

1. РЕГУЛЯРИЗАЦИЯ В ML - НАКЛАДЫВАНИЕ ШТРАФА НА СЛИШКОМ “ВЫБИВАЮЩИЕСЯ” ИЗ ОБЩЕГО РЯДА ВЕСА СЛОЯ. КОНКРЕТНО В СЛУЧАЕ МОДЕЛИ В ДАННОЙ ЛР - L2 РЕГУЛЯРИЗАЦИЯ, ЧЕМ ВЫШЕ ПОСТАВИТЬ КОЭФФИЦИЕНТ - ТЕМ БОЛЬШЕ БУДЕТ ШТРАФ.
2. ПОСТРОИВ ГИСТОГРАММУ АКТИВАЦИЙ ДЛЯ ОПРЕДЕЛЕННОГО СЛОЯ МОЖНО УВИДЕТЬ РАСПРЕДЕЛЕНИЕ ВЫХОДОВ ЭТОГО СЛОЯ НА КОНКРЕТНОМ БАТЧЕ, ТО ЕСТЬ ЗНАЧЕНИЯ ВЫХОДОВ СЛОЯ И КАК ОНИ МЕНЯЮТСЯ СО ВРЕМЕНЕМ.
3. ДУМАЮ, НЕТ. ЭТО СКОРЕЕ ЗАВИСИТ ОТ ГИПЕРПАРАМЕТРОВ “МАКС. КОЛ-ВО СЛОВ В ОБЗОРЕ” И “МАКС. РАЗМЕР СЛОВАРЯ СЛОВ”. КОГДА Я ПРОЕКТИРОВАЛ СЕТЬ, ТО НЕ СМОГ НАЙТИ ЧЕТКУЮ ФУНКЦИЮ ЗАВИСИМОСТИ ТОЧНОСТИ ОТ ЭТИХ ПАРАМЕТРОВ, ПОЭТОМУ ПОДБИРАЛ ИХ ИНТУИТИВНО.
4. КАК Я ПОНЯЛ, АНСАМБЛЬ НС В САМОМ ПРОСТОМ ВАРИАНТЕ ЭТО ПРОСТО НЕСКОЛЬКО МОДЕЛЕЙ, И РЕЗУЛЬТАТ РАБОТЫ - ВЗВЕШЕННАЯ СУММА ВЫХОДОВ СЕТЕЙ. Я ПРЕДПОЛОЖИЛ, ЧТО ЭТО НЕ СИЛЬНО ПОВЛИЯЕТ НА КОНЕЧНЫЙ РЕЗУЛЬТАТ, А ВРЕМЕНИ НА ОБУЧЕНИЕ НЕСКОЛЬКИХ МОДЕЛЕЙ ПОТРАТИТСЯ МНОГО И РЕШИЛ СОВЕРШЕНСТВОВАТЬ ОДНУ МОДЕЛЬ.
5. ЧЕСТНО ГОВОРЯ, НЕ ОЧЕНЬ ПОНЯЛ ВОПРОС. МОЯ МОДЕЛЬ САМА ПО СЕБЕ НЕ РЕКУРРЕНТНАЯ, А ПОСЛЕДОВАТЕЛЬНАЯ (SEQUENTIAL), РЕКУРРЕНТНЫМ ЯВЛЯЕТСЯ ОДИН СЛОЙ GRU.