Default parameters, exploring distance measures with cross entropy model:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Model** | **Similarity measure components** | **Least** | **Most** | **Overall** |
| Cross entropy | 1.0/dist\_diff | 30.3 | 32.4 | 31.3 |
| Cross entropy | 1.0/ absolute\_euclid\_dist\_diff | **32.2** | **34.9** | **33.5** |
| Cross entropy | pairwise\_cossim | **32.2** | **34.9** | **33.5** |
| Cross\_entropy | diff\_cossim | 28.6 | 33.3 | 30.9 |
| Cross\_entropy | 1.0/abs(avg\_cossim - pairwise\_cossim) | 30.4 | 31.9 | 31.2 |
| Cross entropy | 1.0/(absolute\_euclid\_dist\_diff \* abs(avg\_cossim - pairwise\_cossim) ) | 30.3 | 32.6 | 31.5 |
| Cross entropy | pairwise\_cossim/(absolute\_euclid\_dist\_diff \* abs(avg\_cossim - pairwise\_cossim) ) | 32.2 | 32.7 | 32.4 |
| Cross entropy | pairwise\_cossim / (absolute\_euclid\_dist\_diff) | **30.4** | **34.8** | **32.6** |

Hyper-Parameter Tuning

Default parameters:

batch\_size=128 skip\_window=4 num\_skips=8 num\_sampled=64 max\_num\_steps=200001 learning\_rate=1.0

**Accuracy after training cross-entropy from scratch on default parameters:**

**Least illustrative accuracy: 30.3 %, Most illustrative accuracy: 32.6 %, Overall Accuracy: 31.5%**

Tuning models after training from scratch

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Model | Hyper parameters | Least | Most | Overall |
| Cross entropy | batch\_size=64 | 30.0 | 31.6 | 30.8 |
| Cross entropy | batch\_size = 32 | **32.8** | **35.6** | **34.2** |
| Cross entropy | batch\_size=256 | 29.6 | 30.4 | 30.0 |
| Cross entropy | skip\_window=2 num\_skips=4 | 31.1 | 34.1 | 32.6 |
| Cross entropy | skip\_window=8 num\_skips=16 | **35.9** | **33.6** | **34.7** |
| Cross entropy | batch\_size=128, skip\_window=4 num\_skips=8 num\_sampled=64 max\_num\_steps=10000 learning\_rate=1.0, **embedding\_size=256** | 30.4 | 31.3 | 30.9 |

Default training parameters (using pretrained), exploring distance measures with cross entropy model:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Model** | **Hyper parameters** | **Least** | **Most** | **Overall** |
| NCE | 1.0/dist\_diff | **32.2** | **35.0** | **33.6** |
| NCE | 1.0/ absolute\_euclid\_dist\_diff | 29.9 | 31.3 | 30.6 |
| NCE | pairwise\_cossim | 29.9 | 31.3 | 30.6 |
| NCE | diff\_cossim | 28.2 | 28.4 | 28.3 |
| NCE | 1.0/abs(avg\_cossim - pairwise\_cossim) | **32.4** | **35.7** | **34.0** |
| NCE | 1.0/dist\_diff \* abs(avg\_cossim - pairwise\_cossim) ) | **32.3** | **35.4** | **33.9** |

Hyper-Parameter Tuning

Default parameters:

batch\_size=128 skip\_window=4 num\_skips=8 num\_sampled=64 max\_num\_steps=200001 learning\_rate=1.0

**Accuracy after training cross-entropy from scratch on default parameters:**

**Least illustrative accuracy: 32.4 %, Most illustrative accuracy: 35.7 %, Overall Accuracy: 34.0%**

Tuning models using pretrained (Using measure: 1.0/(abs(avg\_cossim - pairwise\_cossim) \* dist\_diff)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Model** | **Hyper parameters** | **Least** | **Most** | **Overall** |
| NCE | batch\_size=64 | 31.5 | 35.2 | 33.4 |
| NCE | batch\_size = 32 (accuracy is pathetic because loss did not converge around 1.x) | 29.8 | 31.4 | 30.6 |
| NCE | batch\_size=256 | 32.1 | 34.7 | 33.4 |
| NCE | skip\_window=2 num\_skips=4 | 31.8 | 35.3 | 33.6 |
| NCE | skip\_window=8 num\_skips=16 | **32.2** | **34.4** | **33.3** |
| NCE | num\_sampled=32 | 33.3 | 35.1 | 34.2 |
| NCE | num\_sampled=16 (loss convergence at 0.22x) | 32.4 | 33.8 | 33.1 |
| NCE | num\_sampled=8 (loss convergence at 0.21x)  (mikolov recommends smaller num\_sampled for large data-sets) | **32.5** | **35.8** | **34.1** |
| NCE | num\_steps=400001 (loss convergence at 1.41)  loss did not converge any better by increasing steps,  so no point in increasing number of steps any more | 32.2 | 33.3 | 34.4 |
| NCE | learning\_rate=0.5  increase in learning rate might overshoot the loss, hence exploring lower learning\_rate than the standard 1.0.  loss did not converge any better (1.45), hence exploring more lower learning rate would not be intuitive | 32.1 | 34.7 | 33.4 |