Yin Junyu G2101985L

**1.a**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | TP | FN | FP | Precision | Recall | F1 | Support |
| O | 21860 | 74 | 1578 | 0.9327 | 0.9966 | 0.9636 | 21934 |
| B-person | 27 | 384 | 37 | 0.4219 | 0.0657 | 0.1137 | 411 |
| B-location | 23 | 120 | 67 | 0.2556 | 0.1608 | 0.1974 | 143 |
| I-person | 16 | 105 | 27 | 0.3721 | 0.1322 | 0.1951 | 121 |
| B-group | 5 | 170 | 9 | 0.3571 | 0.0286 | 0.0529 | 175 |
| I-location | 7 | 36 | 19 | 0.2692 | 0.1628 | 0.2029 | 43 |
| B-corporation | 0 | 150 | 0 | 0 | 0 | 0 | 150 |
| I-creative-work | 14 | 223 | 22 | 0.3889 | 0.0591 | 0.1026 | 237 |
| I-product | 0 | 77 | 12 | 0 | 0 | 0 | 77 |
| I-group | 4 | 62 | 5 | 0.4444 | 0.0606 | 0.1067 | 66 |
| B-product | 0 | 121 | 4 | 0 | 0 | 0 | 121 |
| B-creative-work | 5 | 226 | 10 | 0.3333 | 0.0216 | 0.0407 | 231 |
| I-corporation | 0 | 42 | 0 | 0 | 0 | 0 | 42 |
| macro/weighted F1 (removing O) | | | 0.0843/0.0866 | | | | |

Table 1. Results of the CRF when removing the features about POS tags.

**1.b**

According to the tqdm results, if we train the model for 10,000 epochs, it takes about 5 hours (See Figure 1). Since it is so time-consuming, we decide to change the number of epochs to 1000 while keeping other configurations untouched. And when plotting the training curves (See Figure 2), we can observe that in fact the loss function starts to converge after about 100 epochs. Therefore, it is reasonable to reduce the number of epochs, as the model converges early in the training stage.



Figure 1. Estimated training time for 10,000 epochs.

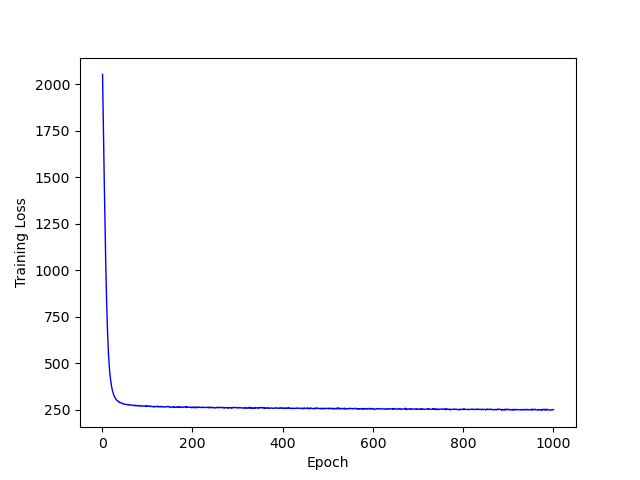


Figure 2. Training curves of the word window classifier.

The experiment results are shown in the following table. we can see that this model suffers a lot from the overfitting problem and the imbalance dataset.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | TP | FN | FP | Precision | Recall | F1 | Support |
| O | 21934 | 0 | 1817 | 0.9235 | 1 | 0.9602 | 21934 |
| B-person | 0 | 411 | 0 | 0 | 0 | 0 | 411 |
| B-location | 0 | 143 | 0 | 0 | 0 | 0 | 143 |
| I-person | 0 | 121 | 0 | 0 | 0 | 0 | 121 |
| B-group | 0 | 175 | 0 | 0 | 0 | 0 | 175 |
| I-location | 0 | 43 | 0 | 0 | 0 | 0 | 43 |
| B-corporation | 0 | 150 | 0 | 0 | 0 | 0 | 150 |
| I-creative-work | 0 | 237 | 0 | 0 | 0 | 0 | 237 |
| I-product | 0 | 77 | 0 | 0 | 0 | 0 | 77 |
| I-group | 0 | 66 | 0 | 0 | 0 | 0 | 66 |
| B-product | 0 | 121 | 0 | 0 | 0 | 0 | 121 |
| B-creative-work | 0 | 231 | 0 | 0 | 0 | 0 | 231 |
| I-corporation | 0 | 42 | 0 | 0 | 0 | 0 | 42 |
| macro/weighted F1 (removing O) | | | 0/0 | | | | |

Table 2. Results of the word window classifier with the default setting except for the number of epochs.

**1.c**

We observe that the batch size of the default configuration is 4. I think that this batch size is too small and may cause the gradient calculated per batch to vary greatly. So I decided to change the batch size to 128 and the learning rate to 0.002 correspondingly (ten times the original).

From 1.b, we can see that in fact the loss function starts to converge after about 100 epochs. And after changing the batch size and the learning rate simultaneously, this number of epochs won’t change a lot, so I decide to use 300 epochs for training here. Figure 3 proves that 300 epochs are sufficient for the loss function to converge.

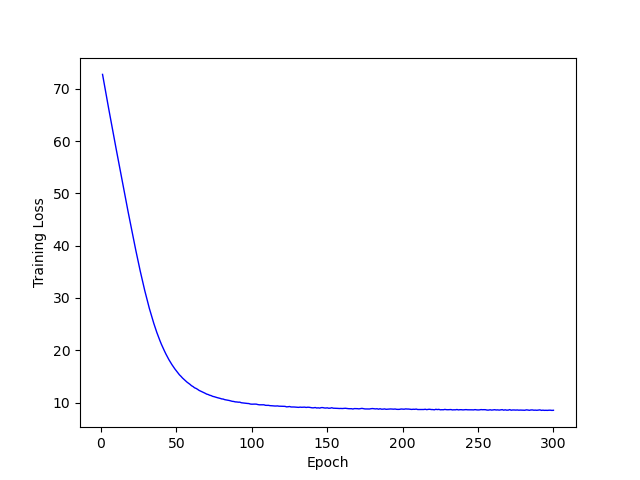


Figure 3. Training curves of the word window classifier with batch size = 128, learning rate = 0.002 and number of epochs = 300.

Then we start our experiments about hyperparameters from the above setting.

Window Size

|  |  |
| --- | --- |
| Window Size | 5 or 3 |
| Embedding Size | 25 |
| Hidden Layer Size | 25 |
| Number of Hidden Layers | 1 |
| Freeze Word Embeddings or Not | No |
| Learning Rate | 0.002 |
| Number of Epochs | 300 |

Table 3. Training configurations.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | TP | FN | FP | Precision | Recall | F1 | Support |
| O | 21934 | 0 | 1817 | 0.9235 | 1 | 0.9602 | 21934 |
| B-person | 0 | 411 | 0 | 0 | 0 | 0 | 411 |
| B-location | 0 | 143 | 0 | 0 | 0 | 0 | 143 |
| I-person | 0 | 121 | 0 | 0 | 0 | 0 | 121 |
| B-group | 0 | 175 | 0 | 0 | 0 | 0 | 175 |
| I-location | 0 | 43 | 0 | 0 | 0 | 0 | 43 |
| B-corporation | 0 | 150 | 0 | 0 | 0 | 0 | 150 |
| I-creative-work | 0 | 237 | 0 | 0 | 0 | 0 | 237 |
| I-product | 0 | 77 | 0 | 0 | 0 | 0 | 77 |
| I-group | 0 | 66 | 0 | 0 | 0 | 0 | 66 |
| B-product | 0 | 121 | 0 | 0 | 0 | 0 | 121 |
| B-creative-work | 0 | 231 | 0 | 0 | 0 | 0 | 231 |
| I-corporation | 0 | 42 | 0 | 0 | 0 | 0 | 42 |
| macro/weighted F1 (removing O) | | | 0/0 | | | | |

Table 4. Results of the word window classifier with window size = 5.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | TP | FN | FP | Precision | Recall | F1 | Support |
| O | 21934 | 0 | 1817 | 0.9235 | 1 | 0.9602 | 21934 |
| B-person | 0 | 411 | 0 | 0 | 0 | 0 | 411 |
| B-location | 0 | 143 | 0 | 0 | 0 | 0 | 143 |
| I-person | 0 | 121 | 0 | 0 | 0 | 0 | 121 |
| B-group | 0 | 175 | 0 | 0 | 0 | 0 | 175 |
| I-location | 0 | 43 | 0 | 0 | 0 | 0 | 43 |
| B-corporation | 0 | 150 | 0 | 0 | 0 | 0 | 150 |
| I-creative-work | 0 | 237 | 0 | 0 | 0 | 0 | 237 |
| I-product | 0 | 77 | 0 | 0 | 0 | 0 | 77 |
| I-group | 0 | 66 | 0 | 0 | 0 | 0 | 66 |
| B-product | 0 | 121 | 0 | 0 | 0 | 0 | 121 |
| B-creative-work | 0 | 231 | 0 | 0 | 0 | 0 | 231 |
| I-corporation | 0 | 42 | 0 | 0 | 0 | 0 | 42 |
| macro/weighted F1 (removing O) | | | 0/0 | | | | |

Table 5. Results of the word window classifier with window size = 3.

Embedding Size

|  |  |
| --- | --- |
| Window Size | 3 |
| Embedding Size | 128 or 256 |
| Hidden Layer Size | 25 |
| Number of Hidden Layers | 1 |
| Freeze Word Embeddings or Not | No |
| Learning Rate | 0.002 |
| Number of Epochs | 300 |

Table 6. Training configurations.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | TP | FN | FP | Precision | Recall | F1 | Support |
| O | 21934 | 0 | 1817 | 0.9235 | 1 | 0.9602 | 21934 |
| B-person | 0 | 411 | 0 | 0 | 0 | 0 | 411 |
| B-location | 0 | 143 | 0 | 0 | 0 | 0 | 143 |
| I-person | 0 | 121 | 0 | 0 | 0 | 0 | 121 |
| B-group | 0 | 175 | 0 | 0 | 0 | 0 | 175 |
| I-location | 0 | 43 | 0 | 0 | 0 | 0 | 43 |
| B-corporation | 0 | 150 | 0 | 0 | 0 | 0 | 150 |
| I-creative-work | 0 | 237 | 0 | 0 | 0 | 0 | 237 |
| I-product | 0 | 77 | 0 | 0 | 0 | 0 | 77 |
| I-group | 0 | 66 | 0 | 0 | 0 | 0 | 66 |
| B-product | 0 | 121 | 0 | 0 | 0 | 0 | 121 |
| B-creative-work | 0 | 231 | 0 | 0 | 0 | 0 | 231 |
| I-corporation | 0 | 42 | 0 | 0 | 0 | 0 | 42 |
| macro/weighted F1 (removing O) | | | 0/0 | | | | |

Table 7. Results of the word window classifier with embedding size = 128.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | TP | FN | FP | Precision | Recall | F1 | Support |
| O | 21934 | 0 | 1817 | 0.9235 | 1 | 0.9602 | 21934 |
| B-person | 0 | 411 | 0 | 0 | 0 | 0 | 411 |
| B-location | 0 | 143 | 0 | 0 | 0 | 0 | 143 |
| I-person | 0 | 121 | 0 | 0 | 0 | 0 | 121 |
| B-group | 0 | 175 | 0 | 0 | 0 | 0 | 175 |
| I-location | 0 | 43 | 0 | 0 | 0 | 0 | 43 |
| B-corporation | 0 | 150 | 0 | 0 | 0 | 0 | 150 |
| I-creative-work | 0 | 237 | 0 | 0 | 0 | 0 | 237 |
| I-product | 0 | 77 | 0 | 0 | 0 | 0 | 77 |
| I-group | 0 | 66 | 0 | 0 | 0 | 0 | 66 |
| B-product | 0 | 121 | 0 | 0 | 0 | 0 | 121 |
| B-creative-work | 0 | 231 | 0 | 0 | 0 | 0 | 231 |
| I-corporation | 0 | 42 | 0 | 0 | 0 | 0 | 42 |
| macro/weighted F1 (removing O) | | | 0/0 | | | | |

Table 8. Results of the word window classifier with embedding size = 256.

Hidden Layer Size

|  |  |
| --- | --- |
| Window Size | 3 |
| Embedding Size | 256 |
| Hidden Layer Size | 128 or 256 |
| Number of Hidden Layers | 1 |
| Freeze Word Embeddings or Not | No |
| Learning Rate | 0.002 |
| Number of Epochs | 300 |

Table 9. Training configurations.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | TP | FN | FP | Precision | Recall | F1 | Support |
| O | 21934 | 0 | 1817 | 0.9235 | 1 | 0.9602 | 21934 |
| B-person | 0 | 411 | 0 | 0 | 0 | 0 | 411 |
| B-location | 0 | 143 | 0 | 0 | 0 | 0 | 143 |
| I-person | 0 | 121 | 0 | 0 | 0 | 0 | 121 |
| B-group | 0 | 175 | 0 | 0 | 0 | 0 | 175 |
| I-location | 0 | 43 | 0 | 0 | 0 | 0 | 43 |
| B-corporation | 0 | 150 | 0 | 0 | 0 | 0 | 150 |
| I-creative-work | 0 | 237 | 0 | 0 | 0 | 0 | 237 |
| I-product | 0 | 77 | 0 | 0 | 0 | 0 | 77 |
| I-group | 0 | 66 | 0 | 0 | 0 | 0 | 66 |
| B-product | 0 | 121 | 0 | 0 | 0 | 0 | 121 |
| B-creative-work | 0 | 231 | 0 | 0 | 0 | 0 | 231 |
| I-corporation | 0 | 42 | 0 | 0 | 0 | 0 | 42 |
| macro/weighted F1 (removing O) | | | 0/0 | | | | |

Table 10. Results of the word window classifier with hidden layer size = 128.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | TP | FN | FP | Precision | Recall | F1 | Support |
| O | 21934 | 0 | 1817 | 0.9235 | 1 | 0.9602 | 21934 |
| B-person | 0 | 411 | 0 | 0 | 0 | 0 | 411 |
| B-location | 0 | 143 | 0 | 0 | 0 | 0 | 143 |
| I-person | 0 | 121 | 0 | 0 | 0 | 0 | 121 |
| B-group | 0 | 175 | 0 | 0 | 0 | 0 | 175 |
| I-location | 0 | 43 | 0 | 0 | 0 | 0 | 43 |
| B-corporation | 0 | 150 | 0 | 0 | 0 | 0 | 150 |
| I-creative-work | 0 | 237 | 0 | 0 | 0 | 0 | 237 |
| I-product | 0 | 77 | 0 | 0 | 0 | 0 | 77 |
| I-group | 0 | 66 | 0 | 0 | 0 | 0 | 66 |
| B-product | 0 | 121 | 0 | 0 | 0 | 0 | 121 |
| B-creative-work | 0 | 231 | 0 | 0 | 0 | 0 | 231 |
| I-corporation | 0 | 42 | 0 | 0 | 0 | 0 | 42 |
| macro/weighted F1 (removing O) | | | 0/0 | | | | |

Table 11. Results of the word window classifier with hidden layer size = 256.

Number of Hidden Layers

|  |  |
| --- | --- |
| Window Size | 3 |
| Embedding Size | 256 |
| Hidden Layer Size (First, Second) | 256, 128 or 256, 64 |
| Number of Hidden Layers | 2 |
| Freeze Word Embeddings or Not | No |
| Learning Rate | 0.002 |
| Number of Epochs | 300 |

Table 12. Training configurations.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | TP | FN | FP | Precision | Recall | F1 | Support |
| O | 21934 | 0 | 1817 | 0.9235 | 1 | 0.9602 | 21934 |
| B-person | 0 | 411 | 0 | 0 | 0 | 0 | 411 |
| B-location | 0 | 143 | 0 | 0 | 0 | 0 | 143 |
| I-person | 0 | 121 | 0 | 0 | 0 | 0 | 121 |
| B-group | 0 | 175 | 0 | 0 | 0 | 0 | 175 |
| I-location | 0 | 43 | 0 | 0 | 0 | 0 | 43 |
| B-corporation | 0 | 150 | 0 | 0 | 0 | 0 | 150 |
| I-creative-work | 0 | 237 | 0 | 0 | 0 | 0 | 237 |
| I-product | 0 | 77 | 0 | 0 | 0 | 0 | 77 |
| I-group | 0 | 66 | 0 | 0 | 0 | 0 | 66 |
| B-product | 0 | 121 | 0 | 0 | 0 | 0 | 121 |
| B-creative-work | 0 | 231 | 0 | 0 | 0 | 0 | 231 |
| I-corporation | 0 | 42 | 0 | 0 | 0 | 0 | 42 |
| macro/weighted F1 (removing O) | | | 0/0 | | | | |

Table 13. Results of the word window classifier with hidden layer size = 256, 128.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | TP | FN | FP | Precision | Recall | F1 | Support |
| O | 21934 | 0 | 1817 | 0.9235 | 1 | 0.9602 | 21934 |
| B-person | 0 | 411 | 0 | 0 | 0 | 0 | 411 |
| B-location | 0 | 143 | 0 | 0 | 0 | 0 | 143 |
| I-person | 0 | 121 | 0 | 0 | 0 | 0 | 121 |
| B-group | 0 | 175 | 0 | 0 | 0 | 0 | 175 |
| I-location | 0 | 43 | 0 | 0 | 0 | 0 | 43 |
| B-corporation | 0 | 150 | 0 | 0 | 0 | 0 | 150 |
| I-creative-work | 0 | 237 | 0 | 0 | 0 | 0 | 237 |
| I-product | 0 | 77 | 0 | 0 | 0 | 0 | 77 |
| I-group | 0 | 66 | 0 | 0 | 0 | 0 | 66 |
| B-product | 0 | 121 | 0 | 0 | 0 | 0 | 121 |
| B-creative-work | 0 | 231 | 0 | 0 | 0 | 0 | 231 |
| I-corporation | 0 | 42 | 0 | 0 | 0 | 0 | 42 |
| macro/weighted F1 (removing O) | | | 0/0 | | | | |

Table 14. Results of the word window classifier with hidden layer size = 256, 64.

Freeze Word Embeddings or Not

|  |  |
| --- | --- |
| Window Size | 3 |
| Embedding Size | 256 |
| Hidden Layer Size | 256 |
| Number of Hidden Layers | 1 |
| Freeze Word Embeddings or Not | No or Yes |
| Learning Rate | 0.002 |
| Number of Epochs | 300 |

Table 15. Training configurations.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | TP | FN | FP | Precision | Recall | F1 | Support |
| O | 21934 | 0 | 1817 | 0.9235 | 1 | 0.9602 | 21934 |
| B-person | 0 | 411 | 0 | 0 | 0 | 0 | 411 |
| B-location | 0 | 143 | 0 | 0 | 0 | 0 | 143 |
| I-person | 0 | 121 | 0 | 0 | 0 | 0 | 121 |
| B-group | 0 | 175 | 0 | 0 | 0 | 0 | 175 |
| I-location | 0 | 43 | 0 | 0 | 0 | 0 | 43 |
| B-corporation | 0 | 150 | 0 | 0 | 0 | 0 | 150 |
| I-creative-work | 0 | 237 | 0 | 0 | 0 | 0 | 237 |
| I-product | 0 | 77 | 0 | 0 | 0 | 0 | 77 |
| I-group | 0 | 66 | 0 | 0 | 0 | 0 | 66 |
| B-product | 0 | 121 | 0 | 0 | 0 | 0 | 121 |
| B-creative-work | 0 | 231 | 0 | 0 | 0 | 0 | 231 |
| I-corporation | 0 | 42 | 0 | 0 | 0 | 0 | 42 |
| macro/weighted F1 (removing O) | | | 0/0 | | | | |

Table 16. Results of the word window classifier with word embeddings not freezed.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | TP | FN | FP | Precision | Recall | F1 | Support |
| O | 21934 | 0 | 1817 | 0.9235 | 1 | 0.9602 | 21934 |
| B-person | 0 | 411 | 0 | 0 | 0 | 0 | 411 |
| B-location | 0 | 143 | 0 | 0 | 0 | 0 | 143 |
| I-person | 0 | 121 | 0 | 0 | 0 | 0 | 121 |
| B-group | 0 | 175 | 0 | 0 | 0 | 0 | 175 |
| I-location | 0 | 43 | 0 | 0 | 0 | 0 | 43 |
| B-corporation | 0 | 150 | 0 | 0 | 0 | 0 | 150 |
| I-creative-work | 0 | 237 | 0 | 0 | 0 | 0 | 237 |
| I-product | 0 | 77 | 0 | 0 | 0 | 0 | 77 |
| I-group | 0 | 66 | 0 | 0 | 0 | 0 | 66 |
| B-product | 0 | 121 | 0 | 0 | 0 | 0 | 121 |
| B-creative-work | 0 | 231 | 0 | 0 | 0 | 0 | 231 |
| I-corporation | 0 | 42 | 0 | 0 | 0 | 0 | 42 |
| macro/weighted F1 (removing O) | | | 0/0 | | | | |

Table 17. Results of the word window classifier with word embeddings freezed.

Learning Rate

|  |  |
| --- | --- |
| Window Size | 3 |
| Embedding Size | 256 |
| Hidden Layer Size | 256 |
| Number of Hidden Layers | 1 |
| Freeze Word Embeddings or Not | No |
| Learning Rate | 0.001 or 0.0005 |
| Number of Epochs | 300 |

Table 18. Training configurations.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | TP | FN | FP | Precision | Recall | F1 | Support |
| O | 21934 | 0 | 1817 | 0.9235 | 1 | 0.9602 | 21934 |
| B-person | 0 | 411 | 0 | 0 | 0 | 0 | 411 |
| B-location | 0 | 143 | 0 | 0 | 0 | 0 | 143 |
| I-person | 0 | 121 | 0 | 0 | 0 | 0 | 121 |
| B-group | 0 | 175 | 0 | 0 | 0 | 0 | 175 |
| I-location | 0 | 43 | 0 | 0 | 0 | 0 | 43 |
| B-corporation | 0 | 150 | 0 | 0 | 0 | 0 | 150 |
| I-creative-work | 0 | 237 | 0 | 0 | 0 | 0 | 237 |
| I-product | 0 | 77 | 0 | 0 | 0 | 0 | 77 |
| I-group | 0 | 66 | 0 | 0 | 0 | 0 | 66 |
| B-product | 0 | 121 | 0 | 0 | 0 | 0 | 121 |
| B-creative-work | 0 | 231 | 0 | 0 | 0 | 0 | 231 |
| I-corporation | 0 | 42 | 0 | 0 | 0 | 0 | 42 |
| macro/weighted F1 (removing O) | | | 0/0 | | | | |

Table 19. Results of the word window classifier with learning rate = 0.001.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | TP | FN | FP | Precision | Recall | F1 | Support |
| O | 20986 | 948 | 1561 | 0.9308 | 0.9568 | 0.9436 | 21934 |
| B-person | 2 | 409 | 18 | 0.1 | 0.0049 | 0.0093 | 411 |
| B-location | 1 | 142 | 54 | 0.0182 | 0.007 | 0.0101 | 143 |
| I-person | 0 | 121 | 8 | 0 | 0 | 0 | 121 |
| B-group | 6 | 169 | 454 | 0.013 | 0.0343 | 0.0189 | 175 |
| I-location | 1 | 42 | 16 | 0.0588 | 0.0233 | 0.0333 | 43 |
| B-corporation | 1 | 149 | 334 | 0.003 | 0.0067 | 0.0041 | 150 |
| I-creative-work | 0 | 237 | 3 | 0 | 0 | 0 | 237 |
| I-product | 1 | 76 | 25 | 0.0385 | 0.013 | 0.0194 | 77 |
| I-group | 4 | 62 | 211 | 0.0186 | 0.0606 | 0.0285 | 66 |
| B-product | 1 | 120 | 33 | 0.0294 | 0.0083 | 0.0129 | 121 |
| B-creative-work | 0 | 231 | 26 | 0 | 0 | 0 | 231 |
| I-corporation | 0 | 42 | 5 | 0 | 0 | 0 | 42 |
| macro/weighted F1 (removing O) | | | 0.0114/ 0.0086 | | | | |

Table 20. Results of the word window classifier with hidden layer size = 0.0005.

Number of Epochs

|  |  |
| --- | --- |
| Window Size | 3 |
| Embedding Size | 256 |
| Hidden Layer Size | 256 |
| Number of Hidden Layers | 1 |
| Freeze Word Embeddings or Not | No |
| Learning Rate | 0.0005 |
| Number of Epochs | 200 or 100 |

Table 21. Training configurations.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | TP | FN | FP | Precision | Recall | F1 | Support |
| O | 21858 | 76 | 1815 | 0.9233 | 0.9965 | 0.9585 | 21934 |
| B-person | 0 | 411 | 12 | 0 | 0 | 0 | 411 |
| B-location | 0 | 143 | 11 | 0 | 0 | 0 | 143 |
| I-person | 0 | 121 | 9 | 0 | 0 | 0 | 121 |
| B-group | 0 | 175 | 4 | 0 | 0 | 0 | 175 |
| I-location | 0 | 43 | 6 | 0 | 0 | 0 | 43 |
| B-corporation | 0 | 150 | 4 | 0 | 0 | 0 | 150 |
| I-creative-work | 0 | 237 | 14 | 0 | 0 | 0 | 237 |
| I-product | 0 | 77 | 2 | 0 | 0 | 0 | 77 |
| I-group | 0 | 66 | 2 | 0 | 0 | 0 | 66 |
| B-product | 0 | 121 | 3 | 0 | 0 | 0 | 121 |
| B-creative-work | 0 | 231 | 7 | 0 | 0 | 0 | 231 |
| I-corporation | 0 | 42 | 4 | 0 | 0 | 0 | 42 |
| macro/weighted F1 (removing O) | | | 0/0 | | | | |

Table 22. Results of the word window classifier with number of epochs = 200.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | TP | FN | FP | Precision | Recall | F1 | Support |
| O | 21183 | 751 | 1679 | 0.9266 | 0.9658 | 0.9458 | 21934 |
| B-person | 9 | 402 | 106 | 0.0783 | 0.0219 | 0.0342 | 411 |
| B-location | 1 | 142 | 44 | 0.0222 | 0.007 | 0.0106 | 143 |
| I-person | 0 | 121 | 50 | 0 | 0 | 0 | 121 |
| B-group | 0 | 175 | 20 | 0 | 0 | 0 | 175 |
| I-location | 0 | 43 | 49 | 0 | 0 | 0 | 43 |
| B-corporation | 0 | 150 | 136 | 0 | 0 | 0 | 150 |
| I-creative-work | 3 | 234 | 146 | 0.0201 | 0.0127 | 0.0155 | 237 |
| I-product | 0 | 77 | 23 | 0 | 0 | 0 | 77 |
| I-group | 0 | 66 | 133 | 0 | 0 | 0 | 66 |
| B-product | 5 | 116 | 65 | 0.0714 | 0.0413 | 0.0524 | 121 |
| B-creative-work | 0 | 231 | 58 | 0 | 0 | 0 | 231 |
| I-corporation | 0 | 42 | 41 | 0 | 0 | 0 | 42 |
| macro/weighted F1 (removing O) | | | 0.0094/0.0141 | | | | |

Table 23. Results of the word window classifier with number of epochs = 100.