APS360 Fundamentals of AI

Lisa Zhang

Lecture 12; March 7, 2019

Agenda

- More Midterm Misconceptions
- Text Generation

Midterm

- Midterm Average: 63%
- With Adjustment: 68%
  - Adjustment: $+4/80$, because of Q5, Q7
  - I will make the adjustment when computing your final grade, not on Quercus.

Remark Request

- Submit your midterm with a cover page explaining the exact issues.
- Your entire midterm will be remarked, so your mark can change in any direction, or stay the same.
- Exception: Q5(b) which was mis-graded, or any counting issue.

Very few remarks on short answer questions were successful. A solution that the grader can't understand is not worth marks.

More Midterm Misconceptions

Supervised vs Unsupervised Learning Problems

- **Supervised learning problem**: building a model to predict a target value
  - **Regression** example: predict a person's age based on their headshot
  - **Classification** example: predict a person's gender based on their headshot
- **Unsupervised learning problem**: learning the structure of data
  - Examples: Clustering, dimension-reduction, learning an embedding, generating new data

Definitions

- parameters
- weights
- activations
- features

GloVe vs word2vec

- The word2vec model describes an **architecture** (like AlexNet)
- GloVe vectors are sets of trained word embeddings (like the AlexNet weights available in `torchvision`)

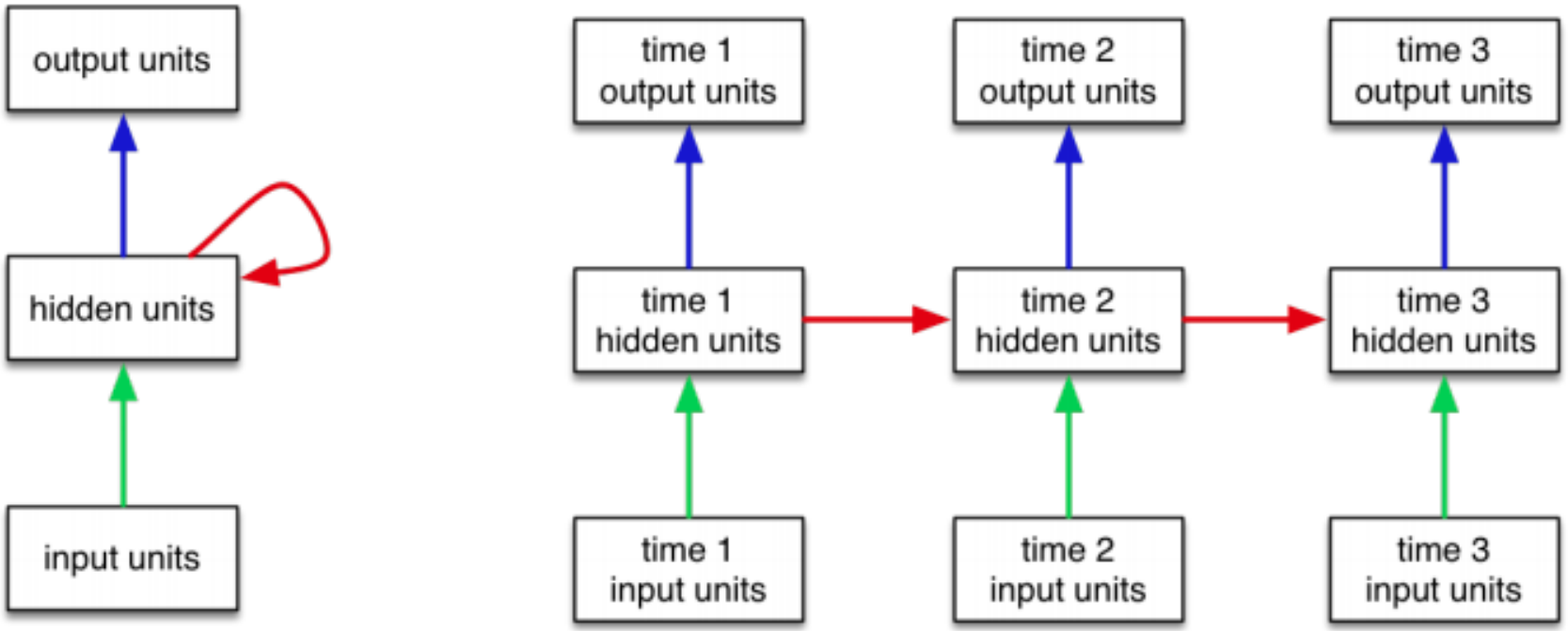
word2vec vs recurrent neural network

A word2vec model is **not** an RNN model

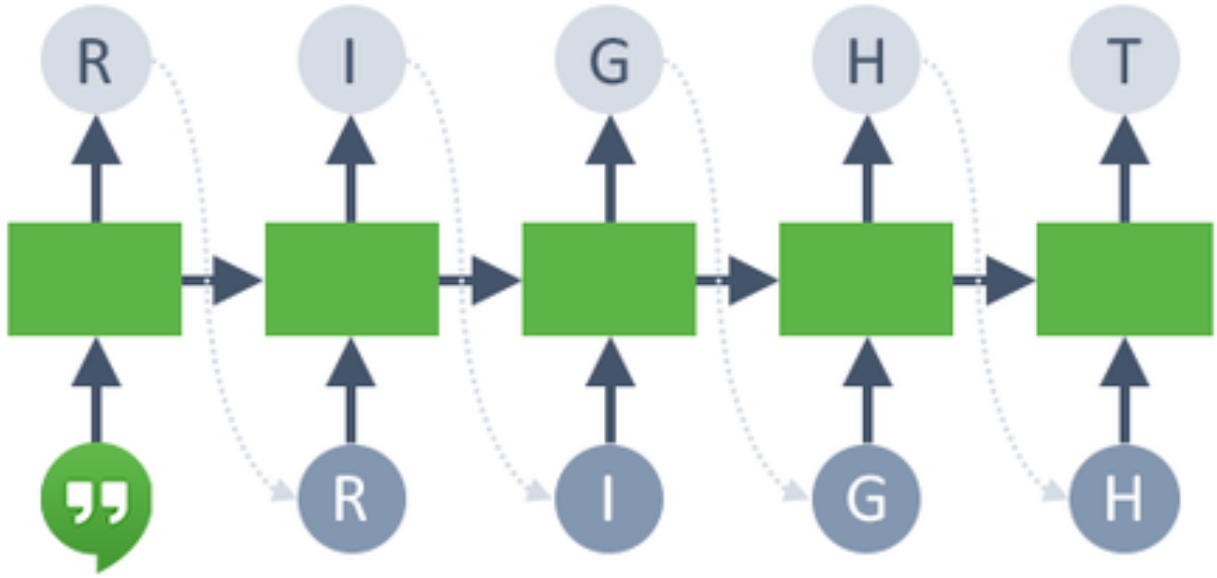
- word2vec model:
  - Fixed-sized input (one-hot encoding of a word)
  - Fixed-sized output (one-hot encodings of a fixed number of context words)
  - Goal:
- RNN model:
  - Can take a variable-sized input
  - Can generate a variable-sized output (not in week 6)

Text Generation with RNN

Recurrent Neural Networks



Text Generation



Jupyter notebook!

More complex models

- Attention
- Transformers

https://towardsdatascience.com/the-fall-of-rnn-lstm-2d1594c74ce0

https://blog.openai.com/better-language-models/