**Ex1 Report**

**Q1:**

1. Let .

We showed that for every , and therefore

1. For :

The result we got means that the softmax and sigmoid functions are equivalent in a way that

1. We can suggest the hyperbolic tangent function to map to :

Then we can use in order to map to .

**Q2:**

1. We can take

That way we get

And we can see that

1. Let’s show that its not possible:

We now require

Also,

And in the same way for we get:

So far, we can conclude that:

Also,

But we know that , so and we can see that

**Q3:**

First random initialization:

Graph of the loss value as a function of the number of epochs:

תמונה שמכילה טקסט, צילום מסך, עלילה, תרשים

התיאור נוצר באופן אוטומטי

Second random initialization:

Graph of the loss value as a function of the number of epochs:

תמונה שמכילה טקסט, צילום מסך, עלילה, תרשים

התיאור נוצר באופן אוטומטי

Looks like an actual solution would be

Let’s check: