

BRAC University: ML Lab - Scaled Dot Product Attention

Zadid Hasan

August 26th, 2023

Here are the tasks that you should complete for today's lab:

1. Use the same titanic dataset we used in the first lab and do the same kind of pre-processing we did previously. You need to centralize each of the features. From each column, subtract the mean of that column. Then divide each column by the standard deviation of that column.
2. There is no training in the attention mechanism.
3. In the testing section, take any test data a .
4. Let's assume that we have n features.
5. We have to figure out whether a survives or not.
6. Let's say, we have m data points in the training dataset. For each $i = \{1, 2, \dots, m\}$, x_i is the i 'th feature vector and y_i is the i 'th label, our output for test data a should be as follows (just another softmax):

$$\text{sign}\{\sum_{i=1}^m \text{score}(x_i, a)y_i\}$$

But how do we calculate the scores? It's as follows:

$$\text{score}(x_i, a) = \frac{e^{\frac{a^T x_i}{\sqrt{n}}}}{\sum_{j=1}^m e^{\frac{a^T x_j}{\sqrt{n}}}}$$

7. In this way, calculate the output for each a in the test set. Then finally calculate the accuracy.