

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL
Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

Natural Language Processing

Deep Learning — Units 5 & 6

Dr. Jon Krohn

jon@untapt.com

November 18th, 2017

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL
Data

Preprocessing

ROC Curve

Sentiment
Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

Outline

- 1 Review Take-Home Exercise
- 2 The Power and Elegance of Deep Learning for NLP
 - Introduction to DL for NLP
 - Computational Representations of NL
 - NLP Applications
- 3 Word Vectors
 - Vector-Space Embedding
 - word2vec
 - Creating Word Vectors with word2vec
- 4 Modeling Natural Language Data
 - Best Practices for Preprocessing NLP Data
 - The Area Under the ROC Curve
 - Sentiment Classification
- 5 Recurrent Neural Networks
 - Simple RNNs
 - LSTMs
- 6 Parallel Network Architectures

- 1 Review Take-Home Exercise
- 2 The Power and Elegance of Deep Learning for NLP
 - Introduction to DL for NLP
 - Computational Representations of NL
 - NLP Applications
- 3 Word Vectors
 - Vector-Space Embedding
 - word2vec
 - Creating Word Vectors
- 4 Modeling Natural Language Data
 - Preprocessing
 - ROC Curve
 - Sentiment Classification
- 5 Recurrent Neural Networks
 - Simple RNNs
 - LSTMs
- 6 Parallel Network Architectures

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL

Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

- 1 Review Take-Home Exercise
- 2 The Power and Elegance of Deep Learning for NLP
 - Introduction to DL for NLP
 - Computational Representations of NL
 - NLP Applications
- 3 Word Vectors
 - Vector-Space Embedding
 - word2vec
 - Creating Word Vectors
- 4 Modeling Natural Language Data
 - Preprocessing
 - ROC Curve
 - Sentiment Classification
- 5 Recurrent Neural Networks
 - Simple RNNs
 - LSTMs
- 6 Parallel Network Architectures

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL

Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL
Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

- 1 Review Take-Home Exercise
- 2 The Power and Elegance of Deep Learning for NLP
 - Introduction to DL for NLP
 - Computational Representations of NL
 - NLP Applications
- 3 Word Vectors
 - Vector-Space Embedding
 - word2vec
 - Creating Word Vectors with word2vec
- 4 Modeling Natural Language Data
 - Best Practices for Preprocessing NLP Data
 - The Area Under the ROC Curve
 - Sentiment Classification
- 5 Recurrent Neural Networks
 - Simple RNNs
 - LSTMs
- 6 Parallel Network Architectures

Outline

- 1 Review Take-Home Exercise
- 2 The Power and Elegance of Deep Learning for NLP
 - Introduction to DL for NLP
 - Computational Representations of NL
 - NLP Applications
- 3 Word Vectors
 - Vector-Space Embedding
 - word2vec
 - Creating Word Vectors with word2vec
- 4 Modeling Natural Language Data
 - Best Practices for Preprocessing NLP Data
 - The Area Under the ROC Curve
 - Sentiment Classification
- 5 Recurrent Neural Networks
 - Simple RNNs
 - LSTMs
- 6 Parallel Network Architectures

Outline

- 1 Review Take-Home Exercise
- 2 The Power and Elegance of Deep Learning for NLP
 - Introduction to DL for NLP
 - Computational Representations of NL
 - NLP Applications
- 3 Word Vectors
 - Vector-Space Embedding
 - word2vec
 - Creating Word Vectors with word2vec
- 4 Modeling Natural Language Data
 - Best Practices for Preprocessing NLP Data
 - The Area Under the ROC Curve
 - Sentiment Classification
- 5 Recurrent Neural Networks
 - Simple RNNs
 - LSTMs
- 6 Parallel Network Architectures

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL

Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

Outline

- 1 Review Take-Home Exercise
- 2 The Power and Elegance of Deep Learning for NLP
 - Introduction to DL for NLP
 - Computational Representations of NL
 - NLP Applications
- 3 Word Vectors
 - Vector-Space Embedding
 - word2vec
 - Creating Word Vectors with word2vec
- 4 Modeling Natural Language Data
 - Best Practices for Preprocessing NLP Data
 - The Area Under the ROC Curve
 - Sentiment Classification
- 5 Recurrent Neural Networks
 - Simple RNNs
 - LSTMs
- 6 Parallel Network Architectures

Take-Home Exercise: VGGNet

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL

Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

Talk through the purpose of every line in the [VGGNet notebook], including all of the following terms:

- ReLU
- cross-entropy
- epoch
- parameters
- hyperparams
- SGD
- learning rate
- batch size
- Adam
- dropout
- batchnorm
- input layer
- dense/FC layer
- convolutional
- max-pooling
- flatten
- softmax layer

Take-Home Exercise: VGGNet

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL

Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

Talk through the purpose of every line in the [VGGNet notebook], including all of the following terms:

- ReLU
- cross-entropy
- epoch
- parameters
- hyperparams
- SGD
- learning rate
- batch size
- Adam
- dropout
- batchnorm
- input layer
- dense/FC layer
- convolutional
- max-pooling
- flatten
- softmax layer

Take-Home Exercise: VGGNet

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL

Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

Talk through the purpose of every line in the [VGGNet notebook], including all of the following terms:

- ReLU
- cross-entropy
- epoch
- parameters
- hyperparams
- SGD
- learning rate
- batch size
- Adam
- dropout
- batchnorm
- input layer
- dense/FC layer
- convolutional
- max-pooling
- flatten
- softmax layer

Take-Home Exercise: VGGNet

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL

Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

Talk through the purpose of every line in the [VGGNet notebook], including all of the following terms:

- ReLU
- cross-entropy
- epoch
- parameters
- hyperparams
- SGD
- learning rate
- batch size
- Adam
- dropout
- batchnorm
- input layer
- dense/FC layer
- convolutional
- max-pooling
- flatten
- softmax layer

Take-Home Exercise: VGGNet

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL

Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

Talk through the purpose of every line in the [VGGNet notebook], including all of the following terms:

- ReLU
- cross-entropy
- epoch
- parameters
- hyperparams
- SGD
- learning rate
- batch size
- Adam
- dropout
- batchnorm
- input layer
- dense/FC layer
- convolutional
- max-pooling
- flatten
- softmax layer

Take-Home Exercise: VGGNet

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL

Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

Talk through the purpose of every line in the [VGGNet notebook], including all of the following terms:

- ReLU
- cross-entropy
- epoch
- parameters
- hyperparams
- SGD
- learning rate
- batch size
- Adam
- dropout
- batchnorm
- input layer
- dense/FC layer
- convolutional
- max-pooling
- flatten
- softmax layer

Take-Home Exercise: VGGNet

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL

Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

Talk through the purpose of every line in the [VGGNet notebook], including all of the following terms:

- ReLU
- cross-entropy
- epoch
- parameters
- hyperparams
- SGD
- learning rate
- batch size
- Adam
- dropout
- batchnorm
- input layer
- dense/FC layer
- convolutional
- max-pooling
- flatten
- softmax layer

Take-Home Exercise: VGGNet

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL

Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

Talk through the purpose of every line in the [VGGNet notebook], including all of the following terms:

- ReLU
- cross-entropy
- epoch
- parameters
- hyperparams
- SGD
- learning rate
- batch size
- Adam
- dropout
- batchnorm
- input layer
- dense/FC layer
- convolutional
- max-pooling
- flatten
- softmax layer

Take-Home Exercise: VGGNet

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL

Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

Talk through the purpose of every line in the [VGGNet notebook], including all of the following terms:

- ReLU
- cross-entropy
- epoch
- parameters
- hyperparams
- SGD
- learning rate
- batch size
- Adam
- dropout
- batchnorm
- input layer
- dense/FC layer
- convolutional
- max-pooling
- flatten
- softmax layer

Take-Home Exercise: VGGNet

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL

Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

Talk through the purpose of every line in the [VGGNet notebook], including all of the following terms:

- ReLU
- cross-entropy
- epoch
- parameters
- hyperparams
- SGD
- learning rate
- batch size
- Adam
- dropout
- batchnorm
- input layer
- dense/FC layer
- convolutional
- max-pooling
- flatten
- softmax layer

Take-Home Exercise: VGGNet

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL

Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

Talk through the purpose of every line in the [VGGNet notebook], including all of the following terms:

- ReLU
- cross-entropy
- epoch
- parameters
- hyperparams
- SGD
- learning rate
- batch size
- Adam
- dropout
- batchnorm
- input layer
- dense/FC layer
- convolutional
- max-pooling
- flatten
- softmax layer

Take-Home Exercise: VGGNet

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL

Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

Talk through the purpose of every line in the [VGGNet notebook], including all of the following terms:

- ReLU
- cross-entropy
- epoch
- parameters
- hyperparams
- SGD
- learning rate
- batch size
- Adam
- dropout
- batchnorm
- input layer
- dense/FC layer
- convolutional
- max-pooling
- flatten
- softmax layer

Take-Home Exercise: VGGNet

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL

Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

Talk through the purpose of every line in the [VGGNet notebook], including all of the following terms:

- ReLU
- cross-entropy
- epoch
- parameters
- hyperparams
- SGD
- learning rate
- batch size
- Adam
- dropout
- batchnorm
- input layer
- dense/FC layer
- convolutional
- max-pooling
- flatten
- softmax layer

Take-Home Exercise: VGGNet

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL

Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

Talk through the purpose of every line in the [VGGNet notebook], including all of the following terms:

- ReLU
- cross-entropy
- epoch
- parameters
- hyperparams
- SGD
- learning rate
- batch size
- Adam
- dropout
- batchnorm
- input layer
- dense/FC layer
- convolutional
- max-pooling
- flatten
- softmax layer

Take-Home Exercise: VGGNet

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

Talk through the purpose of every line in the [VGGNet notebook], including all of the following terms:

- ReLU
- cross-entropy
- epoch
- parameters
- hyperparams
- SGD
- learning rate
- batch size
- Adam
- dropout
- batchnorm
- input layer
- dense/FC layer
- convolutional
- max-pooling
- flatten
- softmax layer

Take-Home Exercise: VGGNet

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

Talk through the purpose of every line in the [VGGNet notebook], including all of the following terms:

- ReLU
- cross-entropy
- epoch
- parameters
- hyperparams
- SGD
- learning rate
- batch size
- Adam
- dropout
- batchnorm
- input layer
- dense/FC layer
- convolutional
- max-pooling
- flatten
- softmax layer

Take-Home Exercise: VGGNet

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

Talk through the purpose of every line in the [VGGNet notebook], including all of the following terms:

- ReLU
- cross-entropy
- epoch
- parameters
- hyperparams
- SGD
- learning rate
- batch size
- Adam
- dropout
- batchnorm
- input layer
- dense/FC layer
- convolutional
- max-pooling
- flatten
- softmax layer

Take-Home Exercise: VGGNet

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

Talk through the purpose of every line in the [VGGNet notebook], including all of the following terms:

- ReLU
- cross-entropy
- epoch
- parameters
- hyperparams
- SGD
- learning rate
- batch size
- Adam
- dropout
- batchnorm
- input layer
- dense/FC layer
- convolutional
- max-pooling
- flatten
- softmax layer

Take-Home Exercise: VGGNet

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

Talk through the purpose of every line in the [VGGNet notebook], including all of the following terms:

- ReLU
- cross-entropy
- epoch
- parameters
- hyperparams
- SGD
- learning rate
- batch size
- Adam
- dropout
- batchnorm
- input layer
- dense/FC layer
- convolutional
- max-pooling
- flatten
- softmax layer

Take-Home Exercise: VGGNet

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

Talk through the purpose of every line in the [VGGNet notebook], including all of the following terms:

- ReLU
- cross-entropy
- epoch
- parameters
- hyperparams
- SGD
- learning rate
- batch size
- Adam
- dropout
- batchnorm
- input layer
- dense/FC layer
- convolutional
- max-pooling
- flatten
- softmax layer

Take-Home Exercise: VGGNet

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

Talk through the purpose of every line in the [VGGNet notebook], including all of the following terms:

- ReLU
- cross-entropy
- epoch
- parameters
- hyperparams
- SGD
- learning rate
- batch size
- Adam
- dropout
- batchnorm
- input layer
- dense/FC layer
- convolutional
- max-pooling
- flatten
- softmax layer

Outline

- 1 Review Take-Home Exercise
- 2 The Power and Elegance of Deep Learning for NLP
 - Introduction to DL for NLP
 - Computational Representations of NL
 - NLP Applications
- 3 Word Vectors
 - Vector-Space Embedding
 - word2vec
 - Creating Word Vectors with word2vec
- 4 Modeling Natural Language Data
 - Best Practices for Preprocessing NLP Data
 - The Area Under the ROC Curve
 - Sentiment Classification
- 5 Recurrent Neural Networks
 - Simple RNNs
 - LSTMs
- 6 Parallel Network Architectures

Outline

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word Vectors

Modeling NL

Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

- 1 Review Take-Home Exercise
- 2 The Power and Elegance of Deep Learning for NLP
 - Introduction to DL for NLP
 - Computational Representations of NL
 - NLP Applications
- 3 Word Vectors
 - Vector-Space Embedding
 - word2vec
 - Creating Word Vectors with word2vec
- 4 Modeling Natural Language Data
 - Best Practices for Preprocessing NLP Data
 - The Area Under the ROC Curve
 - Sentiment Classification
- 5 Recurrent Neural Networks
 - Simple RNNs
 - LSTMs
- 6 Parallel Network Architectures

Two Core Concepts

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL

Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

- 1 Deep Learning
- 2 Natural Language Processing (NLP)

Two Core Concepts

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL

Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

- 1 Deep Learning
- 2 Natural Language Processing (NLP)

TML vs Deep Learning

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL
Data

Preprocessing

ROC Curve

Sentiment

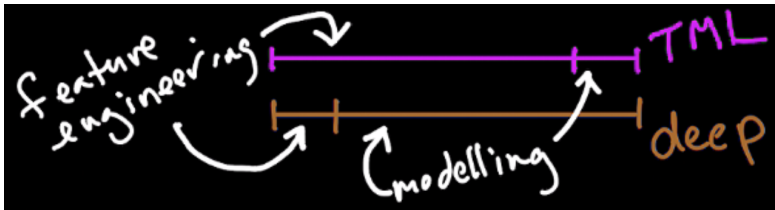
Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets



Two Core Concepts

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL

Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

- 1 Deep Learning
- 2 Natural Language Processing (NLP)

Two Core Concepts

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL

Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

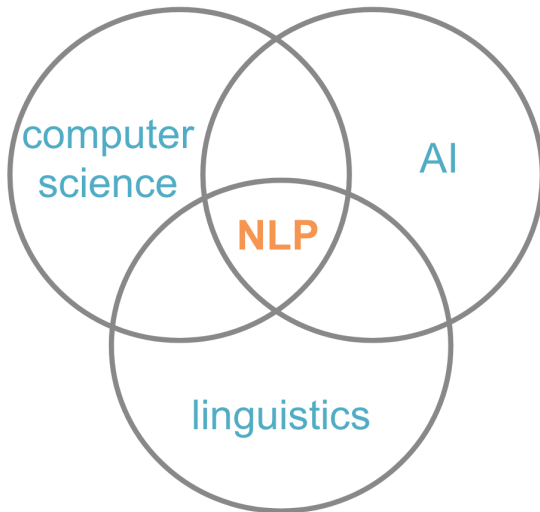
Simple RNNs

LSTMs

Parallel Nets

- 1 Deep Learning
- 2 Natural Language Processing (NLP)

Natural Language Processing



Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL
Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

Examples of NLP in Industry

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL

Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

- **speech recognition (Echo, Siri, Cortana)**
- search (typed into omnibox, spoken)
- classifying documents
- language translation
- chatbots

Examples of NLP in Industry

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL

Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

- speech recognition (Echo, Siri, Cortana)
- search (typed into omnibox, spoken)
 - classifying documents
 - language translation
 - chatbots

Examples of NLP in Industry

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL

Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

- speech recognition (Echo, Siri, Cortana)
- search (typed into omnibox, spoken)
- classifying documents
- language translation
- chatbots

Examples of NLP in Industry

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL

Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

- speech recognition (Echo, Siri, Cortana)
- search (typed into omnibox, spoken)
- classifying documents
- language translation
- chatbots

Examples of NLP in Industry

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL

Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

- speech recognition (Echo, Siri, Cortana)
- search (typed into omnibox, spoken)
- classifying documents
- language translation
- chatbots

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL

Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

- 1 Review Take-Home Exercise
- 2 The Power and Elegance of Deep Learning for NLP
 - Introduction to DL for NLP
 - Computational Representations of NL**
 - NLP Applications
- 3 Word Vectors
 - Vector-Space Embedding
 - word2vec
 - Creating Word Vectors with word2vec
- 4 Modeling Natural Language Data
 - Best Practices for Preprocessing NLP Data
 - The Area Under the ROC Curve
 - Sentiment Classification
- 5 Recurrent Neural Networks
 - Simple RNNs
 - LSTMs
- 6 Parallel Network Architectures

One-Hot Word Representations

| <u>word</u> | The | cat | sat | on | the | mat. |
|-------------|-----|-----|-----|----|-----|------|
| the | 1 | 0 | 0 | 0 | 1 | 0 |
| cat | 0 | 1 | 0 | 0 | 0 | 0 |
| on | 0 | 0 | 0 | 1 | 0 | 0 |
| ⋮ | | | | | | |
| ⋮ | | | | | | |
| ⋮ | | | | | | |

Unique-words

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL

Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

Outline

- 1 Review Take-Home Exercise
- 2 The Power and Elegance of Deep Learning for NLP
 - Introduction to DL for NLP
 - Computational Representations of NL
 - NLP Applications**
- 3 Word Vectors
 - Vector-Space Embedding
 - word2vec
 - Creating Word Vectors with word2vec
- 4 Modeling Natural Language Data
 - Best Practices for Preprocessing NLP Data
 - The Area Under the ROC Curve
 - Sentiment Classification
- 5 Recurrent Neural Networks
 - Simple RNNs
 - LSTMs
- 6 Parallel Network Architectures

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

- **spell checking**
- synonym suggestions
- keyword search

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

- spell checking
- synonym suggestions
- keyword search

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

- spell checking
- synonym suggestions
- keyword search

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL
Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

Intermediate

- reading level
- extracting information
- predicting next words
- classification
- sequence generation
- time-series analysis

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL
Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

Intermediate

- reading level
- extracting information
- predicting next words
- classification
- sequence generation
- time-series analysis

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL
Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

Intermediate

- reading level
- extracting information
- predicting next words
- classification
- sequence generation
- time-series analysis

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL
Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

Intermediate

- reading level
- extracting information
- predicting next words
- classification
- sequence generation
- time-series analysis

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL
Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

Intermediate

- reading level
- extracting information
- predicting next words
- classification
- sequence generation
- time-series analysis

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL
Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

Intermediate

- reading level
- extracting information
- predicting next words
- classification
- sequence generation
- time-series analysis

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL
Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

Complex

- machine translation
- question-answering
- chatbots

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL
Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

Complex

- machine translation
- question-answering
- chatbots

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL
Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

Complex

- machine translation
- question-answering
- chatbots

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL

Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

Outline

- 1 Review Take-Home Exercise
- 2 The Power and Elegance of Deep Learning for NLP
 - Introduction to DL for NLP
 - Computational Representations of NL
 - NLP Applications
- 3 **Word Vectors**
 - Vector-Space Embedding
 - word2vec
 - Creating Word Vectors with word2vec
- 4 Modeling Natural Language Data
 - Best Practices for Preprocessing NLP Data
 - The Area Under the ROC Curve
 - Sentiment Classification
- 5 Recurrent Neural Networks
 - Simple RNNs
 - LSTMs
- 6 Parallel Network Architectures

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space
Embedding

word2vec

Creating Word
Vectors

Modeling NL
Data

Preprocessing

ROC Curve

Sentiment
Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

Outline

- 1 Review Take-Home Exercise
- 2 The Power and Elegance of Deep Learning for NLP
 - Introduction to DL for NLP
 - Computational Representations of NL
 - NLP Applications
- 3 **Word Vectors**
 - Vector-Space Embedding**
 - word2vec
 - Creating Word Vectors with word2vec
- 4 Modeling Natural Language Data
 - Best Practices for Preprocessing NLP Data
 - The Area Under the ROC Curve
 - Sentiment Classification
- 5 Recurrent Neural Networks
 - Simple RNNs
 - LSTMs
- 6 Parallel Network Architectures

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space
Embedding

word2vec

Creating Word
Vectors

Modeling NL
Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

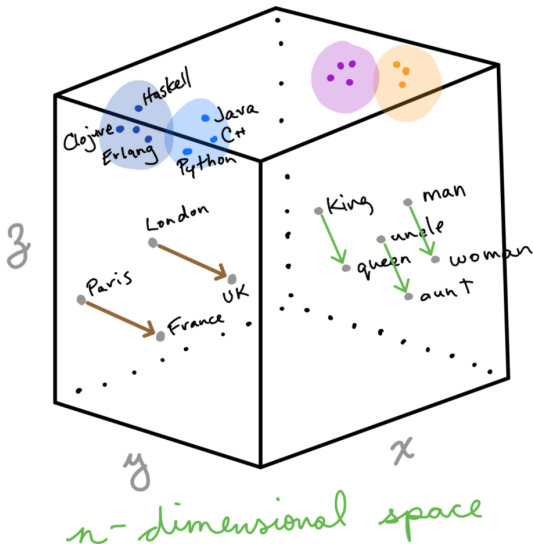
LSTMs

Parallel Nets

JR Firth (1957)

“You shall know a word by the company it keeps”

Vector Representations of Words



Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL

Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

Word Vector Arithmetic

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space
Embedding

word2vec

Creating Word
Vectors

Modeling NL

Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

$$V_{\text{king}} - V_{\text{man}} + V_{\text{woman}} = V_{?}$$

$$V_{\text{jeff_bezos}} - V_{\text{amazon}} + V_{\text{facebook}} = V_{?}$$

$$V_{\text{windows}} - V_{\text{microsoft}} + V_{\text{google}} = V_{?}$$

$$V_{\text{cu}} - V_{\text{copper}} + V_{\text{gold}} = V_{?}$$

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space
Embedding

word2vec

Creating Word
Vectors

Modeling NL
Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

[word2viz demo]

Word Representations

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space
Embedding

word2vec

Creating Word
Vectors

Modeling NL
Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

One-Hot

lack nuance

handle new words poorly

subjective

laborious, manual taxonomies

word similarity ignored

unwieldy with large vocabulary

Vector-Based

extremely **nuanced**

seamlessly incorporate **new words**

driven by natural language **data**

fully-**automatic**

word similarity = closeness in space

accommodate **large vocabularies**

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL

Data

Preprocessing

ROC Curve

Sentiment
Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

Outline

- 1 Review Take-Home Exercise
- 2 The Power and Elegance of Deep Learning for NLP
 - Introduction to DL for NLP
 - Computational Representations of NL
 - NLP Applications
- 3 **Word Vectors**
 - Vector-Space Embedding
 - word2vec**
 - Creating Word Vectors with word2vec
- 4 Modeling Natural Language Data
 - Best Practices for Preprocessing NLP Data
 - The Area Under the ROC Curve
 - Sentiment Classification
- 5 Recurrent Neural Networks
 - Simple RNNs
 - LSTMs
- 6 Parallel Network Architectures

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL
Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

JR Firth (1957)

“You shall know a word by the company it keeps”

Word Representations

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL
Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

| | predicts | relative strengths |
|-----------------------|----------------------|--|
| Skip-Gram (SG) | context given target | <ul style="list-style-type: none">● small data set● rare words |
| CBOW | target given context | <ul style="list-style-type: none">● many times faster● slightly better for frequent words |

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL
Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

Evaluating Word Vectors

- 1 intrinsic
- 2 extrinsic

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL
Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

Evaluating Word Vectors

- 1 intrinsic
- 2 extrinsic

word2vec Hyperparameters

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL

Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

- 1 n dimensions
- 2 window size (SG ~10, CBOW ~5)
- 3 n iterations
- 4 data set size

word2vec Hyperparameters

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL

Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

- 1 n dimensions
- 2 window size (SG ~10, CBOW ~5)
- 3 n iterations
- 4 data set size

word2vec Hyperparameters

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL

Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

- 1 n dimensions
- 2 window size (SG ~10, CBOW ~5)
- 3 n iterations
- 4 data set size

word2vec Hyperparameters

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL

Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

- 1 n dimensions
- 2 window size (SG ~10, CBOW ~5)
- 3 n iterations
- 4 data set size

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL

Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

Outline

- 1 Review Take-Home Exercise
- 2 The Power and Elegance of Deep Learning for NLP
 - Introduction to DL for NLP
 - Computational Representations of NL
 - NLP Applications
- 3 **Word Vectors**
 - Vector-Space Embedding
 - word2vec
 - Creating Word Vectors with word2vec**
- 4 Modeling Natural Language Data
 - Best Practices for Preprocessing NLP Data
 - The Area Under the ROC Curve
 - Sentiment Classification
- 5 Recurrent Neural Networks
 - Simple RNNs
 - LSTMs
- 6 Parallel Network Architectures

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

**Creating Word
Vectors**

Modeling NL
Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

[*creating word vectors* notebook]

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL
Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

Outline

- 1 Review Take-Home Exercise
- 2 The Power and Elegance of Deep Learning for NLP
 - Introduction to DL for NLP
 - Computational Representations of NL
 - NLP Applications
- 3 Word Vectors
 - Vector-Space Embedding
 - word2vec
 - Creating Word Vectors with word2vec
- 4 Modeling Natural Language Data
 - Best Practices for Preprocessing NLP Data
 - The Area Under the ROC Curve
 - Sentiment Classification
- 5 Recurrent Neural Networks
 - Simple RNNs
 - LSTMs
- 6 Parallel Network Architectures

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL

Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

Outline

- 1 Review Take-Home Exercise
- 2 The Power and Elegance of Deep Learning for NLP
 - Introduction to DL for NLP
 - Computational Representations of NL
 - NLP Applications
- 3 Word Vectors
 - Vector-Space Embedding
 - word2vec
 - Creating Word Vectors with word2vec
- 4 Modeling Natural Language Data
 - Best Practices for Preprocessing NLP Data**
 - The Area Under the ROC Curve
 - Sentiment Classification
- 5 Recurrent Neural Networks
 - Simple RNNs
 - LSTMs
- 6 Parallel Network Architectures

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL

Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

Best Practices for Preprocessing NLP Data

[*NL preprocessing best practices* notebook]

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL

Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

Outline

- 1 Review Take-Home Exercise
- 2 The Power and Elegance of Deep Learning for NLP
 - Introduction to DL for NLP
 - Computational Representations of NL
 - NLP Applications
- 3 Word Vectors
 - Vector-Space Embedding
 - word2vec
 - Creating Word Vectors with word2vec
- 4 Modeling Natural Language Data
 - Best Practices for Preprocessing NLP Data
 - The Area Under the ROC Curve
 - Sentiment Classification
- 5 Recurrent Neural Networks
 - Simple RNNs
 - LSTMs
- 6 Parallel Network Architectures

The Area Under the ROC Curve

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL
Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

whiteboard

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL

Data

Preprocessing

ROC Curve

Sentiment
Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

Outline

- 1 Review Take-Home Exercise
- 2 The Power and Elegance of Deep Learning for NLP
 - Introduction to DL for NLP
 - Computational Representations of NL
 - NLP Applications
- 3 Word Vectors
 - Vector-Space Embedding
 - word2vec
 - Creating Word Vectors with word2vec
- 4 Modeling Natural Language Data
 - Best Practices for Preprocessing NLP Data
 - The Area Under the ROC Curve
 - Sentiment Classification**
- 5 Recurrent Neural Networks
 - Simple RNNs
 - LSTMs
- 6 Parallel Network Architectures

Dense Net Classification

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL

Data

Preprocessing

ROC Curve

Sentiment
Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

[*dense sentiment classifier* notebook]

ConvNet Classification

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL

Data

Preprocessing

ROC Curve

Sentiment
Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

[*convolutional sentiment classifier notebook*]

Outline

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL

Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

- 1 Review Take-Home Exercise
- 2 The Power and Elegance of Deep Learning for NLP
 - Introduction to DL for NLP
 - Computational Representations of NL
 - NLP Applications
- 3 Word Vectors
 - Vector-Space Embedding
 - word2vec
 - Creating Word Vectors with word2vec
- 4 Modeling Natural Language Data
 - Best Practices for Preprocessing NLP Data
 - The Area Under the ROC Curve
 - Sentiment Classification
- 5 Recurrent Neural Networks
 - Simple RNNs
 - LSTMs
- 6 Parallel Network Architectures

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL

Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

Outline

- 1 Review Take-Home Exercise
- 2 The Power and Elegance of Deep Learning for NLP
 - Introduction to DL for NLP
 - Computational Representations of NL
 - NLP Applications
- 3 Word Vectors
 - Vector-Space Embedding
 - word2vec
 - Creating Word Vectors with word2vec
- 4 Modeling Natural Language Data
 - Best Practices for Preprocessing NLP Data
 - The Area Under the ROC Curve
 - Sentiment Classification
- 5 Recurrent Neural Networks
 - Simple RNNs
 - LSTMs
- 6 Parallel Network Architectures

RNN Theory

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL
Data

Preprocessing

ROC Curve

Sentiment

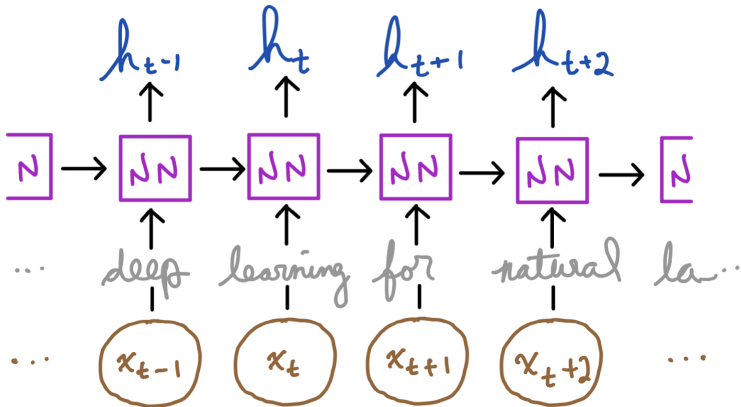
Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets



Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL
Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

RNNs in Practice

[*rnn* notebook]

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL

Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

Outline

- 1 Review Take-Home Exercise
- 2 The Power and Elegance of Deep Learning for NLP
 - Introduction to DL for NLP
 - Computational Representations of NL
 - NLP Applications
- 3 Word Vectors
 - Vector-Space Embedding
 - word2vec
 - Creating Word Vectors with word2vec
- 4 Modeling Natural Language Data
 - Best Practices for Preprocessing NLP Data
 - The Area Under the ROC Curve
 - Sentiment Classification
- 5 Recurrent Neural Networks
 - Simple RNNs
 - LSTMs
- 6 Parallel Network Architectures

LSTM Theory

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL

Data

Preprocessing

ROC Curve

Sentiment

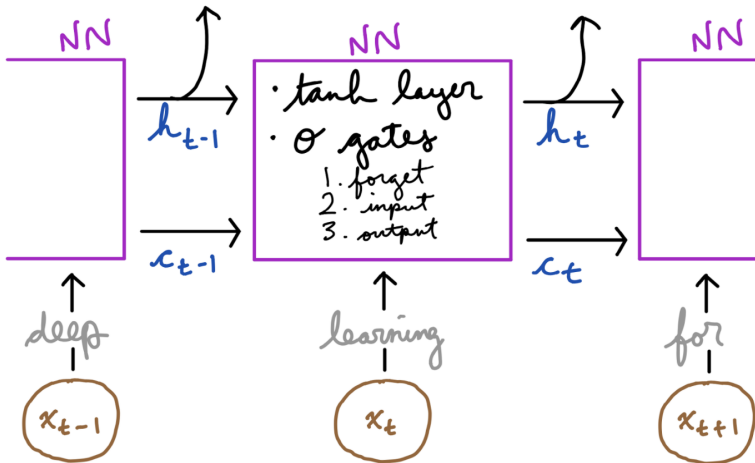
Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets



Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL
Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

LSTMs in Practice

[*vanilla LSTM* and *GRU* notebooks]

Bi-Directional LSTMs

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL
Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

[*Bi-LSTM* notebook]

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL
Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

Stacked LSTMs

[*stacked LSTM* and *ye olde stackeroo* notebooks]

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL

Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

Outline

- 1 Review Take-Home Exercise
- 2 The Power and Elegance of Deep Learning for NLP
 - Introduction to DL for NLP
 - Computational Representations of NL
 - NLP Applications
- 3 Word Vectors
 - Vector-Space Embedding
 - word2vec
 - Creating Word Vectors with word2vec
- 4 Modeling Natural Language Data
 - Best Practices for Preprocessing NLP Data
 - The Area Under the ROC Curve
 - Sentiment Classification
- 5 Recurrent Neural Networks
 - Simple RNNs
 - LSTMs
- 6 Parallel Network Architectures

LSTMs in Practice

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL
Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

[*multi-ConvNet* notebook]

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL
Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

Assessing Your Deep Learning Project III



Assessing

Your Deep Learning Project III

1 split your data

- training set (80% — for optimizing parameters)
- validation set (10% — for hyperparameters)
- test set (10% — don't touch yet!)

2 build and assess architecture

- get above chance (simplifying problem, if necessary)
- do existing performance benchmarks exist?
- if not, use a simple architecture as benchmark

3 “teamwork makes the dream work” (?)

Assessing

Your Deep Learning Project III

- 1 split your data
 - training set (80% — for optimizing parameters)
 - validation set (10% — for hyperparameters)
 - test set (10% — don't touch yet!)
- 2 build and assess architecture
 - get above chance (simplifying problem, if necessary)
 - do existing performance benchmarks exist?
 - if not, use a simple architecture as benchmark
- 3 “teamwork makes the dream work” (?)

Assessing

Your Deep Learning Project III

- 1 split your data
 - training set (80% — for optimizing parameters)
 - validation set (10% — for hyperparameters)
 - test set (10% — don't touch yet!)
- 2 build and assess architecture
 - get above chance (simplifying problem, if necessary)
 - do existing performance benchmarks exist?
 - if not, use a simple architecture as benchmark
- 3 “teamwork makes the dream work” (?)

Assessing

Your Deep Learning Project III

- 1 split your data
 - training set (80% — for optimizing parameters)
 - validation set (10% — for hyperparameters)
 - test set (10% — don't touch yet!)
- 2 build and assess architecture
 - get above chance (simplifying problem, if necessary)
 - do existing performance benchmarks exist?
 - if not, use a simple architecture as benchmark
- 3 “teamwork makes the dream work” (?)

Assessing

Your Deep Learning Project III

- 1 split your data
 - training set (80% — for optimizing parameters)
 - validation set (10% — for hyperparameters)
 - test set (10% — don't touch yet!)
- 2 build and assess architecture
 - get above chance (simplifying problem, if necessary)
 - do existing performance benchmarks exist?
 - if not, use a simple architecture as benchmark
- 3 “teamwork makes the dream work” (?)

Assessing

Your Deep Learning Project III

- 1 split your data
 - training set (80% — for optimizing parameters)
 - validation set (10% — for hyperparameters)
 - test set (10% — don't touch yet!)
- 2 build and assess architecture
 - get above chance (simplifying problem, if necessary)
 - do existing performance benchmarks exist?
 - if not, use a simple architecture as benchmark
- 3 “teamwork makes the dream work” (?)

Assessing

Your Deep Learning Project III

- 1 split your data
 - training set (80% — for optimizing parameters)
 - validation set (10% — for hyperparameters)
 - test set (10% — don't touch yet!)
- 2 build and assess architecture
 - get above chance (simplifying problem, if necessary)
 - do existing performance benchmarks exist?
 - if not, use a simple architecture as benchmark
- 3 “teamwork makes the dream work” (?)

Assessing

Your Deep Learning Project III

- 1 split your data
 - training set (80% — for optimizing parameters)
 - validation set (10% — for hyperparameters)
 - test set (10% — don't touch yet!)
- 2 build and assess architecture
 - get above chance (simplifying problem, if necessary)
 - do existing performance benchmarks exist?
 - if not, use a simple architecture as benchmark
- 3 “teamwork makes the dream work” (?)

Assessing

Your Deep Learning Project III

- 1 split your data
 - training set (80% — for optimizing parameters)
 - validation set (10% — for hyperparameters)
 - test set (10% — don't touch yet!)
- 2 build and assess architecture
 - get above chance (simplifying problem, if necessary)
 - do existing performance benchmarks exist?
 - if not, use a simple architecture as benchmark
- 3 “teamwork makes the dream work” (?)

Up Next: TensorFlow

December 2nd

Review

DL for NLP

Intro

Representations

NLP Applications

Word Vectors

Vector-Space

Embedding

word2vec

Creating Word
Vectors

Modeling NL
Data

Preprocessing

ROC Curve

Sentiment

Classification

RNNs

Simple RNNs

LSTMs

Parallel Nets

