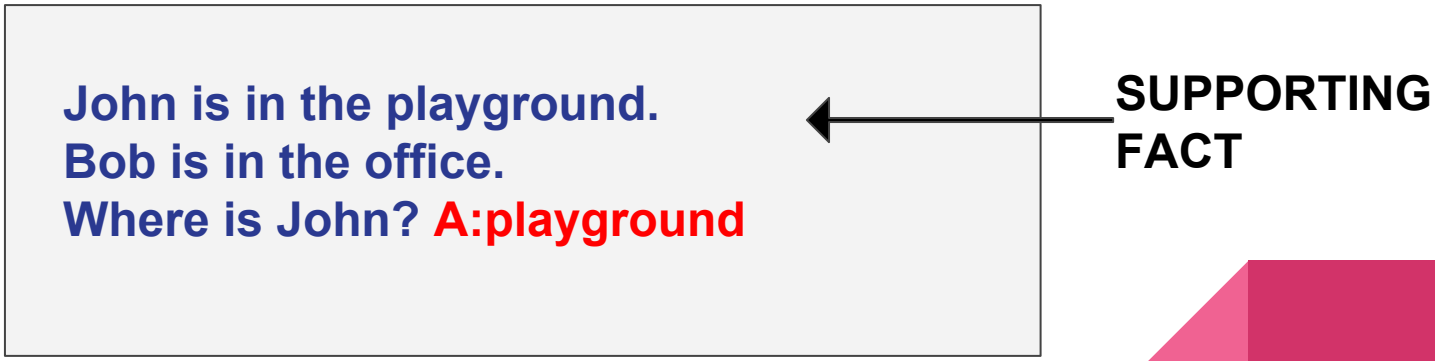
Question Answering System using End to End Memory Networks

Prashil Tumbade

Problem Statement

- Question Answering Task
- Facebook Babi Dataset
- Example: Task with single supporting fact

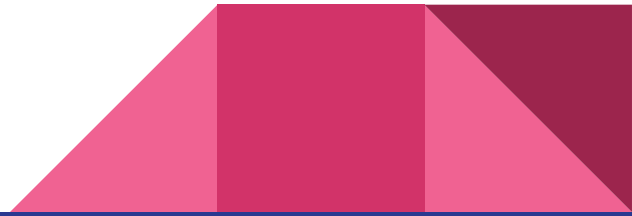


Memory Networks

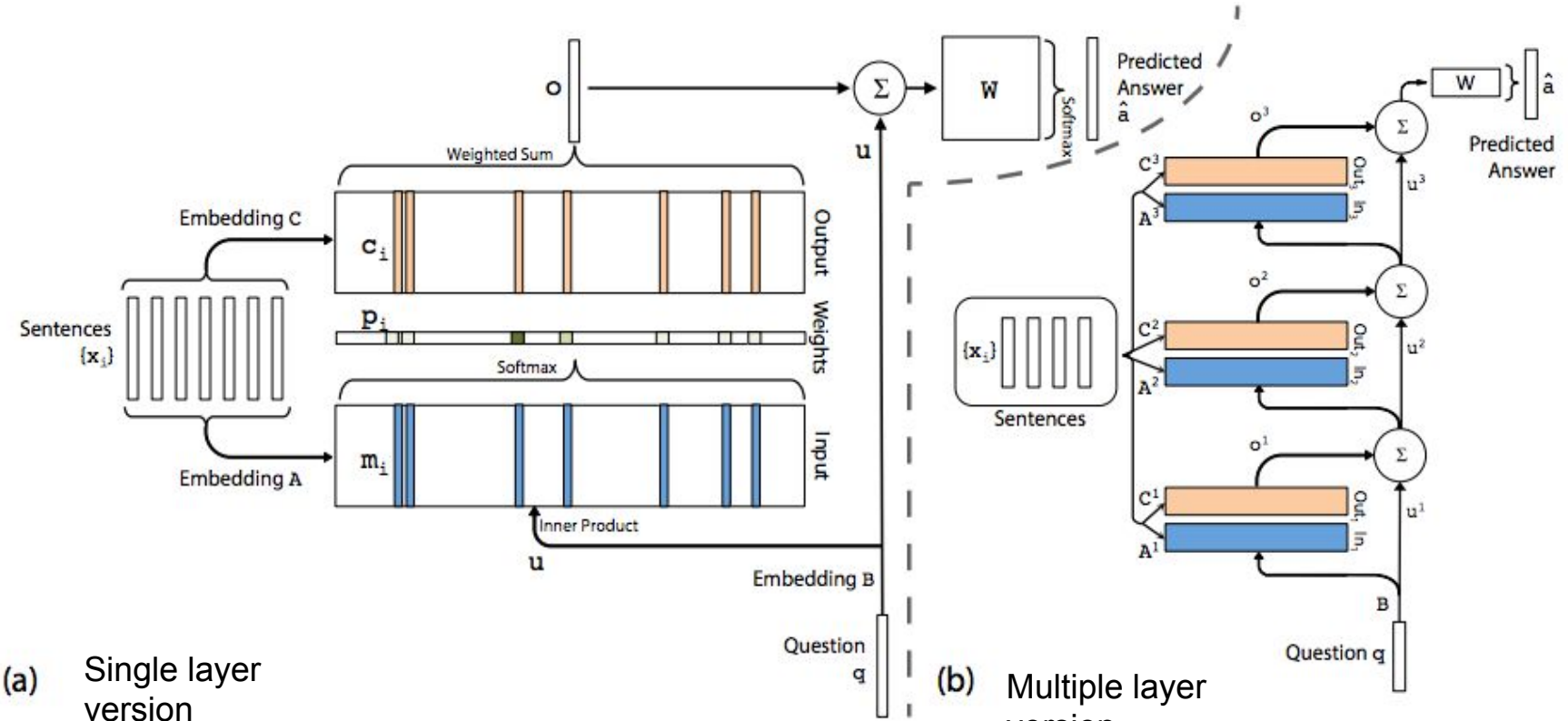
- Memory Networks were introduced in 2014 as a work to be presented in International Conference on Learning Representations (ICLR) 2016.
- Class of models that combine large memory with learning component that can read and write to it.
- Authors of the paper : J. Weston, S. Chopra and A. Bordes
- Simple Concept:
  - Store the text in memory
  - Transform them by using embedding matrices.
  - Now do computations over them to train the model.

End to End Memory Networks

- Neural Network model with external memory.
- Reads the memory with soft attention.
- It accesses memory multiple times, each step being called a hop.
- Uses back propagation to update the model.



Architecture of End to End Memory Networks



Model Details

- Input Sentences: $x_1, x_2, \ldots, x_n$ is taken
- Sentences are embedded into memory vectors m_i and c_i by using embedding matrices A and C.
- Question q is embedded into internal state u.
- Matching is performed between u and m_i with Softmax function.
- Output is calculated by the relation: $o = \sum p_i c_i$
- Another Softmax is used to produce final predictions after summing up o with u.

Results/Observations

Layers	Dropouts	Batch-size	Epochs	Results
LSTM(32)	0.3	32	100	94.6%
LSTM(64)	0.3	32	100	96.5%
LSTM(32), LSTM(32)	(0.5,0.5)	32	100	92.4%
LSTM(32), LSTM(32)	(0.5, 0.5)	32	200	96.9%
GRU(32)	0.3	32	100	86.4%
GRU(64)	0.3	32	100	87.4%
GRU(32), GRU(32)	(0.5,0.5)	64	100	52.6%

Conclusion

- The models with two or more layers required more training since there are more parameters that need to be set, but then have greater accuracies than the other models once trained completely.
- Overall, LSTM based models performed better than GRU based models for this task.

DEMO

Thank you