**I. Pen-and-paper**

1. Forward Propagation:

This part objective is to transform the input no output , for achieving this we apply recursively these functions

Applying forward propagation to the data given in the assignment, the following activation and net input vectors were obtained:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| (input) |  |  |  |  |  |  |
| 1 | 6 | 0.999988 | 3.76157 | 0.99892 | 0 | 0 |
| 1 | 1 | 0.761594 | 3.76157 | 0.99892 | 0 | 0 |
| 1 | 6 | 0.999988 |  | | | |
| 1 |  | | | | | |
| 1 |

Back Propagation:

After performing forward propagation, the stochastic gradient descent is applied in order get new weights and biases that better perform according to the loss function

The derivative needed to calculate SGD can be obtained using the chain rule:

Calculating the

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | |  | |  |  | |  |
| 0 | 0 | 0.00216 | 0 | 0 | 1 | 0 | 0.999988 |
| 0 | 0 | 0 | 0.00216 | 0 | 0 | 1 | 0.761594 |
|  |  |  |  |  |  |  | 0.999988 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  | |  |  | |  |
| -1 | 1 | 0 | -1 | 1 | 0 | 0.99892 |
| 1 | 0 | 1 | 1 | 0 | 1 | 0.99892 |

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | |  | | |  |  | | |  |
| 1 | 1 |  | 0 | 0 | 0 | 1 | 0 | 0 | 1 |
| 1 | 1 | 0 | 0.419974 | 0 | 0 | 0 | 1 | 0 | 1 |
| 1 | 1 | 0 | 0 |  | 0 | 0 | 0 | 1 | 1 |
|  |  |  |  |  |  |  |  |  | 1 |
|  |  |  |  |  |  |  |  |  | 1 |

Updates:

Agora possuímos todos os valores necessários para calcular o SGD, podemos atualizar os pesos e os biases

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | |  | |  |  |
| -0.9989 | -0.9989 | 0.09989 | 0.09989 | -1 | 0.1 |
| 0.9989 | 0.9989 | -0.09989 | -0.09989 | 1 | -0.1 |

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | | | |  | | | | |  |  |
| 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 1 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 1 |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | | |  | | |  |  |
| 0 | 0 | 0 | 1 | 1 | 1 | 0 | 1 |
| 0 | 0 | 0 | 1 | 1 | 1 | 0 | 1 |

1. O problema é semelhante ao anterior, apenas temos uma nova função de ativação na camada de output (softmax) e a loss function (cross-entropy), desta forma os valores do execicio anterior vão ser semenlhantes pois as os dados são semelhantes, apenas vai ser alterado:

Forward Propagation:

Back Propagrtion:

Com estas novas derivadas obtemos

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | |  |
| -2 | 0.25 | -0.25 | -0.5 |
| 0 | -0.25 | 0.25 | 0.5 |

Updates:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | |  | |  |  |
| -0.4995 | -0. 4995 | 0.04995 | 0.04995 | -0.5 | 0.05 |
| 0. 4995 | 0. 4995 | -0.04995 | -0.04995 | 0.5 | -0.05 |

Foi omitido alguns dados que a sua obtenção é igual a do exercício anterior

**II. Programming and critical analysis**

1. Answer 5
2. Answer 6
3. Answer 7
4. Answer 8

**III. APPENDIX**

Paste your programming code here using Consolas 9pt or 10pt.

Use **highlighting** or colored text to facilitate the analysis by your faculty hosts.

**END**