

Event Extraction via Deep Semantic CNN





Event Extraction

- Example:

Obama **beats** McCain -> **election**

Tyson **beats** his opponent -> **attack**

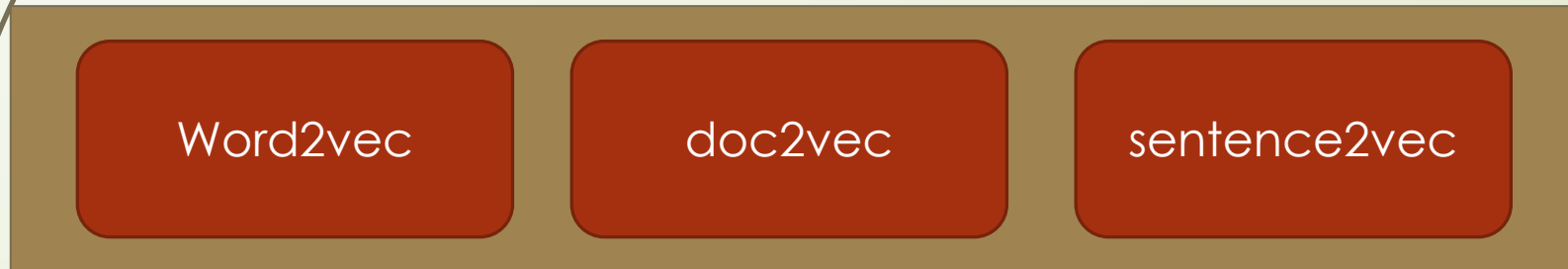
A cameraman **died** when a man **fired** a gun on the hotel in Baghdad -> **die** & **attack**

- Dataset – ACE 2005 corpus
- Examples from Chen et al

Initial Approach

- Each word is a training example represented by word2vec and doc2vec
- Predictions made with Feed-Forward MLP

Obama **beats** McCain



RNN

- Predict sequences of labels over sentences with LSTM
- Problem: Huge class imbalance ~97% no event
- Keras won't allow for class weights on sequences

A cameraman **died** when a man **fired** a gun ...

↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓

Nil Nil Die Nil Nil Nil Attack Nil Nil

Final Approach

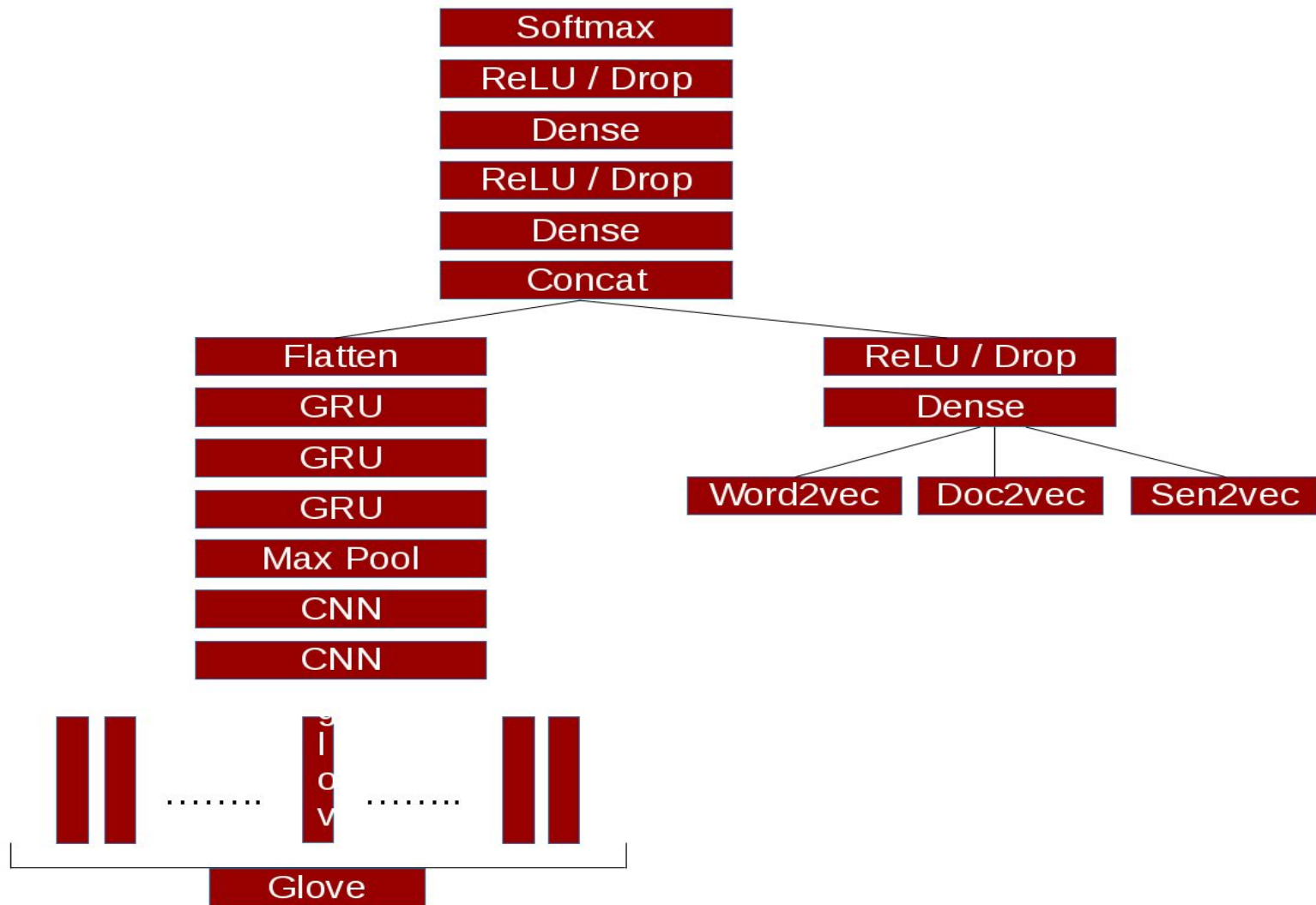
- Make a fixed size window around each candidate word
- The window has one label
- Fit RNN & CNN on each window
- Also include the initial model

... A cameraman **died** when a ...

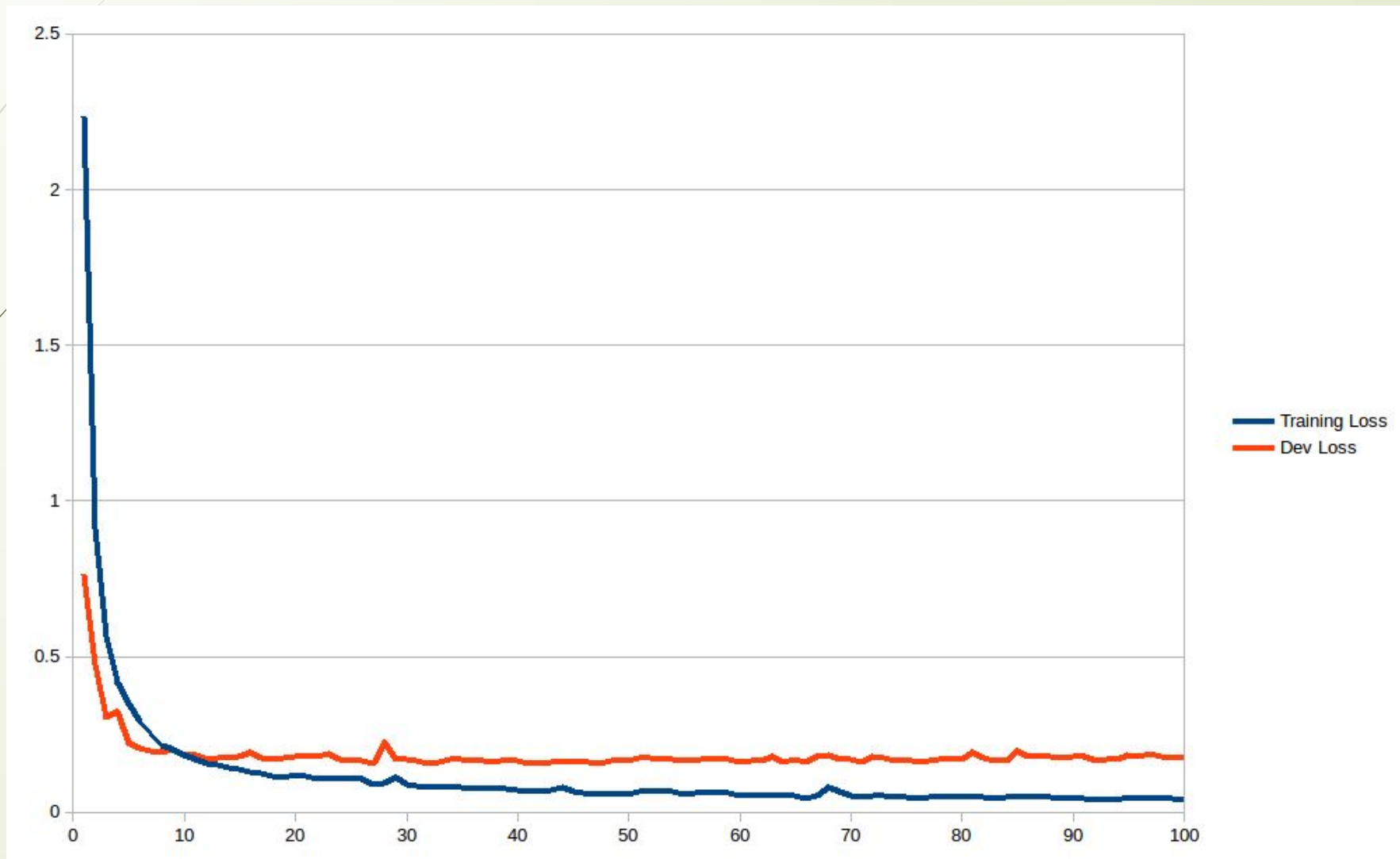


Attack

Final Approach



Observation

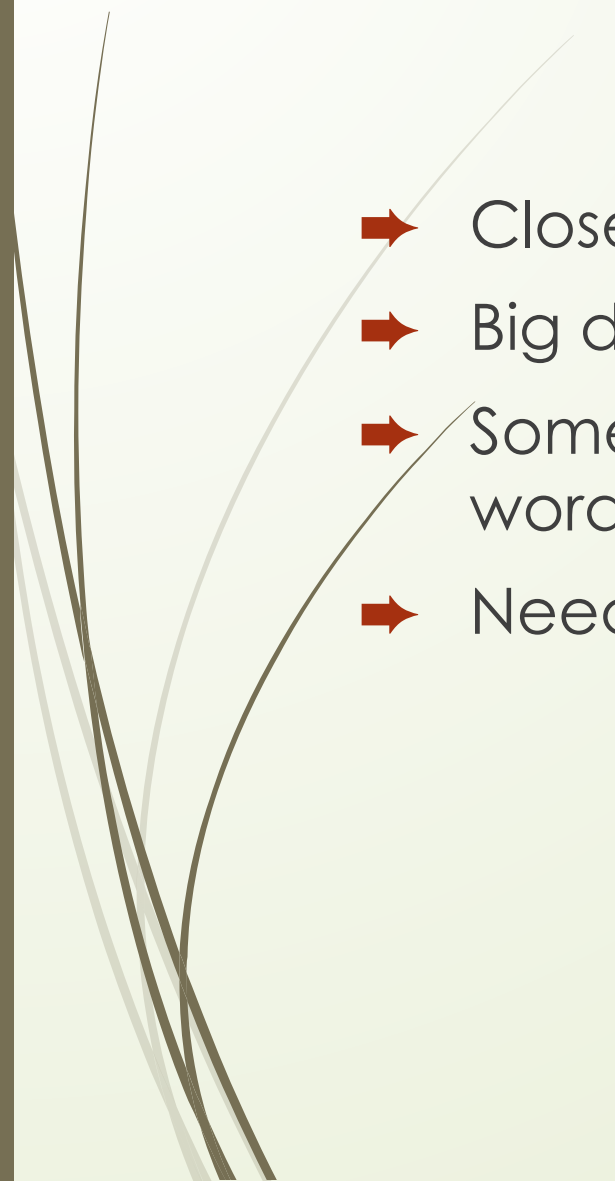


Comparison

Models	Approach	dev F1	test F1
Event extraction via Dynamic Multi-Pooling CNN	CNN + dynamic Multiple pooling	-	69.1
Joint event extraction via Structured Prediction With Global Features	joint framework - structured perceptron with beam search	67.9	67.5
Our model	Expanded multilayer perceptron - multiple CNN + RNN	71.7	65.3
Initial model	Multilayer perceptron	70.3	63.1

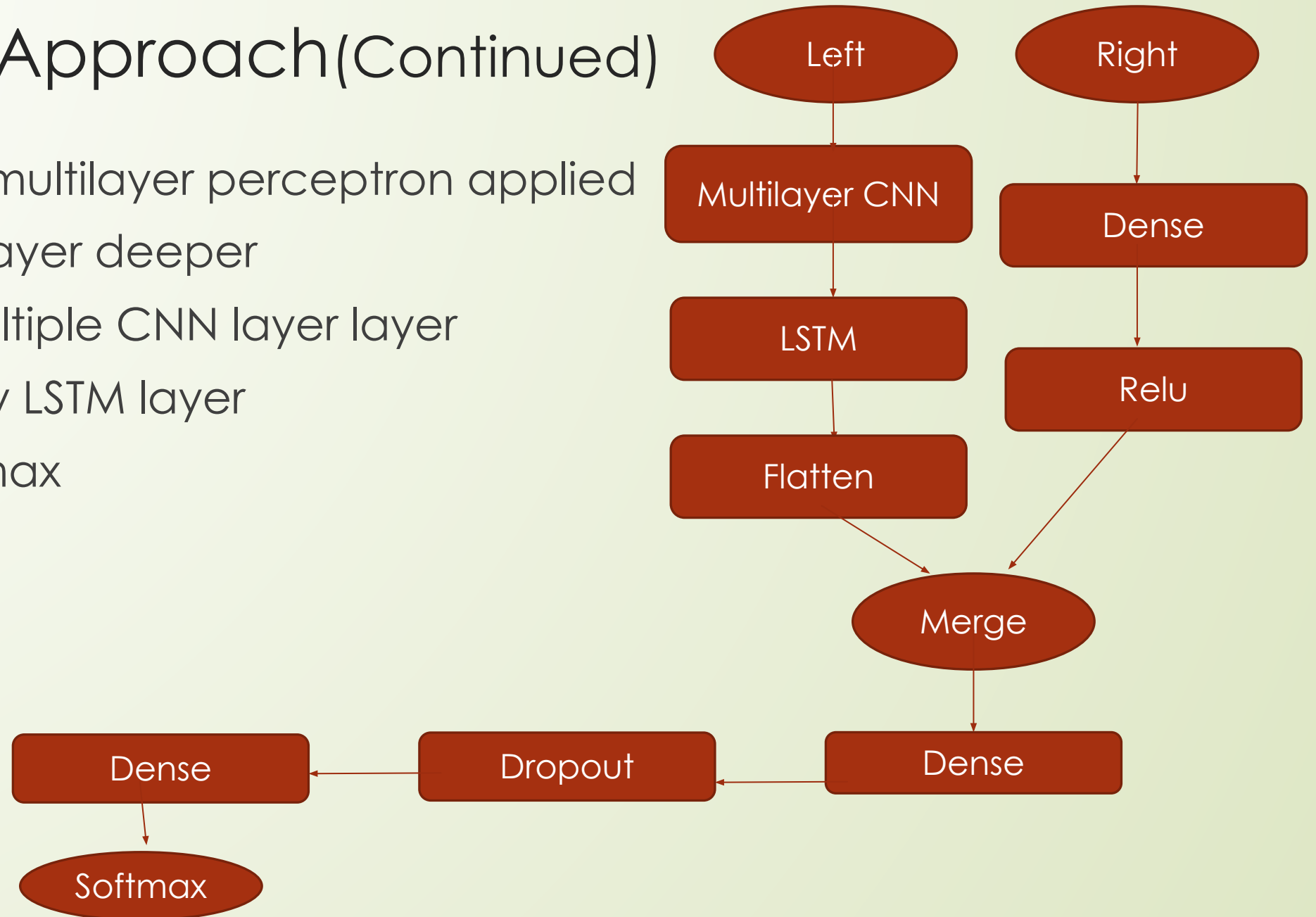


Conclusion

- ➔ Close to state of the art
 - ➔ Big drop going from dev to test set
 - ➔ Some preprocessing can be improved, hyphenated words have no embedding, e.g war-stricken
 - ➔ Need to select best model based on F1
- 

Final Approach(Continued)

- Expanded multilayer perceptron applied
- Made the layer deeper
- Applied multiple CNN layer layer
- Followed by LSTM layer
- Finally softmax



Observation

Epoch	Train loss	Val loss	Train Accuracy	Test Accuracy
1	0.303037	0.295576	0.969469	0.957925
2	0.221091	0.28522	0.970812	0.957925
3	0.215771	0.281765	0.970812	0.957925
4	0.213559	0.279003	0.970812	0.957925
5	0.211956	0.278182	0.970812	0.957925
6	0.210918	0.275801	0.970812	0.957925
7	0.209373	0.275282	0.970812	0.957925
8	0.208322	0.274235	0.970812	0.957925
9	0.206721	0.272353	0.970812	0.957925
10	0.205771	0.270906	0.970812	0.957925

Intermediate Approach

- Model has been divided in two subparts
- Left subpart consist of window around each candid event
- Right subpart consist of doc2vec applied on candid sentence token and documents
- Merging both the subparts
- Applied multilayer perceptron – dense with relu

