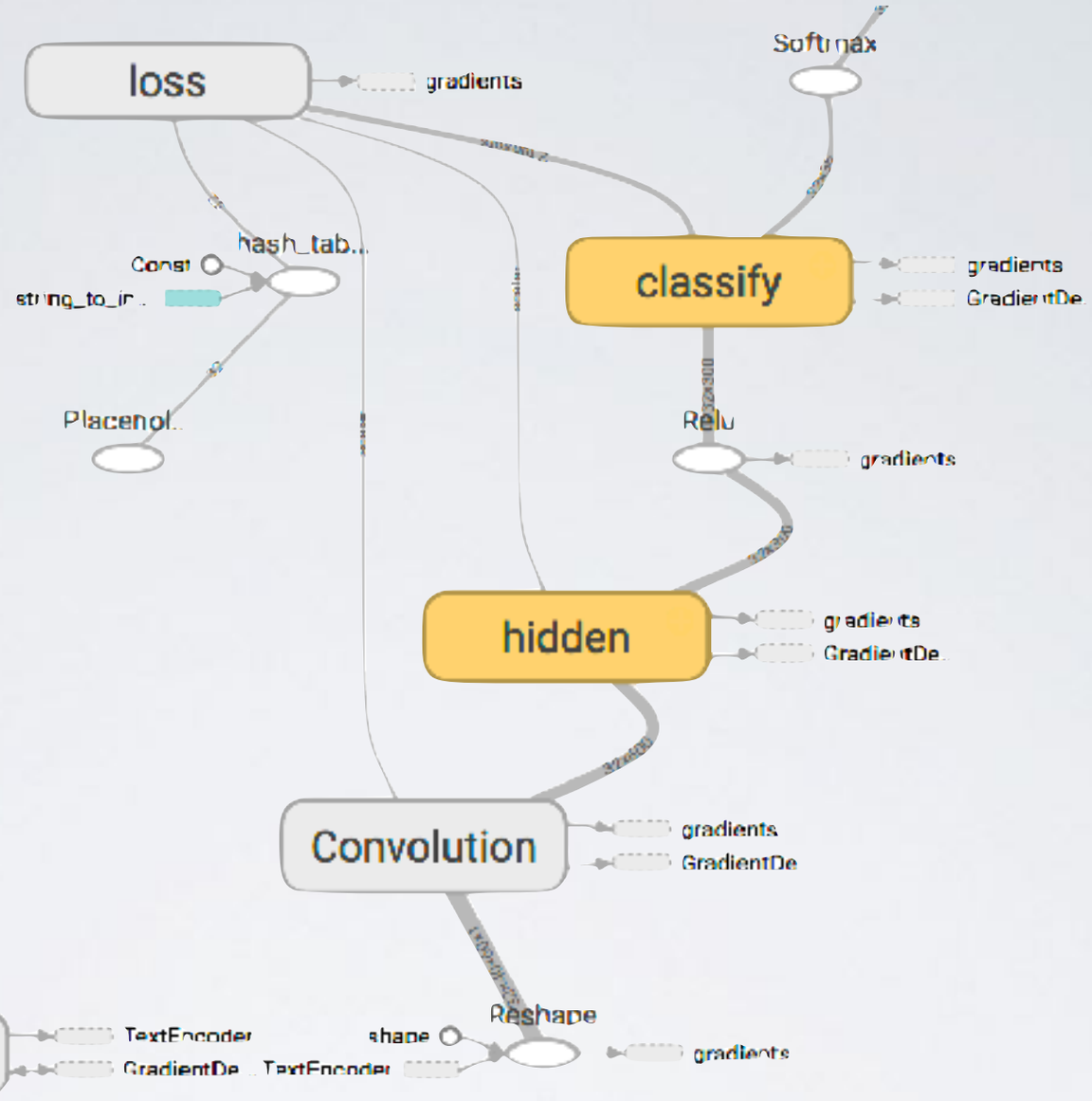


NLP FROM SCRATCH IN TF

Neal Lewis

OVERVIEW

- (Some) NLP Tasks
- Prior to NN for NLP
- NLP (Almost) From Scratch
- Text Representation for MLP and ConvNet
- https://github.com/nrlewis/nlpfromscratch-tf

HELPFUL REFERENCES

- Natural Language Processing (almost) from Scratch
  - Collobert et. al
  - https://arxiv.org/abs/1103.0398
- A Primer on Neural Network Models for Natural Language Processing
  - Yoav Goldberg
  - http://u.cs.biu.ac.il/~yogo/nnlp.pdf
- Implementing CNN for Text Classification
  - http://www.wildml.com/2015/12/implementing-a-cnn-for-text-classification-in-tensorflow/
  - Good for CNNs

NOTES ABOUT CODE

- Implements MLP and Sentence Convolution Approach
- Contains training data from CONLL for POS Tagging and Chunking
- Tensorboard
  - Writes the vocab for embedding PCA
  - Has metrics for Precision, Recall, and F1 Measures
- Please check out!
- If problems, please let me know!

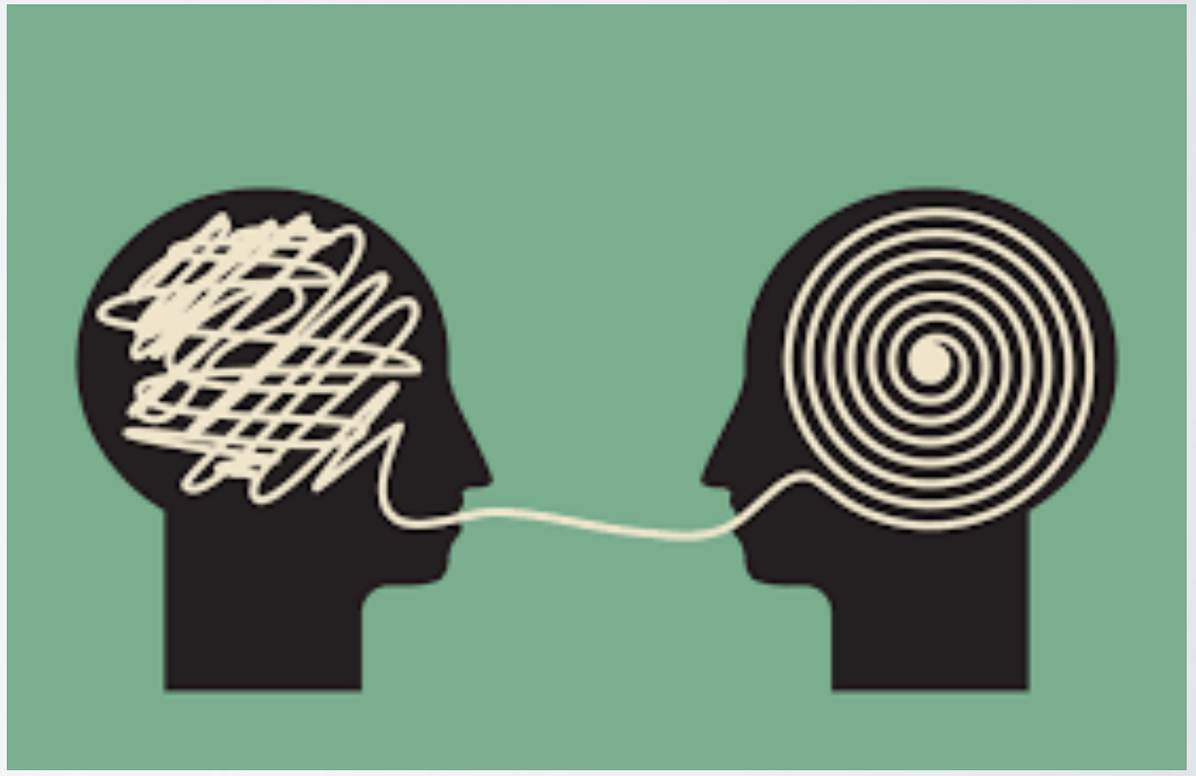
NATURAL LANGUAGE PROCESSING

Processing natural language corpora in a way to extract meaningful information

NATURAL LANGUAGE PROCESSING

Processing natural language corpora in a way to extract meaningful information

- Language Translation



NATURAL LANGUAGE PROCESSING

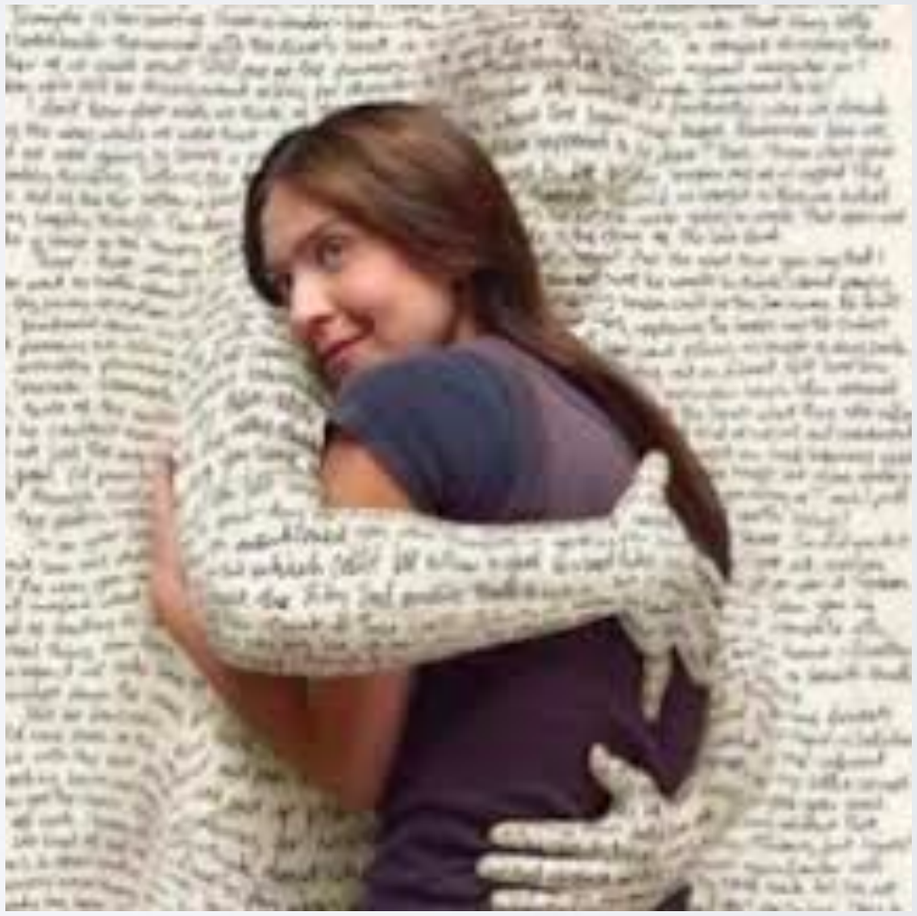
Processing natural language corpora in a way to extract meaningful information

- Language Translation
- Speech Recognition

NATURAL LANGUAGE PROCESSING

Processing natural language corpora in a way to extract meaningful information

- Language Translation
- Speech Recognition
- Text Understanding



TOKEN CLASSIFICATION

"Apple CEO Steve Jobs unveiled the iPhone."

TOKEN CLASSIFICATION

ORG

"Apple CEO Steve Jobs unveiled the iPhone."

TOKEN CLASSIFICATION

ORG

PERSON

"Apple CEO Steve Jobs unveiled the iPhone."

TOKEN CLASSIFICATION

PERSON

ORG

B-PER

I-PER

"Apple CEO Steve Jobs unveiled the iPhone."

TOKEN CLASSIFICATION

PERSON

ORG

B-PER

I-PER

PRODUCT

"Apple CEO Steve Jobs unveiled the iPhone."

TOKEN CLASSIFICATION

PERSON

ORG O B-PER I-PER O O PRODUCT

"Apple CEO Steve Jobs unveiled the iPhone."

TOKEN CLASSIFICATION

Named Entity Recognition (NER)

PERSON

ORG O B-PER I-PER O O PRODUCT

"Apple CEO Steve Jobs unveiled the iPhone."

TOKEN CLASSIFICATION

Named Entity Recognition (NER)

PERSON

ORG O B-PER I-PER O O PRODUCT

"Apple CEO Steve Jobs unveiled the iPhone."

NNP NNP NNP NNP

TOKEN CLASSIFICATION

Named Entity Recognition (NER)

PERSON

ORG O B-PER I-PER O O PRODUCT

"Apple CEO Steve Jobs unveiled the iPhone."

NNP NNP NNP NNP VERB

TOKEN CLASSIFICATION

Named Entity Recognition (NER)

PERSON

ORG O B-PER I-PER O O PRODUCT

"Apple CEO Steve Jobs unveiled the iPhone."

NNP NNP NNP NNP VERB DET

TOKEN CLASSIFICATION

Named Entity Recognition (NER)

PERSON

ORG O B-PER I-PER O O PRODUCT

"Apple CEO Steve Jobs unveiled the iPhone."

NNP NNP NNP NNP VERB DET NNP

TOKEN CLASSIFICATION

Named Entity Recognition (NER)

PERSON

ORG O B-PER I-PER O O PRODUCT

"Apple CEO Steve Jobs unveiled the iPhone."

NNP NNP NNP NNP VERB DET NNP

Part Of Speech Tagging (POS)

TOKEN CLASSIFICATION

Named Entity Recognition (NER)

PERSON

ORG O B-PER I-PER O O PRODUCT

"Apple CEO Steve Jobs unveiled the iPhone."

NNP NNP NNP NNP VERB DET NNP

Part Of Speech Tagging (POS)

MORE CLASSIFICATION

"Apple CEO Steve Jobs unveiled the iPhone."

More Tasks

- Chunking
- Semantic Role Labeling
- Syntax Parsing
- Dependency Parsing

MORE NLP CLASSIFICATION

"Apple CEO Steve Jobs unveiled the iPhone."

More Tasks

- Chunking
- Semantic Role Labeling
- Syntax Parsing
- Dependency Parsing…

Prior Approaches

- Support Vector Machines
- Conditional Random Fields
- Logistic Regression
- Ensemble… And More!

FEATURE ENGINEERING

- Each task had a separate state of the art algorithm
- Each separate state of the art algorithm had many, many features

FEATURE ENGINEERING

- Each task had a separate state of the art algorithm
- Each separate state of the art algorithm had many, many features

- POS tags
- Word Suffixes
- Word Prefixes
- Word Positions relative *
- External Data (WordNet)
- Gazeteer
- Active / Passive Voice
- Word Shapes
- Lemmas
- Stems
- …… And More!

FEATURE ENGINEERING

- Each task had a separate state of the art algorithm
- Each separate state of the art algorithm had many, many features

- POS tags
- Word Suffixes
- Word Prefixes
- Word Positions relative *
- External Data (WordNet)

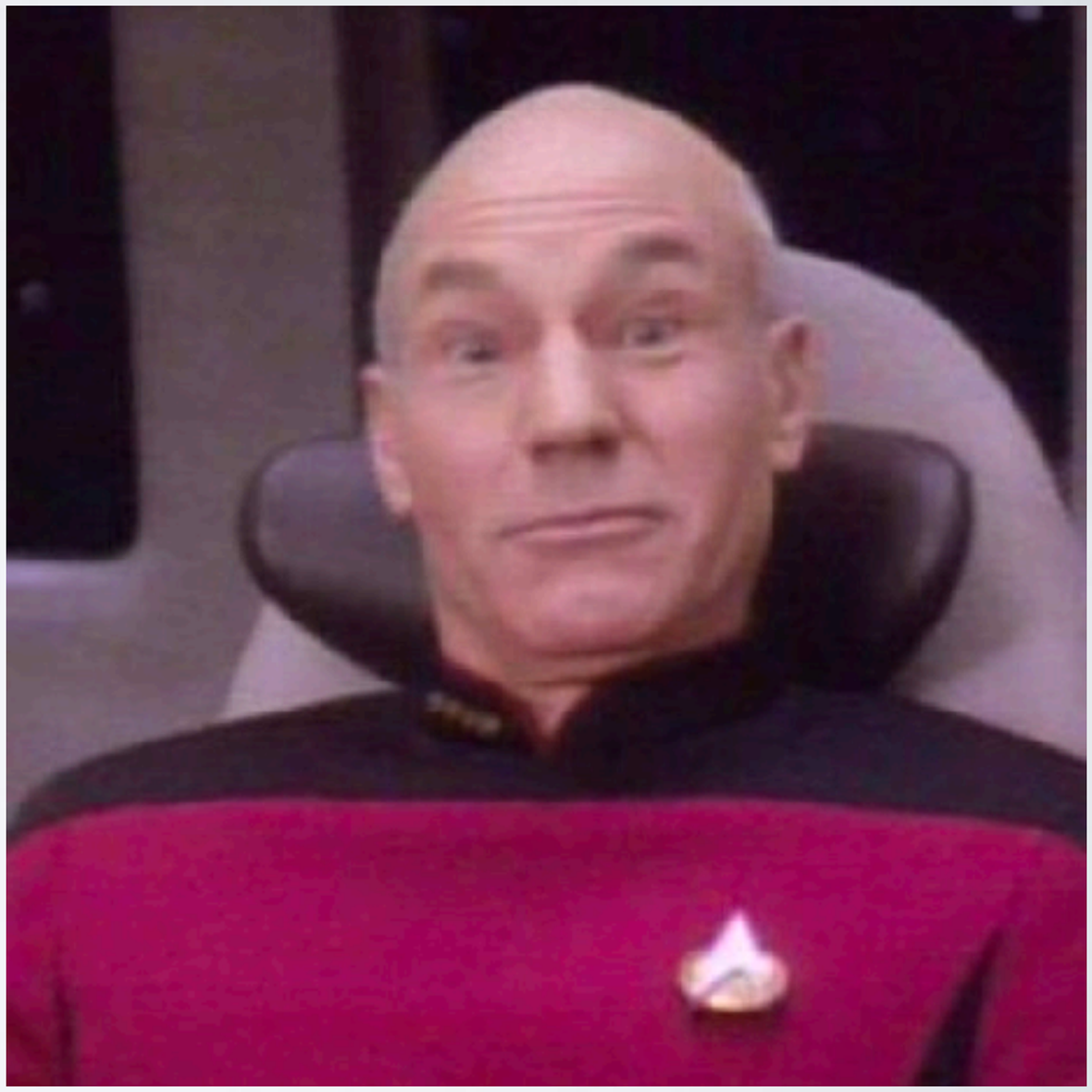
- Gazeteer
- Active / Passive Voice
- Word Shapes
- Lemmas
- Stems
- …… And More!

NLP (ALMOST) FROM SCRATCH!

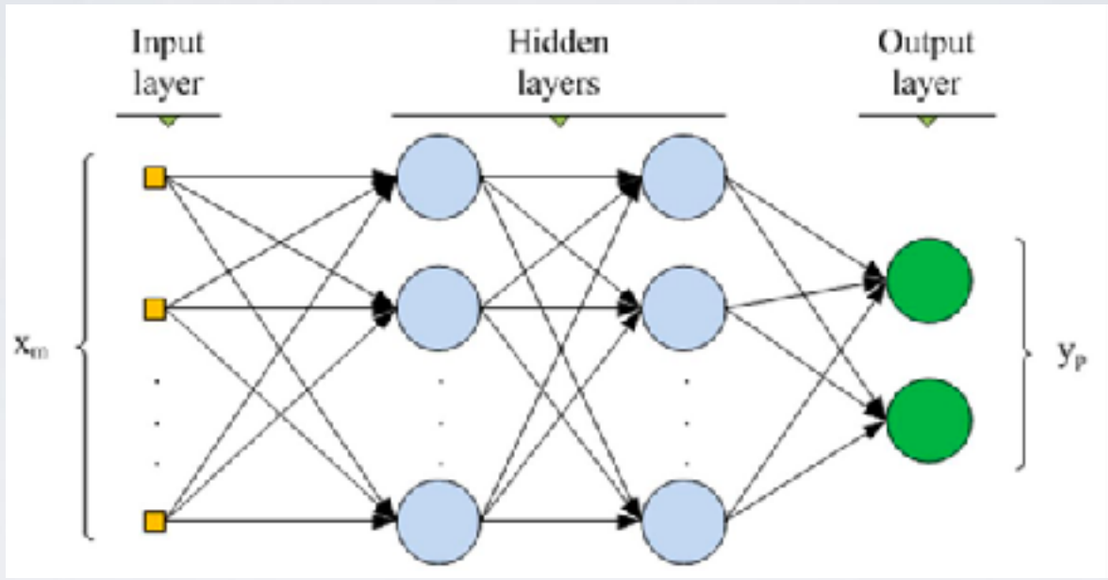
- 2011
- Two architectures for all tasks
- Performed near state the art
- No (almost) feature engineering!
- Using Neural Networks!

NLP (ALMOST) FROM SCRATCH!

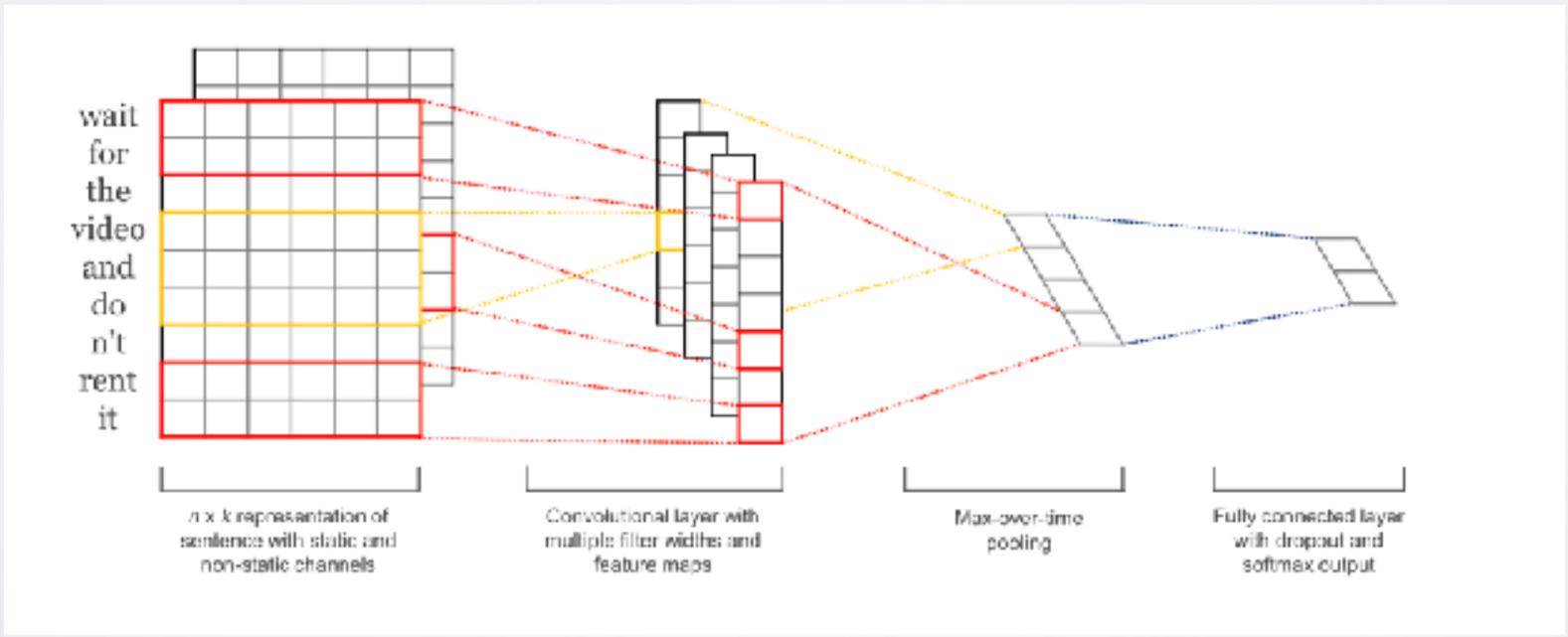
- 2011
- Two architectures for all tasks
- Performed near state the art
- (Almost) no feature engineering!
- Using Neural Networks!



Multilayer Perceptron

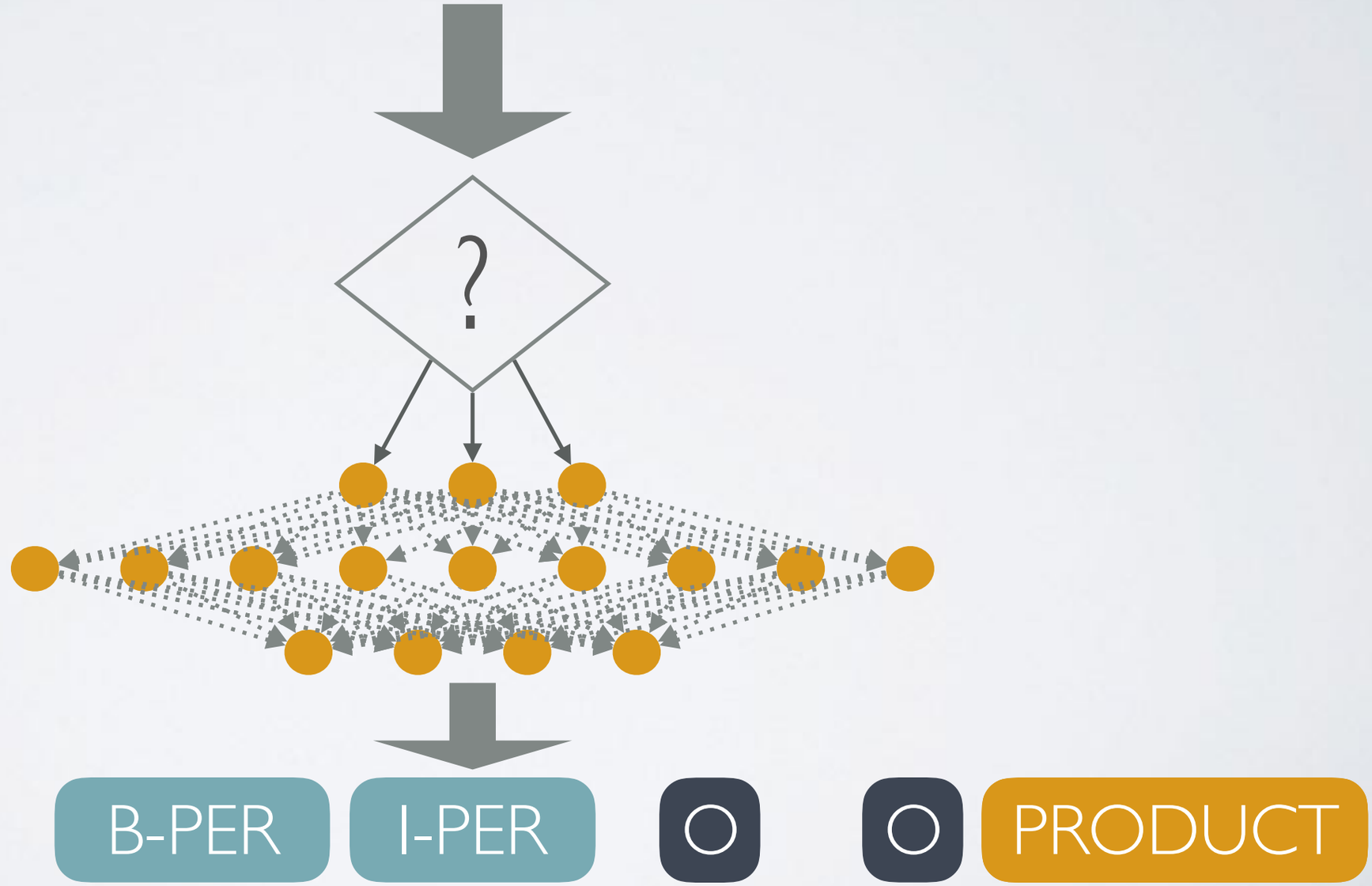


Convolutional Neural Network



HOW DO WE REPRESENT INPUT?

"Apple CEO Steve Jobs unveiled the iPhone."



DISCRETE INPUT REPRESENTATION

Original			
Apple			
CEO			
Steve			
Jobs			
unveiled			
the			
iPhone			
.			

DISCRETE INPUT REPRESENTATION

Original			LABEL
Apple			B-ORG
CEO			O
Steve			B-PER
Jobs			I-PER
unveiled			O
the			O
iPhone			B-PROD
.			O

MINIMAL PREPROCESSING

	INPUT		
Original	Token	Caps	LABEL
Apple	apple	TITLE	B-ORG
CEO	ceo	UPPER	O
Steve	steve	TITLE	B-PER
Jobs	jobs	TITLE	I-PER
unveiled	unvieled	LOWER	O
the	the	LOWER	O
iPhone	iphone	MIXED	B-PROD
.	.	PUNCT	O

DENSE FEATURE VECTORS LOOKUP TABLES

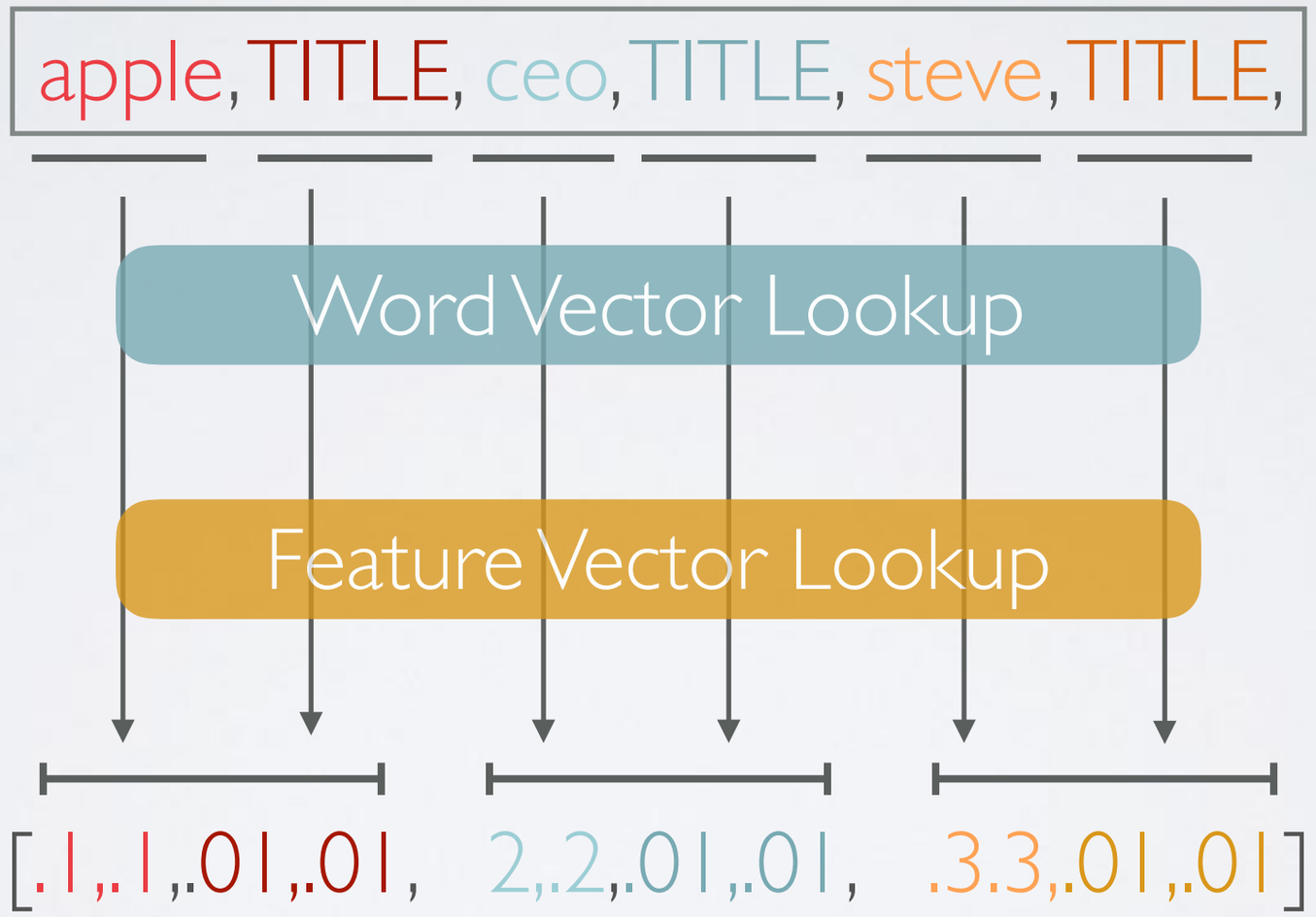
Token (Key)	Word Vectors (Value)
apple	[.1, .1, .1, .1, .1, .1, .1, .1, .1, .1]
ceo	[.2, .2, .2, .2, .2, .2, .2, .2, .2, .2]
steve	[.3, .3, .3, .3, .3, .3, .3, .3, .3, .3]
jobs	[.4, .4, .4, .4, .4, .4, .4, .4, .4, .4,]
unvieled	[.5, .5, .5, .5, .5, .5, .5, .5, .5, .5]
the	[.6, .6, .6, .6, .6, .6, .6, .6, .6, .6]
iphone	[.7, .7, .7, .7, .7, .7, .7, .7, .7, .7]

Feature (Key)	Feature Vectors (Value)
TITLE	[.01, .01,.01]
LOWER	[.02, .02,.02]
pos_-1	[.03, .03,.03]
pos_0	[.04, .04,.04]
pos_1	[.05, .05,.05]

LOOKUP AND CONCAT

apple,TITLE, ceo,TITLE, steve,TITLE,

LOOKUP AND CONCAT



WINDOW AND CONCAT

Token	apple	ceo	steve	jobs	unvield	the	iphone
Caps	TITLE	TITLE	TITLE	TITLE	LOWER	LOWER	MIXED
Label	B-ORG	O	B-PER	I-PER	O	O	B-PROD

Input #	w_-1	f_-1	w_0	f_0	w_1	f_1	Label
1	PAD	PAD	**apple**	**TITLE**	ceo	TITLE	B-ORG
2	apple	TITLE	**ceo**	**TITLE**	steve	TITLE	O
3	ceo	TITLE	**steve**	**TITLE**	jobs	TITLE	B-PER
4	steve	TITLE	**jobs**	**TITLE**	unveile	LOWER	I-PER
5	jobs	TITLE	**unveile**	**LOWER**	the	LOWER	O

WINDOW INPUT

Token	apple	ceo	steve	jobs	unvield	the	iphone
Caps	TITLE	TITLE	TITLE	TITLE	LOWER	LOWER	MIXED
Label	B-ORG	O	B-PER	I-PER	O	O	B-PROD

Input #	w_-1	f_-1	w_0	f_0	w_1	f_1	Label
1	PAD	PAD	**apple**	**TITLE**	ceo	TITLE	B-ORG
2	apple	TITLE	**ceo**	**TITLE**	steve	TITLE	O
3	ceo	TITLE	**steve**	**TITLE**	jobs	TITLE	B-PER
4	steve	TITLE	**jobs**	**TITLE**	unveile	LOWER	I-PER
5	jobs	TITLE	**unveile**	**LOWER**	the	LOWER	O

WINDOW INPUT

Token	apple	ceo	steve	jobs	unvield	the	iphone
Caps	TITLE	TITLE	TITLE	TITLE	LOWER	LOWER	MIXED
Label	B-ORG	O	B-PER	I-PER	O	O	B-PROD

Input #	w_-1	f_-1	w_0	f_0	w_1	f_1	Label
1	PAD	PAD	**apple**	**TITLE**	ceo	TITLE	B-ORG
2	apple	TITLE	**ceo**	**TITLE**	steve	TITLE	O
3	ceo	TITLE	**steve**	**TITLE**	jobs	TITLE	B-PER
4	steve	TITLE	**jobs**	**TITLE**	unveile	LOWER	I-PER
5	jobs	TITLE	**unveile**	**LOWER**	the	LOWER	O

WINDOW INPUT

Token	apple	ceo	steve	jobs	unvield	the	iphone
Caps	TITLE	TITLE	TITLE	TITLE	LOWER	LOWER	MIXED
Label	B-ORG	O	B-PER	I-PER	O	O	B-PROD

Input #	w_-1	f_-1	w_0	f_0	w_1	f_1	Label
1	PAD	PAD	**apple**	**TITLE**	ceo	TITLE	B-ORG
2	apple	TITLE	**ceo**	**TITLE**	steve	TITLE	O
3	ceo	TITLE	**steve**	**TITLE**	jobs	TITLE	B-PER
4	steve	TITLE	**jobs**	**TITLE**	unveile	LOWER	I-PER
5	jobs	TITLE	**unveile**	**LOWER**	the	LOWER	O

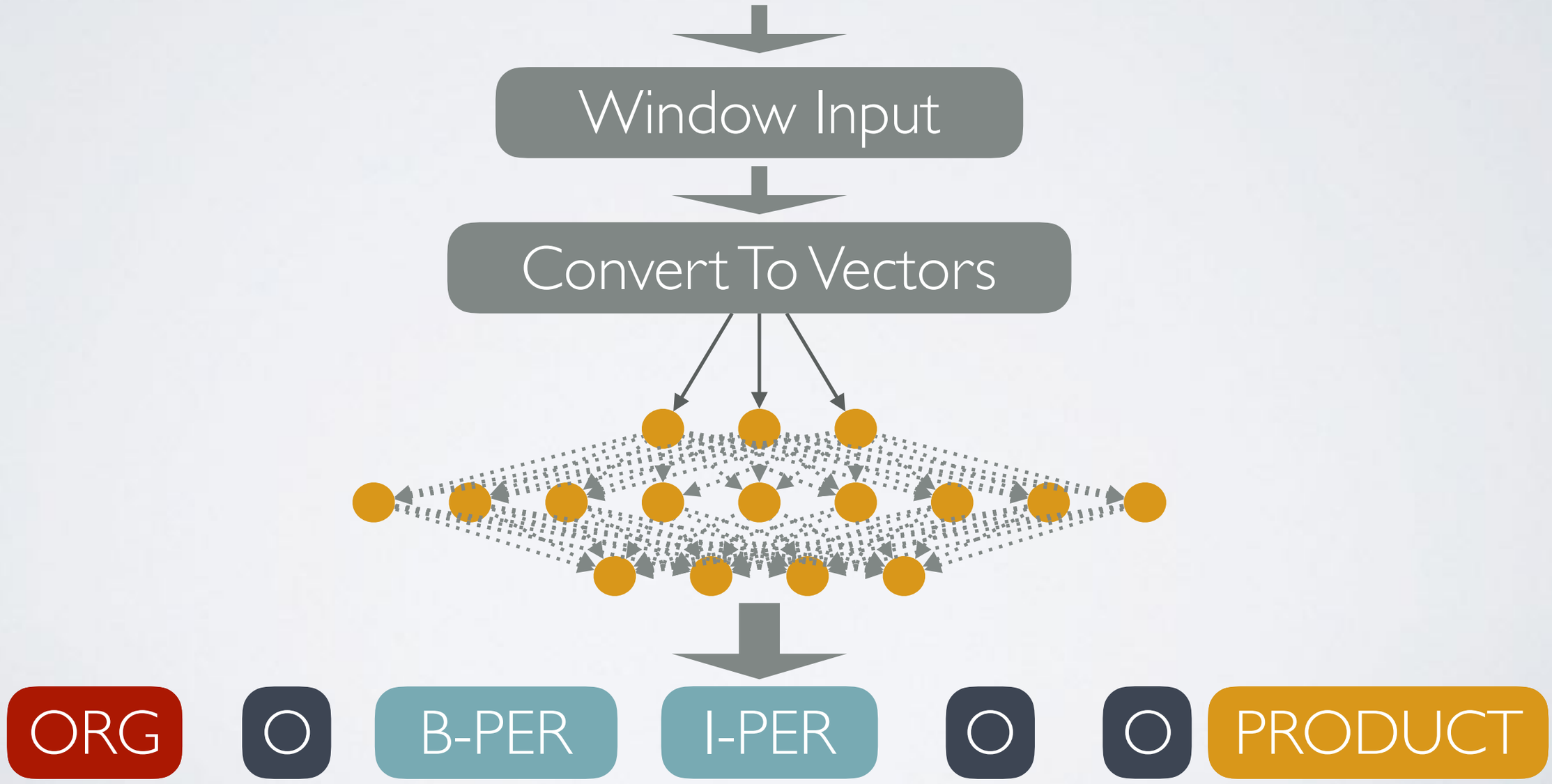
CONVERT TO VECTOR

Token	apple	ceo	steve	jobs	unvield	the	iphone
Caps	TITLE	TITLE	TITLE	TITLE	LOWER	LOWER	MIXED
Label	B-ORG	O	B-PER	I-PER	O	O	B-PROD

Input #	w_-1; f_-1; w_0; f_0; w_1; f_1	Label
1	[0.0, 0.0; 0.0, 0.0; **.1, .1; .01, .01;** .2, .2; .01, .01]	B-ORG
2	[.2, .2; .01, .01; **.3, .3; .01, .01;** .4, .4; .01, .01]	O
3	[.3, .3; .01, .01; **.4, .4; .01, .01;** .5. ,5; .02, .02]	B-PER
4	[.4, .4; .01, .01; **.5, .5; .02, .02;** .6, .6; .02, .02]	I-PER
5	[.5, .5; .02, .02; **.6, .6; .02, .02;** .7, .7; .03, .03]	O

WINDOWING INTO MLP

"Apple CEO Steve Jobs unveiled the iPhone."



SENTENCE CONVOLUTION INPUT

Token	apple	ceo	steve	jobs	unvield	the	iphone
Caps	TITLE	TITLE	TITLE	TITLE	LOWER	LOWER	MIXED
Label	B-ORG	O	B-PER	I-PER	O	O	B-PROD

Input #	w_0	f_0,0	f_0,1	w_1	f_1,0	f_1,1	w_2	f_2,0	f_2,1	Label
1	**apple**	**TITLE**	**pos_0**	ceo	TITLE	pos_1	steve	TITLE	pos_2	B-ORG
2	apple	TITLE	pos_-1	**ceo**	**TITLE**	**pos_0**	steve	TITLE	pos_1	O
3	apple	TITLE	pos_-2	ceo	TITLE	pos_-1	**steve**	**TITLE**	**pos_0**	B-PER

SENTENCE CONVOLUTION INPUT

Token	apple	ceo	steve	jobs	unvield	the	iphone
Caps	TITLE	TITLE	TITLE	TITLE	LOWER	LOWER	MIXED
Label	B-ORG	O	B-PER	I-PER	O	O	B-PROD

Relative Position Vectors

Input #	w_0	f_0,0	f_0,1	w_1	f_1,0	f_1,1	w_2	f_2,0	f_2,1	Label
1	**apple**	**TITLE**	**pos_0**	ceo	TITLE	pos_1	steve	TITLE	pos_2	B-ORG
2	apple	TITLE	pos_-1	**ceo**	**TITLE**	**pos_0**	steve	TITLE	pos_1	O
3	apple	TITLE	pos_-2	ceo	TITLE	pos_-1	**steve**	**TITLE**	**pos_0**	B-PER

SENTENCE CONVOLUTION INPUT

Token	apple	ceo	steve	jobs	unvield	the	iphone
Caps	TITLE	TITLE	TITLE	TITLE	LOWER	LOWER	MIXED
Label	B-ORG	O	B-PER	I-PER	O	O	B-PROD

Input	w_0; f_0,0; f_0,1; w_1; f_1,0; f_1,1; w_2; f_2,0; f_2,1	Label
1*	[**.1, .1; .01, .01; .04, .04;** 2, .2; 01,.01; .05, .05; .3, .3; .01,.01;.06, .06]	B-ORG
2*	[.1, .1; .01, .01; .-05, .-05; **2, .2; .01, .01; .04,.04;**.3.3;.01,.01;.05,.05]	O
3*	[.1,.1; .01, .01; .-06,.-06; .2, .2; .01, .01; .-05, .05, **.3.3, .01,.01,.04,.04**]	B-PER

RESHAPE FOR CONVOLUTION

apple, **TITLE**, **pos_0**, ceo,TITLE, pos_1, steve,TITLE, pos_2,…

[**.1,.1,.01,.01,.04,.04,** 2,.2,.02,.02,.05, .05,.3.3,.03,.03,.06,.06,…]

(words for clarity)	2D Input Sentence
apple; TITLE; pos_0	**.1,.1; .01,.01; .04,.04**
ceo;TITLE; pos_1	.2,.2; .01,.01; .05, .05
steve;TITLE; pos_2	.3.3; .01,.01; .06,.06
jobs;TITLE; pos_3	.4.4; .01,.01; .07,.07
unveiled; LOWER; pos_4	.5.5; .02,.02; .08,.08

RESHAPE FOR CONVOLUTION

apple, **TITLE**, **pos_0**, ceo,TITLE, pos_1, steve,TITLE, pos_2,…

[**.1,.1,.01,.01,.04,.04,** 2,.2,.02,.02,.05, .05,.3.3,.03,.03,.06,.06,…]

(words for clarity)	2D Input Sentence
apple; TITLE; pos_0	**.1,.1; .01,.01; .04,.04**
ceo;TITLE; pos_1	.2,.2; .01,.01; .05, .05
steve;TITLE; pos_2	.3.3; .01,.01; .06,.06
jobs;TITLE; pos_3	.4.4; .01,.01; .07,.07
unveiled; LOWER; pos_4	.5.5; .02,.02; .08,.08

Convolution

Output

0.342

RESHAPE FOR CONVOLUTION

apple, **TITLE**, **pos_0**, ceo,TITLE, pos_1, steve,TITLE, pos_2,…

[**.1,.1,.01,.01,.04,.04,** 2,.2,.02,.02,.05, .05,.3.3,.03,.03,.06,.06,…]

(words for clarity)	2D Input Sentence
apple; TITLE; pos_0	**.1,.1; .01,.01; .04,.04**
ceo;TITLE; pos_1	.2,.2; .01,.01; .05, .05
steve;TITLE; pos_2	.3.3; .01,.01; .06,.06
jobs;TITLE; pos_3	.4.4; .01,.01; .07,.07
unveiled; LOWER; pos_4	.5.5; .02,.02; .08,.08

Convolution

Output

0.342
0.673

RESHAPE FOR CONVOLUTION

apple, **TITLE**, **pos_0**, ceo,TITLE, pos_1, steve,TITLE, pos_2,…

[**.1,.1,.01,.01,.04,.04,** 2,.2,.02,.02,.05, .05,.3.3,.03,.03,.06,.06,…]

(words for clarity)	2D Input Sentence
apple; TITLE; pos_0	**.1,.1; .01,.01; .04,.04**
ceo;TITLE; pos_1	.2,.2; .01,.01; .05, .05
steve;TITLE; pos_2	.3.3; .01,.01; .06,.06
jobs;TITLE; pos_3	.4.4; .01,.01; .07,.07
unveiled; LOWER; pos_4	.5.5; .02,.02; .08,.08

Convolution

Output

0.342
0.673
0.734

RESHAPE FOR CONVOLUTION

apple, **TITLE**, **pos_0**, ceo,TITLE, pos_1, steve,TITLE, pos_2,…

[**.1,.1,.01,.01,.04,.04,** 2,.2,.02,.02,.05, .05,.3.3,.03,.03,.06,.06,…]

(words for clarity)	2D Input Sentence
apple; TITLE; pos_0	**.1,.1; .01,.01; .04,.04**
ceo;TITLE; pos_1	.2,.2; .01,.01; .05, .05
steve;TITLE; pos_2	.3.3; .01,.01; .06,.06
jobs;TITLE; pos_3	.4.4; .01,.01; .07,.07
unveiled; LOWER; pos_4	.5.5; .02,.02; .08,.08

Convolution Output

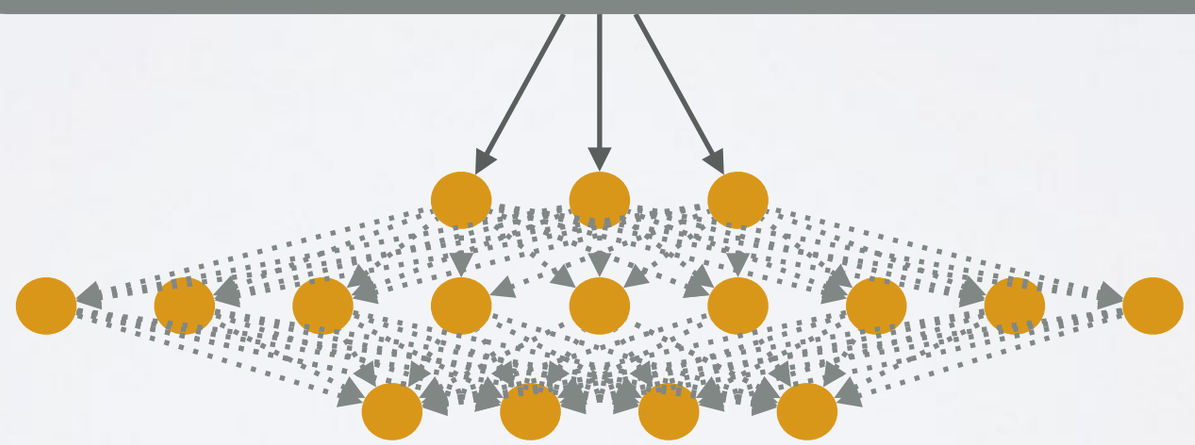
0.342
0.673
0.734
0.845

SENTENCE CONVOLUTION

"Apple CEO Steve Jobs unveiled the iPhone."

Expand Input and Convert to Vectors

Convolution and Max Pooling



ORG O B-PER I-PER O O PRODUCT

Questions?