NeuralNetwork

lucassauvage

November 2024

Contents

Definitions

Loss Function : A loss function is a mathematical function that measures the difference between the predicted output of a machine learning model and the actual output.

1 Introduction

This document summarizes all the mathematics used in my Neural Network project. This project is an AI currently trained to recognize numbers. Coding was not the hardest part of this project, as there is a lot of code already available online. The challenge lies in understanding the mathematics required to build a program that can learn on its own.

2 Cross Entropy

The cross entropy between two probability distribution p and q over the same underlying set of events, measures the average number of bits needed to identify an event drawn from the set when the coding scheme used for the set is optimized for an estimated probability distribution q, rather than the true distribution p

2.1 For discrete probability

For discrete probability distribution p and q this means :

$$H(p, q) = -\sum_{x} p(x) \log q(x)$$

$H(p, q)$ is often used for the cross entropy and the joint entropy of p and q.

2.2 Cross-entropy minimization

Cross-entropy is often used in the maximum likelihood framework. In my project, cross entropy's used to for the loss function. Using the cross-entropy formulation between the real (or "true") data distribution p and the estimated distribution q, we actually minimize the cross-entropy to find the optimal parameters. Minimize cross entropy means maximising model likelihood.

3 Sources

Wikipedia : https://en.wikipedia.org/wiki/Cross-entropy