

1. В чем плюс и минусы того, если модель предсказывала слова, а не символы?

Минус: Для предсказания слов требуется больше памяти

Плюс: Для предсказания текста работают лучше.

2. Что такое Стэкинг?

Staking – один из самых популярных способов ансамблирования моделей, т.е. использования нескольких моделей для решения одной задачи машинного обучения.

Выборку разбивают на части (фолды), затем последовательно перебирая фолды модель обучается на всех фолдах, кроме одного, а на оставшемся получают ответы моделей и трактуют их как значения соответствующих признаков (метапризнаков) на этом фолде. Для получения метапризнаков объектов тестовой выборки базовые алгоритмы обучают на всей обучающей выборке и берут их ответы на тестовой.