.nn.Sigmoid`
- ☐ передаётся как параметр линейным модулям `torch.nn.Linear(activation='sigmoid')`
- ☐ нет явной реализации

♦ Сколько нейронов с пороговой активацией достаточно для реализации XOR?

- ☐ 4
- ☐ 3
- ☐ 2
- ☐ 1

♦ Теорема об универсальной аппроксимации Хорника о... (м.б. несколько вариантов)

- ☐ реализации любых функций
- ☐ реализации непрерывных функций
- ☐ реализации гладких функций
- ☐ однослойных сетях
- ☐ двухслойных сетях
- ☐ глубоких сетях

♦ В теореме аппроксимации Пинкуса... (м.б. несколько вариантов)

- ☐ полиномиальные активации в 1м слое
- ☐ полиномиальные активации везде
- ☐ неполиномиальные активации в 1м слое
- ☐ тождественные активации во 2м слое
- ☐ однослойные сети
- ☐ двухслойные сети
- ☐ глубокие сети

♦ Что означает черта в конце названия метода в pytorch?

- ☐ inplace
- ☐ hidden
- ☐ private
- ☐ forbidden

♦ При реализации softmax(a) на практике, применяют приём, в котором

- ☐ вычитается $\max(a)$
- ☐ прибавляется $\max(a)$
- ☐ делится на $\max(a)$
- ☐ умножается на $\max(a)$
- ☐ вычитается $\min(a)$
- ☐ прибавляется $\min(a)$
- ☐ делится на $\min(a)$
- ☐ умножается на $\min(a)$

♦ При вычислении `loss = F.cross_entropy(input, target)` тензор input имеет размеры 3×5. Какие размеры у target?

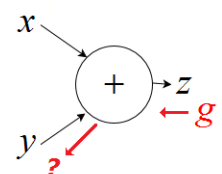
- ☐ 3×5
- ☐ 5×3
- ☐ 3×3
- ☐ 5×5
- ☐ 1
- ☐ 3
- ☐ 5
- ☐ любые

♦ Какая сложность нахождения всех мёртвых нейронов (Dead Neurons) от длины выборки m?

- ☐ m
- ☐ sqrt(m)
- ☐ m²
- ☐ m log(m)
- ☐ m² log(m)
- ☐ m³

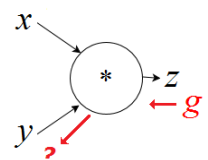
♦ Что при обратном распространении градиента передаётся в переменную y гейта сложения?

- ☐ g
- ☐ zg
- ☐ xg
- ☐ yg
- ☐ x
- ☐ y
- ☐ 0



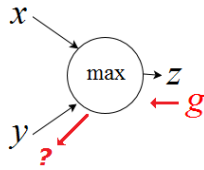
♦ Что при обратном распространении градиента передаётся в переменную y гейта умножения?

- ☐ g
- ☐ zg
- ☐ xg
- ☐ yg
- ☐ x
- ☐ y
- ☐ 0



♦ Что при обратном распространении градиента передаётся в переменную y max-гейта при $x > y$?

- ☐ g
- ☐ zg
- ☐ xg
- ☐ yg
- ☐ x
- ☐ y
- ☐ 0



♦ Какие из перечисленных функций активаций немонотонные?

- | | |
|--|--------------------------------|
| <input type="checkbox"/> сигмоида | <input type="checkbox"/> swish |
| <input type="checkbox"/> тождественная | <input type="checkbox"/> ELU |
| <input type="checkbox"/> ReLU | <input type="checkbox"/> GELU |

♦ Пусть x – тензор размера 2×2 , что выдаст $x.\text{unsqueeze}(\text{dim}=1).\text{shape}$?

- | | |
|---------------------------------|---|
| <input type="checkbox"/> [2, 2] | <input type="checkbox"/> [1, 2, 2] |
| <input type="checkbox"/> [4] | <input type="checkbox"/> [2, 1, 2] |
| <input type="checkbox"/> [4, 1] | <input type="checkbox"/> [2, 2, 1] |
| <input type="checkbox"/> [1, 4] | <input type="checkbox"/> нет верного ответа |

♦ Пусть x – тензор размера 2×2 , что выдаст $x.\text{squeeze}().\text{shape}$?

- | | |
|---------------------------------|---|
| <input type="checkbox"/> [2, 2] | <input type="checkbox"/> [1, 2, 2] |
| <input type="checkbox"/> [4] | <input type="checkbox"/> [2, 1, 2] |
| <input type="checkbox"/> [4, 1] | <input type="checkbox"/> [2, 2, 1] |
| <input type="checkbox"/> [1, 4] | <input type="checkbox"/> нет верного ответа |

♦ Что выдаст следующий код?

```
x = torch.Tensor([1, 2, 3])
w = torch.tensor([1., 1, 1], requires_grad=True)
z = w @ x
z.backward()
print(x.grad)
```

- | | |
|------------------------------------|------------------------------------|
| <input type="checkbox"/> ошибку | <input type="checkbox"/> [1, 1, 1] |
| <input type="checkbox"/> None | <input type="checkbox"/> [0, 0, 0] |
| <input type="checkbox"/> [1, 2, 3] | <input type="checkbox"/> 0 |

♦ Строчки кода ниже являются признаком...

```
from torchvision import models
resnet18 = models.resnet18(pretrained=True)
```

- ☐ обучения с учителем
- ☐ обучения без учителя
- ☐ трансферного обучения
- ☐ обучения с масками
- ☐ самообучения

♦ Инициализация Ксавьера (`nn.init.xavier_uniform`) выведена в предположении...

- ☐ тождественной активации
- ☐ однослойности сети
- ☐ большой глубины сети
- ☐ равномерного распределения сигналов
- ☐ отсутствия смещения (`bias=False`)

♦ Техника **Early Stopping** применяется (отметьте все верные варианты):

- ☐ при обучении бустинга
- ☐ при обучении нейросетей
- ☐ при работе сегментаторов
- ☐ при отсутствии стоп-токенов

♦ В каких методах оптимизации градиент вычисляется не в текущей точке?

- | | |
|------------------------------------|----------------------------------|
| <input type="checkbox"/> SGD | <input type="checkbox"/> Adagrad |
| <input type="checkbox"/> momentum | <input type="checkbox"/> RMSprop |
| <input type="checkbox"/> Нестерова | <input type="checkbox"/> Adam |

♦ В каких методах оптимизации шаг делается не строго в сторону текущего градиента?

- | | |
|------------------------------------|----------------------------------|
| <input type="checkbox"/> SGD | <input type="checkbox"/> Adagrad |
| <input type="checkbox"/> momentum | <input type="checkbox"/> RMSprop |
| <input type="checkbox"/> Нестерова | <input type="checkbox"/> Adam |

♦ В каких методах оптимизации шаг нормируется (градиент делится на что-то)?

- | | |
|------------------------------------|----------------------------------|
| <input type="checkbox"/> SGD | <input type="checkbox"/> Adagrad |
| <input type="checkbox"/> momentum | <input type="checkbox"/> RMSprop |
| <input type="checkbox"/> Нестерова | <input type="checkbox"/> Adam |

♦ Что означает буква D в названии метода оптимизации Adam?

- | | |
|----------------------------------|---|
| <input type="checkbox"/> Double | <input type="checkbox"/> Dummy |
| <input type="checkbox"/> Descent | <input type="checkbox"/> нет верного ответа |

♦ Где в Pytorch указывается необходимость перемешивания выборки (`shuffle=True`)?

- | | |
|--|---|
| <input type="checkbox"/> TensorDataset | <input type="checkbox"/> backward |
| <input type="checkbox"/> DataLoader | <input type="checkbox"/> при инициализации HC |
| <input type="checkbox"/> optimizer | <input type="checkbox"/> нет верного ответа |

♦ Где в Pytorch указывается размер батча (`batch_size=512`)?

- | | |
|--|---|
| <input type="checkbox"/> TensorDataset | <input type="checkbox"/> backward |
| <input type="checkbox"/> DataLoader | <input type="checkbox"/> при инициализации HC |
| <input type="checkbox"/> optimizer | <input type="checkbox"/> нет верного ответа |

♦ Где в коде меняются параметры нейросети?

- ☐ optimizer.zero_grad()
- ☐ output = model(input)
- ☐ loss = loss_fn(output, target)
- ☐ loss.backward()
- ☐ optimizer.step()

♦ Где в коде вычисляются градиенты в вычислительном графе нейросети?

- ☐ optimizer.zero_grad()
- ☐ output = model(input)
- ☐ loss = loss_fn(output, target)
- ☐ loss.backward()
- ☐ optimizer.step()

♦ Где используется указанная формула?

- ☐ Standard Dropout
- ☐ Inverted Dropout
- ☐ DropConnect
- ☐ Gradient clipping
- ☐ BatchNorm

$$y = \frac{1}{1-p} f(Wx) \circ m$$

♦ После какой строки логично делать обрезку градиента (clip_grad_norm)?

- ☐ optimizer.zero_grad()
- ☐ output = model(input)
- ☐ loss = loss_fn(output, target)
- ☐ loss.backward()
- ☐ optimizer.step()

♦ При обосновании какого метода упоминался эффект Covariate shift (смещения распределений)?

- ☐ Standard Dropout
- ☐ Gradient clipping
- ☐ Inverted Dropout
- ☐ BatchNorm
- ☐ DropConnect
- ☐ аугментация

♦ Какие слои по-разному работают в режимах train() и eval()?

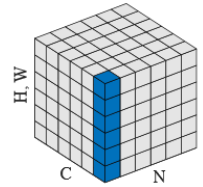
- ☐ Dropout
- ☐ Linear
- ☐ BatchNorm
- ☐ conv2d
- ☐ ReLU
- ☐ нет верного ответа

♦ Какой параметр передаётся при инициализации программатора оптимизации torch.optim.lr_scheduler?

- ☐ нейросеть
- ☐ градиенты нейросети
- ☐ параметры нейросети
- ☐ оптимизатор

♦ При какой нормировке статистики считаются по указанному региону?

- ☐ BatchNorm
- ☐ Layer Norm
- ☐ Instance Norm
- ☐ Group Norm
- ☐ Batch renormalization
- ☐ нет верного ответа



♦ Что делается в Snapshot Ensemble?

- ☐ стекинг нейросетей
- ☐ усреднение нейронки и другого метода
- ☐ усреднение сетей с разной инициализацией
- ☐ усреднение сетей с разными гиперпараметрами
- ☐ усреднение сети на разных эпохах

♦ Тензор x имеет размер [1, 1, 28, 28], что выведет код?

```
f = nn.Conv2d(in_channels=1, out_channels=1, kernel_size=3)
```

```
F.max_pool2d(f(x), kernel_size=2).shape
```

- ☐ [1, 1, 28, 28]
- ☐ [1, 1, 26, 26]
- ☐ [1, 1, 14, 14]
- ☐ [1, 1, 13, 13]
- ☐ [1, 1, 11, 11]
- ☐ [1, 1, 5, 5]

♦ Какая сеть считается первой свёрточной?

- ☐ LeNet-5
- ☐ AlexNet
- ☐ VGG
- ☐ GoogLeNet
- ☐ ResNet
- ☐ нет верного ответа

♦ В какой сети впервые использовались ReLU и maxpool?

- ☐ LeNet-5
- ☐ AlexNet
- ☐ VGG
- ☐ GoogLeNet
- ☐ ResNet
- ☐ нет верного ответа

♦ В какой сети впервые использовались последовательности («каскад») 3×3-свёрток?

- ☐ LeNet-5
- ☐ AlexNet
- ☐ VGG
- ☐ GoogLeNet
- ☐ ResNet
- ☐ нет верного ответа

♦ В какой сети использовалась модульная архитектура и было всего 5M параметров?

- ☐ LeNet-5
- ☐ AlexNet
- ☐ VGG
- ☐ GoogLeNet
- ☐ ResNet
- ☐ нет верного ответа

♦ В какой сети впервые использовали BatchNorm и отказались от Dropout?

- ☐ LeNet-5
- ☐ AlexNet
- ☐ VGG
- ☐ GoogLeNet
- ☐ ResNet
- ☐ нет верного ответа

♦ Выберите верные утверждения для «Squeeze-and-Excitation» (SE) блока:

- ☐ не меняет размеры тензора
- ☐ делает перекалибровку каналов
- ☐ использует двухслойную нейросеть
- ☐ использует «узкое горло»

♦ Тензор x имеет размер [1, 20, 30, 10], что выведет код?

```
ct = nn.ConvTranspose2d(in_channels=20,  
                        out_channels=10, kernel_size=2)
```

ct(x).shape

- ☐ [20, 10, 29, 9]
- ☐ [20, 10, 30, 10]
- ☐ [20, 10, 31, 11]
- ☐ [1, 10, 29, 9]
- ☐ [1, 10, 30, 10]
- ☐ [1, 10, 31, 11]

♦ Перед какими слоями разумно делать свёртку без смещения (bias=False)?

- ☐ ReLU
- ☐ Dropout
- ☐ BatchNorm
- ☐ maxpool

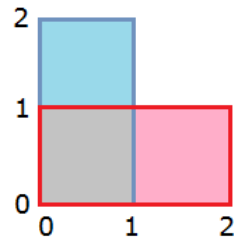
♦ Как называется задача, в которой ответ представляется в виде приведённого рисунка?

- ☐ Локализация
- ☐ Семантическая сегментация
- ☐ Детектирование объектов
- ☐ Сегментация объектов



♦ Подсчитайте IoU для прямоугольников на рисунке.

- ☐ 0
- ☐ 1/4
- ☐ 1/3
- ☐ 1/2
- ☐ 2/3
- ☐ 3/4
- ☐ нет верного ответа



♦ Вычислите Average Precision (AP) для указанного случая (рассматривается конкретный класс, ответ округлить до десятых).

- ☐ 0.5
- ☐ 0.6
- ☐ 0.7
- ☐ 0.8
- ☐ нет верного ответа

conf	IoU>th
0.6	true
0.9	true
0.7	false
0.8	true

♦ Выберите верные утверждения для R-CNN:

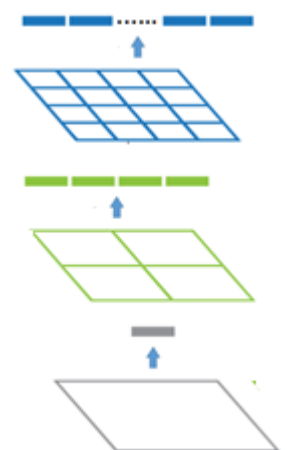
- ☐ использовался Selective Search
- ☐ использовался SVM
- ☐ использовалась bounding-box-перпексия
- ☐ использовался ROI Pooling слой
- ☐ использовалась Region Proposal сеть (RPN)

♦ Для чего нужна процедура NMS (Non Maximum Suppression)?

- ☐ для сокращения числа регионов
- ☐ для уменьшения числа классов в ответе
- ☐ для уменьшения числа параметров сети
- ☐ для уменьшения уверенности в регионах-ответах

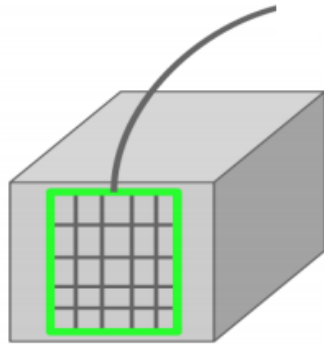
♦ Пусть тензор имеет размеры 10×100×100 (10 – число каналов), какой размер будет у ответа Spatial Pyramid Pooling слоя (см. рис.)?

- ☐ 10
- ☐ 4
- ☐ 16
- ☐ 32
- ☐ 160
- ☐ 200
- ☐ 210
- ☐ 1600
- ☐ нет верного ответа



♦ Пусть тензор имеет размеры $10 \times 100 \times 100$ (10 – число каналов), какой размер будет после ROI Pooling слой с сеткой 5×5 (см. рис.)?

- ☐ 10
- ☐ 25
- ☐ 5×5
- ☐ 5×10
- ☐ 100×100
- ☐ $1 \times 5 \times 5$
- ☐ $10 \times 5 \times 5$
- ☐ $10 \times 100 \times 100$
- ☐ нет верного ответа



♦ Выберите верные утверждения для Fast R-CNN:

- ☐ использовался Selective Search
- ☐ использовался SVM
- ☐ использовалась bounding-box-регрессия
- ☐ использовался ROI Pooling слой
- ☐ использовалась Region Proposal сеть (RPN)

♦ Выберите верные утверждения для Faster R-CNN:

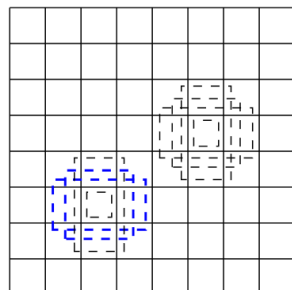
- ☐ использовался Selective Search
- ☐ использовался SVM
- ☐ использовалась bounding-box-регрессия
- ☐ использовался ROI Pooling слой
- ☐ использовалась Region Proposal сеть (RPN)

♦ Сеть YOLO строится для сетки 2×2 , 10 классов, 2 региона для каждого класса. Сколько элементов в выходном тензоре?

- ☐ 4
- ☐ 10
- ☐ 20
- ☐ 40
- ☐ 80
- ☐ нет верного ответа

♦ Из статьи про какую сеть эта иллюстрация?

- ☐ R-CNN
- ☐ Fast R-CNN
- ☐ YOLO
- ☐ SSD
- ☐ нет верного ответа



(b) 8×8 feature map

♦ Выберите верные утверждения для CornerNet:

- ☐ безякорная сеть
- ☐ использовался особый вид пулинга
- ☐ использовалась Region Proposal сеть (RPN)

♦ Пусть в SSD на масштабе (тензоре) $3 \times 3 \times 256$ используется 3 якоря, 10 классов. Какой размер тензора классификации на этом масштабе?

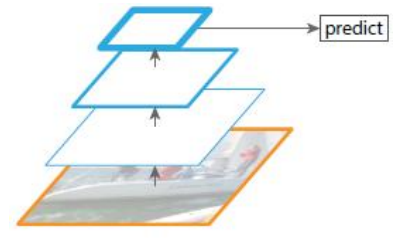
- ☐ 3×10
- ☐ 3×11
- ☐ 3×30
- ☐ 3×33
- ☐ $3 \times 3 \times 10$
- ☐ $3 \times 3 \times 11$
- ☐ $3 \times 3 \times 30$
- ☐ $3 \times 3 \times 33$
- ☐ нет верного ответа

♦ Пусть в SSD на масштабе (тензоре) $3 \times 3 \times 256$ используется 3 якоря, 10 классов. Какой размер тензора BV-регрессии на этом масштабе?

- ☐ 3×10
- ☐ 3×11
- ☐ 3×12
- ☐ $3 \times 3 \times 10$
- ☐ $3 \times 3 \times 11$
- ☐ $3 \times 3 \times 12$
- ☐ $3 \times 3 \times 33$
- ☐ нет верного ответа

♦ Какой архитектуре соответствует рисунок?

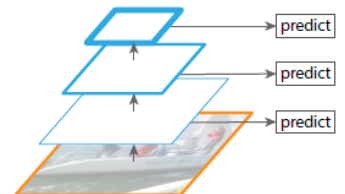
- ☐ R-CNN
- ☐ Fast R-CNN
- ☐ YOLO
- ☐ SSD
- ☐ FPN
- ☐ U-net



(b) Single feature map

♦ Какой архитектуре соответствует рисунок?

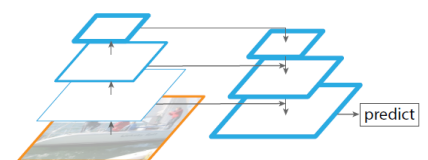
- ☐ R-CNN
- ☐ Fast R-CNN
- ☐ YOLO
- ☐ SSD
- ☐ FPN
- ☐ U-net



(c) Pyramidal feature hierarchy

♦ Какой архитектуре соответствует рисунок?

- ☐ R-CNN
- ☐ Fast R-CNN
- ☐ YOLO
- ☐ SSD
- ☐ FPN
- ☐ U-net



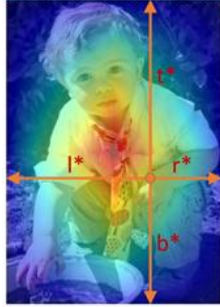
(e) Similar Structure with (d)

♦ В FCOS два целевых региона пересекаются. Для точки в пересечении какой регион считается целевым?

- ☐ оба
- ☐ новый регион - пересечение
- ☐ случайный
- ☐ наибольший
- ☐ наименьший
- ☐ никакой

♦ Из статьи про какую сеть эта иллюстрация?

- ☐ R-CNN
- ☐ Fast R-CNN
- ☐ YOLO
- ☐ SSD
- ☐ CornerNet
- ☐ FCOS
- ☐ нет верного ответа



♦ Какой размер выходного тензора-регрессии FCOS на масштабе 100×100 ?

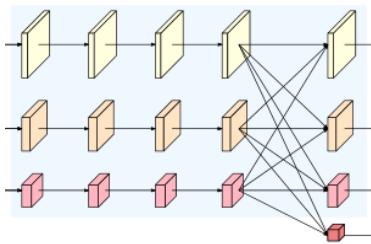
- ☐ $10 \times 10 \times 2$
- ☐ $100 \times 100 \times 2$
- ☐ $10 \times 10 \times 4$
- ☐ $100 \times 100 \times 4$
- ☐ 100×100
- ☐ нет верного ответа

♦ Что использовалось в указанных сетях?

- ☐ «обратный» пулинг в SegNet
- ☐ «расширенные» свёртки в DeepLab
- ☐ RPN в YOLO

♦ Из статьи про какую сеть эта иллюстрация?

- ☐ SegNet
- ☐ Unet
- ☐ DeepLab
- ☐ PSP-Net
- ☐ SharpMask
- ☐ HR-Representations
- ☐ нет верного ответа



♦ Что делалось в Unet (оригинальная работа)?

- ☐ вырезание подтензоров (crop)
- ☐ конкатенация после прокидывания связи
- ☐ «обратный» пулинг
- ☐ 3×3 -свёртки

♦ Что верно для Mask R-CNN?

- ☐ заменили RoIPool на RoIAlign
- ☐ добавили «обратный» пулинг
- ☐ добавили 3×3 -свёртки
- ☐ фактически сеть решает 3 задачи CV (ошибка равна сумме трёх)

♦ Из статьи про какую сеть эта иллюстрация?

- ☐ Mask R-CNN
- ☐ TensorMask
- ☐ YOLO
- ☐ PolarMask
- ☐ CornerNet
- ☐ FCOS
- ☐ нет верного ответа

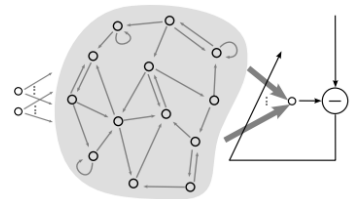


♦ Какой размер выходного тензора Mask-регрессии у PolarMask на масштабе 100×100 , 10 классах, 8 лучах?

- ☐ 10×4
- ☐ $100 \times 100 \times 8$
- ☐ 10×8
- ☐ $100 \times 100 \times 10$
- ☐ 10×16
- ☐ $100 \times 100 \times 16$
- ☐ $100 \times 100 \times 4$
- ☐ нет верного ответа

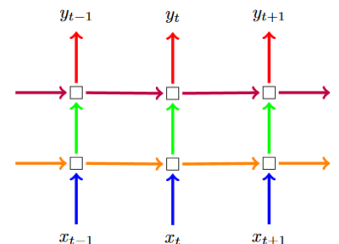
♦ Из статьи про какую концепцию эта иллюстрация?

- ☐ LSTM
- ☐ peephole connections
- ☐ резервуарные вычисления
- ☐ динамический dropout
- ☐ нет верного ответа



♦ Из статьи про какую концепцию эта иллюстрация?

- ☐ многослойные LSTM
- ☐ peephole connections
- ☐ variational dropout
- ☐ языковые модели
- ☐ нет верного ответа



♦ В какой ячейке меньше гейтов?

- ☐ LSTM
- ☐ LSTM with peephole connections
- ☐ GRU

♦ В приведённом коде, какой размер у hidden?

```
lstm = nn.LSTM(input_size=3, hidden_size=3,  
               num_layers=1, bias=True,  
               batch_first=False, bidirectional=False)
```

out, hidden = lstm(input, hidden)

- ☐ 1×1×3
- ☐ 1×1×6
- ☐ [1×1×3, 1×1×3]
- ☐ [1×1×6, 1×1×6]
- ☐ [1×1×3, 1×1×3, 1×1×3]
- ☐ нет верного ответа

♦ Какие активации используются в LSTM-ячейке?

- ☐ сигмоида
- ☐ гиперболический тангенс
- ☐ ReLU
- ☐ softmax
- ☐ ELU

♦ Какие активации используются в GRU-ячейке?

- ☐ сигмоида
- ☐ гиперболический тангенс
- ☐ ReLU
- ☐ softmax
- ☐ ELU

♦ В каком подходе используется дискриминатор?

- ☐ teacher forcing
- ☐ professor forcing
- ☐ free running
- ☐ scheduled sampling
- ☐ нет верного ответа

♦ Что из перечисленного может использоваться в RNN?

- ☐ BN
- ☐ DO
- ☐ grad clipping
- ☐ teacher forcing

♦ Что в аббревиатуре ТВРТТ (для RNN) обозначает первая буква Т?

- ☐ tensor
- ☐ test
- ☐ time
- ☐ tick
- ☐ truncated
- ☐ нет верного ответа

♦ Что из перечисленного верно для ELMo?

- ☐ предтренировки без учителя
- ☐ представление слова = л/к скрытых состояний
- ☐ двунаправленная языковая модель (biLM)
- ☐ не учитывает контекст слова

♦ В каком из этих представлений производится разложение матрицы логарифмов частот?

- ☐ w2v
- ☐ fasttext
- ☐ GloVe
- ☐ ELMo

♦ В каком из этих представлений представление слова равно сумме представлений n-грамм?

- ☐ w2v
- ☐ fasttext
- ☐ GloVe
- ☐ ELMo

♦ Что верно для (или используется в) word2vec?

- ☐ учитывается контекст слова
- ☐ Negative Sampling
- ☐ Hierarchical Softmax
- ☐ нужна размеченная выборка текстов

♦ Что означает буква С в аббревиатуре CBOW?

- ☐ count
- ☐ continuous
- ☐ circle
- ☐ нет верного ответа