Basic MLP with manually-derived Backprop

Teo Asinari

October 7, 2023

1 Introduction

Goal: To design, train and use a simple 3-layer MLP for binary classification of size-2 vectors.

Design: of the form

$$[(layer\_size, Activation)...]$$

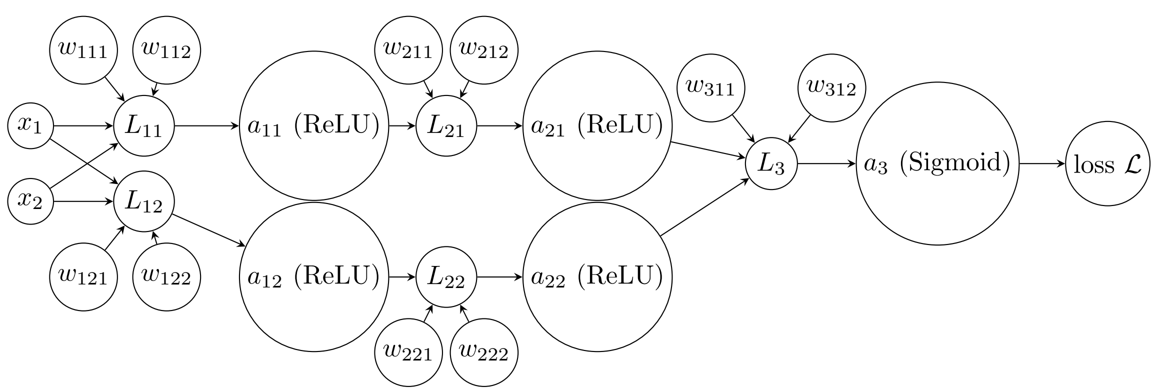
: [(2, ReLU), (2, ReLU), (1, Sigmoid)]

1.1 Diagrams

1.1.1 Vectorized Diagram

$$\boldsymbol{x} \longrightarrow \overset{\boldsymbol{w_1}\ \searrow}{\boldsymbol{L_1}} \to \boldsymbol{a_1} \to \overset{\boldsymbol{w_2}\ \searrow}{\boldsymbol{L_2}} \to \boldsymbol{a_2} \to \overset{\boldsymbol{w_3}\ \searrow}{\boldsymbol{L_3}} \to \boldsymbol{a_3} \longrightarrow \mathcal{L}$$

1.1.2 Expanded Diagram



1.2 Definitions

1.2.1 Remark on weight notation

$w_{i,j,k}$ is to say the weight at the i-th layer, j-th neuron, k-th weight. Hence w_{111} is the first weight of the first neuron in the first layer, etc.

1.2.2 Remark on layer notation

This is a sub-case of the weight notation. I.e., L_{ij} is the j-th neuron at the i-th layer, etc.

1.2.3 Neuron firing calculation

This is just a straightforward dot-product. We have:

$$L_i = \boldsymbol{w}_i \boldsymbol{x}_i$$

1.2.4

1.3 BackPropagation Derivation