Units 5 and 6 — NLP

Review

L for NLI

Representation

Word Vectors

Vector-Space Embedding

word2vec Creating Word Vectors

Modeling NL

Preprocessin ROC Curve Sentiment Classification

RNNs Simple RNN

2011115

Natural Language Processing Deep Learning — Units 5 & 6

Dr. Jon Krohn
jon@untapt.com

Slides available at jonkrohn.com/talks

August 4th, 2018



eview 1

DL for NL

Intro
Representations
NI P Application

Word Vectors

Vector-Space

Embedding word2vec Creating Word Vectors

Modeling N Data

Preprocessin ROC Curve Sentiment Classification

KININS
Simple RNI

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

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

L for NL

Intro
Representations
NLP Applications

Word Vectors
Vector-Space
Embedding
word2vec
Creating Word

Modeling N Data

Preprocessing ROC Curve Sentiment Classification

KNNS Simple RNN

Davallal Nate

- Review Take-Home Exercise
 The Power and Elegance of D
- 2 The Power and Elegance of Deep Learning for NLP Introduction to DL for NLP Computational Representations of NL NLP Applications
- Word Vectors Vector-Space Embedding word2vec Creating Word Vectors with
- 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

L for NL

Intro
Representations
NLP Applications

Word Vector
Vector-Space
Embedding
word2vec
Creating Word
Vectors

Modeling N Data

Preprocessing ROC Curve Sentiment Classification

RNNs Simple RNN

Parallal Nate

- 2 The Power and Elegance of Deep Learning for NLP Introduction to DL for NLP Computational Representations of NL
- Word Vectors Vector-Space Embedding word2vec Creating Word Vectors with word2vec

Review Take-Home Exercise

NLP Applications

- 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

L for NLP

Representations
NLP Applications

Vector-Space Embedding word2vec Creating Word Vectors

Modeling N Data

Preprocessing ROC Curve Sentiment Classification

RNNs Simple RNNs

Parallal Note

- The Power and Elegance of Deep Learning for NLP Introduction to DL for NLP Computational Representations of NL NLP Applications
- Word Vectors Vector-Space Embedding word2vec Creating Word Vectors with word2vec

Review Take-Home Exercise

- 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

L for NLP

Word Vector
Vector-Space
Embedding
word2vec

Modeling N Data

Preprocessing ROC Curve Sentiment Classification

RNNs Simple RNNs

Parallel Nets

- Review Take-Home Exercise
- The Power and Elegance of Deep Learning for NLP Introduction to DL for NLP Computational Representations of NL NLP Applications
- 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

L for NLP

Intro
Representations
NLP Application

Word Vector
Vector-Space
Embedding
word2vec
Creating Word
Vectors

Modeling N Data

Preprocessing ROC Curve Sentiment Classification

RNNs Simple RNNs LSTMs

Parallel Nets

- 1 Review Take-Home Exercise
- The Power and Elegance of Deep Learning for NLP Introduction to DL for NLP Computational Representations of NL NLP Applications
- 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

L for NI

ro presentations

Word Vectors
Vector-Space
Embedding

Embedding word2vec Creating Word Vectors

Modeling N Data

Preprocessing ROC Curve Sentiment Classification

RNNs Simple F

Review Take-Home Exercise

- The Power and Elegance of Deep Learning for NLP Introduction to DL for NLP Computational Representations of NL NLP Applications
- Vectors
 Vector-Space Embedding
 word2vec
 Creating Word Vectors with word2vec
- 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



Units 5 and 6 — NLP

Review

DL for NL

Intro Representations

Word Vectors

Embedding word2vec Creating Word

Modeling NL

Preprocessing ROC Curve Sentiment Classification

Simple RNN

Parallel Nets

Take-Home Exercise: VGGNet

- cross-entropy
- epoch
- parameters
- hyperparams

- learning rate
- batch size
- Adam
 - dropout
- batchnorm

- input laver
- dense/FC layer
- convolutional
- max-pooling
- flatten
- softmax layer



I for NI

Intro Representation

Word Vectors

Vector-Space Embedding word2vec Creating Word

Modeling NI

Preprocessing ROC Curve Sentiment Classification

RNNs Simple RN

LSTMs

Parallel Nets

Take-Home Exercise: VGGNet

- ReLU
- cross-entropy
- epoch
- parameters
- hyperparams

- SGD
- loovoloo voto
- a hatab ciza
- Adam
- dropout
- batchnorm

- input laver
- dense/FC layer
- convolutional
- max-pooling
- flatten
- softmax layer



Take-Home Exercise: VGGNet

- ReLU



I for NI

Intro
Representations

Word Vectors

Embedding word2vec Creating Word

Modeling NI Data

Preprocessing ROC Curve Sentiment Classification

Simple RNN

Parallel Nets

Take-Home Exercise: VGGNet

- ReLU
 * SGD * input laye
 - cross-entropy
 - batch size
 convolutiona
 - Adam
 max-pooling
 - dropout flatten
 - batchnormsoftmax laye



I for NI

Intro
Representations

Word Vectors

Embedding word2vec Creating Word

Modeling N

Preprocessing ROC Curve Sentiment Classification

Simple RNN

Parallel Nets

Take-Home Exercise: VGGNet

- ReLU
 * SGD * input laye
- cross-entropy
- epochbatch sizeconvolutions
- Adam
 max-pooling
- dropoutflatten
 - batchnorm
 softmax layer



I for NI

Intro
Representations

Word Vectors
Vector-Space
Embedding
word2vec

Modeling N

Preprocessing ROC Curve Sentiment Classification

Simple RNN

Parallel Nets

Take-Home Exercise: VGGNet

- cross-entropy
- epoch

 * batch size

 convolutions
- parameters
- dropout flatten
- batchnorm
 softmax laye



I for NI

Intro
Representations

Word Vectors
Vector-Space
Embedding
word2vec
Creating Word

Modeling N

Preprocessing ROC Curve Sentiment Classification

Simple RNN

Parallel Nets

Take-Home Exercise: VGGNet

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

- ReLU
 * SGD * input laye
- cross-entropy
- epoch
- parameters
- hyperparams

softmax layer



I for NI

Intro
Representations

Word Vector:
Vector-Space
Embedding
word2vec
Creating Word

Modeling NI

Preprocessing ROC Curve Sentiment Classification

KNNS Simple R

LSTMs

Parallel Nets

Take-Home Exercise: VGGNet

- ReLU
- cross-entropy
- epoch
- parameters
- hyperparams

- SGD
- learning rate
- batch size
- Adam
- dropout
- batchnorn

- input laver
- dense/FC layer
- convolutional
- max-pooling
- flatten
- softmax layer



L for NL

Intro
Representations
NI P Application

Word Vectors

Vector-Space
Embedding

word2vec

Modeling N

Preprocessing ROC Curve Sentiment Classification

KININS Simple

LSTMs

Parallel Nets

Take-Home Exercise: VGGNet

- ReLU
- cross-entropy
- epoch
- parameters
- hyperparams

- SGD
- learning rate
- batch size
- Adam
- dropout
- batchnorm

- input laver
 - dense/FC layer
- convolutional
- max-pooling
- flatten
- softmax layer



I for NI

Intro
Representation:

Word Vectors

Vector-Space
Embedding
word2vec
Creating Word

Modeling N

Preprocessing ROC Curve Sentiment Classification

KNNS Simple R

LSTMs

Parallel Nets

Take-Home Exercise: VGGNet

- ReLU
- cross-entropy
- epoch
- parameters
- hyperparams

- SGD
- learning rate
- batch size
- Adam
- dropout
- batchnorn

- input layer
 - dense/FC laye
 - convolutional
- max-pooling
- flatten
- softmax layer



L for NL

Intro
Representations
NLP Applications

Word Vectors
Vector-Space
Embedding
word2vec
Creating Word

Modeling N

Preprocessing ROC Curve Sentiment Classification

Simple RNI

Parallel Nets

Take-Home Exercise: VGGNet

- ReLU
- cross-entropy
- epoch
- parameters
- hyperparams

- SGD
- learning rate
- batch size
- Adam
- dropout
- batchnorr

- input layer
 - dense/FC laye
 - convolutional
 - max-pooling
- flatten
- softmax layer



L for NL

Intro
Representations
NI P Application

Word Vectors
Vector-Space
Embedding
word2vec
Creating Word

Modeling N

Preprocessing ROC Curve Sentiment Classification

Simple RN

LSIMs

Parallel Nets

Take-Home Exercise: VGGNet

- ReLU
- cross-entropy
- epoch
- parameters
- hyperparams

- SGD
- learning rate
- batch size
- Adam
- dropout
- batchnorn

- input layer
 - dense/FC layer
 - convolutional
- max-pooling
- softmax laver
- softmax layer



L for NL

Intro
Representations

Word Vector
Vector-Space
Embedding
word2vec
Creating Word

Modeling N

Preprocessing ROC Curve Sentiment Classification

KNNS Simple RN

LSTMs

Parallel Nets

Take-Home Exercise: VGGNet

- ReLU
- cross-entropy
- epoch
- parameters
- hyperparams

- SGD
- learning rate
- batch size
- Adam
- dropout
- batchnorm

- dense/FC lave
- dense/FC layer
- may pooling





L for NL

Intro
Representations
NI P Applications

Word Vectors
Vector-Space
Embedding
word2vec
Creating Word

Modeling N

Preprocessing ROC Curve Sentiment Classification

Simple RI

LSTMs

Parallel Nets

Take-Home Exercise: VGGNet

- ReLU
- cross-entropy
- epoch
- parameters
- hyperparams

- SGD
- learning rate
- batch size
- Adam
- dropout
- batchnorm



L for NL

Intro
Representations
NLP Applications

Word Vectors
Vector-Space
Embedding
word2vec
Creating Word

Modeling NI Data

Preprocessing ROC Curve Sentiment Classification

RNNs Simple RI

Simple RNNs LSTMs

Parallel Nets

Take-Home Exercise: VGGNet

- 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



L for NL

Intro
Representations
NI P Application

Word Vectors
Vector-Space
Embedding
word2vec
Creating Word

Modeling NI Data

Preprocessing ROC Curve Sentiment Classification

Simple RN

Parallel Nets

Take-Home Exercise: VGGNet

- 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



L for NL

Intro
Representations
NI P Application

Word Vectors

Vector-Space
Embedding

word2vec

Creating Word

Modeling NI Data

Preprocessing ROC Curve Sentiment Classification

KNNS Simple RN

LSTMs

Parallel Nets

Take-Home Exercise: VGGNet

- 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



L for NL

Intro
Representations
NLP Application

Word Vectors
Vector-Space
Embedding
word2vec
Creating Word

Modeling NI

Preprocessing ROC Curve Sentiment Classification

RNNs Simple F

LSTMs

Parallel Nets

Take-Home Exercise: VGGNet

- 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



L for NL

Intro
Representations
NI P Application

Word Vector
Vector-Space
Embedding
word2vec
Creating Word

Modeling NI

Preprocessing ROC Curve Sentiment Classification

Simple RNN

Parallel Nets

Take-Home Exercise: VGGNet

- 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



L for NL

Intro
Representations
NLP Application

Word Vectors

Vector-Space
Embedding

word2vec

Creating Word

Modeling NI Data

Preprocessing ROC Curve Sentiment Classification

RNNs Simple F

LSTMs

Parallel Nets

Take-Home Exercise: VGGNet

- ReLU
- cross-entropy
- epoch
- parameters
- hyperparams

- SGD
- learning rate
- batch size
- Adam
 - dropout
- batchnorm

- input layer
- dense/FC layer
- convolutionalmax-pooling
- flatten
- softmax layer



L for NL

Intro
Representations
NLP Applications

Word Vector: Vector-Space Embedding word2vec Creating Word

Modeling NI

Preprocessing ROC Curve Sentiment Classification

RNNs Simple F

Simple RNNs LSTMs

Parallel Net

Take-Home Exercise: VGGNet

- 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



DI for NI P

Intro
Representations

Word Vectors

Vector-Space Embedding word2vec Creating Word Vectors

Modeling N Data

Preprocessing ROC Curve Sentiment Classification

RNN

Simple RNNs LSTMs

Parallel Nets

Outline

- 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
- 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
- 5 Recurrent Neural Networks Simple RNNs LSTMs
- 6 Parallel Network Architectures



Review

L for NL

Representation:

Word Vectors
Vector-Space
Embedding
word2vec

Modeling NI

Preprocessing ROC Curve Sentiment Classification

Simple RNN:

Parallel Nets

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

Word Vectors
Vector-Space Embedding

word2vec

wordzvec

Creating Word Vectors with word2vec

Modeling Natural Language Data
 Root Practices for Proprocessing

Best Practices for Preprocessing NLP Data
The Area Under the BOC Curve

Sentiment Classification

5 Recurrent Neural Networks
Simple RNNs
LSTMs

6 Parallel Network Architectures



Two Core Concepts

- Deep Learning



OL for NI

Intro

P Application

Word Vector

word2vec
Creating Word

Modeling NI

Preprocessin
ROC Curve
Sentiment
Classification

Classificatio

Simple RNNs LSTMs

Parallel Nets

Two Core Concepts

- Deep Learning
- 2 Natural Language Processing (NLP)



Units 5 and 6 — NLP

Review

OL for NLF

Intro

representation

Word Vectors

Word vectors

Embeddin

word2vec

Vectors Vectors

Modeling NI

Data

Sentiment

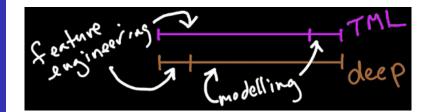
Classification

RNNs

Simple RNN LSTMs

Parallel Nets

TML vs Deep Learning





Two Core Concepts

- Deep Learning



DL for NL

Panragantatio

Word Vector

vvoid vector:

word2vec Creating Word

Modeling NI

Data Preprocessing

ROC Curve Sentiment Classification

Simple RNN

LSTMs

Parallel Net

Two Core Concepts

- Deep Learning
- 2 Natural Language Processing (NLP)



Units 5 and 6 — NLP

Review

OL for NII

Representation

Word Vector

Embedding word2vec Creating Wor

Modeling NL

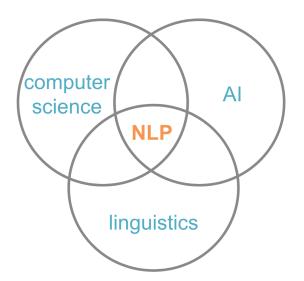
Preprocessing ROC Curve Sentiment

RNNs

Simple RNN: LSTMs

Parallel Nets

Natural Language Processing





OL for NI

Intro

epresentation P Application

Word Vectors

Embedding word2vec

Creating Work Vectors

Modeling NI

Preprocessing
ROC Curve

ROC Curve Sentiment Classification

Simple RNN

LSTWS

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



OI for NI

Representation

Word Vectors

Embedding word2vec Creating Word Vectors

Modeling NI

Preprocessing ROC Curve Sentiment Classification

Simple RNN

LSTMs

Parallel Net

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



OI for NI

Representation

NLP Applicatio

Word Vector

Embedding word2vec Creating Word Vectors

Modeling NI

Preprocessing ROC Curve Sentiment Classification

Simple RNN

LSTMs

Parallel Net

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



I for NI

Representation

Word Vector

vector-Space Embedding word2vec Creating Word Vectors

Modeling NI

Preprocessing ROC Curve Sentiment Classification

Simple RNN

Parallel Net

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



I for NI

Representation

Word Vector

vector-Space Embedding word2vec Creating Word Vectors

Modeling NI

Preprocessing ROC Curve Sentiment Classification

Simple RNN

LSTMs

Parallel Net

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



Outline

Review

L for NL

Representations
NLP Applications

Word Vector: Vector-Space Embedding word2vec Creating Word

Modeling N

Preprocessing ROC Curve Sentiment Classification

RNNs Simple RNN

Parallel Nets

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

Vord Vectors
Vector-Space Embedding
word2vec
Creating Word Vectors with word2vec

Modeling Natural Language Data Best Practices for Preprocessing NLP Data The Area Under the ROC Curve Sentiment Classification

Recurrent Neural Networks Simple RNNs LSTMs





Units 5 and 6 — NLP

Review

L for NL

Representations

Word Vectors

Vector-Space Embedding word2vec

Modeling N

Preprocessing ROC Curve Sentiment

RNNs Simple RNI

Parallel Nets

One-Hot Word Representations

	The	cat	sat	oh	the	mat.
the	1	٥	0	٥	1	
cat	0	0	0		Ø	
:						

Nunique_words



Outline

Review

L for NL

Representations

NLP Applications

Word Vectors
Vector-Space
Embedding
word2vec
Creating Word

Modeling N Data

Preprocessing ROC Curve Sentiment Classification

RNNs Simple RNN:

Parallel Nets

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

Vord Vectors

Vector-Space Embedding

word2vec

Creating Word Vectors with word2vec

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





Easy

Review

DI for NI

Intro

NLP Applications

word vector

Vector-Space

word2vec

Creating Work

Modeling NL

Data Preprocessing

ROC Curve Sentiment Classification

Simple RNN

LSTMs

- spell checking
- synonym suggestions
- keyword search



Easy

Review

OL for NI

Intro

NLP Applications

Word Vectors

Vector-Space

word2voc

Creating Wor

Modeling NI

Data

ROC Curve Sentiment Classification

KINNS Simple BNN

Simple RNNs LSTMs

- spell checking
- synonym suggestions
- keyword search



Easy

Review

OI for NI

IIIIIO

NLP Applications

Word Vector

Vector-Space Embedding word2vec

Modeling N

Data

ROC Curve Sentiment

Classificatio

Simple RNNs

- spell checking
- synonym suggestions
- keyword search



DI for NI

Intro

NLP Applications

Word Vector

.....

LIIIDOGGGII

Creating Wor

Modeling NL

Preprocessir ROC Curve Sentiment

DAINI-

Simple RNN: LSTMs

Parallel Nets

Intermediate

reading level

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



Intermediate

Review

DL for NL

Intro

NLP Applications

Word Vectors

vvoid vector

Embeddii

word2vec

Creating Work Vectors

Modeling NI

Preprocessing ROC Curve

Sentiment Classificatio

CININS

LSTMs

Parallel Nets

reading level

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



NLP Applications

- reading level
- extracting information
- predicting next words



NLP Applications

- reading level
- extracting information
- predicting next words
- classification



I for NI

Intro

NLP Applications

Word Vector

Vector-Space Embedding

word2vec Creating Wor

Modeling NI

ROC Curve Sentiment Classification

Simple RNN

LSTMs

Parallel Net

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



I for NI

Intro

NLP Applications

Word Vector

Vector-Space Embedding word2vec Creating Word Vectors

Modeling N

Preprocessin ROC Curve Sentiment Classification

Simple RNI

LSTMs

Parallel Net

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



DI for NI

Intro

NLP Applications

word vector

Vector-Space

word2vec

Creating Wor Vectors

Modeling NL

ROC Curve Sentiment

Sentiment Classification

Simple RNN

Parallal Not

Complex

- machine translation
- question-answering
- chatbots



OI for NI

Intro

NLP Applications

Word Vecto

Vector-Space

....

Creating Wor

Modeling NL

Preprocessin ROC Curve Sentiment

Sentiment Classificatio

KNNs

LSTMs

Parallel Net

Complex

- machine translation
- question-answering
- chatbots



DL for NL

intro

NLP Applications

Word Vector

Ventor Conne

.

Creating Work

Modeling NL

Preprocessing ROC Curve

Sentiment Classification

Circula Di

Simple RNNs LSTMs

Parallel Nets

Complex

- machine translation
- question-answering
- chatbots



Outline

Review

I for NI

Representation:

Word Vectors

Vector-Space Embedding word2vec Creating Word Vectors

Modeling N Data

Preprocessing ROC Curve Sentiment Classification

KININS Simple RNN

Simple RNNs LSTMs

Parallel Nets

The Power and Flagance of F

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





Outline

Review

L for NL

Representations

Word Vectors

Vector-Space
Embedding

word2vec
Creating Word
Vectors

Modeling N Data

Preprocessing ROC Curve Sentiment Classification

KNNS Simple RNN

Parallel Nets

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

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

6 Recurrent Neural Networks
Simple RNNs
LSTMs





Units 5 and 6 — NLP

Review

OI for NI

Intro

II D A--E--E-

Word Vector

Vector-Space Embedding

word2vec

Creating Work

Modeling NL

Data

ROC Curve

Sentiment Classificatio

HNNS

LSTMs

Parallel Nets

JR Firth (1957)

"You shall know a word by the company it keeps"



Units 5 and 6 — NLP

Review

N. for NII

Intro Representation

Word Vectors

Vector-Space Embedding

word2vec Creating Wor

Modeling NL

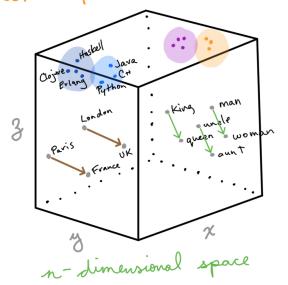
Preprocessir ROC Curve Sentiment

RNNs

Simple RNN LSTMs

Parallel Nets

Vector Representations of Words





Word Vector Arithmetic

Review

L for NL

Intro
Representations

Word Vectors Vector-Space Embedding

vector-space Embedding word2vec Creating Word Vectors

Modeling NL Data

Preprocessing ROC Curve Sentiment Classification

RNNs Simple Rt

Simple RNNs LSTMs

$$V_{\text{king}} - V_{\text{man}} + V_{\text{woman}} = V_{?}$$
 $V_{\text{jeff_bezos}} - V_{\text{amazon}} + V_{\text{facebook}} = V_{?}$
 $V_{\text{windows}} - V_{\text{microsoft}} + V_{\text{google}} = V_{?}$
 $V_{\text{cu}} - V_{\text{copper}} + V_{\text{gold}} = V_{?}$



Units 5 and 6 - NLP

Vector-Space

Embedding

[word2viz demo]



Units 5 and 6 — NLP

Review

I for NI

Representation:

Word Vectors Vector-Space

Embedding word2vec Creating Word

Modeling NL

ROC Curve Sentiment Classification

Simple RNt

Parallel Net

Word Representations

One-Hot

Vector-Based

lack nuance

handle new words poorly

subjective

laborious, manual taxonomies

word similarity ignored

unwieldy with large vocabulary

extremely nuanced

seamlessly incorporate new words

driven by natural language data

fully-automatic

word similarity = closeness in space

accommodate large vocabularies



Outline

Review

I for NI

Intro Representations

Word Vectors

Vector-Space

word2vec

Creating Work Vectors

Modeling N Data

Preprocessing ROC Curve Sentiment Classification

Simple RNN

Parallel Nets

1 Review Take-Home Exercise

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

- 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



Units 5 and 6 — NLP

Review

OL for NI

Intro

II D Applicatio

Word Vector

Vector-Space

word2vec

Creating Word

Modeling NL

Preprocessing ROC Curve

Sentiment Classification

RNNs

LSTMs

Parallel Nets

JR Firth (1957)

"You shall know a word by the company it keeps"



I for NI

Representation

Word Vectors

vvoid vectors

Embedding

word2vec

Creating Work

Modeling NI

Preprocessin ROC Curve Sentiment

Classificatio

Simple RNN LSTMs

Parallel Ne

Word Representations

	predicts	relative strengths	
Skip-Gram (SG)	context given target	small data setrare words	
CBOW	target given context	many times fasterslightly better for frequent words	



I for NI

Intro

u o a u u

Word Vector

Vector-Space

word2vec

Creating Work

Modeling N

Data

ROC Curve Sentiment

Sentiment Classification

TININS

LSTMs

Parallel Nets

Evaluating Word Vectors

- 1 intrinsic
 - extrinsic



word2vec

Evaluating Word Vectors

- intrinsic
- 2 extrinsic



DL for NL

Intro

JI D Application

Word Vector

Vector Space

Vector-Space Embodding

word2vec

Creating Wor Vectors

Modeling N

Data

ROC Curve Sentiment

Sentiment Classification

Simple RNf

LSTMs

Parallel Nets

word2vec Hyperparameters

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



L for NL

Representation

Word Vectors

Venter Spane

word2vec

Creating Work

Modeling N

Preprocessin ROC Curve Sentiment

Sentiment Classification

Simple RNN

LSTMs

Parallel Net

word2vec Hyperparameters

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



L for NL

Intro Representations

Word Vectors

Vector-Space Embedding

word2vec Creating Wor

Modeling N

Data Preprocessing

ROC Curve Sentiment Classification

Simple RNN

LSTMs

Parallel Net

word2vec Hyperparameters

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



Review

L for NL

Intro Representations

Word Vectors

Vector-Space Embedding

word2vec

Creating Work

Modeling N

Preprocessing ROC Curve Sentiment

Sentiment Classification

Simple RNN LSTMs

Parallel Net

word2vec Hyperparameters

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



Review

L for NL

Representations
NLP Application

Word Vectors
Vector-Space
Embedding
word2vec
Creating Word

Modeling NI Data

Vectors

Preprocessing ROC Curve Sentiment Classification

Simple RNN:

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

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





Review

OL for NI

Intro

NII D Assissants

Word Vector

Vector-Space Embedding

Creating Word Vectors

Modeling NL

Preprocessing ROC Curve Sentiment

RNNs

Simple RNNs

Parallel Nets

[creating word vectors notebook]



Review

DI for NI

Intro
Representations
NI P Application

Word Vector: Vector-Space Embedding

Embedding word2vec Creating Word Vectors

Modeling NL Data

Preprocessing ROC Curve Sentiment Classification

Simple RNN

Parallal Nata

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

L for NL

Intro
Representations
NLP Applications

Word Vectors
Vector-Space
Embedding
word2vec
Creating Word

Modeling N Data

Preprocessing ROC Curve Sentiment Classification

RNNs Simple RNN:

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
- 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

The Area Under the ROC Curve Sentiment Classification

- 5 Recurrent Neural Networks
 Simple RNNs
 LSTMs
- 6 Parallel Network Architectures



Review

DL for NL

Intro Representatio

Word Vector

vvoid vector

word2vec

Creating Work

Modeling NL

Data Preprocessing

ROC Curve

Sentiment Classification

Simple RNN

LSTMs

Parallel Net

Best Practices for Preprocessing NLP Data

[NL preprocessing best practices notebook]



Review

L for NL

Representations

Word Vectors
Vector-Space
Embedding
word2vec

Modeling Ni Data

ROC Curve Sentiment

RNNs Simple RNNs

Parallel Nets

1 Review Take-Home Exercise

The Power and Elegance of Deep Learning for NLP Introduction to DL for NLP Computational Representations of NL NLP Applications

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





Review

I for NI

Intro Representations

Mard Vactors

Vector-Space Embedding word2vec Creating Word

Modeling NI

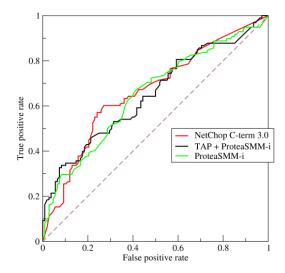
Preprocessing

Sentiment Classification

RNNs Simple RNN

Parallel Nets

The Area Under the ROC Curve





Review

OI for NI

Intro
Representations
NLP Application

Word Vectors

Vector-Space
Embedding

word2vec

Creating Word

Modeling N Data

Preprocessir ROC Curve Sentiment

Classification

Simple RNNs LSTMs

Parallal Note

- 2 The Power and Elegance of Deep Learning for NLP Introduction to DL for NLP Computational Representations of NL NLP Applications
 - 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

I for NI

Intro

...- . ..

Word Vectors

Vector-Space

Embedding

Propting Word

Madalina NI

Data

Preprocess

Continued

Sentiment Classification

RNN:

Simple RNNs

Parallel Net

Dense Net Classification

[dense sentiment classifier notebook]



Review

OL for NI

Intro

nepresentation

Word Vector

Vector-Space

Embedding

Creating Word

Modeling NL

Data

ROC Curve

Sentiment

Classification

HININS

LSTMs

Parallel Nets

ConvNet Classification

[convolutional sentiment classifier notebook]



RNNs

- Introduction to DL for NLP NLP Applications
 - **Vector-Space Embedding** Creating Word Vectors with word2vec
- Best Practices for Preprocessing NLP Data The Area Under the ROC Curve
- Recurrent Neural Networks



Review

OI for NI

Representations

Word Vectors

Vector-Space
Embedding

word2vec

Modeling N

Preprocessing ROC Curve Sentiment Classification

KININS Simple RNNs

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

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

6 Parallel Network Architectures



RNN Theory

Review

I for NI

Intro Representations

Word Vectors

Vector-Space Embedding word2vec Creating Word

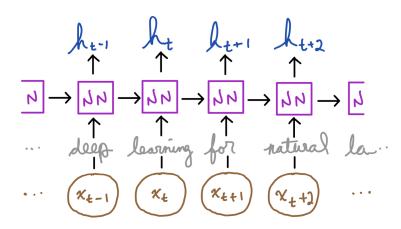
Modeling NI

Preprocessing ROC Curve Sentiment Classification

RNN:

Simple RNNs LSTMs

Parallel Nets





Review

DI for NI

Intro

nepresentation

Word Vectors

Vector-Space Embedding

word2vec

Creating Word

Modeling NL

Data

ROC Curve

Sentiment Classification

RNNs

Simple RNNs

Parallal Nate

RNNs in Practice

[rnn notebook]



Review

L for NL

Intro
Representations
NI P Application

Word Vectors

Vector-Space
Embedding

Embedding word2vec Creating Word Vectors

Modeling N Data

Preprocessing ROC Curve Sentiment Classification

Simple RNN

Dawallal Mata

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

5 Recurrent Neural Networks
Simple RNNs

LSTMs





LSTM Theory

Review

I for NI

Intro Representations

Mord Vootoro

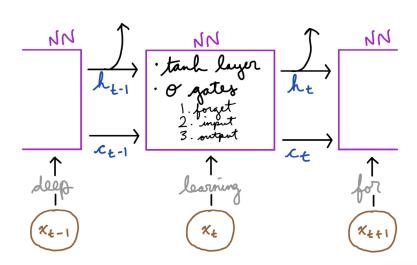
Vector-Space Embedding word2vec Creating Word

Modeling NL

Preprocessing ROC Curve Sentiment Classification

RNNs Simple RNI LSTMs

Parallal Nate





Review

OL for NI

Intro

.. - . ..

Word Vector

Vector-Space

word?woo

Creating Word

Modeling NL

Preprocessin ROC Curve

Sentiment Classification

Simple RNI

LSTMs

Parallel Nets

LSTMs in Practice

[vanilla LSTM and GRU notebooks]



Review

I for NI

Intro

Representation

Word Vectors

Vector-Space Embedding

Creating Word

Modeling NI

Data Preprocessing

ROC Curve Sentiment

RNNs

Simple RNNs LSTMs

Parallel Nets

Bi-Directional LSTMs

[Bi-LSTM notebook]



Review

OL for NI

Representation

Word Vecto

Vector-Space Embedding word2vec

Creating Word Vectors

Modeling NL

Preprocessing ROC Curve Sentiment

RNNs

.

Stacked LSTMs

[stacked LSTM and ye olde stackeroo notebooks]



Review

DL for NL

Intro
Representations
NI P Applications

Word Vectors

Vector-Space
Embedding

Embedding word2vec Creating Word Vectors

Modeling N Data

Preprocessing ROC Curve Sentiment Classification

KNNS Simple RNN

Parallel Nets

- 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
- 5 Recurrent Neural Networks Simple RNNs LSTMs
- 6 Parallel Network Architectures



Review

I for NI

Intro

.. -

Word Vecto

TTOTA TOOLS

vector-Space

ward Ouan

Creating Wor

Modeling NL

Data

ROC Curve

Sentiment Classificatio

Circula D

Parallel Nets

Parallel Network Architectures

[multi-ConvNet notebook]



Review

I for NIL

Intro Representation

Word Voctor

Embedding word2vec Creating Wor

Modeling NL

Preprocessing ROC Curve Sentiment Classification

RNNs

Simple RNNs LSTMs

Parallel Nets

Assessing Your Deep Learning Project III





Review

DL for NL

Representation

Word Vector

.. -

Embeddir

word2vec

vectors

Data

Preprocessin ROC Curve Sentiment

RNNs Simple RNN

Parallel Nets

Assessing

- 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" (?)



Review

DL for NL

Representation

Word Vector

Vector-Space Embedding

word2vec Creating Word

Modeling NL Data

Preprocessing ROC Curve Sentiment

RNNs Simple RNN

LSTMs

Parallel Nets

Assessing

- 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" (?)



Review

DL for NL

Representation

Word Vectors

Vector-Space Embedding word2vec Creating Word

Modeling NL Data

Preprocessing ROC Curve Sentiment Classification

RNNs Simple RNN

LSTMs

Parallel Nets

Assessing

- 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" (?)



Review

L for NL

Intro
Representation

Word Vectors

Vector-Space Embedding word2vec Creating Word Vectors

Modeling NL Data

Preprocessing ROC Curve Sentiment Classification

Simple RNN

Parallel Nets

Assessing

- 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" (?)



Review

L for NL

Representation

Word Vector

Vector-Space Embedding word2vec Creating Word Vectors

Modeling NI Data

Preprocessing ROC Curve Sentiment Classification

Simple RNN

LSTMs

Parallel Nets

Assessing

- 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" (?)



Review

L for NL

Intro
Representation
NLP Application

Word Vector

Vector-Space Embedding word2vec Creating Word Vectors

Modeling NI Data

Preprocessing ROC Curve Sentiment Classification

KNNS Simple RNN

Parallel Nets

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" (?)



Review

L for NL

Representation

Word Vector

Vector-Space Embedding word2vec Creating Word Vectors

Modeling NI Data

Preprocessing ROC Curve Sentiment Classification

KNNS Simple RNN

Parallel Nets

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" (?)



Review

L for NL

Intro
Representation

Word Vector

Vector-Space Embedding word2vec Creating Word Vectors

Modeling NI Data

Preprocessing ROC Curve Sentiment Classification

Simple RNN

LSTMs

Parallel Nets

Assessing

- 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" (?)



Review

L for NL

Representations

Word Vector

Vector-Space Embedding word2vec Creating Word Vectors

Modeling NI Data

Preprocessing ROC Curve Sentiment Classification

Simple RNN

LSTMs

Parallel Nets

Assessing

- 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" (?)



Review

OL for NI

. . .

Representation

Word Vector

Vector-Space

word2vec

Modeling NI

Data

ROC Curve

Classificati

RNN:

Simple RNNs

Parallel Nets

Up Next: TensorFlow

