

Задание 6 по теме “Переобучение нейронных сетей и методы регуляризации.”

6.1 Нейронная сеть с двумя промежуточными слоями. Реализовать алгоритм обучения нейронной сети с двумя промежуточными слоями для оценки длительности гамма-всплеска T_{90} по данным каталога FERMIGBRST в условиях задачи 3.2.

- а) построить кривую обучения в зависимости от числа нейронов в каждом из скрытых слоев
- б) сравнить точность и скорость обучения для сырых, нормированных и декоррелированных данных;
- в) какой будет точность определения T_{90} вблизи длительности 2 секунды?
- г) для какого размера сети можно увидеть эффект переобучения?
- д) применить регуляризацию и проверить присутствие эффекта переобучения.

6.2 Классификация гамма-всплесков.

- а) Используя методы задачи 6.1, построить классификатор гамма-всплесков, не использующий длительность в числе входных параметров. Для тренировки классификатора считать короткими всплески короче 1 секунды, а длинными – дольше 10 секунд.
- б) Насколько уверенно выполняется классификация всплесков длительностью от 1 до 10 секунд? Есть ли основания предполагать существование третьего класса гамма-всплесков с другой физической природой?

6.3 Поиск корреляций между параметрами Planck. Используя алгоритм из задачи один выполнить оценку одного из космологических параметров Planck на основе остальных, используя данные марковских цепочек Planck, описанные в условии задачи 4.1.

- а) сравнить точность и скорость обучения для сырых, нормированных и декоррелированных данных;
- б) построить кривую обучения в зависимости от размера батча, используемого для обучения нейронной сети;
Замечание: кривую в этом случае строить как функцию процессорного времени, а не числа эпох.

6.4 Модификация капельной модели ядра. Фреймворки для машинного обучения позволяют находить не только оптимальные веса многослойных нейросетей, но и вообще произвольных сложных функций. Вам предлагается поэкспериментировать с этим и внести собственные модификации капельной модели ядра, усложнив модель на свой вкус. Файл с данными энергии связи можно взять по ссылке: http://ppc.inr.ac.ru/grisha_data/A_Z_deltaE.dat.
Возможные идеи:

- а) Добавить ещё одно слагаемое $a_6 F(A, Z)$, где функцию F вы придумали самостоятельно, а множитель a_6 ищется по эмпирическим данным.
- б) Сделать показатель степени одного из слагаемых произвольным параметром модели. Фреймворк машинного обучения должен самостоятельно вычислить градиент и найти оптимальный показатель степени в процессе обучения модели. По возможности придумайте физическое объяснение, почему показатель степени отличается от аналогичного в формуле Вайцеккера.
- в) Построить многослойную нейросеть, принимающую на вход слагаемые капельной модели ядра.