

1. Интерпретировал как “Плюсы и минусы подхода, когда сеть предсказывает слова, а не символы”:

Плюсы:

- 1) Большая осмысленность текста.
- 2) Установление связей между словами, в результате чего легче генерировать осмысленные фразы.

Минусы:

- 1) Ограниченный словарь.
- 2) Может возникнуть переобучения из – за размера обучающей выборки.

2. Блендинг – простейшая схема стеккинга (ансамблирования алгоритмов путём получения среднего или голосования по большинству): обучающую выборку делят на две части. На первой обучают базовые алгоритмы. Затем получают их ответы на второй части и на тестовой выборке. Понятно, что ответ каждого алгоритма можно рассматривать как новый признак (т.н. «метапризнак»). На метапризнаках второй части обучения настраивают метаалгоритм. Затем запускают его на метапризнаках теста и получают ответ.

3. Бустинг: мы создаем семейство моделей, которые объединяются, чтобы получить сильного ученика, который лучше работает. Бустинг — это метод, который заключается в том, чтобы адаптировать последовательно нескольких слабых учеников следующим способом: каждая модель в последовательности подбирается, что придает большее значение объектам в датасете, которые плохо обрабатывались предыдущими моделями в последовательности. Интуитивно, каждая новая модель фокусирует свои усилия на самых сложных объектах выборки при обучении предыдущих моделей, чтобы мы получили в конце процесса сильного ученика с более низким смещением.