

Отчет по практическому заданию №2 "**Применение линейных моделей для
определения токсичности комментария**".

Логистическая регрессия и градиентный спуск.

Содержание

1	Введение	2
2	Эксперименты	2
2.1	Исследование поведения градиентного спуска	2
2.1.1	Параметр размера шага step_alpha	2
2.1.2	Параметр размера шага step_beta	2

1 Введение

В данном документе представлен отчет о проделанных экспериментах по практическому заданию №2, анализ результатов. Краткое описание задания: необходимо реализовать линейный классификатор с произвольной функцией потерь.

2 Эксперименты

В этом блоке приведены все обязательные эксперименты, которые изложены в формулировке задания. Все эксперименты проводились на упрощенном датасете (рассматривается задача бинарной классификации) из соревнования **Toxic Comment Classification Challenge**, в котором нужно определить токсичность комментария.

Стандартный дизайн эксперимента:

- Оценка качества и подбор параметров модели проводились на каждой эпохе с помощью отложенной тренировочной выборки (30%). Все графики ниже построены по значениям ассигасу, посчитанным на отложенной выборке.
- В тренировочную выборку был добавлен признак, состоящий из всех единиц, который позволяет учитывать смещение (**bias**). Было решено не использовать смещение в $L2$ -регуляризации, чтобы даже при плохом выборе коэффициента регуляризации решающая гиперплоскость не вырождалась в 0.
- В стохастическом градиентном спуске проверяется критерий останова на каждой эпохе (не итерации).

2.1 Исследование поведения градиентного спуска

Обновления весов модели при использовании градиентного спуска происходит по следующей формуле:

$$w_t = w_{t-1} - \frac{\alpha}{t^\beta} \times \frac{1}{N} \times \sum_{i=1}^N \nabla_w L(x_i, y_i | w_{t-1}), \quad (1)$$

где t - номер итерации, β - **step_beta**, $\nabla_w L(x_i, y_i | w_{t-1})$ - градиент функции потерь.

2.1.1 Параметр размера шага **step_alpha**

Параметр **step_alpha** (α) используется в градиентном спуске при обновлении весов в формуле 1. Рассмотрим следующие зависимости при разных значениях параметра **step_alpha**:

1. зависимость значения функции потерь от реального времени работы метода
2. зависимость значения функции потерь от итерации метода
3. зависимость точности (ассигасу) от реального времени работы метода
4. зависимость точности (ассигасу) от итерации метода

Соответствующие графики приведены на

2.1.2 Параметр размера шага **step_beta**

Параметр **step_beta** (β) используется в градиентном спуске при обновлении весов в формуле 1. Аналогично предыдущему пункту рассмотрим зависимости из 2.1.1 при разных значениях параметра **step_beta** и проанализируем соответствующие графики: