2.1 Embeding size

Constant parameters:

Skip window = 1

Learning rate = 1.0

Optimizer = AdamOptimizer

Number of steps = 9

Table: Average loss per Embeding size

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Embdeing  Size | Step 1 | Step 2 | Step 3 | Step 4 | Step 5 | Step 6 | Step 7 | Step 8 | Step9 |
| 4 | 339.8 | 236.3 | 114.6 | 110.0 | 213.2 | 142.7 | 94.36 | 173.8 | 122.5 |
| 8 | 1081.2 | 678.3 | 253.4 | 273.5 | 572.7 | 263.4 | 187.8 | 400.8 | 267.4 |
| 16 | 2137.7 | 1774.9 | 863.9 | 782.3 | 1119.4 | 726.6 | 576.7 | 999.1 | 622.4 |
| 32 | 2953.5 | 3099.5 | 1989.5 | 1476.2 | 2167.2 | 1838.2 | 1412.1 | 2209.3 | 1392.8 |

Conclusion: As the embeding size increases the average loss also increases. The best result was obtained when the embeding size was 4.

2.2 Skip window

Constant parameters:

Embeding size = 4

Learning rate = 1.0

Optimizer = AdamOptimaizer

Number of steps = 9

Table: Average loss per skip window

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Skip  window | Step 1 | Step 2 | Step 3 | Step 4 | Step 5 | Step 6 | Step 7 | Step 8 | Step9 |
| 1 | 339.8 | 236.3 | 114.6 | 110.0 | 213.2 | 142.7 | 94.36 | 173.8 | 122.5 |
| 2 | 308.4 | 231.9 | 161.7 | 154.4 | 143.5 | 118.0 | 89.41 | 255.79 | 127.1 |
| 3 | 285.4 | 247.2 | 186.3 | 123.7 | 180.6 | 137.3 | 98.52 | 158.8 | 114.3 |
| 4 | 320.3 | 249.1 | 120.1 | 130.2 | 221.4 | 88.22 | 81.86 | 157.9 | 120.5 |
| 5 | 363.5 | 222.9 | 116.2 | 131.2 | 177.2 | 137.0 | 91.68 | 173.2 | 107.1 |
| 6 | 393.99 | 244.7 | 161.5 | 129.4 | 166.2 | 104.5 | 126.8 | 129.8 | 113.5 |
| 7 | 347.8 | 180.3 | 154.16 | 159.3 | 158.37 | 109.8 | 99.10 | 169.1 | 148.3 |

Conclusion: As the skip window increases the average loss decreases and then increases again. The best result was obtained wen the window size was 5.

2.3 Learning rate

Constant parameters:

Embeding size = 4

Skip window = 5

Optimizer = AdamOptimaizer

Number of steps = 9

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Learning  rate | Step 1 | Step 2 | Step 3 | Step 4 | Step 5 | Step 6 | Step 7 | Step 8 | Step9 |
| 1.0 | 363.5 | 222.9 | 116.2 | 131.2 | 177.2 | 137.0 | 91.68 | 173.2 | 107.1 |
| 0.1 | 4.49 | 3.7 | 3.36 | 3.59 | 3.56 | 3.48 | 3.38 | 3.95 | 3.59 |
| 0.01 | 10.7 | 3.92 | 3.22 | 3.29 | 3.18 | 3.194 | 3.09 | 3.46 | 3.30 |
| 0.001 | 40.7 | 16.6 | 9.18 | 7.37 | 6.10 | 5.15 | 4.58 | 4.53 | 4.2 |
| 0.0001 | 56.16 | 50.74 | 45.01 | 39.54 | 34.03 | 30.3 | 26.2 | 25.7 | 23.3 |

Conclusion: As the learning rate decreases the average loss decreases and then increases again. The best result was obtained wen the learning rate was 0.01.

2.4 Optimizer

Constant parameters:

Embeding size = 4

Skip window = 5

Learning rate = 0.01

Number of steps = 9

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Optimizer | Step 1 | Step 2 | Step 3 | Step 4 | Step 5 | Step6 | Step7 | Step8 | Step9 |
| AdagradOptimizer | 44.4 | 31.7 | 24.1 | 20.2 | 17.1 | 16.3 | 14.3 | 15.1 | 14.2 |
| GradientDecentOptimizer | 17.8 | 6.94 | 4.48 | 4.24 | 3.8 | 3.73 | 3.44 | 3.83 | 3.67 |
| AdamOptimizer | 10.7 | 3.92 | 3.22 | 3.29 | 3.18 | 3.194 | 3.09 | 3.46 | 3.30 |
| RMSPropOptimizer | 12.7 | 4.08 | 3.60 | 4.07 | 4.21 | 4.35 | 4.42 | 5.16 | 4.73 |

Conclusion: The best result was obtained for AdamOptimizer.

2.4 Number of steps

Constant parameters:

Embeding size = 4

Skip window = 5

Learning rate = 0.01

Optimizer = AdamOptimizer

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Number of  steps | Average loss per steps | | | |
| 5 | 9 | 7 | 15 |
| 1 | 11.2 | 11.19 | 11.1 | 11.1 |
| 2 | 3.8 | 3.92 | 3.96 | 4.08 |
| 3 | 3.22 | 3.23 | 3.22 | 3.27 |
| 4 | 3.28 | 3.30 | 3.28 | 3.29 |
| 5 |  | 3.21 | 3.18 | 3.19 |
| 6 |  | 3.19 | 3.19 | 3.2 |
| 7 |  | 3.09 | 3.09 | 3.09 |
| 8 |  | 3.45 |  | 3.46 |
| 9 |  | 3.31 |  | 3.306 |
| 10 |  |  |  | 3.28 |
| 11 |  |  |  | 3.24 |
| 12 |  |  |  | 3.39 |
| 13 |  |  |  | 3.19 |
| 14 |  |  |  | 3.06 |
| 15 |  |  |  | 3.25 |

Conclusion : The best result was obtained when the number of steps was 7.

Finally, The best result for this data set can be obtained when the parameters are as follows:

Embeding size = 4

Skip window = 5

Learning rate = 0.01

Optimizer = AdamOptimizer

Number of steps = 7

