

Исследование необходимого размера скрытого слоя на примере датасета Iris

Илья Гридасов

20 февраля 2019 г.

1 Цель

Целью исследования являлось нахождения оптимального числа количества нейронов скрытого слоя для решения задачи классификации датасета Iris.

2 Выбранные модели и метрики

Согласно описанию данных, всего в датасете находятся описания 3 сортов ириса. Одним из методов решения задачи классификации является нейронные сети.

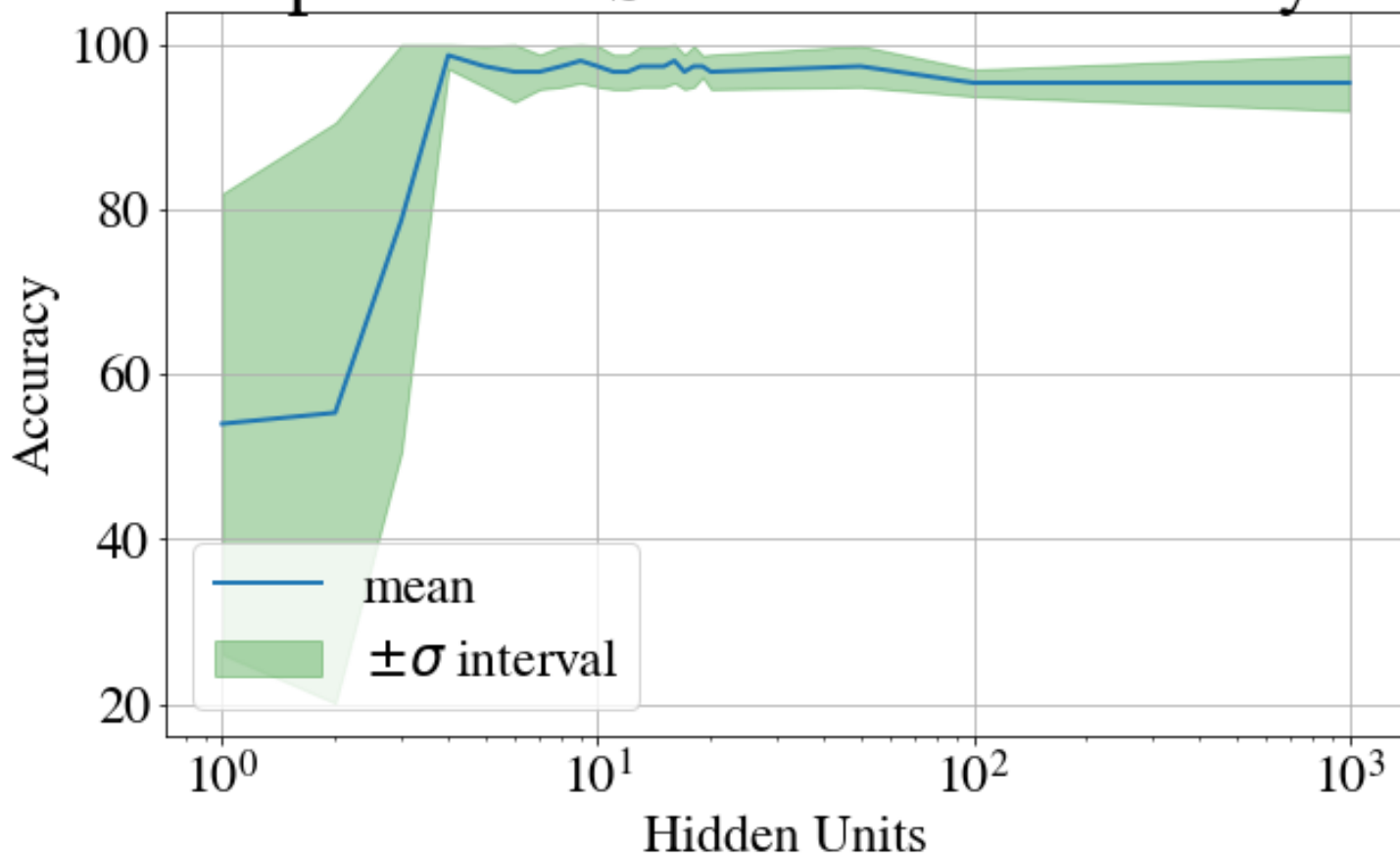
Зафиксируем рассматриваемую архитектуру нейронной сети. Входной слой будет состоять из 4-х нейронов - признаков каждого объекта из датасета. Далее, скрытый полносвязный слой с функцией активации *ReLU*, количество нейронов в котором и будет изучаться в данной задаче. Затем, выходной слой с активацией *softmax*, количество нейронов в котором соответствует количеству изучаемых классов ириса, то есть 3.

Функция ошибки будет *categorical crossentropy* - наиболее частый выбор в задачах классификации. Конечная метрика будет *accuracy* - доля верно предсказанных классов.

Оценивание модели будет происходить посредством кросс-валидации по 5 разбиениям, с подсчётом среднего *accuracy* и его дисперсии.

3 Результаты

Dependence Score from hidden layer



4 Выводы

При слишком малом кол-ве нейронов в скрытом слое, меньше, чем кол-во нейронов во входном слое, качество модели сильно деградирует, так как *mean accuracy* падает, так же она становится крайне нестабильной, так как *accuracy std* крайне велико.

Дальнейшее, же увеличение кол-ва нейронов в скрытом слое не оказывает влияние на *score* модели и на её стабильность. Интересной осо-

бенностью является, что даже при очень большом размере скрытого слоя, относительно входного и выходного, переобучения не возникает.