ATCS Practical

Natural Language Inference

Lab session 1

TAs: Phillip Lippe, Verna Dankers

March 31, 2020

Premise: "Bob is in his PJs and in spite of the thunder and lightning outside, he is in dreamland."

Premise: "Bob is in his PJs and in spite of the thunder and lightning outside, he is in dreamland."

Hypothesis: "Bob is asleep."

Premise: "Bob is in his PJs and in spite of the thunder and lightning outside, he is in dreamland."

Hypothesis: "Bob is asleep."

Premise: "Bob is in his PJs and in spite of the thunder and lightning outside, he is in dreamland."

Hypothesis: "Bob lives with his parents."

Premise: "Bob is in his PJs and in spite of the thunder and lightning outside, he is in dreamland."

Hypothesis: "Bob lives with his parents."

?

Premise: "Bob is in his PJs and in spite of the thunder and lightning outside, he is in dreamland."

Hypothesis: "Bob is wearing a tuxedo."

Premise: "Bob is in his PJs and in spite of the thunder and lightning outside, he is in dreamland."

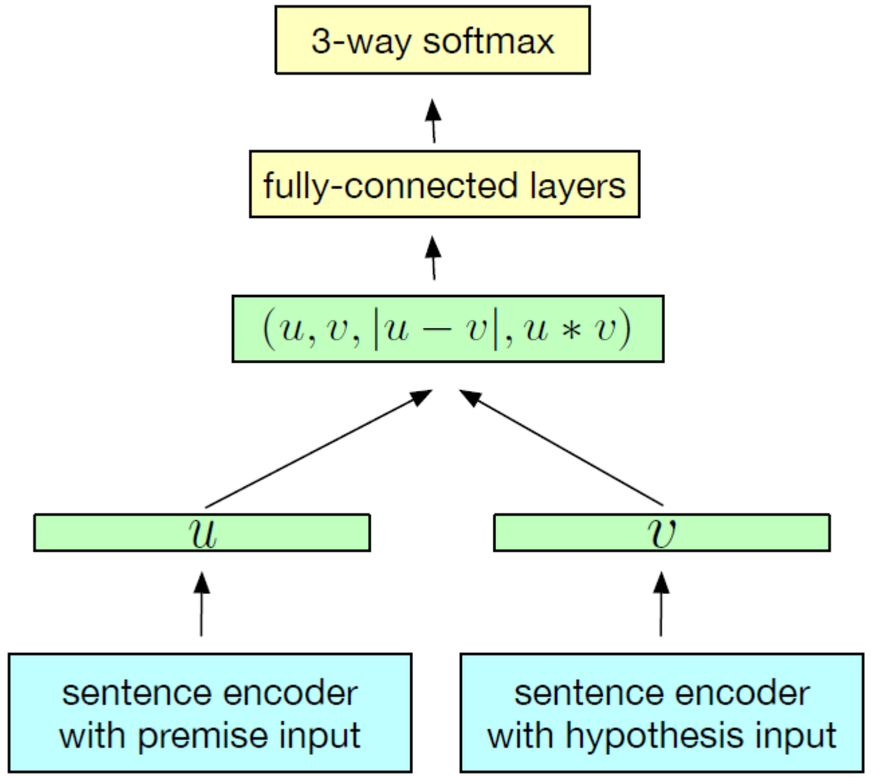
Hypothesis: "Bob is wearing a tuxedo."

Assignment: SNLI Corpus of Bowman et al. (2015)

- Data Stanford Natural Language Inference (SNLI) Corpus;
- Size 570k sentence pairs;
- Labels entailment, contradiction, neutral.

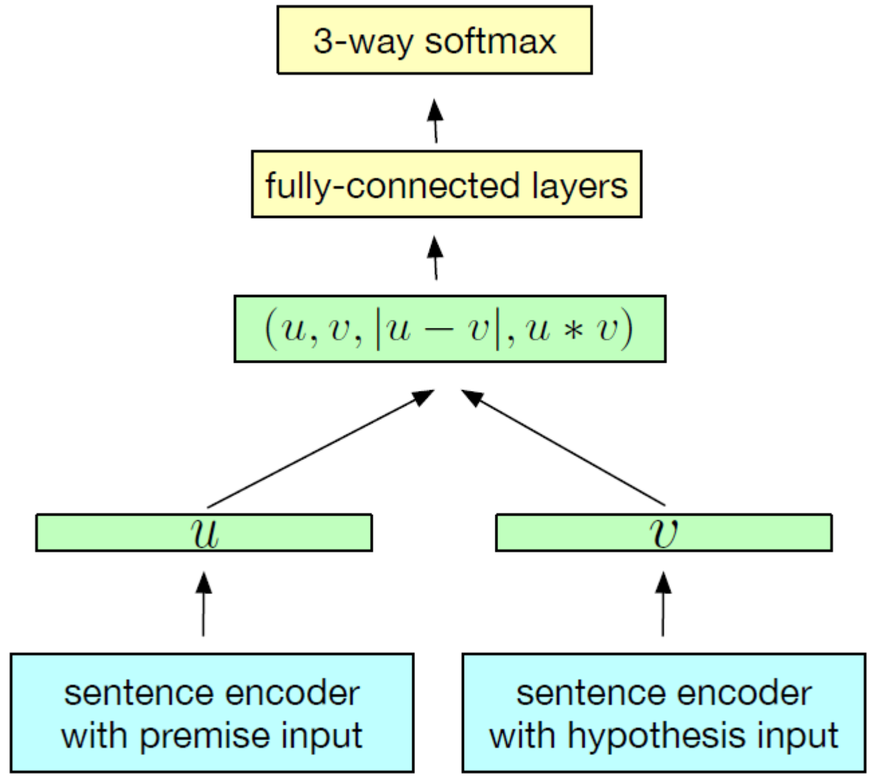
Assignment: Model of Conneau et al. (2017)

1. Embed the words of P and H with GloVe word embeddings;
2. Encode P and H with the same encoder and pool words;
3. Classify with MLP.



Assignment: Model of Conneau et al. (2017)

1. Embed the words of P and H with GloVe word embeddings;
2. Encode P and H with the same encoder and pool words;
    a. Average word embeddings;
    b. Uni-LSTM, use last hidden state;
    c. Bi-LSTM, use first and last hidden state;
    d. Bi-LSTM, use max pooling over words.
3. Classify with MLP.



Assignment: Evaluation

- Regular testing using the SNLI test set (Bowman et al., 2015);
- Transfer testing using the SentEval library (Conneau & Kiela, 2018).

Assignment: Practicalities

- Read the papers before starting;
- Implement in PyTorch, use Torchtext for preprocessing SNLI and GloVe;
- Use a Tensorboard;
- Follow a tutorial for using SentEval;
- Use Lisa to train!

Deliverables

- Code Python files for training and evaluation;
- Documentation A ReadMe describing code with instructions for running;
- Pretrained models The final checkpoint for each model;
- Demo and error analysis A Jupyter notebook containing
  - example inferences,
  - a result overview (SNLI & SentEval),
  - error analysis.

Deadline is Friday, April 17, midnight.

Grading

- In-person evaluation through Zoom screen-sharing:
  - You demonstrate your results and analysis in the notebook;
  - We shortly inspect your code.
- Scheduled to take place Tuesday, April 21.

It's Q&A time: submit questions through the chat.

References

- S. R. Bowman, G. Angeli, C. Potts, and C. D. Manning. A large annotated corpus for learning natural language inference. arXiv preprint arXiv:1508.05326, 2015.

- A. Conneau and D. Kiela. Senteval: An evaluation toolkit for universal sentence representations. In Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC-2018), 2018.

- A. Conneau, D. Kiela, H. Schwenk, L. Barrault, and A. Bordes. Supervised learning of universal sentence representations from natural language inference data. In Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, pages 670-680, Copenhagen, Denmark, September 2017. Association for Computational Linguistics.