

Опишите общий алгоритм того, как Вы подбирали архитектуру сети и параметры обучения? Чем руководствовались в том или ином случае?

Сначала посмотрим, как поведет себя сеть при количестве слоев и нейронов в них по умолчанию (заданных в методе). Если нет переобучения, то оставляем как есть. Если же есть, то пробуем изменить число слоев, число нейронов в них. Так делаем пока переобучение не перестанет наступать слишком рано. Как только добьемся этого, то ограничим число эпохи, до эпохи, в которой точность на обучающих данных продолжает расти, а на тренировочных уже падает. В результате получим оптимальную архитектуру.