Yeshiva University

AIM 5007 Neural Networks and Deep Learning
Problem Set 4

Introduction

In this problem set, we will explore some different possible activation functions. Previously, we have used the sigmoid function as our activation function. Here, we will consider a few more possible choices. For each of the activation functions given below, you need to differentiate it in order to have what is needed for gradient descent.

For each of the following functions, do the following:

1. Sketch the function
2. Differentiate the function (give the formula for $\frac{df}{dx}$)
3. Identify any tricky points in taking the derivative (i.e., discontinuities) and propose a reasonable solution to the difficulty (i.e., assign a value for that point)
4. Sketch the derivative

You may sketch by hand or use software to produce the sketches. You should show your steps for the differentiation.

Problem 1: Hyperbolic Tangent

$$f(x) = \tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

Problem 2: Linear Function

$$f(x) = kx$$

Problem 3: Step Function

$$f(x) = \begin{cases} 0 \; if \; x \leq 0 \\ 1 \; if \; x > 0 \end{cases}$$

Problem 4: Rectified Linear Unit (ReLU)

$$f(x) = \begin{cases} 0 \; if \; x \leq 0 \\ x \; if \; x > 0 \end{cases}$$

Problem 5: Softplus (Not Softmax!)

$$f(x) = \ln(1 + e^x)$$