

Lec 15

13 декабря 2024 г. 9:56

1. AI - the ability of some digital computer to perform tasks that are usually performed by intelligent beings
2. ML - AI subsection that is studying how systems learn based on statistics and optimization theory

Машинное обучение (3)

- Том Митчел (1998):
 - Говорят, что компьютерная программа обучается на опыте E решению некоторой задачи T в соответствии с измерением качества решения этой задачи P , если качество решения T измеренное при помощи P , улучшается при накоплении опыта E

- 3.
4. Model learns on data, not on given rules. It should form complex rules by itself. But if rules are really simple than ML is not needed
5. Supervised - learning on labeled data
6. Unsupervised learning on just data
7. Reinforcement learning - data, action and score on how system performs this action
8. Semi-supervised - when data is partially labeled
9. It's quite trending nowadays to build multimodal models that can do multiple tasks
10. Deep Learning - neural networks

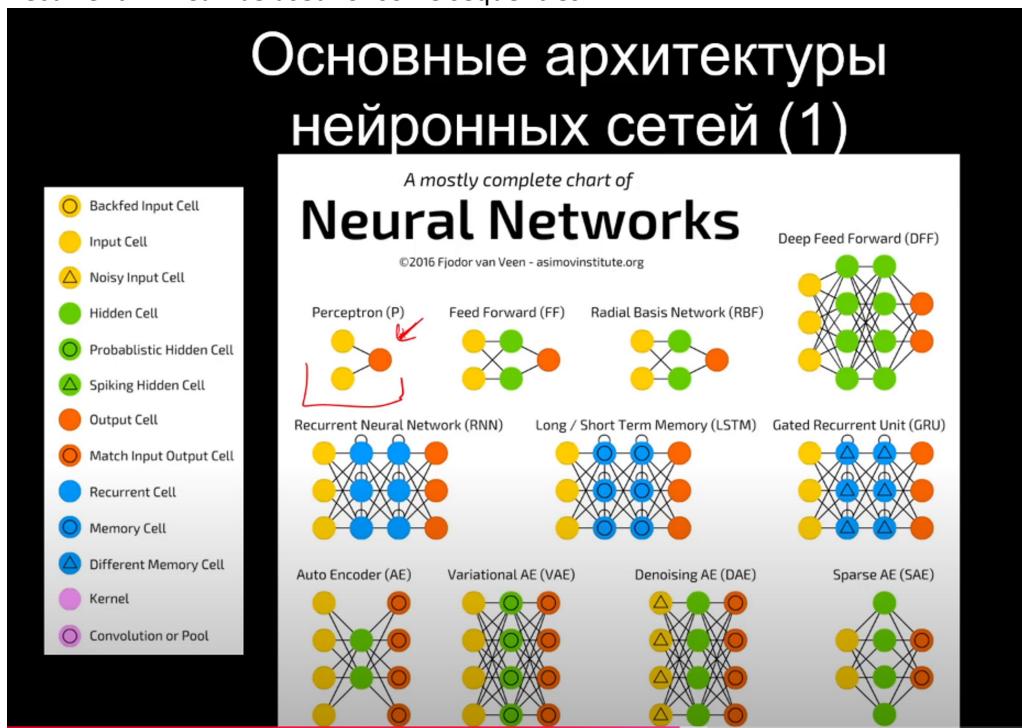


- 11.
12. SNN - more brain related NN.
13. ---
14. **Neural Networks**
15. Neurons are receivers and senders of electrical impulses
16. The learning part is that the web - the set of connections is changeable. New connections can

be formed with new eights.

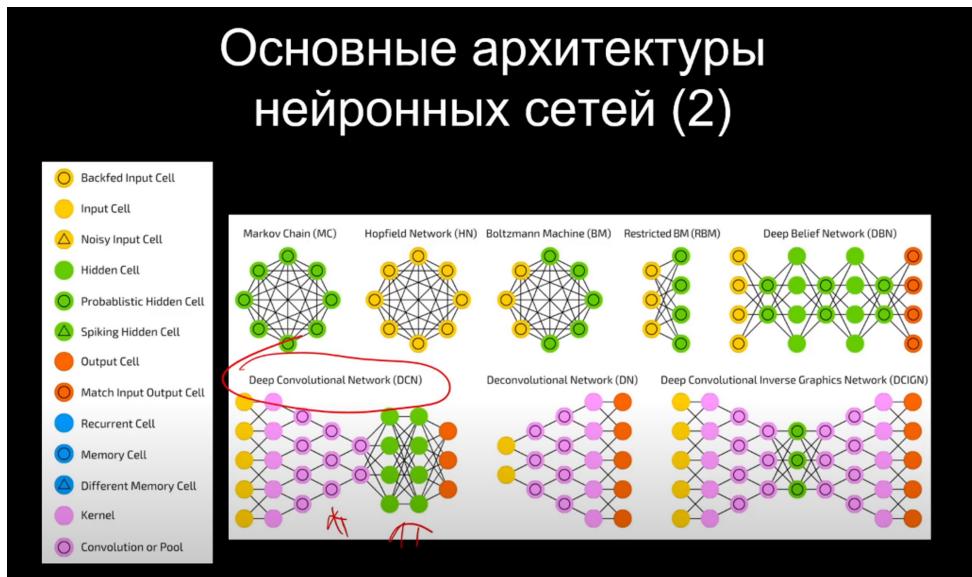
17. But these are Impulsive neural networks
18. Neuron have inputs with weights which combined linearly and put in some function $h(ax+by+cz+\dots)$
19. The main goal is to tune weights in a way to get certain output
20. H - activation function
21. Deep NN - NN with multiple layers
22. Feed-forward NN - signal propagates only from input to output. They are stateless. They don't remember what it been given to recognize
23. Recurrent NN. Can be used for some sequences

24.



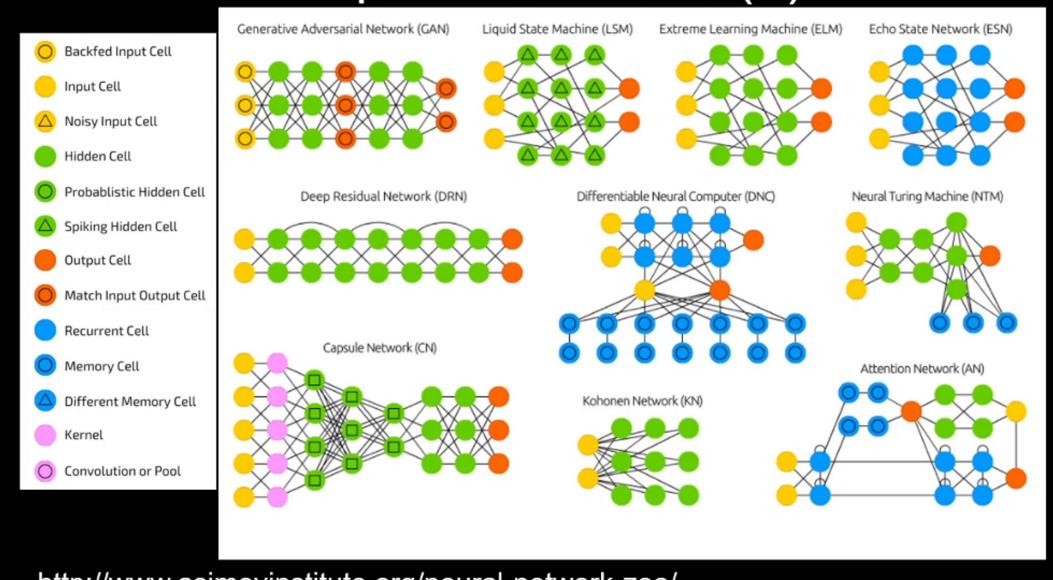
25. Auto-encoders are quite useful for detection of anomalies, cause we will see much difference on different data

26.



Основные архитектуры нейронных сетей (3)

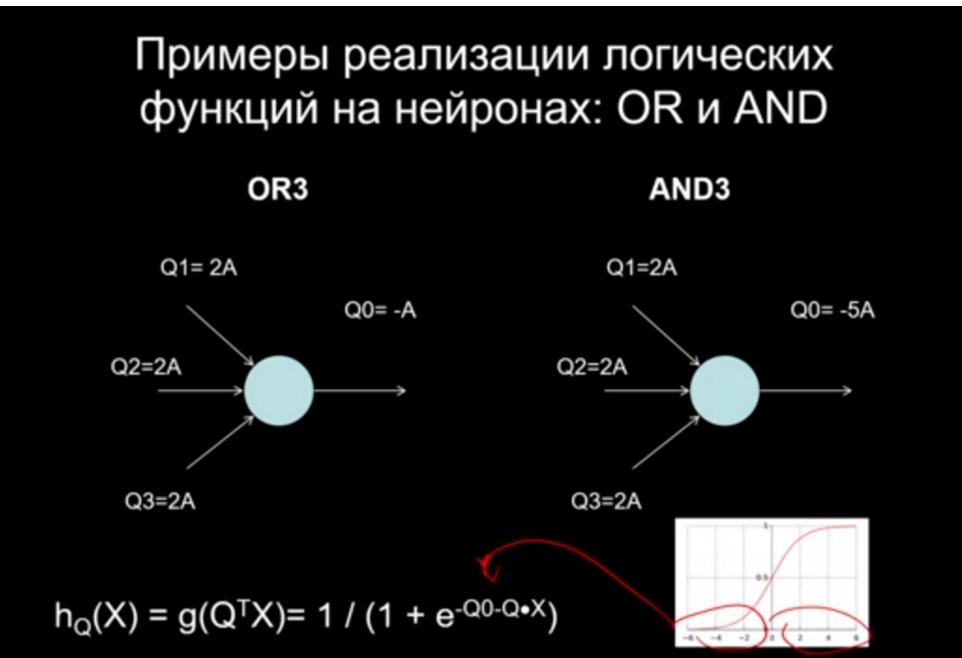
27.



28. GAN - generative adversarial network - basically parallel learning of two neural networks. One is becoming better at generation, other one in recognition. Or cycleGan
29. Feed-forward - activation function from linear combination of params and their weights
30. Cool

Примеры реализации логических функций на нейронах: OR и AND

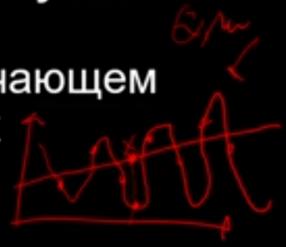
31.



32. But how do they learn? How do they choose coefficients
33. Supervised - classics, unsupervised - auto-encoders, GAN, reinforcement
34. Loss function - as in regression - the goal is to minimize it.
35. Most popular are MSE - mean square error, MAE - mean absolute error and binary cross entropy (with logarithms).
- 36.
37. **Back propagation error**
38. To watch
39. **Start initialization**
40. Weights are usually initialized randomly. But not symmetrical, cause otherwise you will have just copies of one neuron. Network won't learn.

Проблема переобучения

- Хорошо работает на обучающем наборе и плохо на тестах
- 41.
- Решения проблемы
 - Штрафы за большие веса (регуляризация)
 - Разделяемые веса
 - Ранняя остановка обучения
 - Взятие среднего по нескольким моделям
 - ...



42. Bad learning curve usually signalises about bad learning rate or whole architecture

43. Drop out?

Tuning model

Number of neurons, number of layers etc.

Выбор параметров модели

- Количество нейронов, количество слоев, конфигурация сети...
- «Пока не попробуешь – не узнаешь»
- Решение:
 - Обучающий набор (~60-80%)
 - Валидационный набор (~10-20%)
 - На нем подбираем параметры сети
 - Тестовый набор (~10-20%)

Main problems of deep neural networks:

1. Vanishing gradient problem. Neurons that are close to inputs are staying almost unchangeable, cause gradient is almost zero. Solution - expresses.
2. Exploding gradient problem. Overshooting. Solution - more accurate choice of activation function.