

МАШИННОЕ ОБУЧЕНИЕ И ГРАДИЕНТНЫЙ СПУСК

Рассматривается градиентный спуск как один из основных методов обучения моделей машинного обучения. Описаны математические основы метода, его разновидности и особенности применения в задачах оптимизации.

Введение

Машинное обучение – подраздел искусственно-го интеллекта, изучающий методы построения алгоритмов, способных обучаться на данных [1]. В отличие от классических алгоритмов, модели машинного обучения самостоятельно выявляют закономерности в информации. Для уменьшения ошибки предсказания используются методы оптимизации, среди которых одним из наиболее распространённых является градиентный спуск [2].

Цель работы – изучить принцип работы градиентного спуска, рассмотреть его основные разновидности и определить значение метода в обучении моделей машинного обучения.

1. Градиентный спуск

Градиентный спуск представляет собой итерационный метод оптимизации, при котором параметры модели изменяются в направлении, противоположном градиенту функции потерь:

$$w_{k+1} = w_k - \eta \nabla f(w_k) \quad (1)$$

Здесь η обозначает шаг обучения, а $\nabla f(w_k)$ показывает направление наибольшего увеличения функции. Переход в противоположную сторону позволяет уменьшать значение ошибки модели.

Метод широко применяется при обучении нейронных сетей, линейной и логистической регрессии, а также других моделей машинного обучения.

II. Виды градиентного спуска

Существует три основных варианта метода (рис. 1), которые различаются способом вычисления градиента, скоростью и стабильностью обучения [3].

– Батчевый градиентный спуск использует всю обучающую выборку на каждом шаге. Он обеспечи-

вает стабильную сходимость, однако требует значительных вычислительных ресурсов.

– Стохастический градиентный спуск (SGD) обновляет параметры по одному примеру. Такой подход ускоряет вычисления, но делает процесс обучения менее стабильным.

– Мини-батч градиентный спуск использует небольшие подвыборки данных и сочетает преимущества двух предыдущих методов.

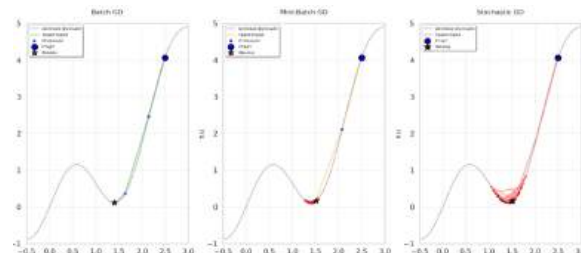


Рис. 1 – Сравнение видов градиентного спуска

III. Заключение

В работе были рассмотрены математические основы градиентного спуска и его основные разновидности. Установлено, что мини-батч градиентный спуск является наиболее эффективным на практике, поскольку обеспечивает баланс между скоростью и точностью обучения моделей.

Список литературы

1. Гасников, А. В. Современные численные методы оптимизации / А. В. Гасников. – М.: МЦНМО, 2021. – С. 45–67.
2. Миронов, В. В. Машинное обучение. Ч. 1: учебное пособие / В. В. Миронов. – Тверь: ТвГТУ, 2019. – С. 112–120.
3. Ворон, О. Л. Универсальный градиентный спуск в задачах оптимизации больших данных / О. Л. Ворон, А. В. Гасников. – М.: ИППИ РАН, 2020. – С. 23–35.

Пищ Тихон Геннадьевич, студент кафедры интеллектуальных информационных технологий БГУИР, elnikovvalentin2@gmail.com.

Яцук Кирилл Михайлович, студент кафедры интеллектуальных информационных технологий БГУИР, y.yatsukwr@gmail.com.

Куличик Владислав Сергеевич, студент кафедры интеллектуальных информационных технологий БГУИР, kulichik.00@bk.ru.

Научный руководитель: Зотов Никита Владимирович, ассистент кафедры интеллектуальных информационных технологий БГУИР, n.zotov@bsuir.by.