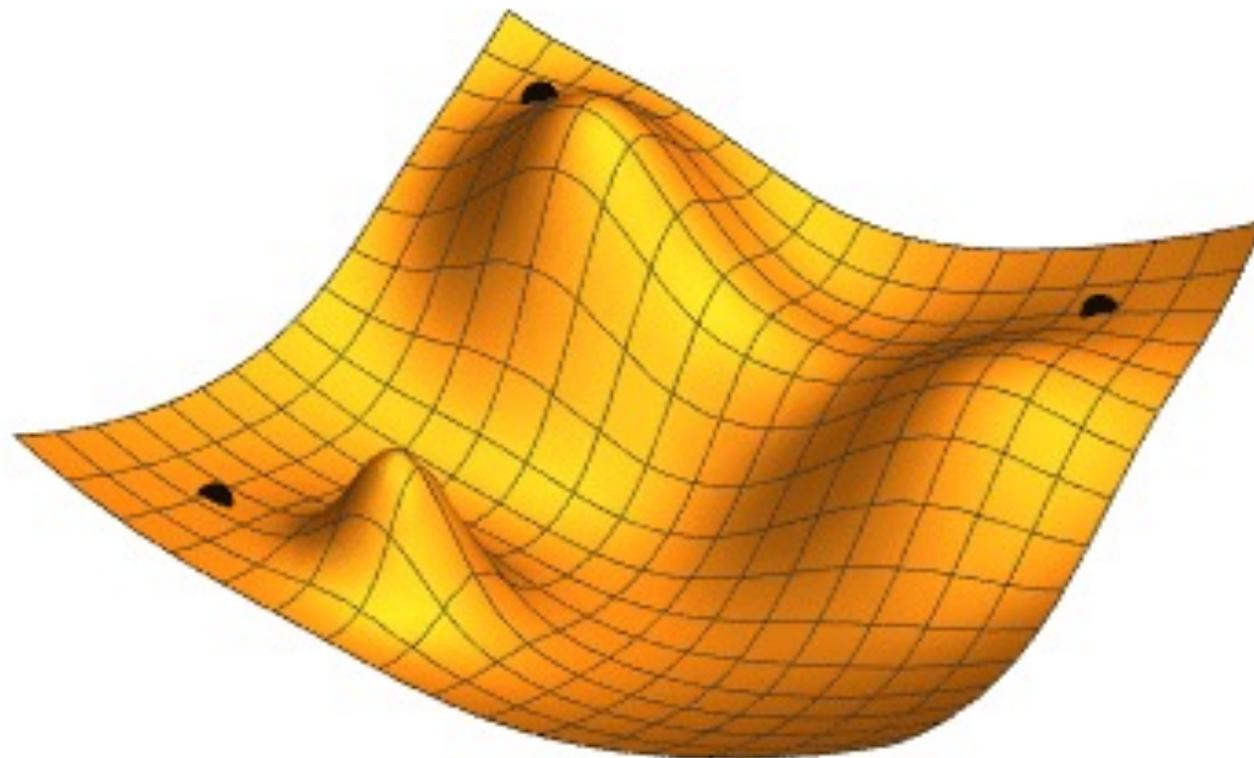
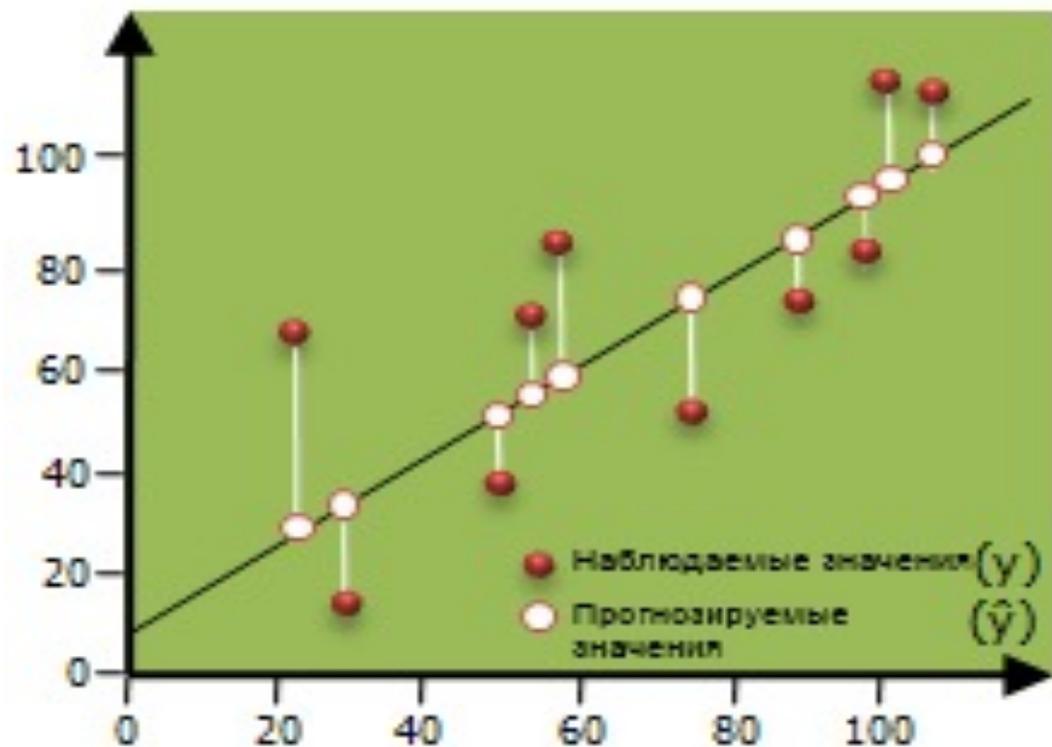
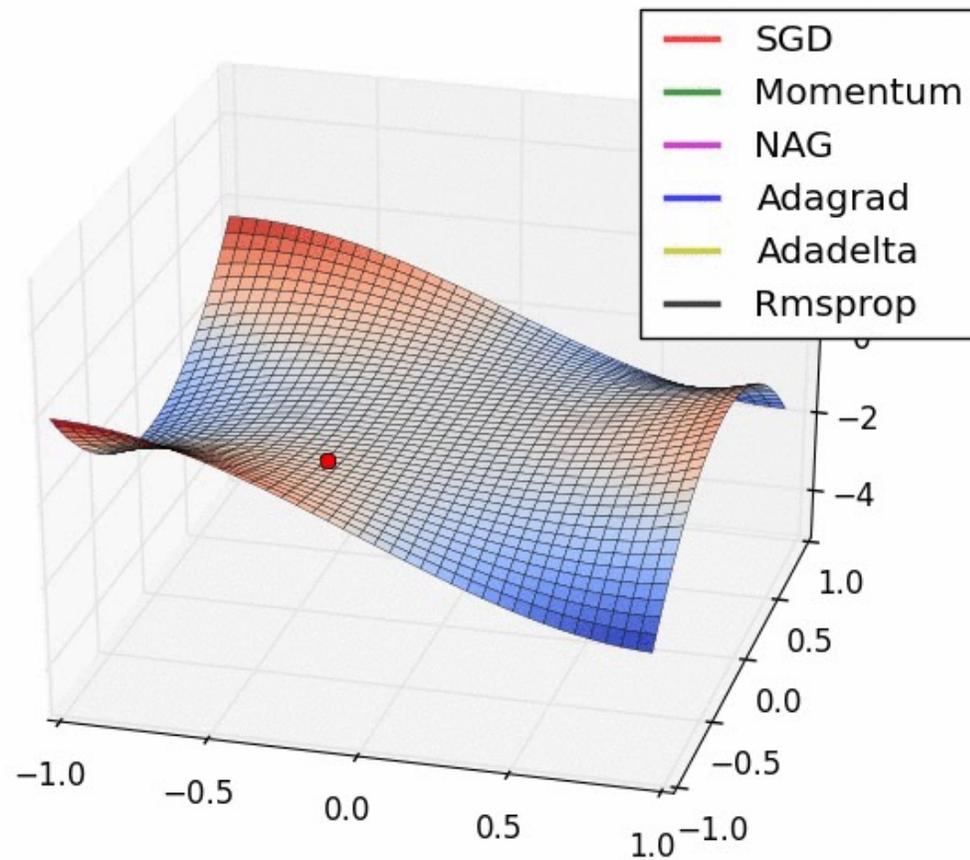
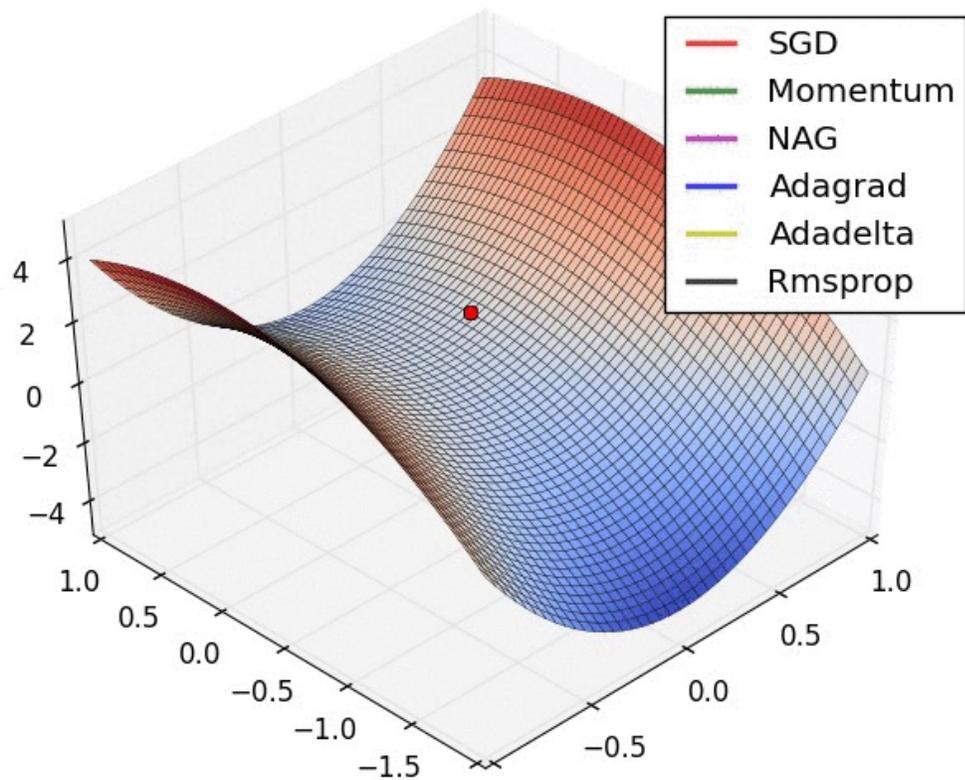


Функция потерь (loss)

- Нам нужен минимум функции
- Метод наименьших квадратов
- Кроссэнтропия



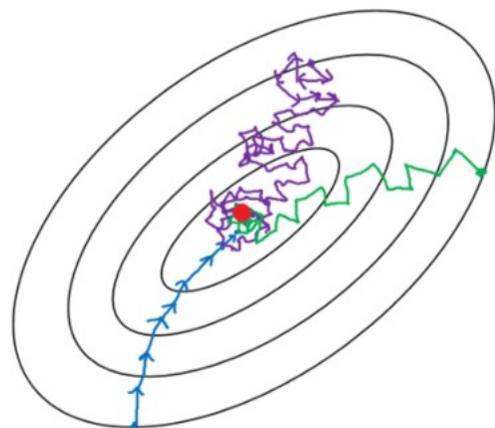
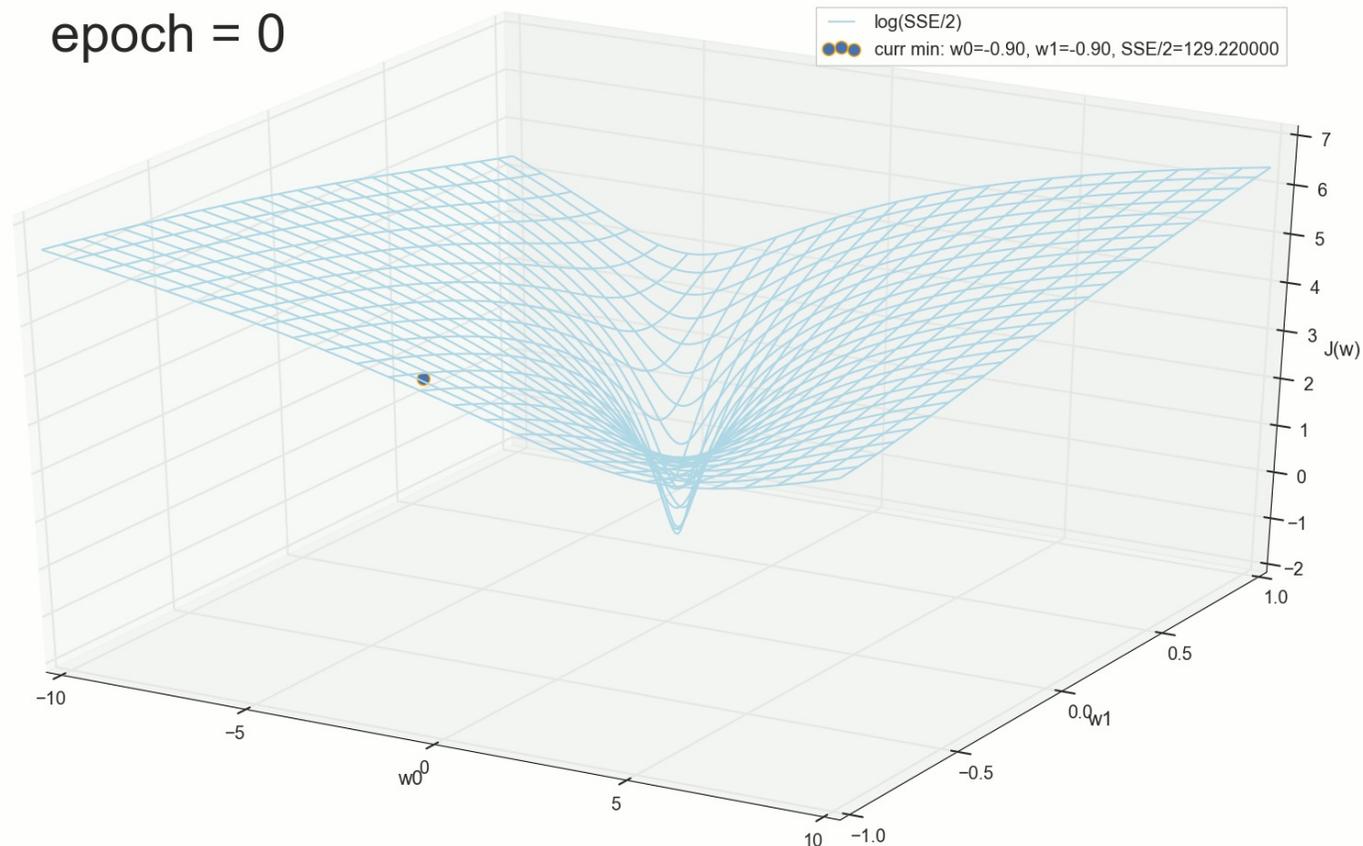
Другие оптимизаторы



Батчи

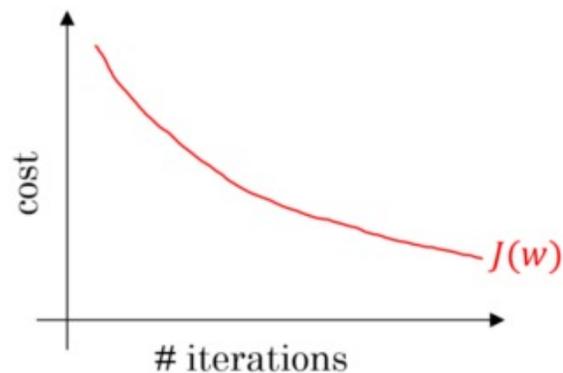
- Функцию потерь считаем по одному примеру, но потом их складываем в одну
- Обучать на всем датасете — долго. Каждый раз берем небольшую порцию данных
- Но обучение становится хаотичнее

epoch = 0

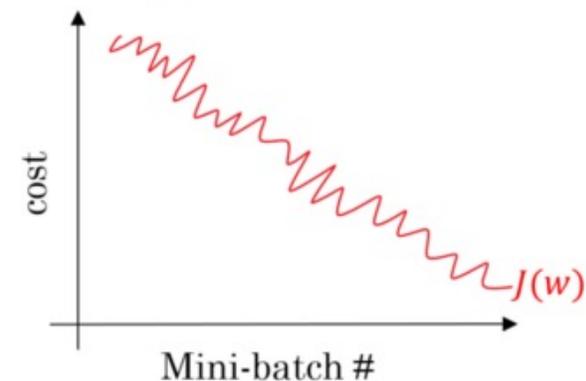


- Batch gradient descent
- Mini-batch gradient Descent
- Stochastic gradient descent

Batch gradient descent

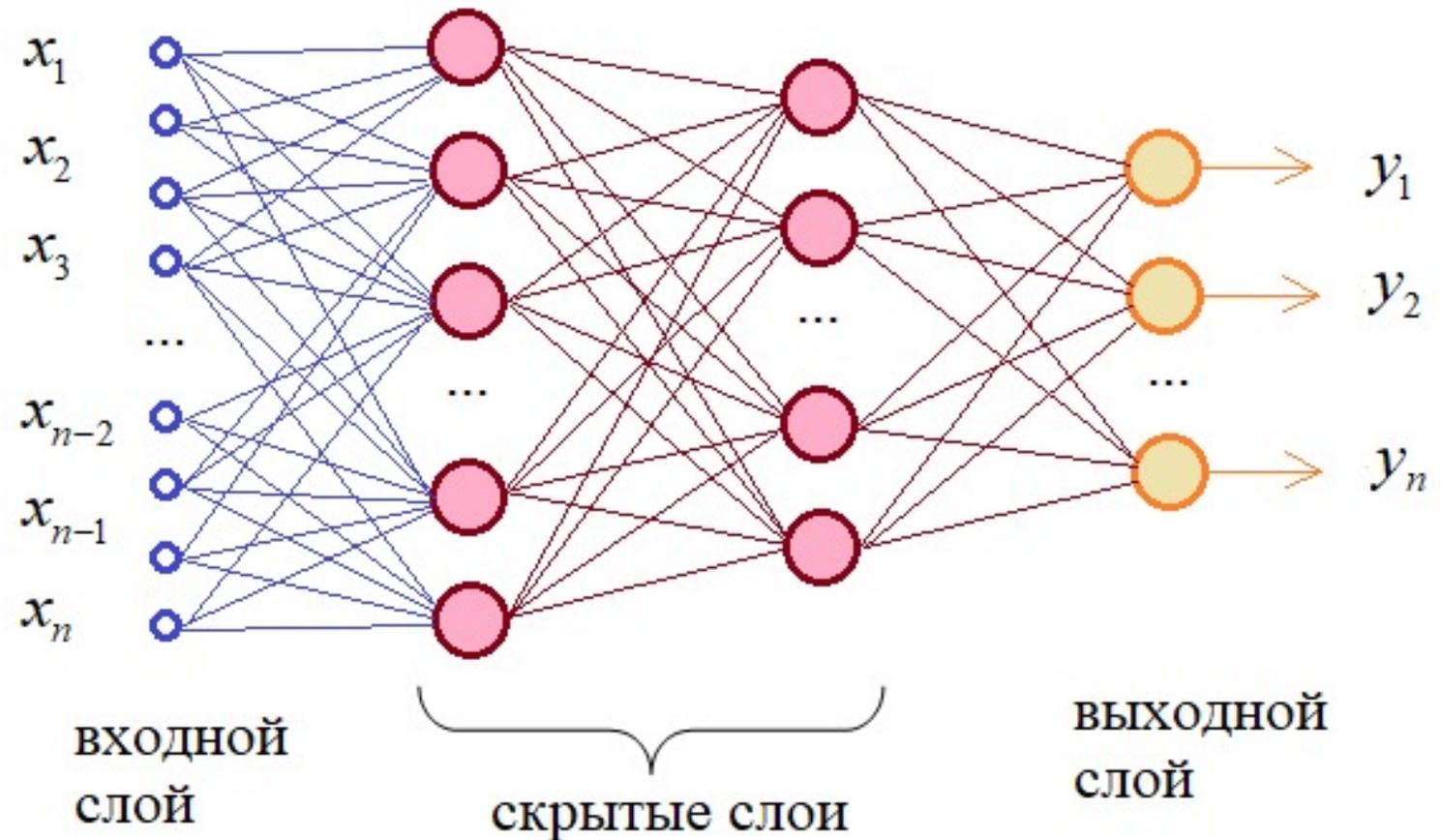


Mini-batch gradient descent



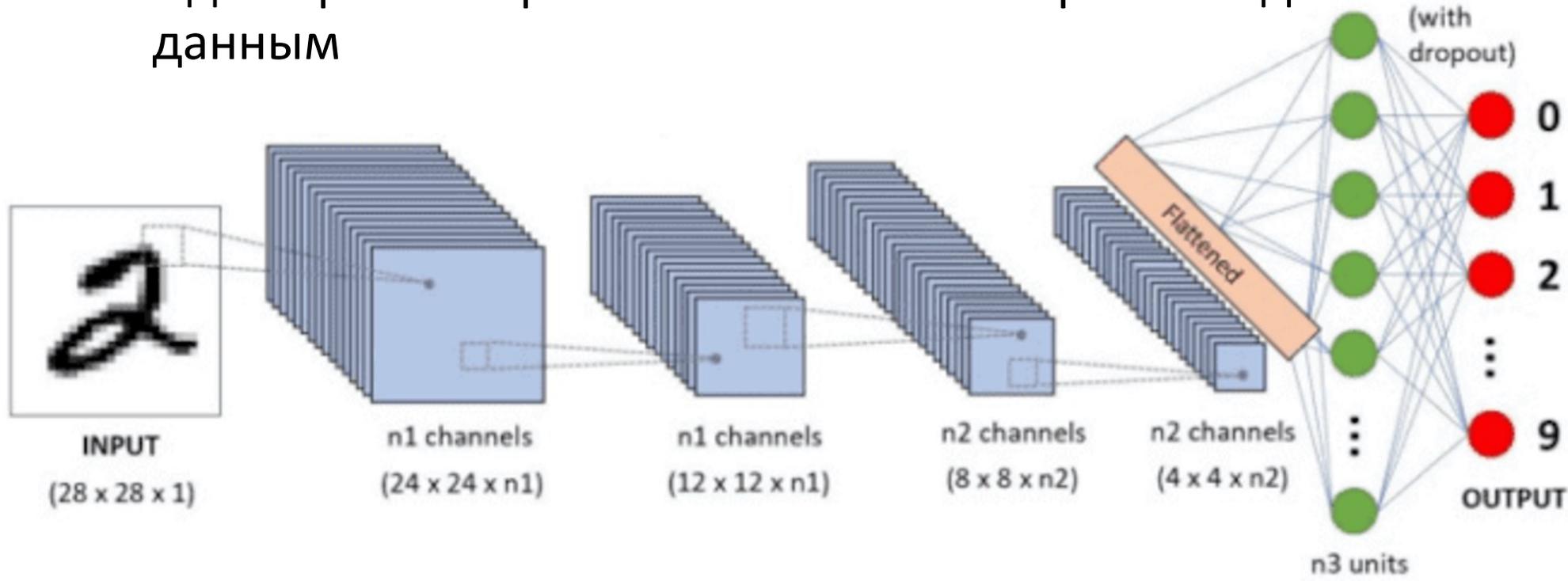
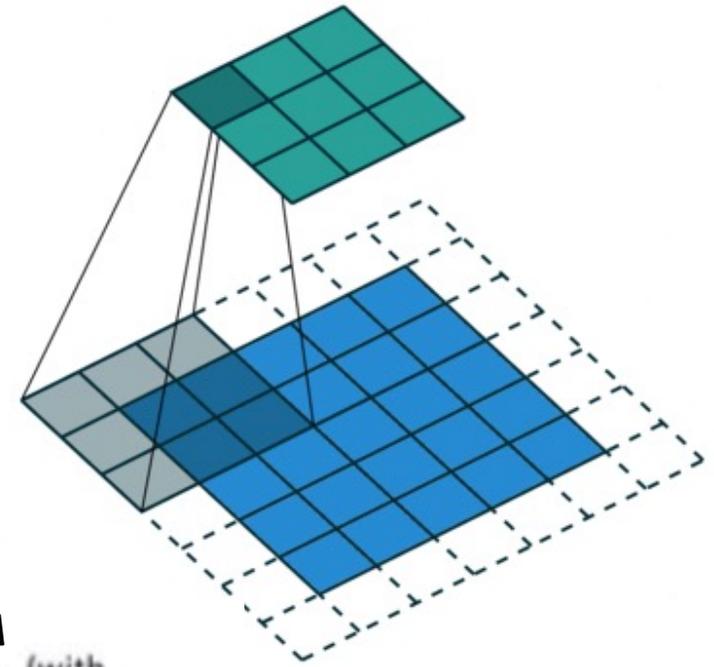
Полносвязная нейросеть

- Хорошая классификация при малом объеме входных данных
- Очень много связей при большом объеме данных



Сверточная нейросеть

- Сверточная нейросеть позволяет выявить признаки
- Обучаются сразу несколько признаков (нейронов) с небольшим числом параметров
- Один признак применяется несколько раз к входным данным



ONNX

Step 3. Select class labels and get predictions

Выбрать файл `cifar100_CNN.onnx` Select ONNX file

Выбрать файл `c25c94fe96_1000.jpg`

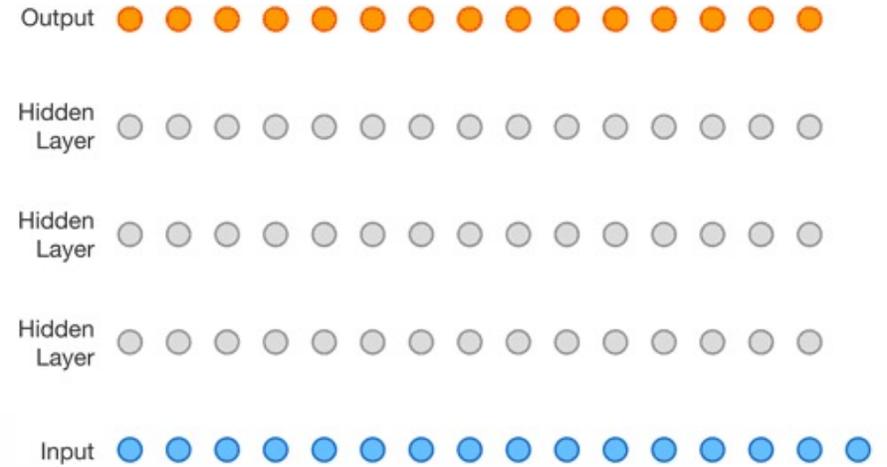
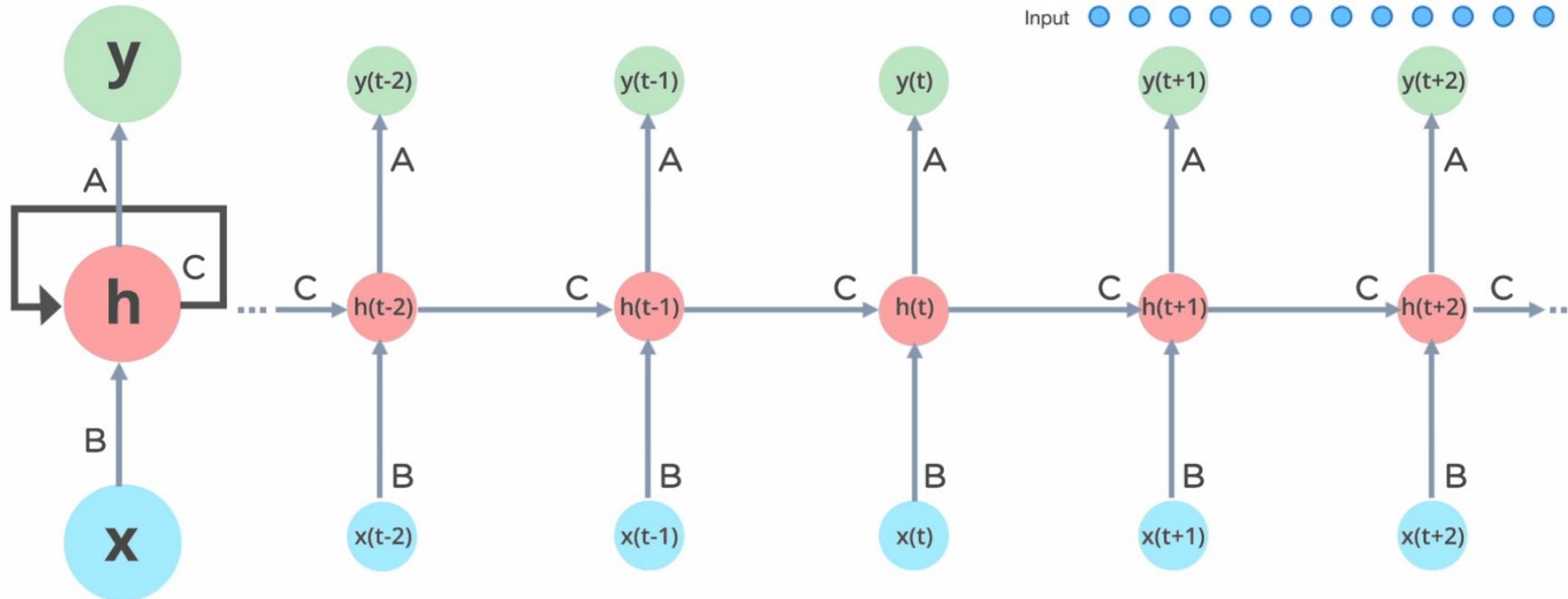
Class label

0,50,54



Рекуррентная нейросеть

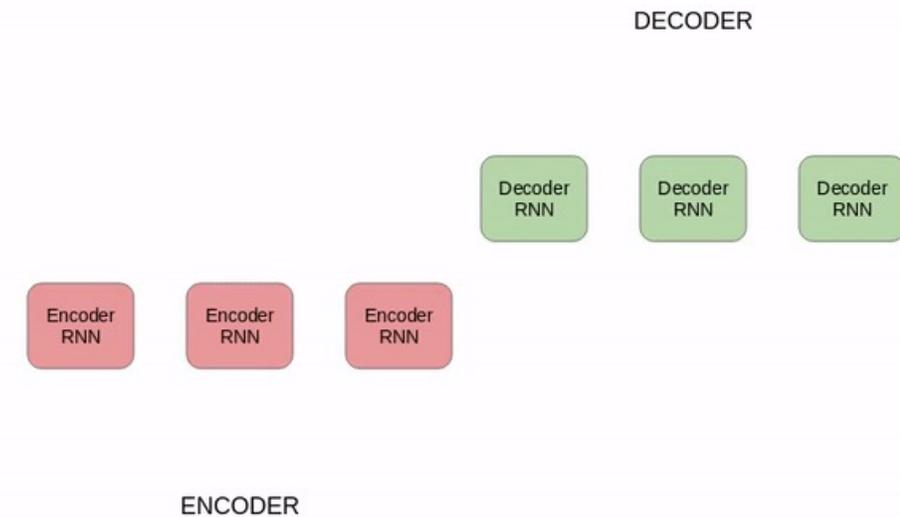
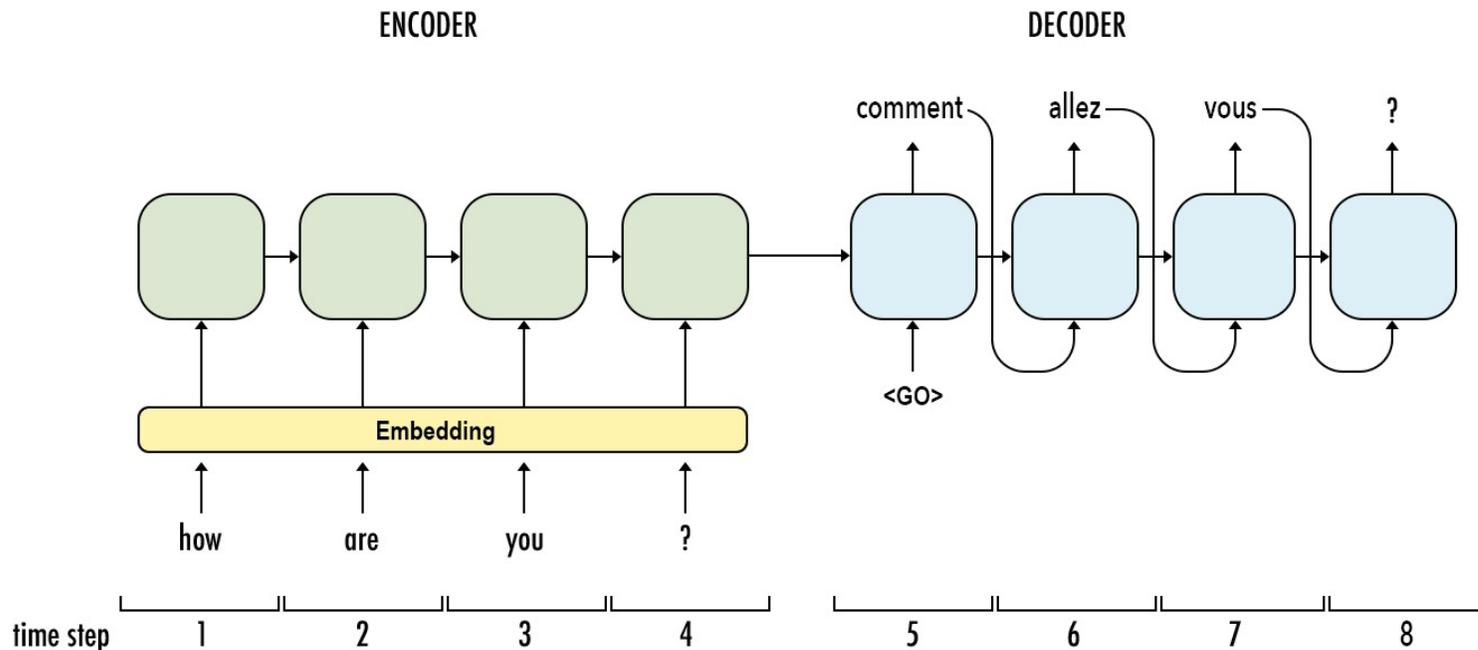
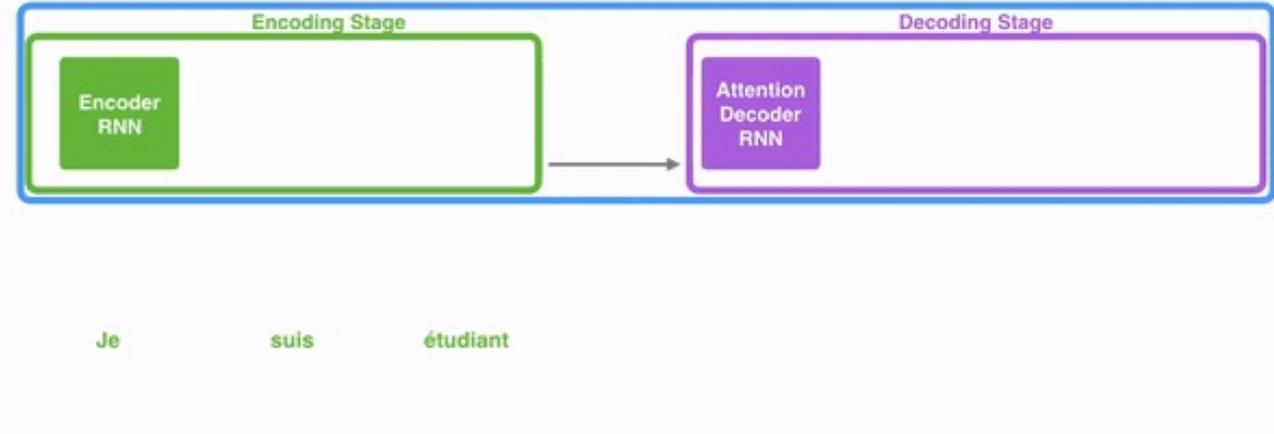
- Рекуррентная нейросеть позволяет реализовать память
- GRU, LSTM



Seq2seq

- Seq2seq для предсказания временных рядов и машинного обучения

Neural Machine Translation SEQUENCE TO SEQUENCE MODEL WITH ATTENTION

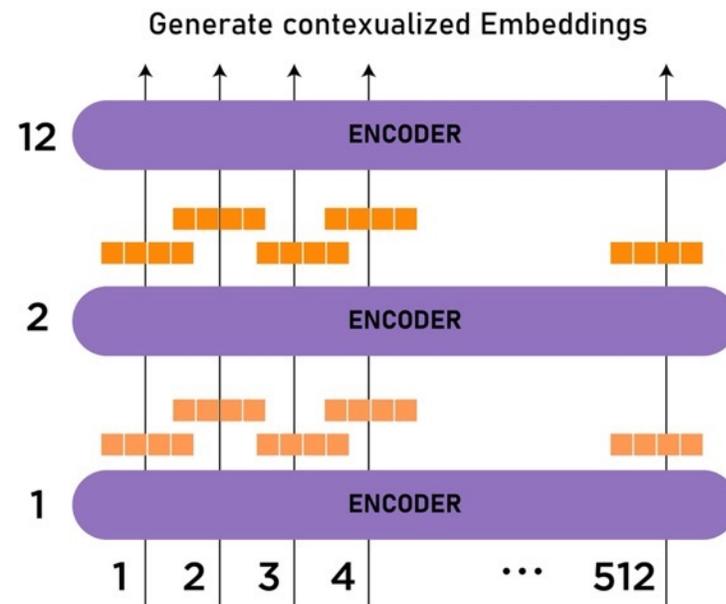


Трансформер

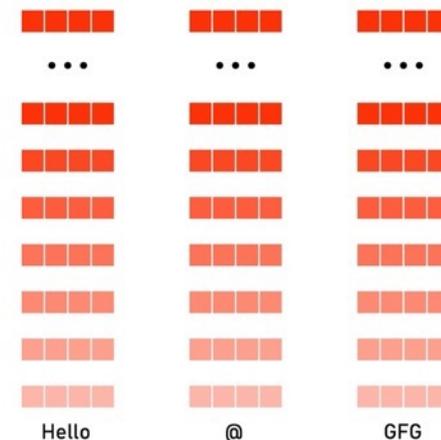
- Трансформер с механизмом ВНИМАНИЯ
- BERT, GPT

Decoding time step: 1 2 3 4 5 6

OUTPUT



The output of each encoder layer can be used to represent the feature for that token



BERT

