



Во всех вопросах может быть несколько правильных ответов

В тестовых заданиях первая галочка — правильный ответ, вторая галочка — выбранный ответ. Цвет обозначает, правильно ли в данном пункте поставлена галочка. Если все пункты верные (галочки совпадают / все пункты зеленые), то за задание ставится полный балл, в противном случае ставится 0 баллов.

1. Выделите преимущества глубоких нейронных сетей по сравнению с неглубокими (3х слойными).

- ☒ ☒ глубокие нейросети автоматически определяют наиболее информативные признаки на промежуточном слое, которыми оперируют последние слои глубокой нейросети
- ☒ ☒ за счет перекрестного использования признаков в промежуточных слоях глубокая нейросеть получает возможность строить более точные и менее параметризованные модели
- ☐ ☐ даже с увеличением числа нейронов сеть с тремя слоями теоретически не способна моделировать настолько же широкий класс зависимостей, как глубокие нейросети.

Балл: 0.75

Комментарий к правильному ответу:

2. Выберите размерность выхода и выходную функцию активации $f(x)$ последнего слоя для задачи многоклассовой классификации:

- ☐ ☐ одномерный выход, $f(x) = x$
- ☒ ☒ многомерный выход, $f(x) = \text{Softmax}(x)$
- ☐ ☐ одномерный выход, $f(x) = \text{sigmoid}(x)$

Балл: 0.75

Комментарий к правильному ответу:

3. Выберите функции активации, приводящие к нулевым или пренебрежимо малым градиентам (с вычислительной точки зрения) по весам в методе обратного распространения ошибки на неограниченном хотя бы с одной из сторон подмножестве значений аргумента функции активации:

- ☒ ☒ ReLU
- ☒ ☒ сигмоида
- ☒ ☒ гиперболический тангенс



☐ ☐ Leaky ReLU

Балл: 0.75

Комментарий к правильному ответу:

4. Выберите архитектуру автокодировщика, в котором всегда число слоёв на внутреннем слое (между кодировщиком и декодировщиком) всегда должно быть ниже, чем число входов (признаков), чтобы он не выучивал тождественное преобразование

☐ ☐ sparse autoencoder

☐ ☐ contractive autoencoder

☒ ☒ undercomplete autoencoder

☐ ☐ denoising autoencoder

Балл: 0.75

Комментарий к правильному ответу:

5. Пусть $z = \text{ReLU}(Wx)$. Тогда Якобиан dz/dx :

☒ ☒ зависит от знака в компонентах вектора Wx

☐ ☐ диагональная матрица

☒ ☒ зависит от W

Балл: 0.75

Комментарий к правильному ответу:

6. Когда в Pytorch надо делать шаг оптимизации `optimizer.step()`?

☐ ☐ после прямого прохода

☐ ☐ после вычисления ошибки (loss)

☒ ☒ после обратного прохода `loss.backward()`

☐ ☐ в самом начале обучения

Балл: 0.75

Комментарий к правильному ответу:



7. Возможно ли, модифицируя их архитектуру, с помощью нейросетей предсказывать не отдельный числовой прогноз, а распределение отклика (целевого значения)?

☒ ☒ да

☐ ☐ нет

Балл: 0.75

Комментарий к правильному ответу:

Например, предсказывая μ , σ или частоту в каждой ячейке гистограммы.

8. Выберите преимущества метода SGD с инерцией (momentum) по сравнению с SGD без инерции. При использовании инерции используемые значения модификации весов:

☒ ☒ используют усредненные значения градиентов по градиентам на текущем минибатче и всем предыдущих минибатчам с экспоненциально затухающими весами, что обеспечивает более стабильную оценку градиента по всем объектам

☐ ☐ точнее отражают текущее значение градиента функции потерь на текущем минибатче

☐ ☐ используют усредненные значения градиентов по градиентам на текущем минибатче и K (фиксированный гиперпараметр) предыдущих минибатчам, что обеспечивает более стабильную оценку градиента по всем объектам

☐ ☐ обрезают градиент по норме, обеспечивая более устойчивую сходимость для нелинейной функции потерь

Балл: 0.75

Комментарий к правильному ответу: