به نام خدا

تغییر وزن‌ها در شبکه عصبی می‌تواند به روش‌های مختلفی انجام شود، که به طور کلی به عنوان روش‌های بهینه‌سازی شناخته می‌شوند. این روش‌ها بر اساس شیوه‌ی محاسبه و به‌روزرسانی وزن‌ها دسته‌بندی می‌شوند. در ادامه، برخی از رایج‌ترین روش‌های بهینه‌سازی و نحوه‌ی تغییر وزن‌ها در هر یک توضیح داده می‌شود:

1. \*\*گرادیان نزولی (Gradient Descent)\*\*

- \*\*تک‌تکه (Batch Gradient Descent):\*\*

در این روش، گرادیان‌ها برای کل داده‌های آموزشی محاسبه می‌شوند و سپس وزن‌ها یک بار به‌روزرسانی می‌شوند. این روش ممکن است بسیار کند باشد، به‌ویژه برای داده‌های بزرگ.

- \*\*گرادیان نزولی تصادفی (Stochastic Gradient Descent - SGD):\*\*

در این روش، به‌روزرسانی وزن‌ها بعد از هر نمونه‌ی آموزشی انجام می‌شود. این روش به دلیل سرعت بالاتر در هر تکرار، محبوب است اما می‌تواند نویز بیشتری در به‌روزرسانی‌ها داشته باشد.

- \*\*گرادیان نزولی دسته‌ای (Mini-Batch Gradient Descent):\*\*

ترکیبی از دو روش بالا است. به‌روزرسانی وزن‌ها برای یک دسته‌ی کوچک (مینی‌بچ) از داده‌ها انجام می‌شود. این روش به تعادل خوبی بین سرعت و پایداری دست می‌یابد.

2. \*\*گرادیان نزولی با مومنتوم (Gradient Descent with Momentum)\*\*

در این روش، سرعت تغییر وزن‌ها (مومنتوم) نیز در نظر گرفته می‌شود. این روش به کاهش نوسانات و تسریع همگرایی کمک می‌کند.

3. \*\*آدام (Adam - Adaptive Moment Estimation)\*\*

آدام یکی از محبوب‌ترین روش‌های بهینه‌سازی است که از دو مفهوم مومنتوم و RMSProp استفاده می‌کند.

1. به‌روزرسانی میانگین گرادیان‌ها (مومنتوم):

2. به‌روزرسانی میانگین مربع گرادیان‌ها:

3. اصلاحات بایاس برای مومنتوم و میانگین مربع گرادیان‌ها:

4. به‌روزرسانی وزن‌ها:

4. \*\*RMSProp (Root Mean Square Propagation)\*\*

این روش از میانگین مربع گرادیان‌ها برای تنظیم نرخ یادگیری استفاده می‌کند.

1. به‌روزرسانی میانگین مربع گرادیان‌ها:

2. به‌روزرسانی وزن‌ها:

این روش‌ها هر کدام مزایا و معایب خاص خود را دارند و بسته به مسئله‌ی مورد نظر، داده‌ها و تنظیمات شبکه عصبی انتخاب می‌شوند.