Литература

1. Зачем нужна математика? [Электронный ресурс]. - Режим доступа: https://logiclike.com. - Дата доступа: 22.09.2023.

2. Математические модели в энергоменеджменте [Электронный ресурс]. - Режим доступа: http://spravochnik.ru. - Дата доступа: 21.05.23.

3. Зачем энергоменеджеру высшая математика? [Электронный ресурс]. - Режим доступа: http://spbu.ru . - Дата доступа: 15.06.23.

4. Энергетическая связь [Электронный ресурс]. - Режим доступа: https://studizba.com - Дата доступа: 17.08.23.

УДК 51-37

МЕТОДЫ ОПТИМИЗАЦИИ И УСКОРЕНИЯ ОБУЧЕНИЯ НЕЙРОННЫХ СЕТЕЙ ДЛЯ МЕТОДА ОБУЧЕНИЯ BACKPROPAGATION

Пустошило А.В.
Научный руководитель – Королёва М.Н., старший преподаватель кафедры «Высшая математика»

Целью научного исследования является анализ существующих методов оптимизации и рассмотрения математического обоснования каждого оптимизатора для понимания их настройки.

Обучение нейросетей реализуется с помощью метода распространения обратной ошибки, в котором расчёт веса нейрона производится формулой:

$$\omega_n = \omega_n - \lambda \cdot \delta \cdot f_{n-1} (1)$$

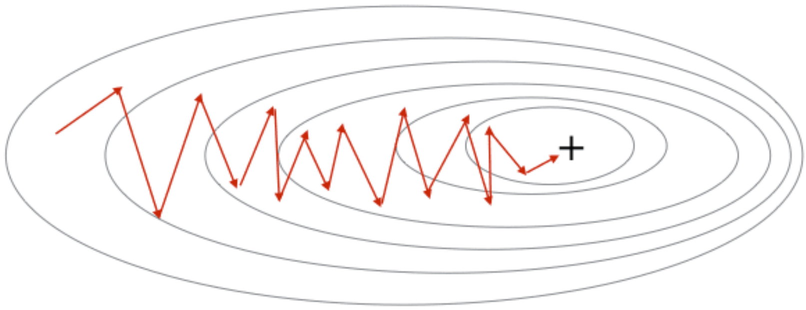
гдеω - вес входной связи нейрона, λ - шаг сходимости, δ - локальный градиент функции активации нейрона, f_{n-1}- функция активации нейрона прошлого слоя. При использовании формулы(1)возникают проблема остановки в локальном минимуме и медленной скорости обучения модели.Задачей метода распространения обратной ошибки является минимизация значения функции потерь, которая отражает разницу между выходным значением нейросети и верным значением из множества данных для обучения, поэтому оптимизаторы базируются на алгоритме градиентного спуска и являются его модификациями.

В данной работе будут рассмотреныпопулярные оптимизаторы, которые решают вышеописанные проблемы:

1. Мини-пакетный стохастический градиентный спуск (SGD).

Данный оптимизатор после каждой тренировочной итерации выбирает, случайное подмножество тренировочного набора данных, этоменяет“ландшафт” функции потерь, что позволяет преодолевать локальные минимумы и увеличивает вычислительную эффективность.

Однако на рисунке сходимости функции потерь наблюдаются большие горизонтальные колебания, которые тормозят обучение модели:



2. Оптимизатор импульсов (Momentum).

Оптимизатор импульсовпозволяет ускорить обучение путём усреднения градиента с помощью, экспоненциально скользящего среднего (2):

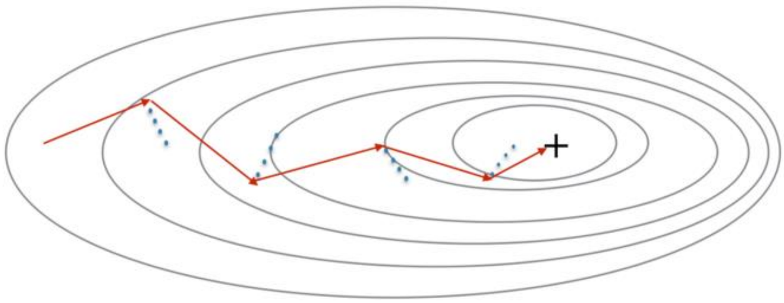
$$\nu = \gamma\nu + (1-\gamma)\cdot\delta \; (2)$$

где v - переменная, которая накапливает среднее значение градиентов, y - константа от 0 до 1, которая указывает на сколько много прошлых градиентов будет учитываться (обычно он равняется 0,9).

Заменив второе слагаемое в формуле (1) на v), сохранив шаг сходимости λ, получимформулудля расчёта весов:

$$\omega_n = \omega_n - \lambda\cdot\nu \; (3)$$

Благодаря появлению «инерции» мы преодолеваем небольшие локальные минимумы и получаем сглаживание колебаний по вертикальной оси, что отражает рисунок ниже:



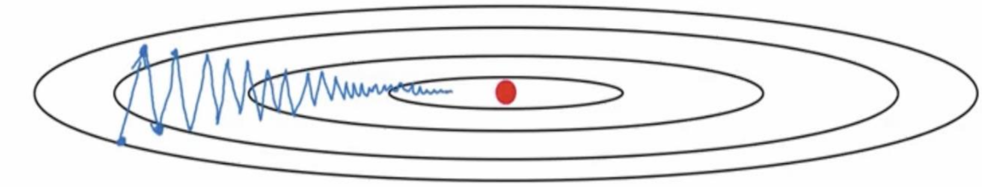
3. Среднеквадратичное распространение (RMSProp).

Градиенты могут сильно отличаться по диапазону значений, что приводит к неравномерной скорости изменения весов. Для равномерного изменения весов в RMSProp применяются следующие формулы:

$G = \gamma_2 G + (1-\gamma_2)\cdot\delta^2$ (4),

$\omega_n = \omega_n - \lambda\cdot\dfrac{\delta}{\sqrt{G+\varepsilon}}$ (5)

где G - переменная, которая накапливает среднее значение градиентов, ε– небольшое значение для избегания деления на 0. На рисунке можно наблюдать затухающие колебания:



Такой оптимизатор позволяет градиентам с маленьким значением расти быстрее, при это более большие градиенты уменьшаются, что позволяет быстрее находить точку минимума функции потерь из-за уменьшения амплитуды колебаний алгоритма.

4. Адаптивная оценка момента (Adam).

На сегодняшний день, Adam является наиболее популярным оптимизатором, он представляет из себя объединениеRMSPropи Momentum, поэтому для его реализации выполняются расчеты по формулам (2), (4) и производится нормировка с использованием следующих формул:

$\hat{v} = v \div (1-\gamma)$ (6),

$\hat{G} = G \div (1-\gamma_2)$ (7)

Это производится для настройки скорости затуханияградиентов коэффициентами γ и γ_2. Для расчета весов применяется формула:

$$\omega_n = \omega_n - \lambda \cdot \frac{\hat{v}}{\sqrt{\hat{G}+\varepsilon}} \quad (8)$$

Такая реализация позволяет этому алгоритму быть наиболее универсальным, однако он не является панацеей.

Таким образом в данной работе были рассмотрены наиболее часто используемые оптимизаторы, их математические формулы, параметры для настройки, и различия между ними.Как показывает практика, выбор оптимизатора будет завесить от параметров модели, знаний инженера и его практики в разработке нейросетей.

Литература:

1. Методы оптимизации нейронных сетей [Электронный ресурс]. – Режим доступа: https://habr.com/ru/articles/318970/ – Дата доступа 10.05.2024

2. Оптимизаторы градиентных алгоритмов: RMSProp, AdaDelta, Adam, Nadam [Электронный ресурс]. – Режим доступа: https://proproprogs.ru/ml/ml-optimizatory-gradientnyh-algoritmov-rmsprop-adadelta-adam-nadam– Дата доступа 13.05.2024

3. Нейросети — это просто (Часть 7): Адаптивные методы оптимизации [Электронный ресурс]. – Режим доступа: https://www.mql5.com/ru/articles/8598#para24/ – Дата доступа 11.05.2024

УДК 004.932

ВИРТУАЛЬНЫЕ ТУРЫ: МАТЕМАТИЧЕСКАЯ МОДЕЛЬ

Живоглод Н.А., Павловский Н. С., Романко С. Ю.
Научный руководитель – Бадак Б.А., старший преподаватель кафедры «Высшая математика»

Актуальность данной работы заключается в том, что изучение математической модели виртуальных туров может существенно упростить их создание, а также повысить их качество.

Целью научного исследования является создание математической модели виртуального тура

Для достижения поставленной цели нами были выработаны следующие математические методы:

I. Анализ и оценка расчета погрешности

В ходе работы был использован штатив и iPhone 13. Угол падения и угол отражения света при фотосъемке с использованием камеры iPhone или любой другой камеры подчиняются закону отражения. Согласно этому закону, угол падения (θ_1) равен углу отражения (θ_2).

θ_1 - угол между падающим лучом и нормалью к поверхности в точке падения, а θ_2 - угол между отраженным лучом и той же нормалью.

При съемке панорамы с использованием штатива возможно смещение головки штатива в горизонтальной плоскости. Это может привести к небольшому съезду при стыке панорамных кадров. Для оценки этого смещения можно использовать следующую формулу:

$$x = fdL$$

где:

(x) - смещение

(d) - расстояние от головки штатива до стыка панорамных кадров (в миллиметрах),

(f) - фокусное расстояние объектива (в миллиметрах),

(L) - длина панорамы (в миллиметрах).

II. Алгоритм обхода всех точек интересов по кратчайшему пути

В работе используется Гамильтонов цикл для обхода всех точек интересов по кратчайшему пути (1⇒2⇒3⇒4⇒5⇒1).