|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Model** | **Epochs** | **Count of input sentences** | **Test Accuracy** | **Best Validation Accuracy** |
| Elmo + BiLSTM + Max-pooling +  2 layer fully-connected  (Ilić, Suzana, et al. , 2018) | 700 | 12 | 0.866 | 0.875  (epoch: 379) |
| Elmo + BiLSTM +  fully-connected layer  (Our modification) | 700 | 12 | 0.860 | 0.869  (epoch : 96) |
| Elmo + dropout + LSTM + dropout +  fully-connected layer  (Николенко, С. И, Кадурин А. А., Архангельская Е. О. *Глубокое обучение*. ,2017.) | 700 | 12 | 0.881 | 0.885  (epoch: 93) |
| Elmo + dropout + BiLSTM + dropout + fully-connected layer  (Our modification) | 700 | 12 | **0.890** | 0.890  (epoch: 554) |
| Elmo + conv1d + Max-pooling1d + LSTM + fully-connected layer  (Николенко, С. И, Кадурин А. А., Архангельская Е. О. *Глубокое обучение*. ,2017.) | 700 | 12 | 0.850 | 0.849 (epoch: 502) |
| **Biattentive Classification Network (BCN)**  2 ReLU feed forward networks +  2 BiLSTM Encoders +  Biattention mechanism (Seo et al. 2017, Xiong et al., 2017) +  2 one-layer BiLSTMs +  2 Pooling layers (Max - pooling, Mean - pooling, Min – pooling, Self – attentive pooling) +  3 – layer batch-normalized Maxout Network (Goodfellow et al., 2013)  (McCann, Bryan, et al., 2017) | 700 | 12 | 0.879 | 0.867  (epoch: 691) |
| BCN modification  2 ReLu feed forward networks +  2 BiLSTM Encoders +  Biattention mechanism (Seo et al. 2017, Xiong et al., 2017) +  2 one-layer BiLSTMs +  2 Pooling layers (Max - pooling, Mean - pooling, Min – pooling, Self – attentive pooling) +  2 – layer ReLU feed forward network  (Peters, Matthew E., et al., 2018) | 700 | 12 | 0.866 | 0.860  (epoch: 554) |

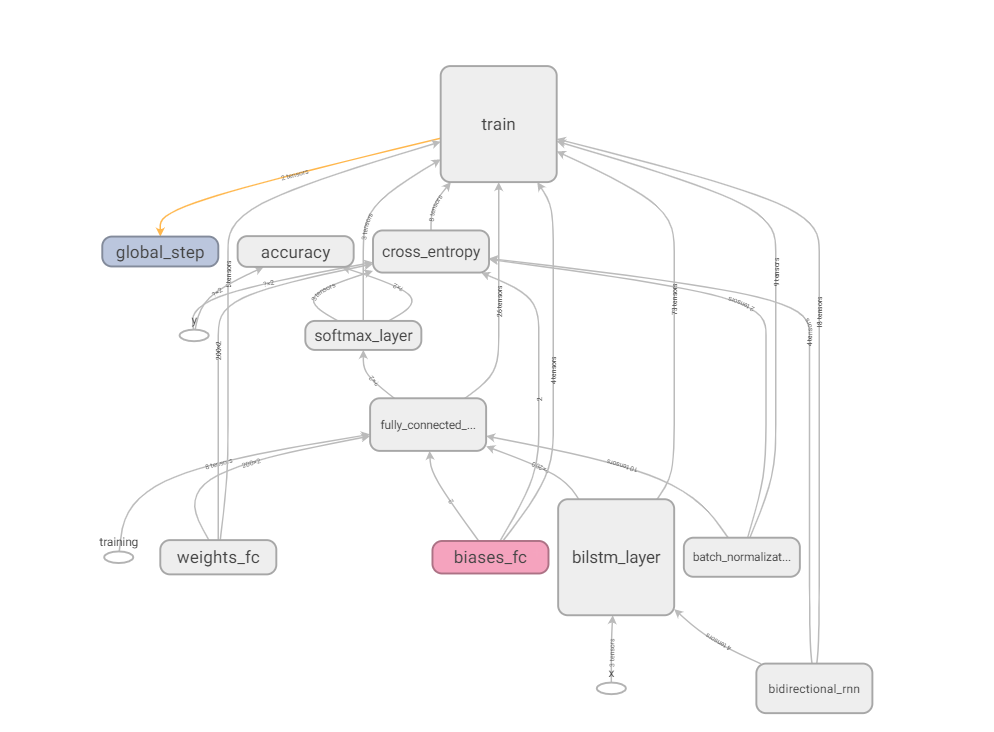
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Model** | **Epochs** | **Count of input sentences** | **Test Accuracy** | **Best Validation Accuracy** |
| Elmo + dropout + BiLSTM + dropout + fully-connected layer  *(with regularization and batch normalization and exponential decay)*  (Our modification) | 700 | 20 | 0.887 | 0.893 (epoch: 210) |
| Elmo + dropout + BiLSTM + dropout + fully-connected layer  (Our modification) | 700 | 60 | 0.883 | 0.890 (epoch: 628) |
| Elmo + dropout + BiLSTM + dropout + fully-connected layer  *(with regularization, batch normalization and exponential decay)*  (Our modification) | 700 | 60 | **0.892** | 0.890 (epoch: 22) |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Model** | **Epochs** | **Count of input sentences** | **Test Accuracy** | **Best Validation Accuracy** |
| Elmo + dropout + BiLSTM + dropout + fully-connected layer  **+** **Task-specific weighting scheme** (Peters, Matthew E., et al. (2018))  *(with regularization, batch normalization and exponential decay)*  (Our modification) | 2 | 4 | 0.758 | 0.770  (epoch: 2) |

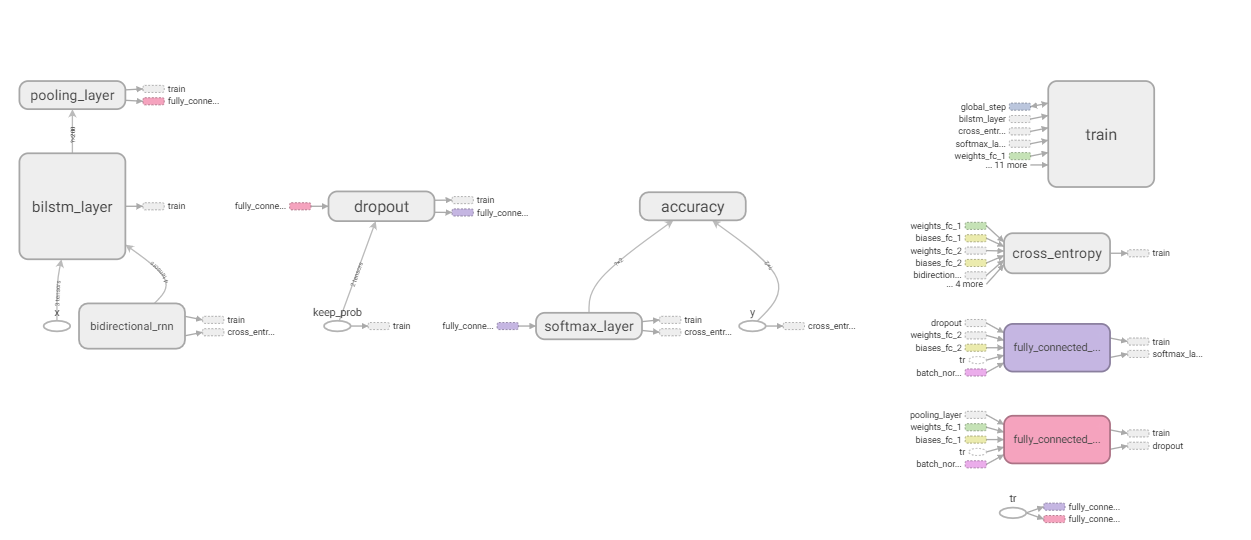
*\*all models are* *with regularization, batch normalization and exponential decay*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Model** | **Epochs** | **Count of input sentences** | **Test Accuracy** | **Best Validation Accuracy** |
| Elmo + BiLSTM + Max-pooling +  2 layer fully-connected  (Ilić, Suzana, et al. , 2018) | 700 | Not fixed | 0.883 | 0.868  (epoch: 700) |
| Elmo + BiLSTM +  fully-connected layer  (Our modification) | 700 | Not fixed | 0.883 | 0.869  (epoch: 696) |
| Elmo + dropout + BiLSTM + dropout + fully-connected layer  ( + Xavier init)  (Our modification) | 700 | Not fixed | **0.902** | 0.891  (epoch: 700) |

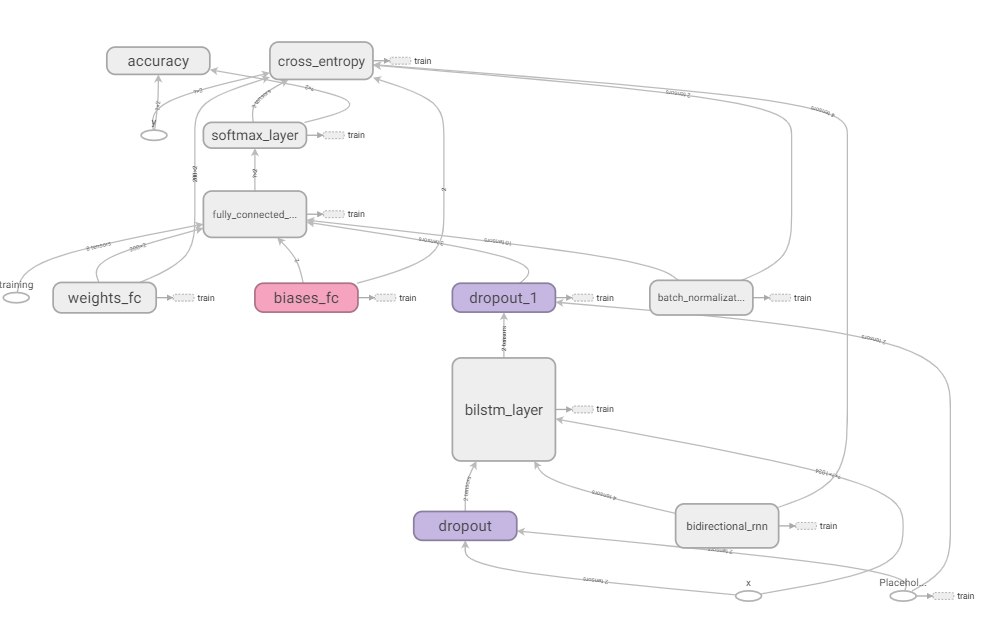
1. Elmo + BiLSTM + fully-connected layer (Our modification)



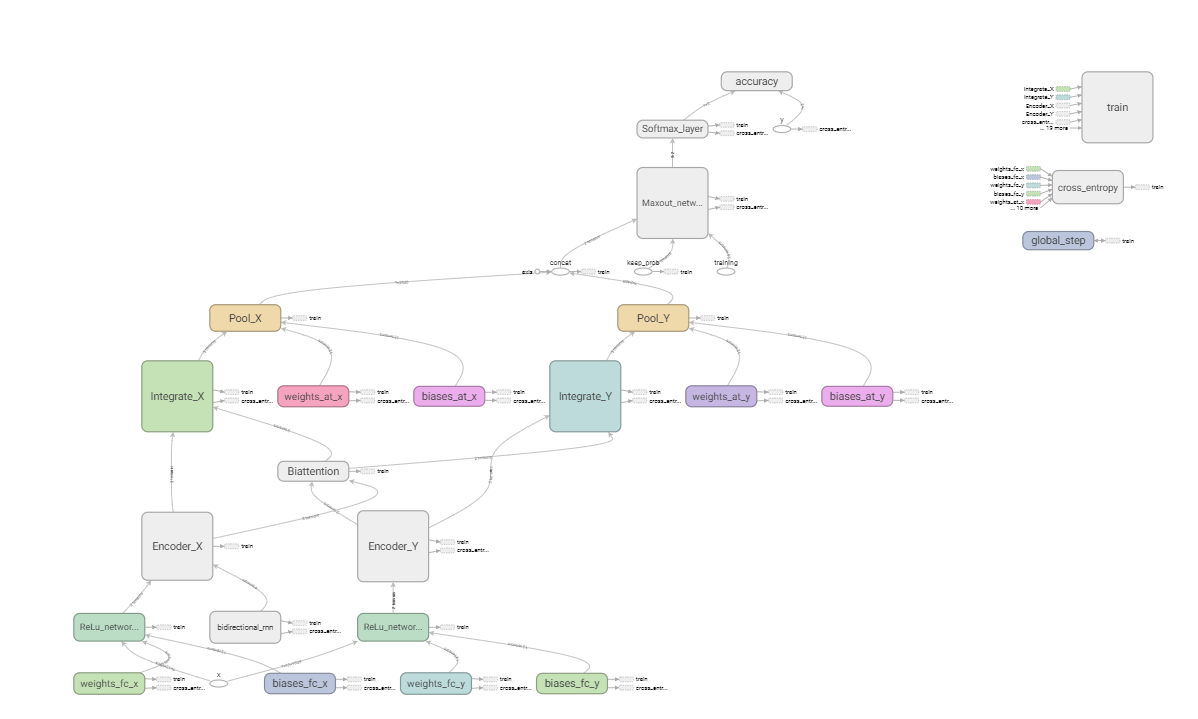
1. Elmo + BiLSTM + Max-pooling + 2 layer fully-connected (Ilić, Suzana, et al. , 2018)



1. Elmo + dropout + BiLSTM + dropout + fully-connected layer (Our modification)



1. BCN (with 3-layer Maxout network)



1. BCN (with 2-layer ReLU network)
2. w/ training scheme
3. Elmo + dropout + LSTM + dropout + fully-connected layer

(Николенко, С. И, Кадурин А. А., Архангельская Е. О. *Глубокое обучение*. ,2017.)

1. Elmo + conv1d + Max-pooling1d + LSTM + fully-connected layer

(Николенко, С. И, Кадурин А. А., Архангельская Е. О. *Глубокое обучение*. ,2017.)