



Deep Natural Language Processing for Search and Recommender Systems



[Weiwei Guo](#)



[Huiji Gao](#)



[Jun Shi](#)



[Bo Long](#)



[Liang Zhang](#)



[Bee-Chung Chen](#)



[Deepak Agarwal](#)

Agenda

- 1 Introduction
- 2 Deep Learning for Natural Language Processing
- 3 Deep NLP in Search and Recommender Systems
- 4 Real World Examples

Agenda

- 1 Introduction
- 2 Deep Learning for Natural Language Processing
- 3 Deep NLP in Search and Recommender Systems
- 4 Real World Examples



Introduction

Bo Long

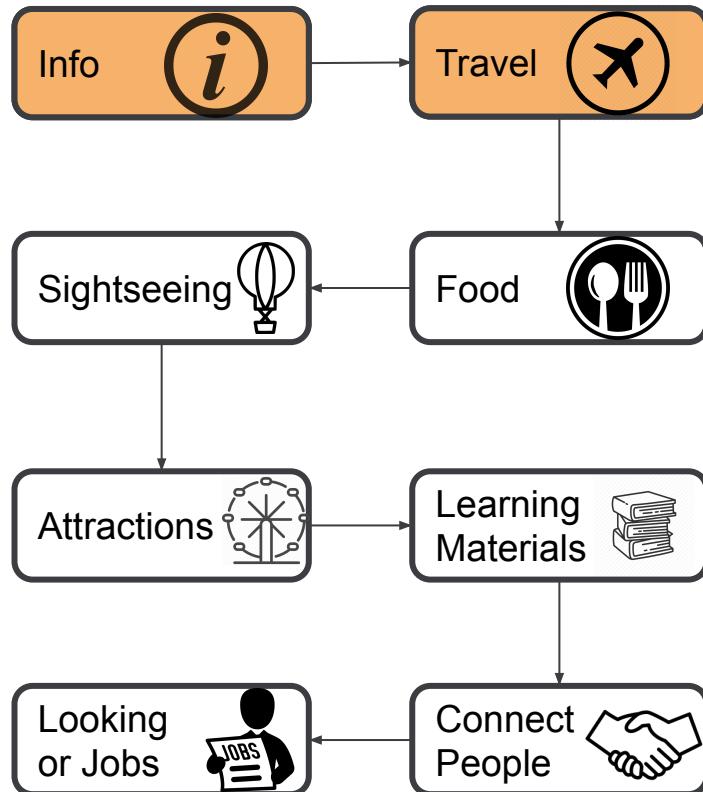
Introduction - Search and Recommender Systems



**Search and
Recommendation
are Everywhere**

Introduction - Search and Recommender Systems

KDD 2019



kdd 2019

USNEWS TRAVEL » Home Vacations Travel Guides Hotels Cruises Rewards Advice

TRAVEL / HOTELS / USA / BEST HOTELS IN ANCHORAGE

Best Hotels in Anchorage

U.S. News & World Report ranks the best hotels in Anchorage based on an analysis of industry awards, hotel star ratings and user ratings. Hotels that scored in the top 10 percent of the [Best Hotels in the USA](#) earned a Gold badge. Hotels that appear after ranked hotels are sorted by hotel class and then by user rating, as provided by TripAdvisor.

[READ THE BEST HOTELS METHODOLOGY »](#)

[See Anchorage Travel Guide »](#)

53 matches

SELECT DATES	HOTEL NAME	4.0-star HOTEL CLASS
08/14/19 → 08/15/19	The Hotel Alyeska Girdwood, AK 31.1 miles to city center [SEE MAP] #1 in Best Hotels in Anchorage TripAdvisor (2529)	Free Parking Free WiFi

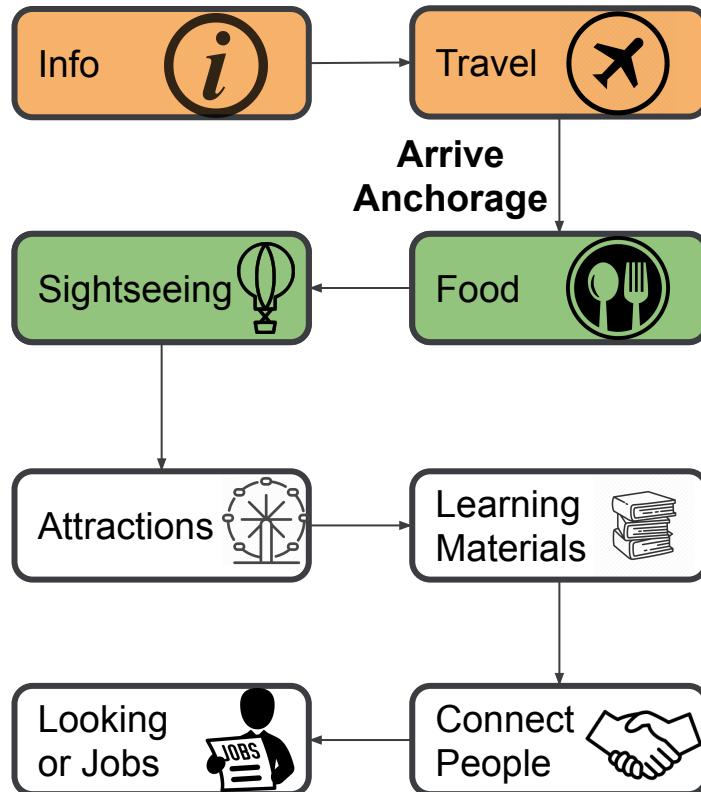
What do you need for KDD?

What is KDD Conference?

BEST HOTELS
U.S. NEWS
RANKINGS
2019

Introduction - Search and Recommender Systems

KDD 2019



A screenshot of a search interface, likely Yelp, showing results for "best aurora views near anchorage".

Search bar: best aurora views near anchorage

Filter Results

Best

- POSTS FROM
 - Anyone
 - You
 - Your Friends
 - Your Groups and Pages
 - Public
 - + Choose a Source...

All Results

Links

The Best Northern Lights Viewing Spots Near Anchorage, Alaska
The glow from streetlights and human infrastructure can rub the edge off an aurora display. But this list of spots on the... alaska.org Nov 28, 2018 · 2 Shares

Dale Reirson
Lieutenant/Firefighter at UTTR DoD · Near Anchorage, Alaska
Oct 20, 2017 · My Skyline View is filled with the Aurora Borealis. This is the best I could get at 4 am near a LED street light outside my house. Cool picture anyway. Good Morning Everyone time for this guy to catch some Zzz's.
1 Comment

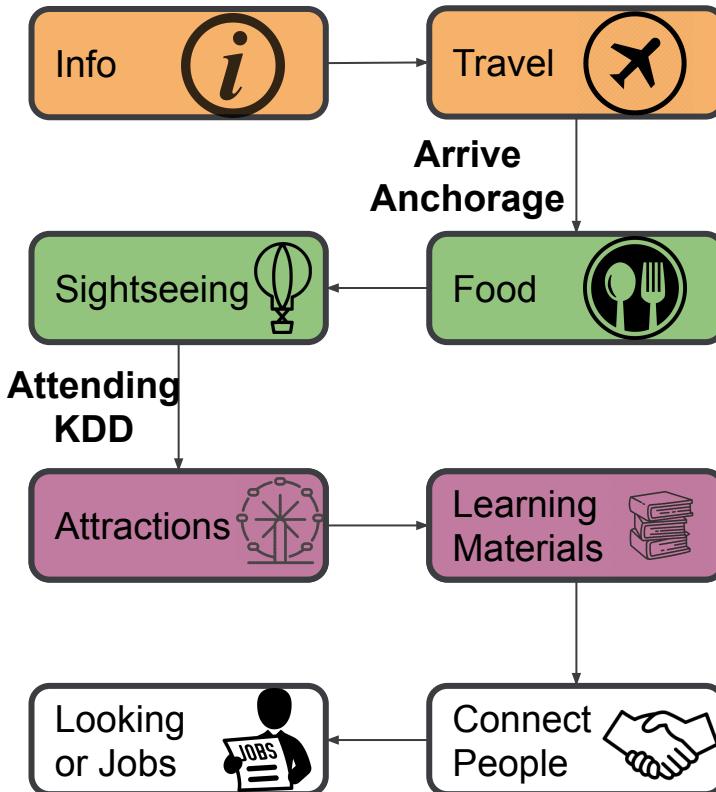
Videos

Phuket (ภูเก็ต), pronounced (roughly) "puh-KET", is... iholiday.com November 28, 2018 · 32 Views

DATE POSTED

Introduction - Search and Recommender Systems

KDD 2019



anchorage attractions 5 days

amazon prime

All deep natural language processing for search and recommender systems

Deliver to Ye Sunnyvale 94085 Your Pickup Location Browsing History Today's Deals Hello, Ye Account & Lists Buy Again Gift Cards EN Orders Prime

1 result for "deep natural language processing for search and recommender systems"

Department Books Kindle Store Condition New

Packt SPONSORED BY PACKT PUBLISHING Real World Data Science with Python Shop now

Python Machine Learning: Machine Learning and Deep Learning with Python, ... ★★★★☆ 51 prime

Artificial Intelligence with Python: A Comprehensive Guide to Building Intelligent ... ★★★★☆ 18 prime

Maturity and Innovation in Digital Libraries: 20th International Conference on Asia-Pacific Digital Libraries, ICDL 2018, Hamilton, New Zealand, by Milena Dobreva, Annika Hinze, et al. | Nov 15, 2018 Paperback \$1735 \$79.99 prime FREE One-Day. Get it Tomorrow, Aug 3 Kindle \$1648 \$79.99

Use fewer keywords or try these instead
deep language processing See all 368 results

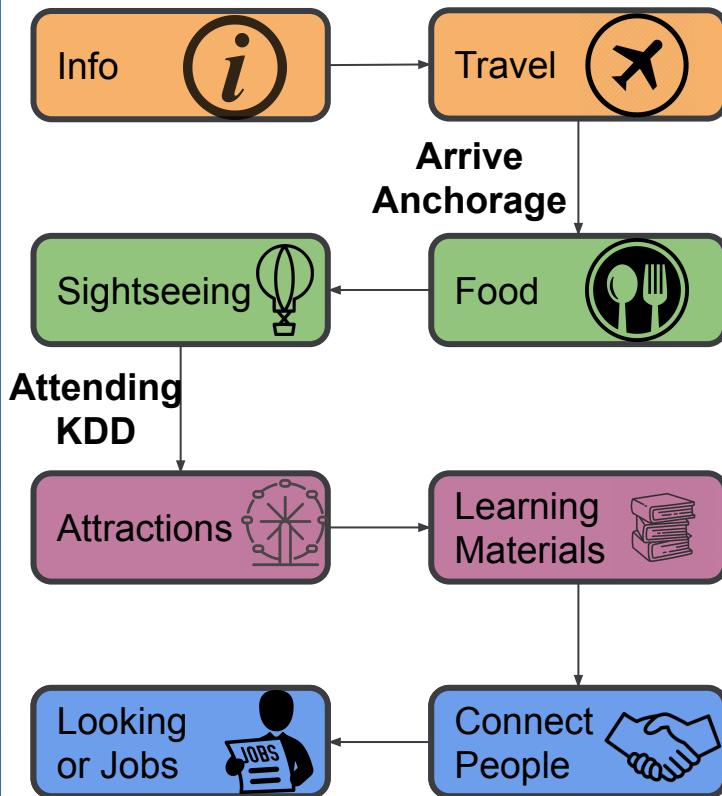
WHAT IS THERE TO DO IN ANCHORAGE FOR THREE?

Feedback

This screenshot shows an Amazon search results page for the query "deep natural language processing for search and recommender systems". The top navigation bar includes delivery information to Sunnyvale, 94085, and links for account settings, orders, and prime. The search bar contains the query "anchorage attractions 5 days". The main search results page displays two books from Packt Publishing: "Real World Data Science with Python" and "Python Machine Learning: Machine Learning and Deep Learning with Python, ...". Below these are results for "Maturity and Innovation in Digital Libraries: 20th International Conference on Asia-Pacific Digital Libraries, ICDL 2018, Hamilton, New Zealand,". The results show the book's title, author (Milena Dobreva, Annika Hinze, et al.), publication date (Nov 15, 2018), format (Paperback, Kindle), price (\$1735, \$1648), and a prime shipping offer. A note at the bottom suggests using fewer keywords.

Introduction - Search and Recommender Systems

KDD 2019



LinkedIn search results for "deep natural language processing" in the United States:

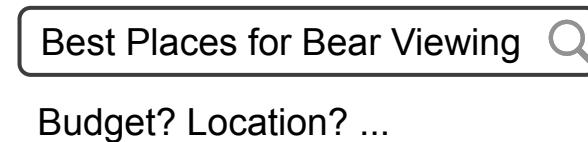
Jobs you may be interested in

- Data Science Manager**
Facebook
Menlo Park, CA, US
1 week ago
- Manager - Machine Learning for Messaging & Marketing**
Netflix
Los Gatos, CA, US
3 weeks ago
- Sr Director, Data Science**
realtor.com
Santa Clara, CA, US
1 week ago
- Technical Program Manager**
Carta
Palo Alto, CA, US
3 days ago
- ML Lead/ Manager**
Apple
Cupertino, CA, US
2 weeks ago

LinkedIn navigation and sidebar are visible on the right.

Introduction - NLP in Search and Recommender Systems

- **Understand User Intention**
 - Search Queries (search)
 - User Profiles (recommendation)
- **Understand Documents**
 - Posts, Reviews, Comments
 - Synonyms, Facets, Semantics, etc.
- **Matching**
 - Query & Doc Matching for Retrieval & Ranking
- **Online Assistance**
 - Query Reformulation



Engineering Manager - Machine Learning and AI
Aug 2018 – Present · 1 yr 1 mo
San Francisco Bay Area

Lead LinkedIn Personalization and Search AI Foundation team - Provide LinkedIn users with intelligent experience through **natural language understanding** (across multiple languages) and personalization powered by Machine Learning and Artificial Intelligence.

Manager, Engineering -
Conversational AI

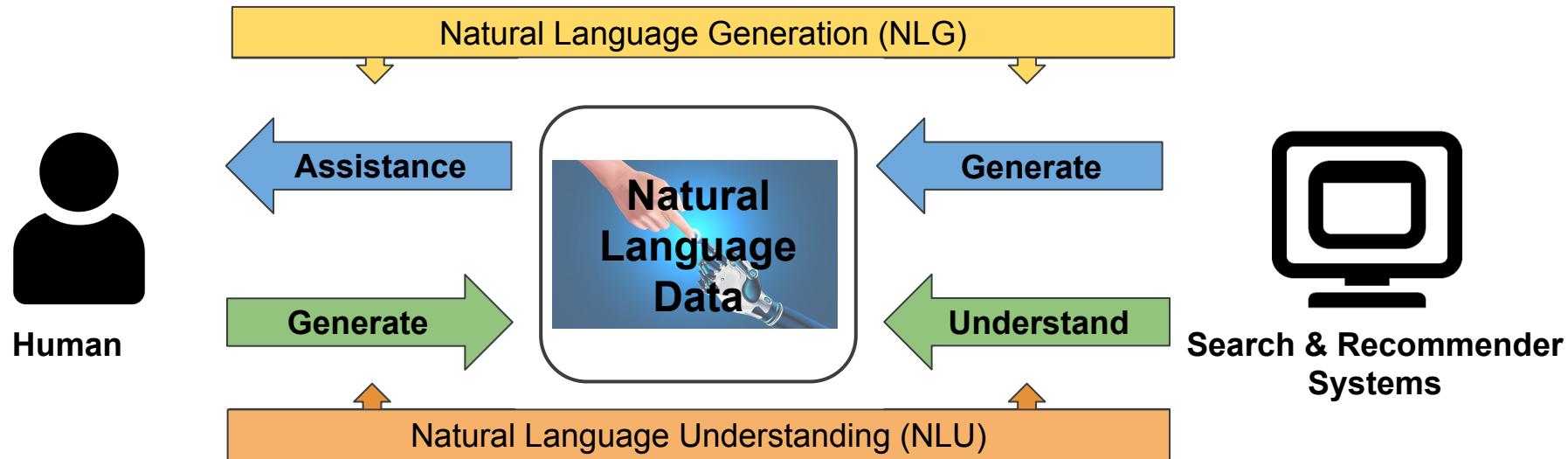
Facebook · Menlo Park, CA, US

Posted 1 week ago · 33 views

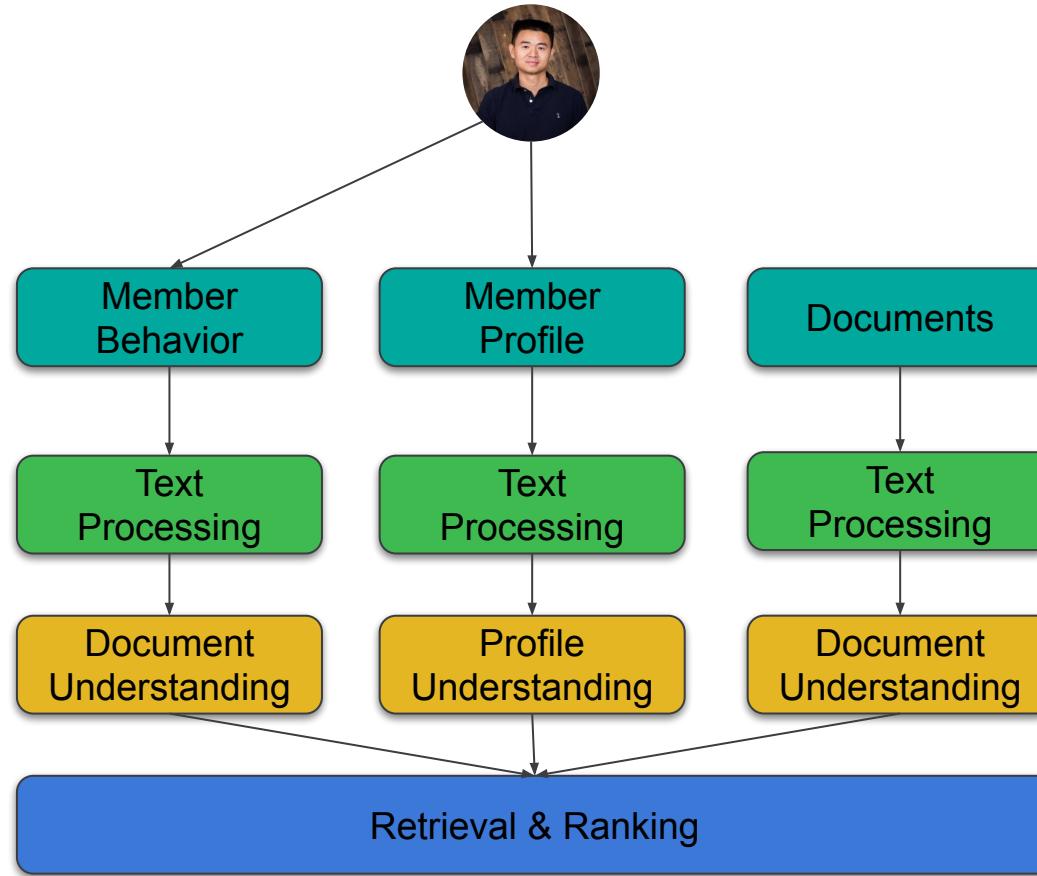
Save Apply

Job	Company	Connections
<ul style="list-style-type: none">1 applicantFull-time	<ul style="list-style-type: none">10,001 employeesInternet	<ul style="list-style-type: none">67 connections586 company alumni

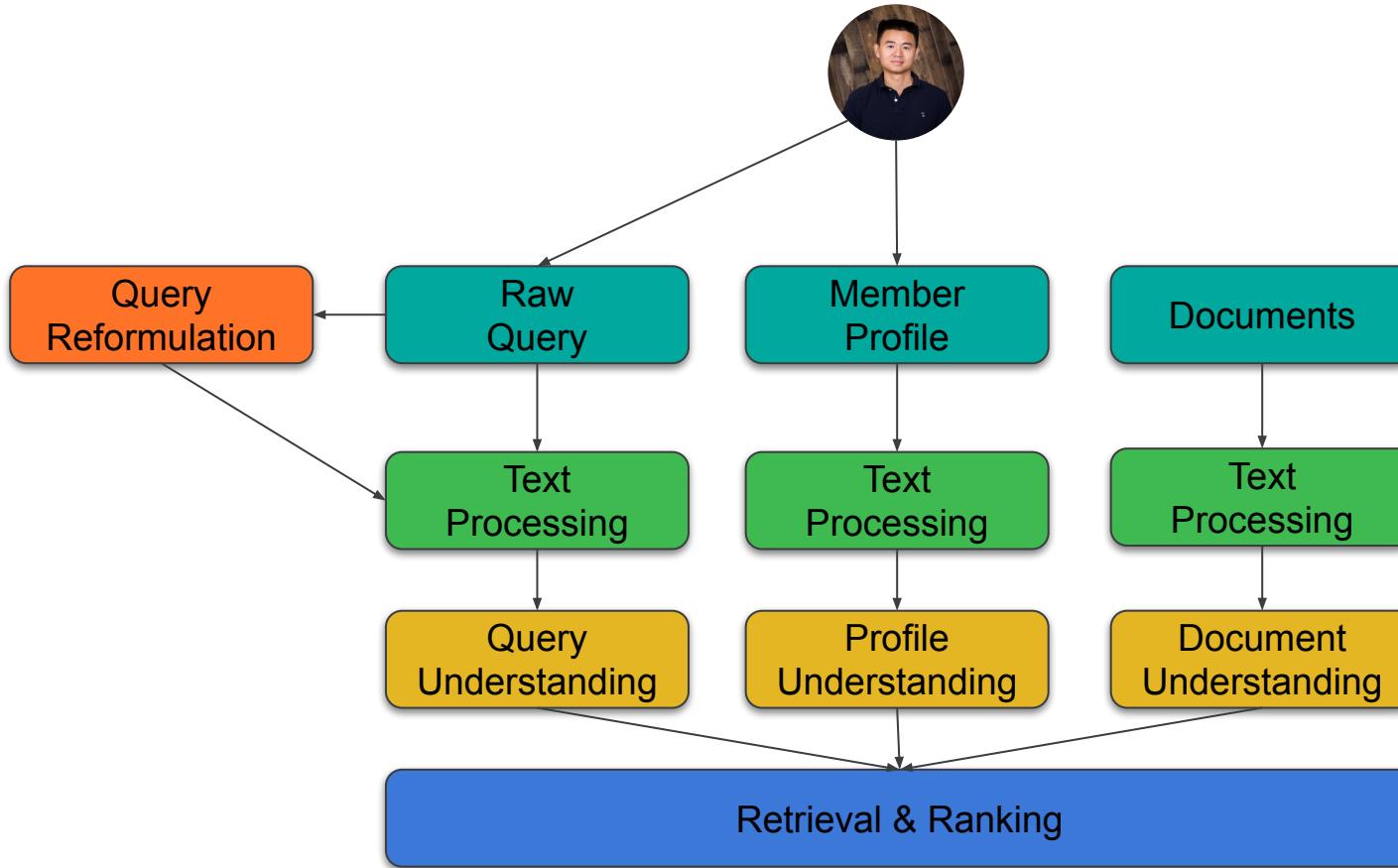
Introduction - NLP in Search and Recommender Systems



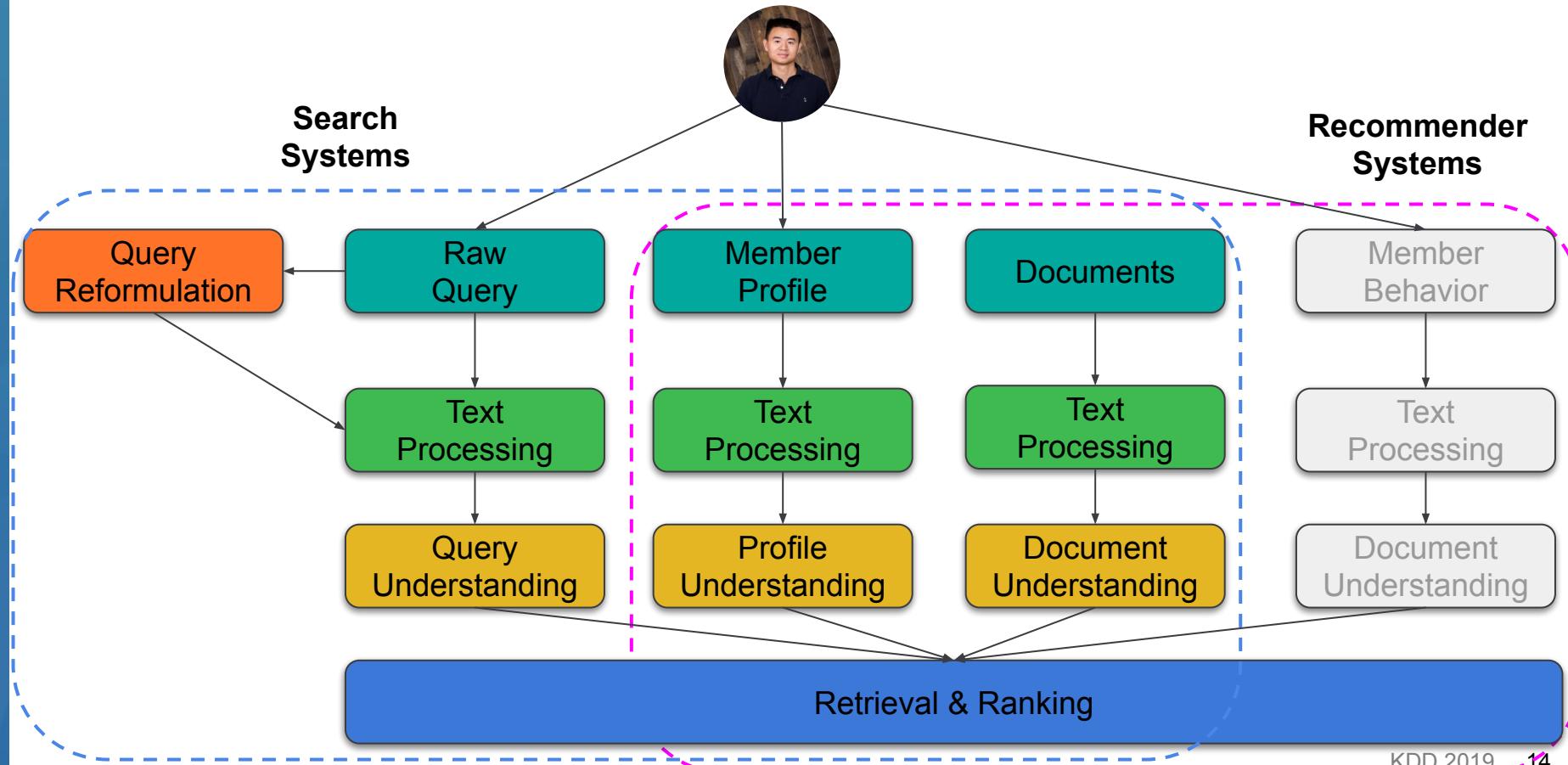
Natural Language Processing in Recommendation Ecosystem



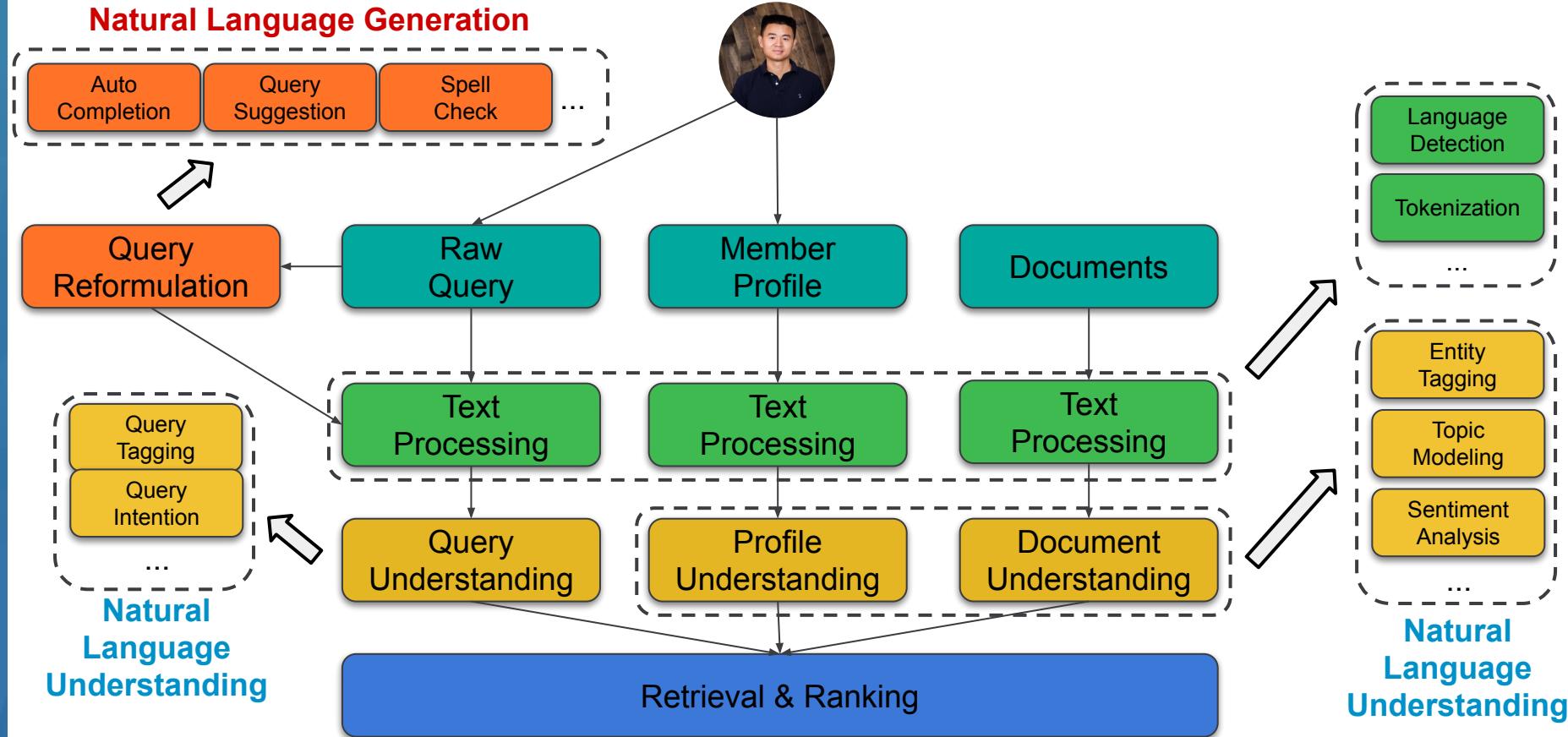
Natural Language Processing in Search Ecosystem



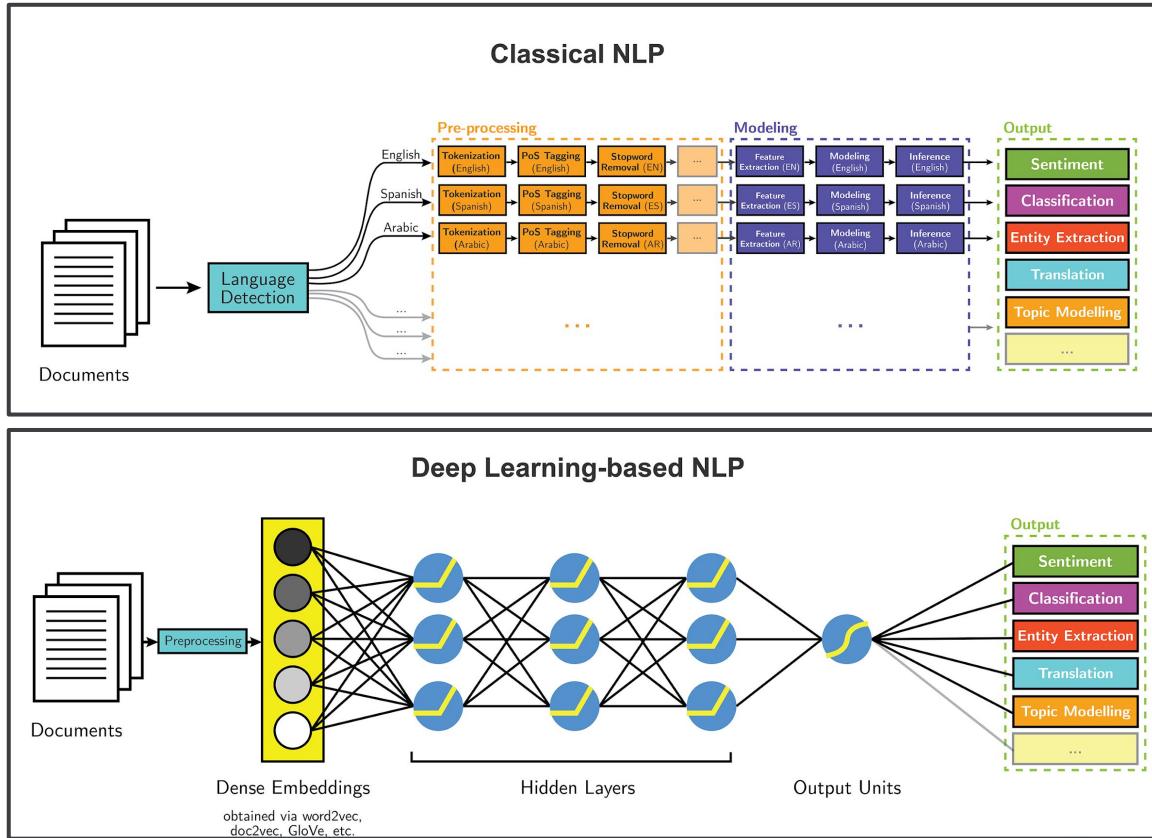
Natural Language Processing in Search and Recommendation



Natural Language Processing in Search and Recommendation



Introduction - Deep Learning for NLP



Opportunities - Deep Learning for NLP in Search and Recommender Systems

Why Deep Learning?

- **Deep Semantics** from High Dimension and Sparse Data
 - Synonymous, Disambiguation, etc.
- **Easy Feature Engineering**
 - Hand Crafted Features V.S. Auto Feature Representation
- **Model Flexibility**
 - Model end-to-end process
 - Various NN components to model and cooperate systematically
- **Multi-level Feature Representation**
 - Hierarchical Representations
character -> token -> word -> phrase -> sentence

[Young et. al. 2018]

Agenda

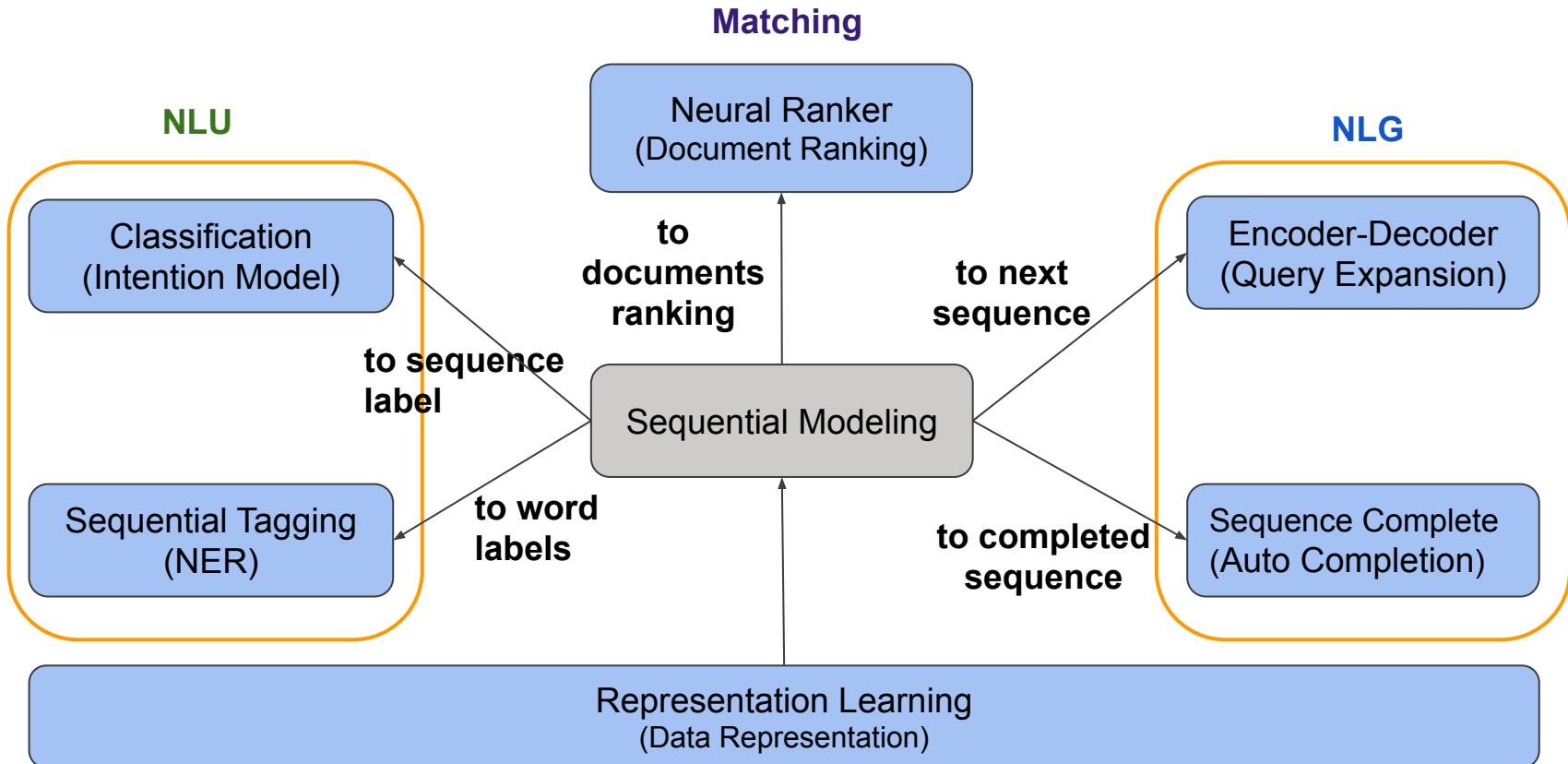
- 1 Introduction
- 2 Deep Learning for Natural Language Processing
- 3 Deep NLP in Search and Recommender Systems
- 4 Real World Examples



Deep Learning for Natural Language Processing

Bo Long

Deep Learning for Natural Language Processing

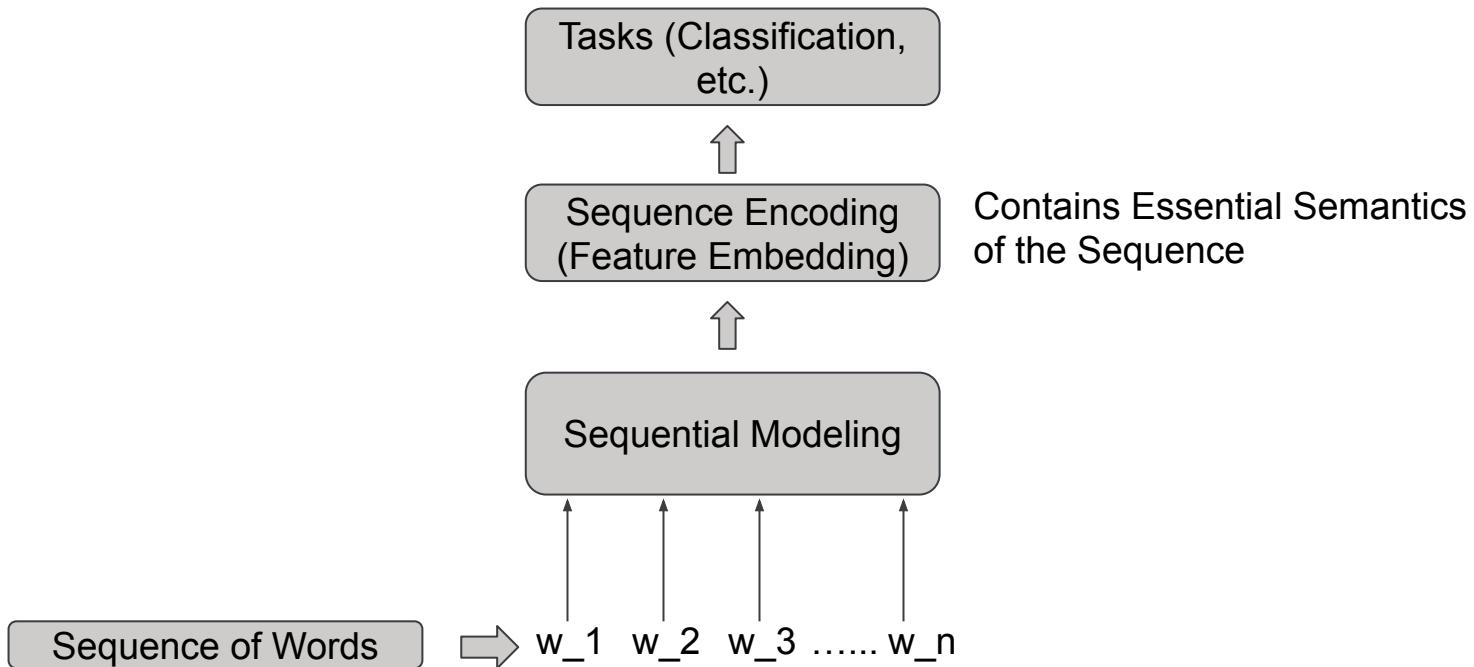


Deep Learning for Natural Language Processing

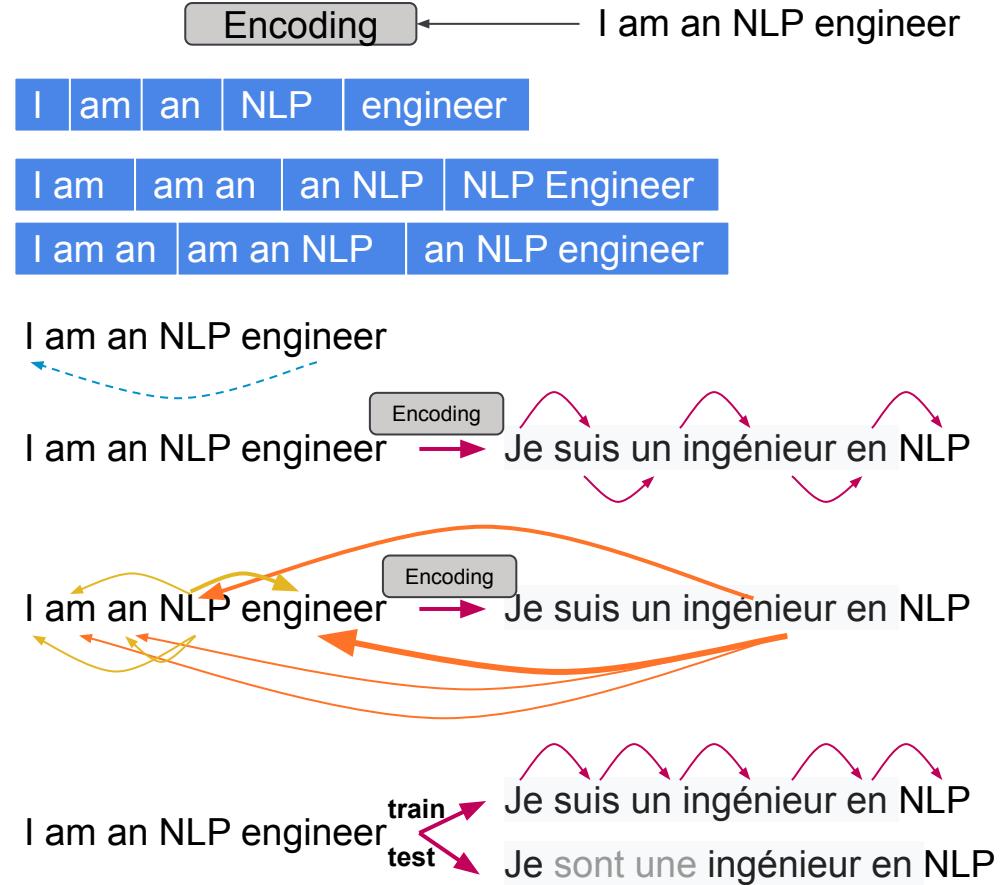
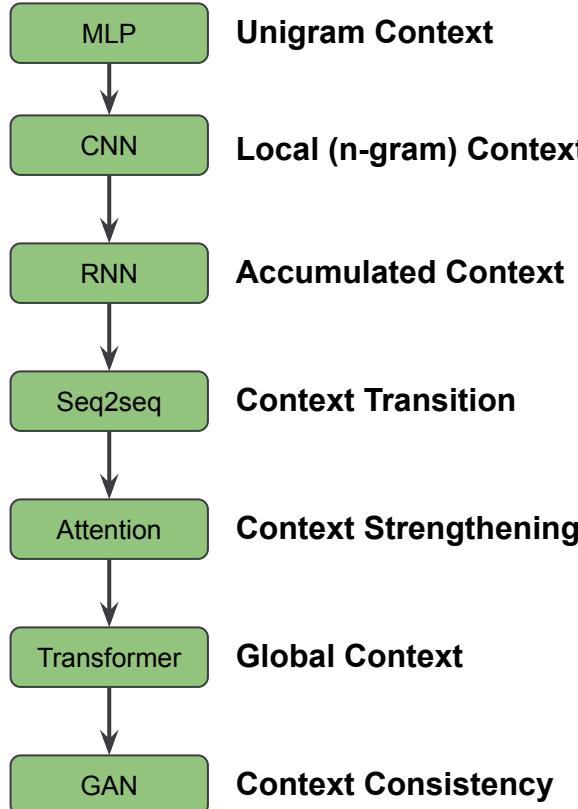
- **Sequential Modeling on Semantic Context**
- Representation Learning for Data Processing

Deep Learning for Natural Language Processing

- **Sequential Modeling on Semantic Context**



Sequential Modeling on Semantic Context



Multilayer Perceptron (MLP)

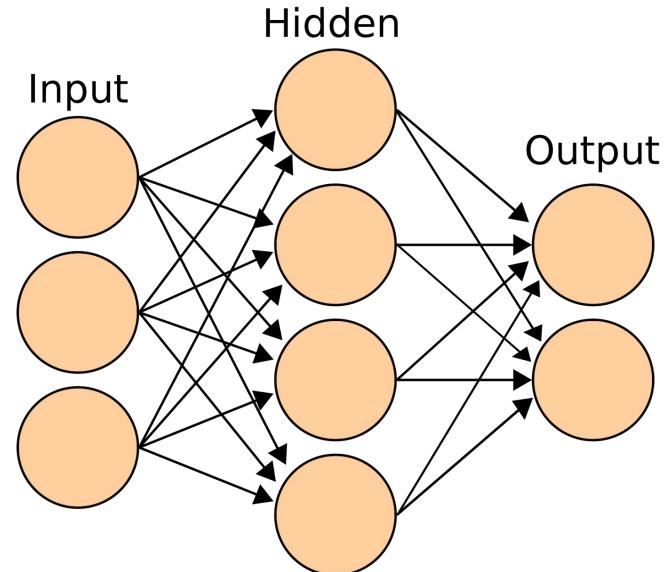
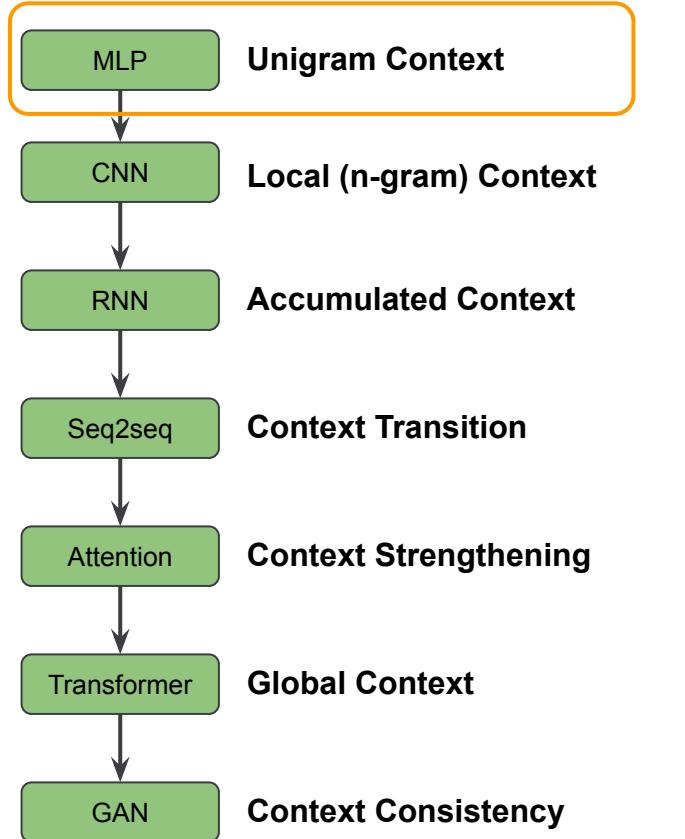
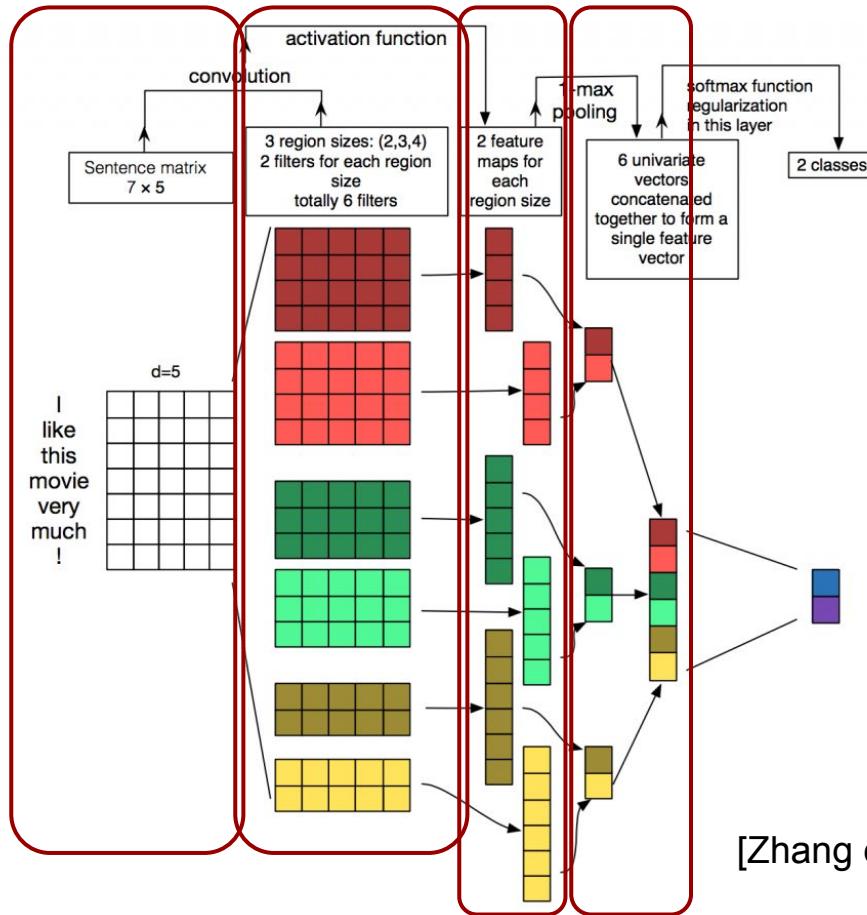
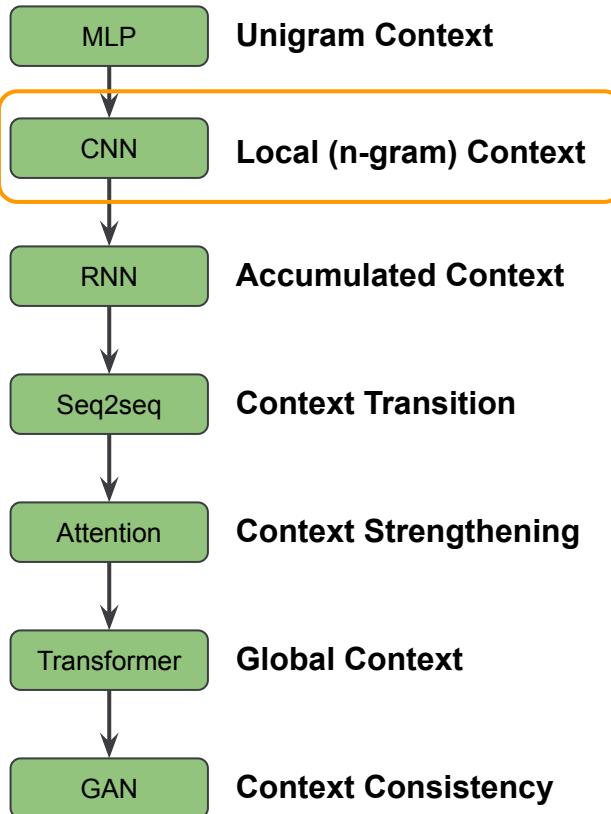
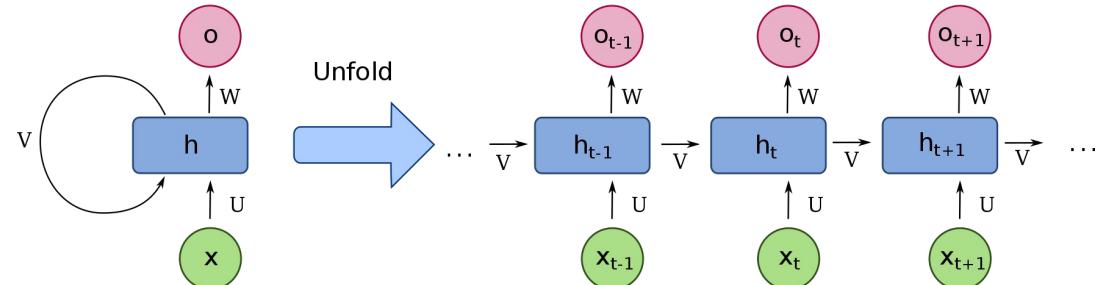
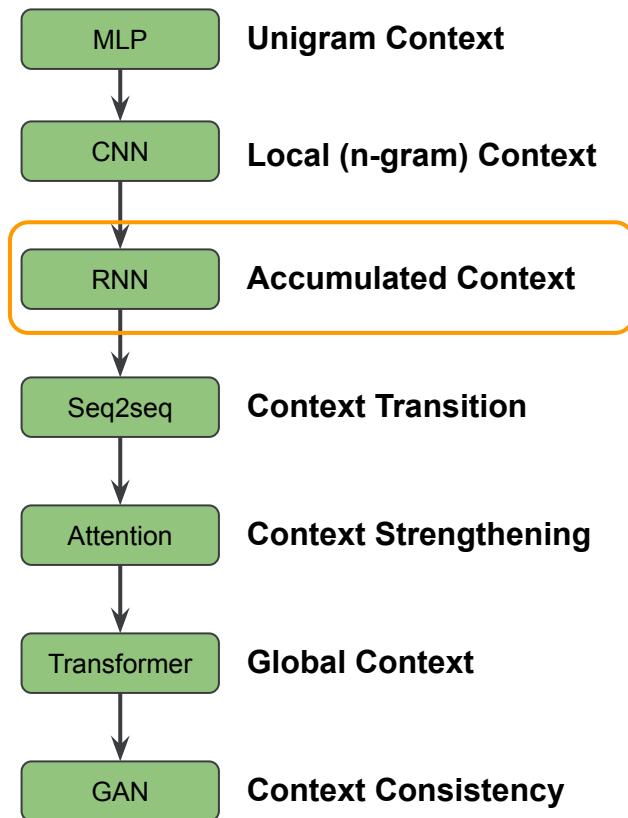


Figure source: <https://stackabuse.com/introduction-to-neural-networks-with-scikit-learn/>

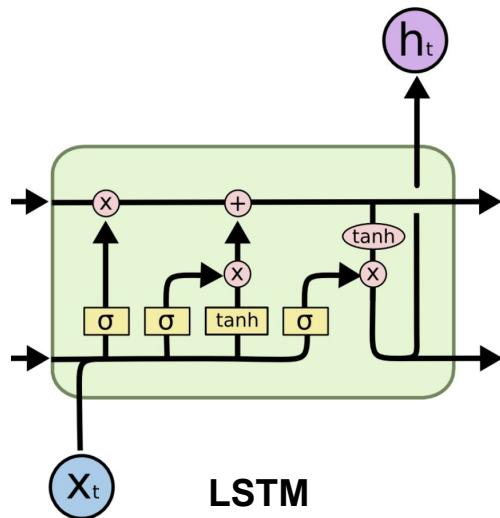
Convolutional Neural Networks (CNN)



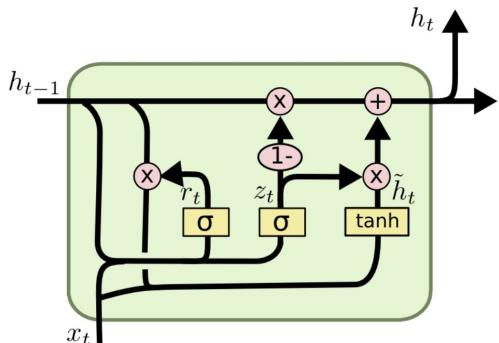
Recurrent Neural Networks (RNN)



[LeCun et. al. 2015]

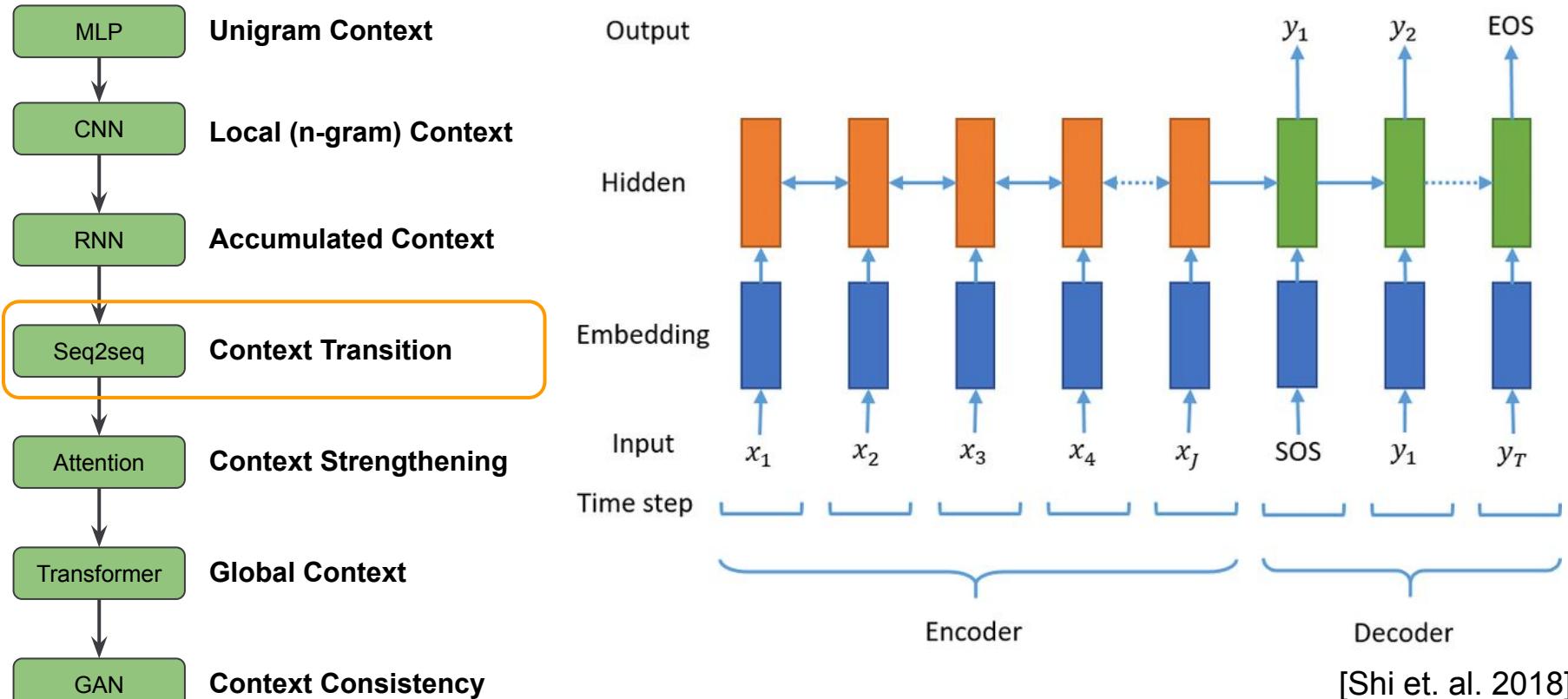


LSTM

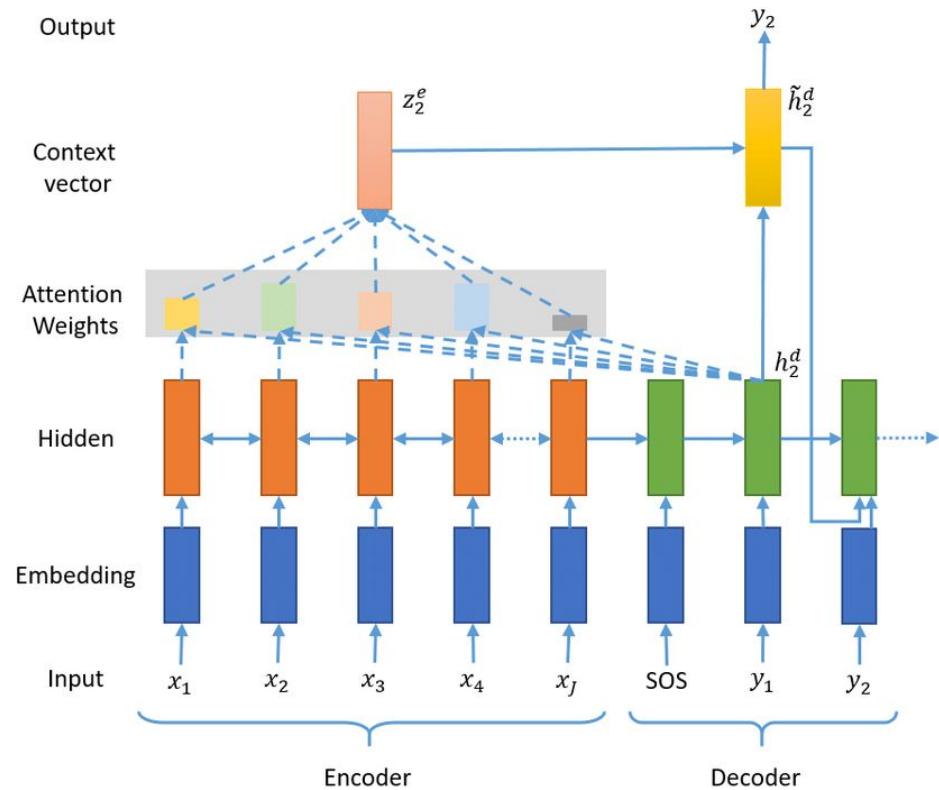
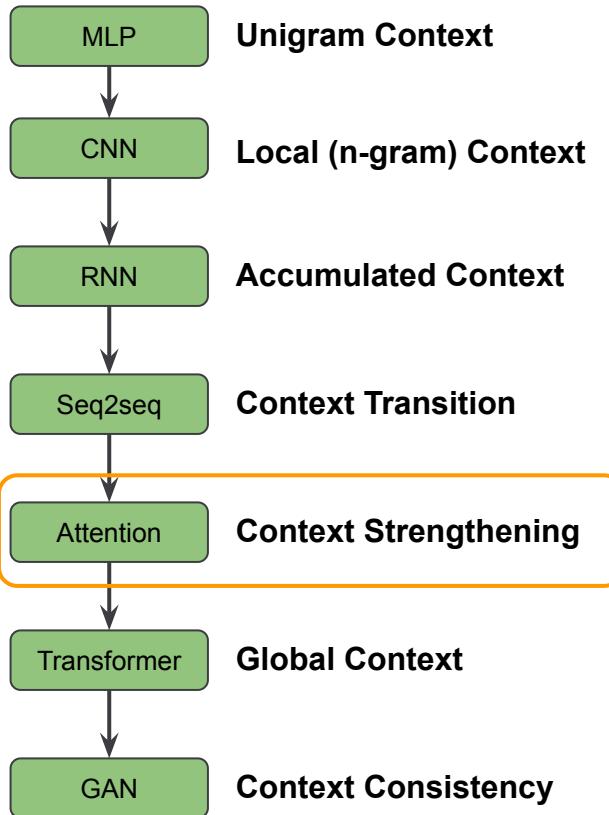


GRU

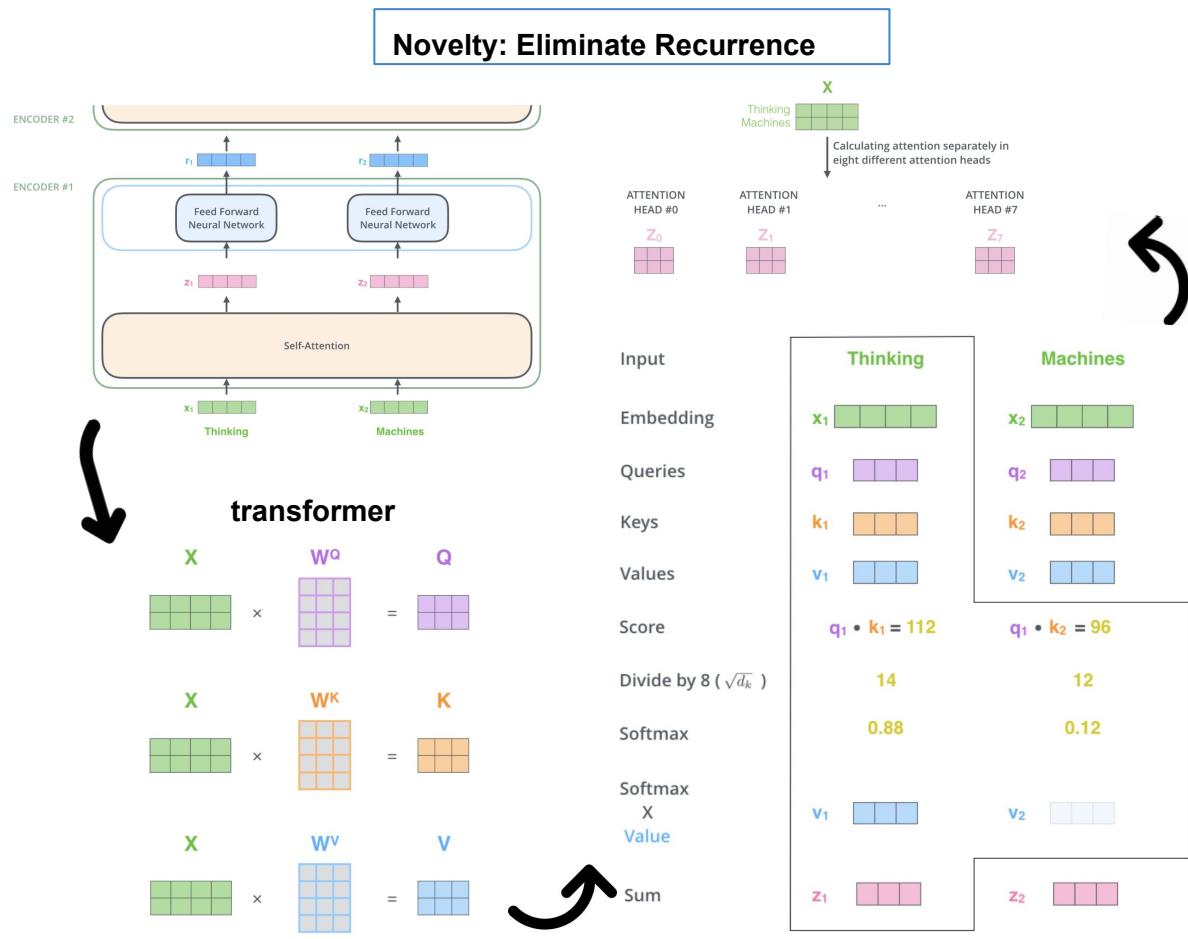
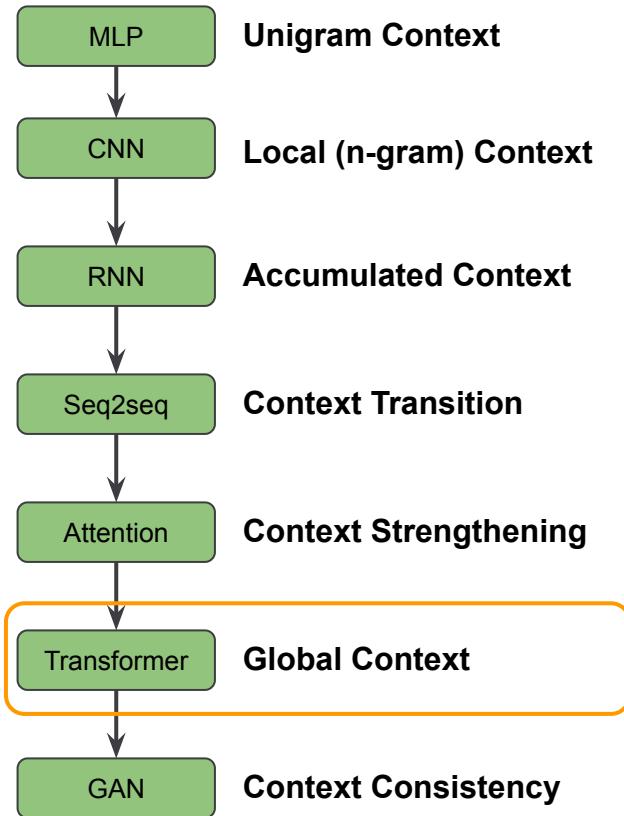
Sequence to Sequence (Encoder - Decoder)



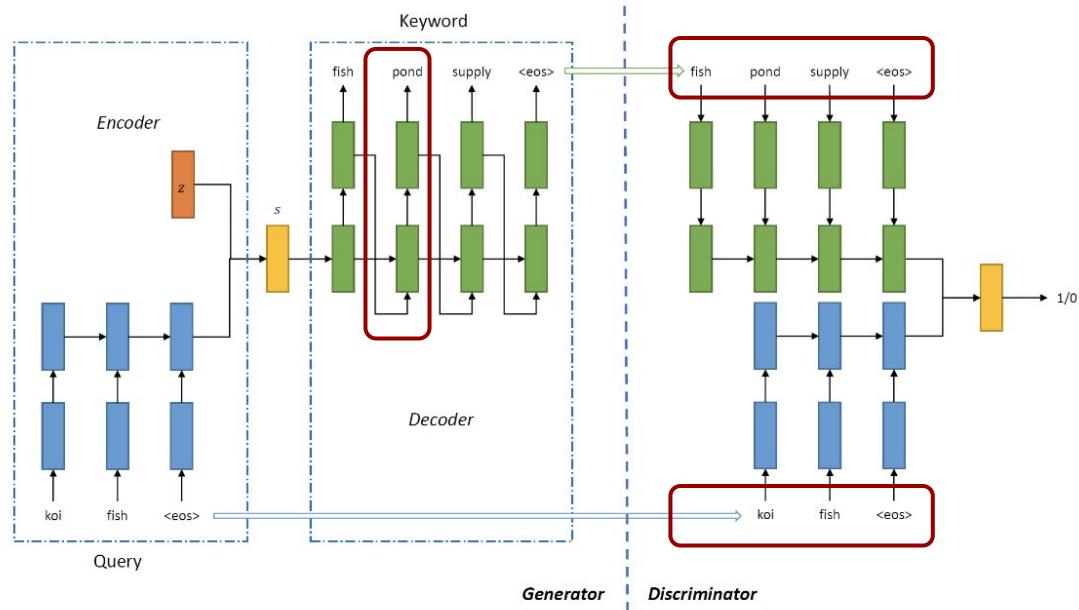
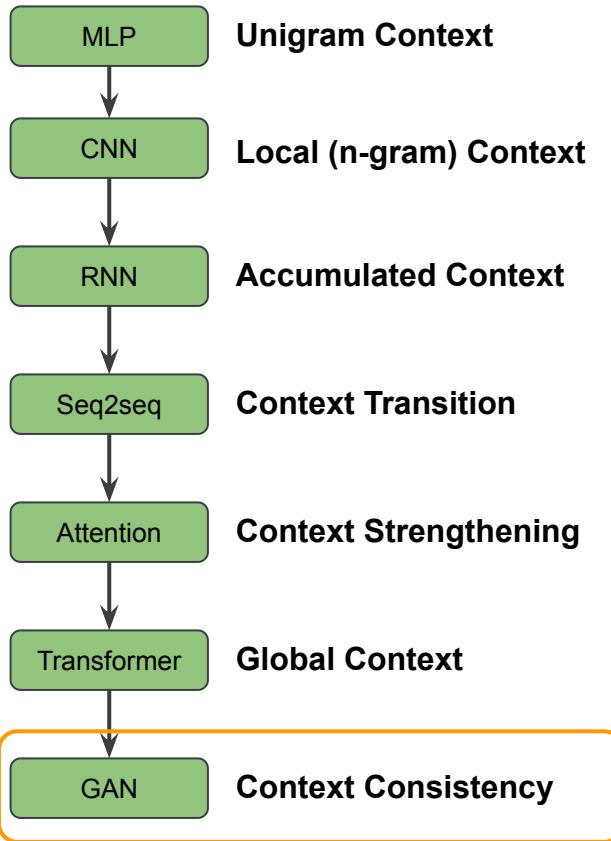
Attention Mechanisms



Transformer

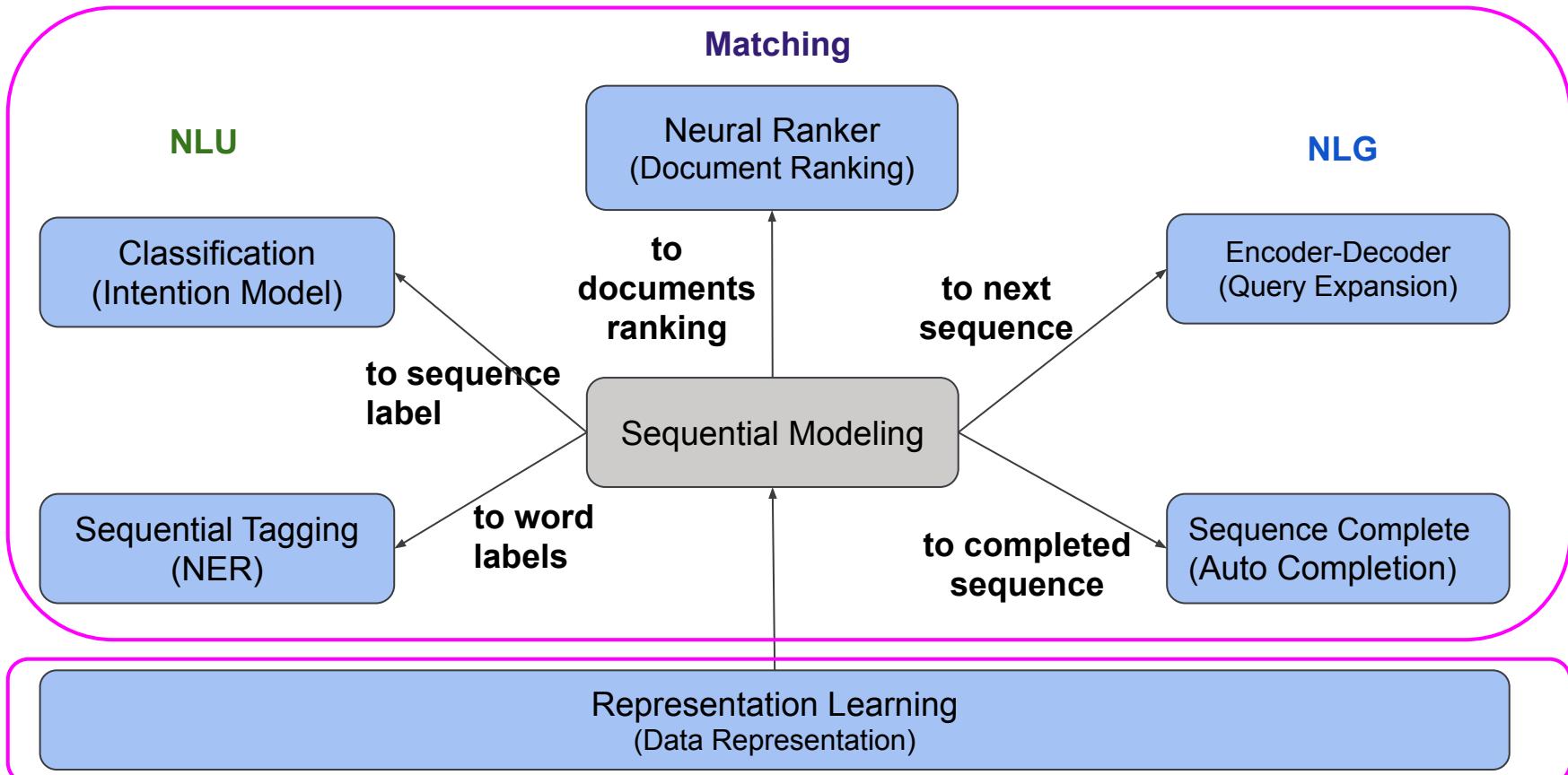


Generative Adversarial Networks (GANs)



[Lee, et al., KDD 2018]

Deep Learning for Natural Language Processing

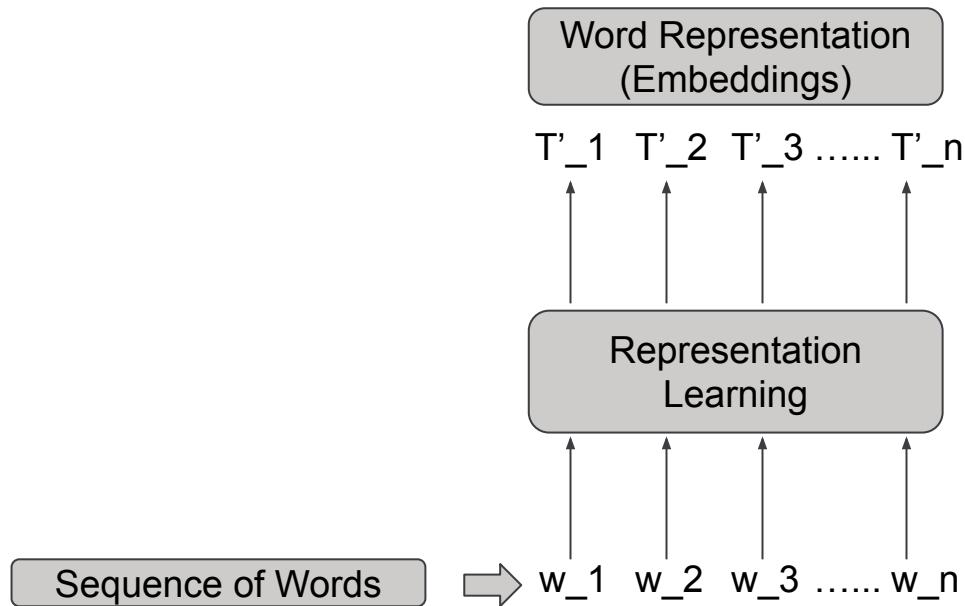


Deep Learning for Natural Language Processing

- Sequential Modeling on Semantic Context
- **Representation Learning for Data Processing**

Deep Learning for Natural Language Processing

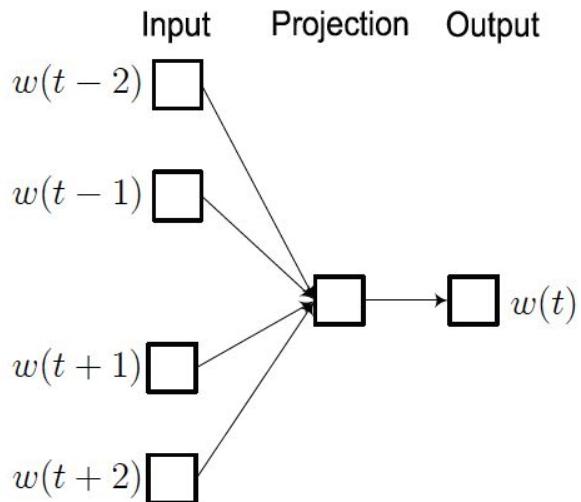
- **Representation Learning**



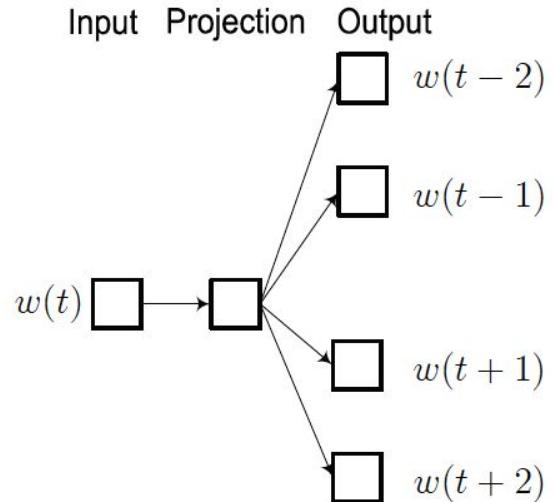
Representation Learning

Word2Vec

Static Word Embedding



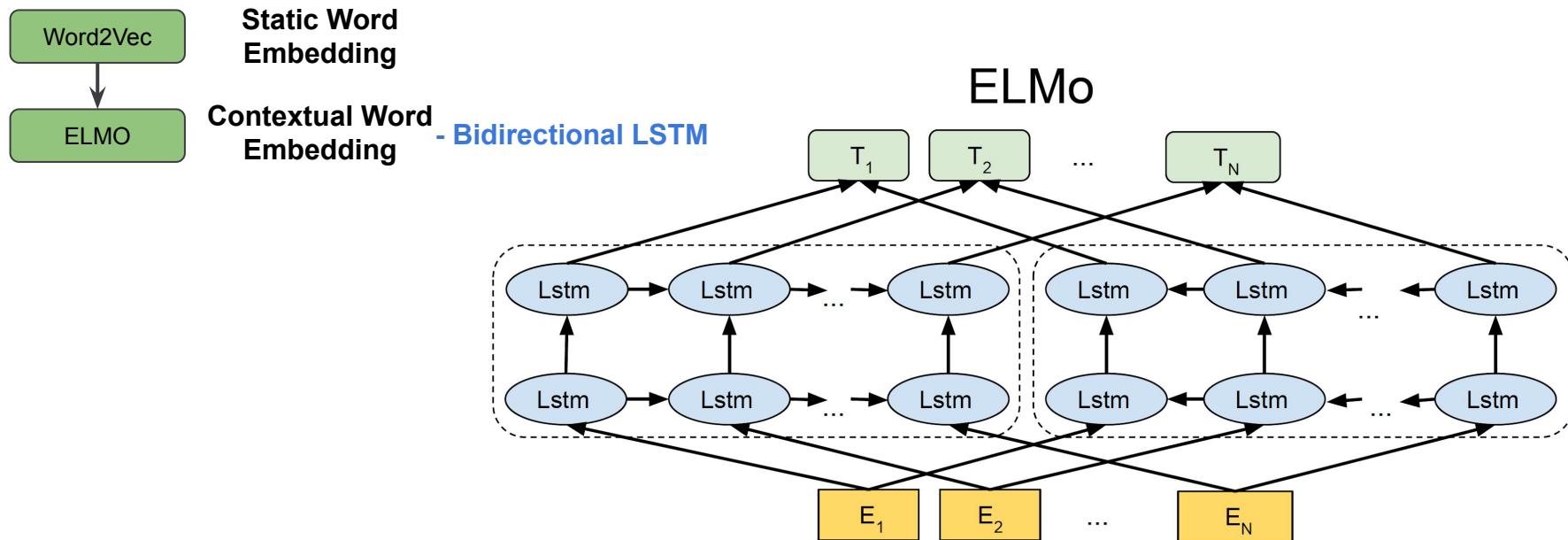
(a) CBOW.



(b) Skip-gram.

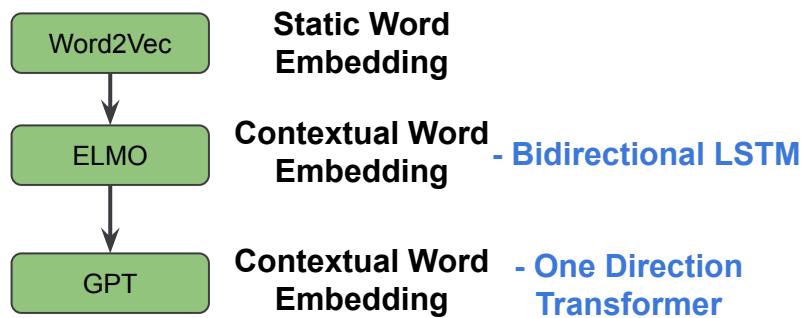
[Mikolov et. al. 2013]

Representation Learning

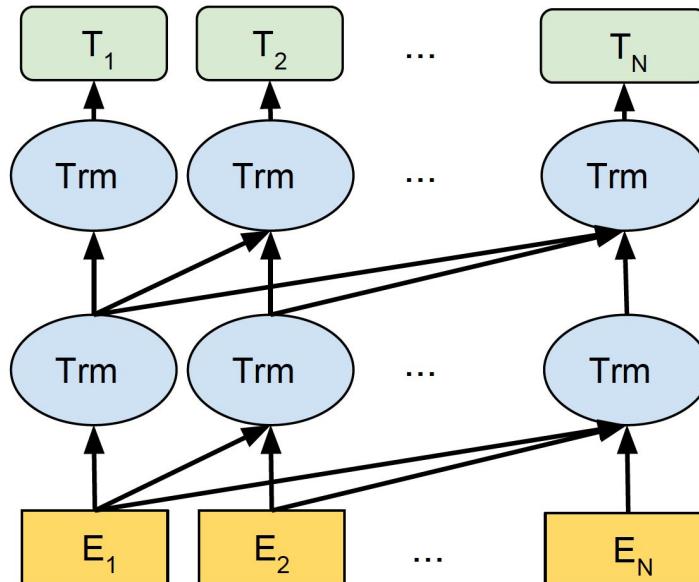


[Peters et. al. 2018]

Representation Learning

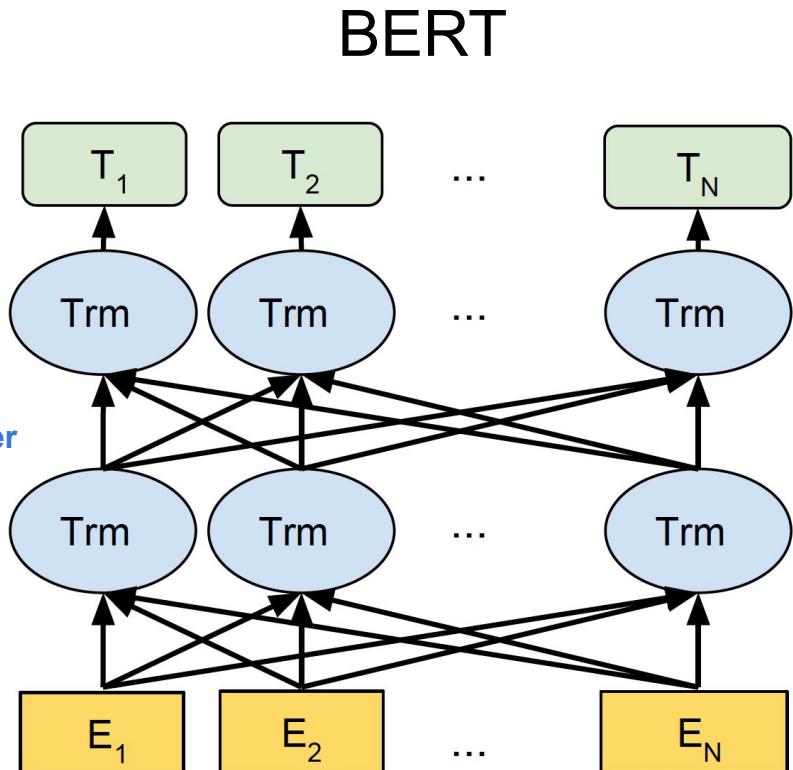
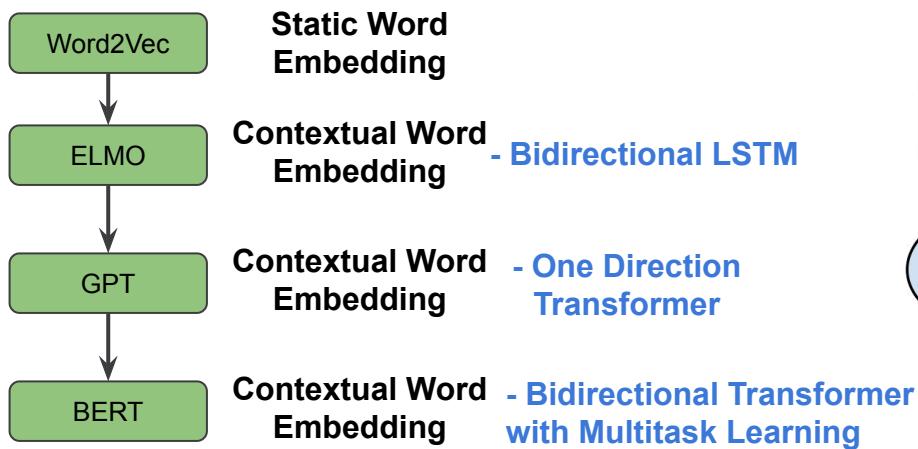


OpenAI GPT

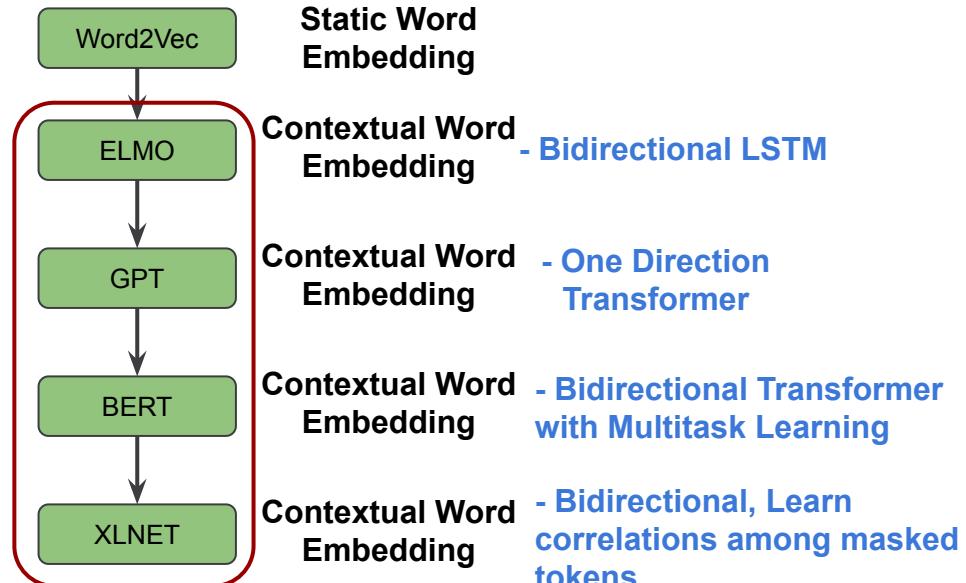


[Radford et. al. 2018]

Representation Learning

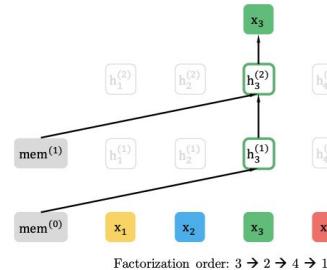


Representation Learning

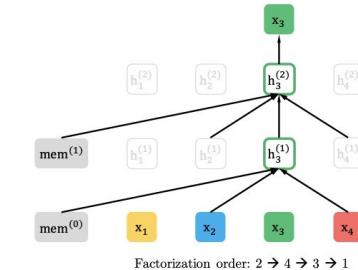


Pre-trained NLP Model

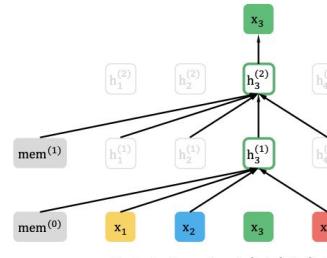
XLNet



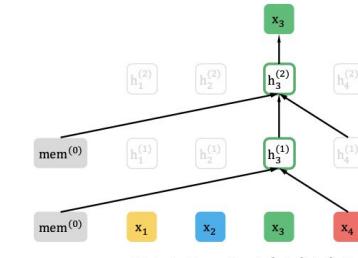
Factorization order: $3 \rightarrow 2 \rightarrow 4 \rightarrow 1$



Factorization order: $2 \rightarrow 4 \rightarrow 3 \rightarrow 1$



Factorization order: $1 \rightarrow 4 \rightarrow 2 \rightarrow 3$



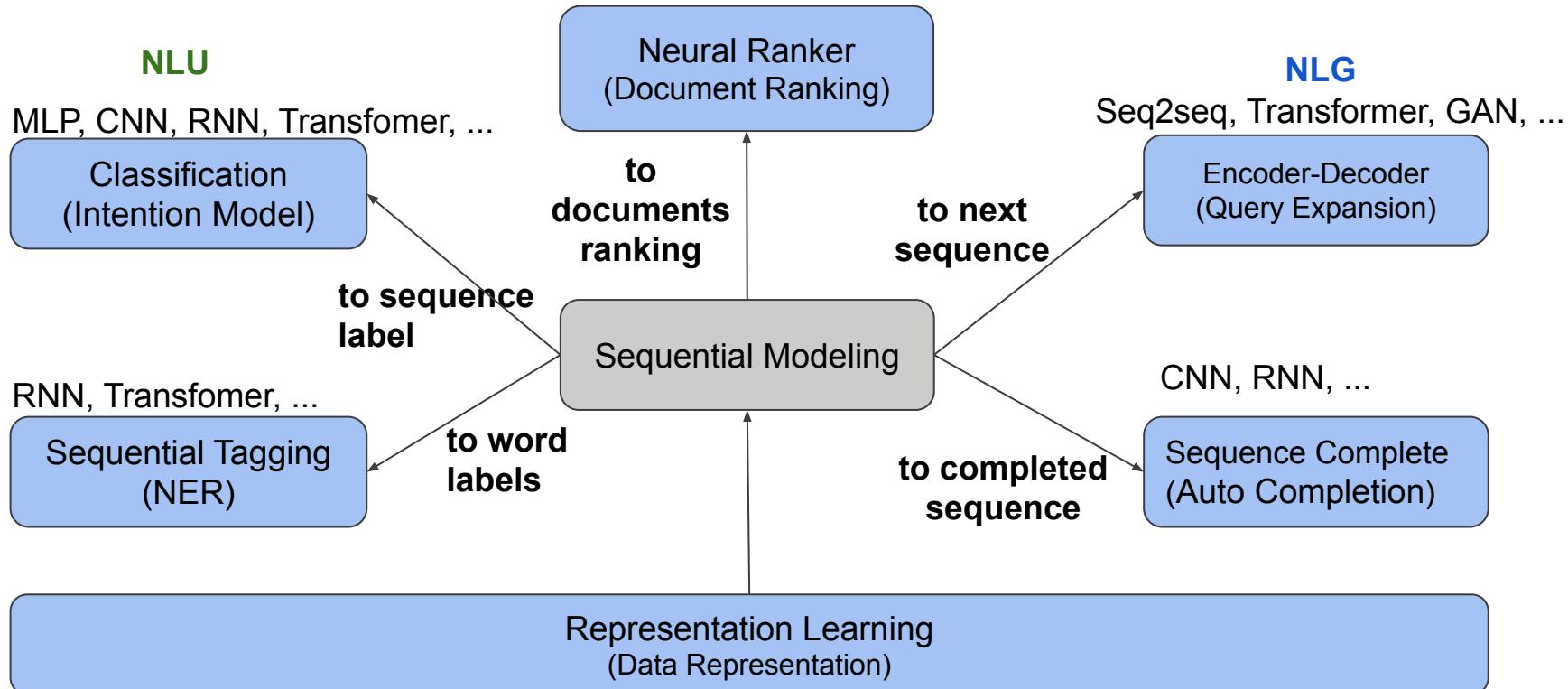
Factorization order: $4 \rightarrow 3 \rightarrow 1 \rightarrow 2$

[Yang et. al. 2019]

Deep Learning for Natural Language Processing

MLP, CNN, RNN, seq2seq, attention, transformer, GAN

Matching



References - Deep Learning for Natural Language Processing

- [Devlin et. al. 2018] Devlin, Jacob and Chang, Ming-Wei and Lee, Kenton and Toutanova, Kristina. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, arXiv preprint arXiv:1810.04805, 2018
- [LeCun et. al, 2015] Deep learning, Nature, vol. 521, no. 7553, pp. 436–444, 2015.
- [Lee et. al, 2018] Rare query expansion through generative adversarial networks in search advertising, In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, pp. 500-508. ACM, 2018.
- [Mikolov et. al. 2013] Distributed representations of words and phrases and their compositionality, in Advances in neural information processing systems, 2013
- [Mikolov et. al. 2013] Efficient estimation of word representations in vector space, arXiv preprint arXiv:1301.3781, 2013.
- [Peters et. al. 2018] Matthew Peters, Mark Neumann, Mohit Iyyer, Matt Gardner, Christopher Clark, Kenton Lee, and Luke Zettlemoyer. 2018a. Deep contextualized word representations. In NAACL.
- [Radford et.al. 2018] Alec Radford, Karthik Narasimhan, Tim Salimans, and Ilya Sutskever. 2018. Improving language understanding with unsupervised learning. Technical report, OpenAI.
- [Robinson, 2018] <https://stackabuse.com/introduction-to-neural-networks-with-scikit-learn>, 2018.
- [Shi et. al, 2018] Neural Abstractive Text Summarization with Sequence-to-Sequence Models, arXiv preprint arXiv:1812.02303v2, 2018.
- [Yang et. al. 2019] XLNet: Generalized Autoregressive Pretraining for Language Understanding, arXiv preprint arXiv:1906.08237v1, 2019
- [Young et. al. 2018] Recent Trends in Deep Learning Based Natural Language Processing, arXiv preprint arXiv:1708.02709v8, 2018
- [Zhang et. al. 2015] A sensitivity analysis of (and practitioners' guide to) convolutional neural networks for sentence classification, arXiv preprint arXiv:1510.03820, 2015.

Agenda

- 1** Introduction
- 2** Deep Learning for Natural Language Processing
- 3** Deep NLP in Search and Recommender Systems
- 4** Real World Examples



Deep NLP in Search and Recommender Systems - Language Understanding

Jun Shi

Deep NLP in Search and Recommender Systems

- **Language Understanding**
 - Entity Tagging: word level prediction
 - Entity Disambiguation: knowledge base entity prediction
 - Intent Classification: sentence level prediction
 - Sentiment Analysis and Opinion Mining
- Document Retrieval and Ranking
 - Efficient Candidate Retrieval
 - Deep Ranking Models
- Language Generation for Assistance
 - Auto Completion
 - Query Reformulation
 - Spell Correction
 - Conversational Recommendation

Entity Tagging

- Problem statement
- Traditional statistical models
 - Hidden Markov Model
 - Maximum Entropy Markov Model
 - (Semi-Markov) Conditional Random Field
- Deep learning models
 - Input layer
 - Context layer
 - Decoder layer
- Training dataset
- Evaluation

Entity Tagging - Problem Statement

- A named entity, a word or a phrase that clearly identifies one item from a set of other items that have similar attributes. [Li et. al. 2018]
- Entity tagging (Named Entity Recognition, NER), the process of locating and classifying named entities in text into predefined entity categories.

Washington D.C. was named after George Washington.

LOCATION

PERSON

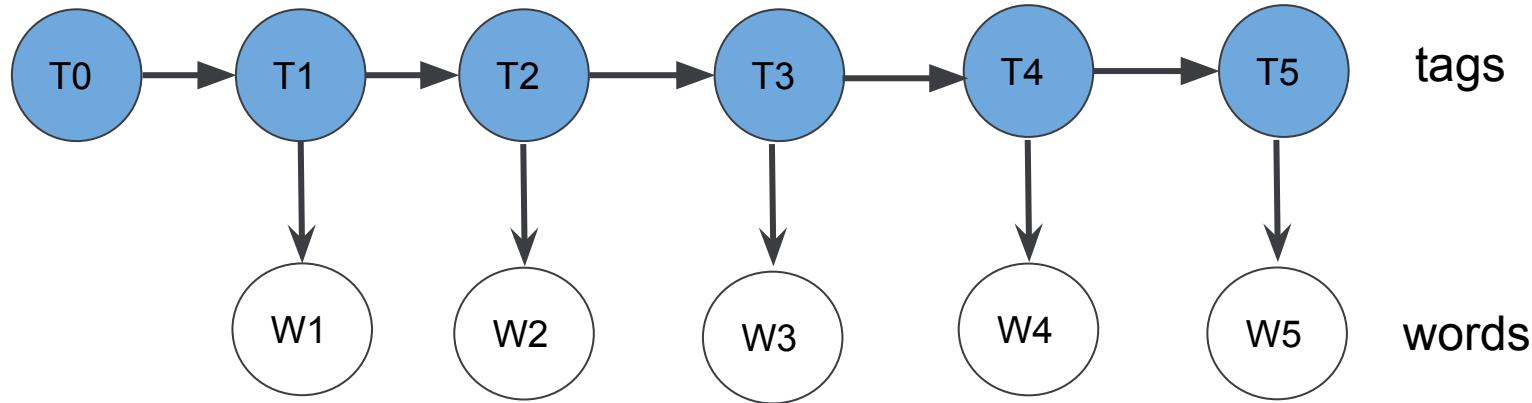
Motivation

- Efficiency
 - Looking only for named entities can speed up search.
- Precision
 - Matching both named entity and tags can increase search precision.
- Quality
 - Ranking retrieved results by considering tags improves search quality.

Traditional Statistical Models

- Generative model
 - Hidden Markov Model [Baum, et. al. 1966]
- Discriminative model
 - Maximum Entropy Markov Model [McCallum, et. al. 2000]
 - (Semi-Markov) Conditional Random Field [Lafferty, et. al. 2001]

Hidden Markov Model

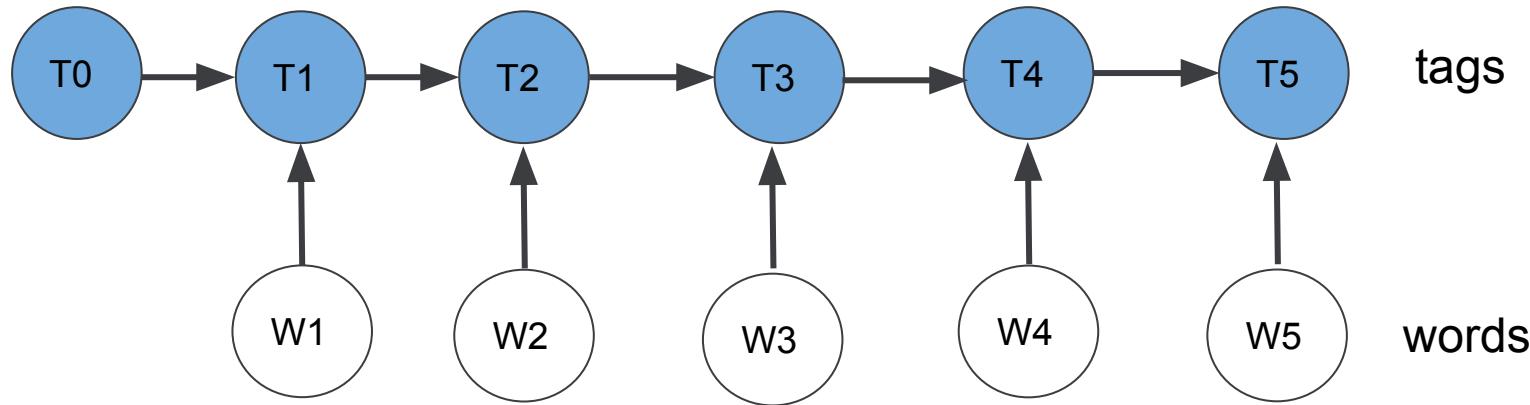


Generative model, model joint probability $\Pr(T, W)$ instead of conditional probability $\Pr(T | W)$.

$$\Pr(\mathbf{T}, \mathbf{W}) = \prod_{i=1}^L (\Pr(w_i | t_i) \Pr(t_i | t_{i-1}))$$

t_0 is a dummy start state.

Maximum Entropy Markov Model

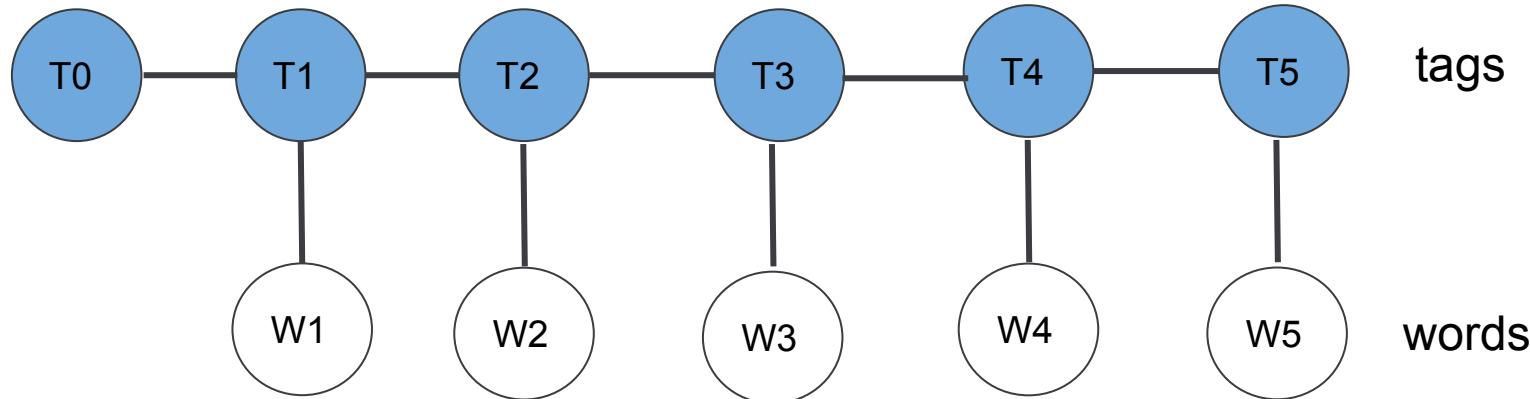


Discriminative model, model conditional probability $\Pr(T | W)$ directly.

$$\Pr(\mathbf{T}|\mathbf{W}) = \prod_{i=1}^L \Pr(t_i|t_{i-1}, w_i) = \prod_{i=1}^L \frac{\exp(\sum_j \beta_j f_j(t_{i-1}, w_i))}{Z(t_{i-1}, w_i)}$$

t_0 is a dummy start state.

Conditional Random Field

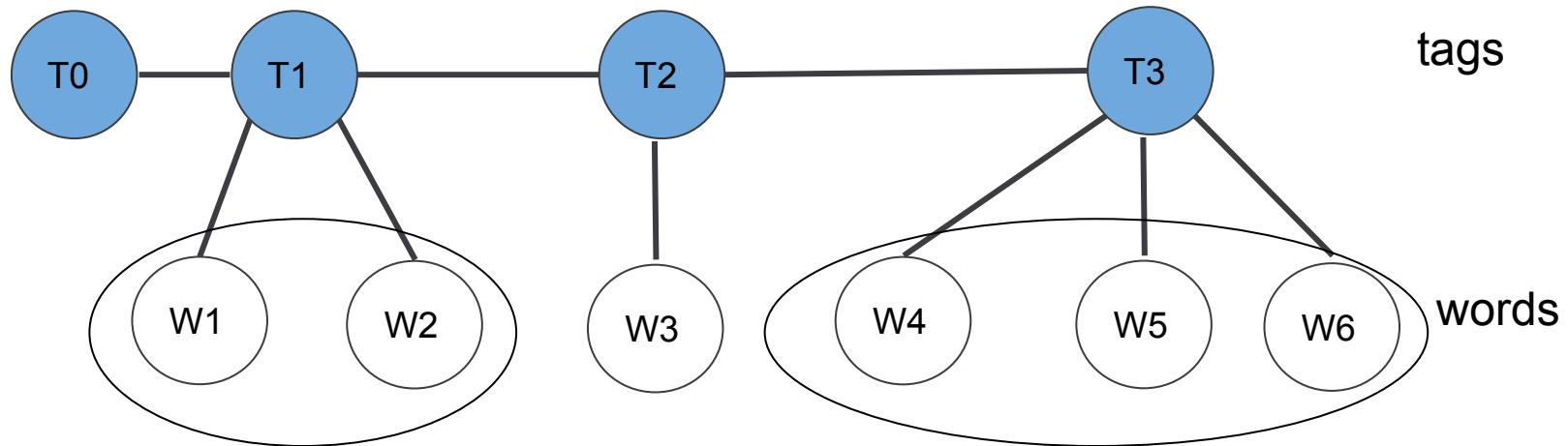


Discriminative model, model conditional probability $\text{Pr}(\mathbf{T} | \mathbf{W})$ directly.

$$\text{Pr}(\mathbf{T} | \mathbf{W}) = \frac{\prod_{i=1}^L \exp(\sum_j \beta_j f_j(t_{i-1}, \mathbf{W}))}{Z(\mathbf{T}, \mathbf{W})}$$

t_0 is a dummy start state.

Semi-Markov Conditional Random Field

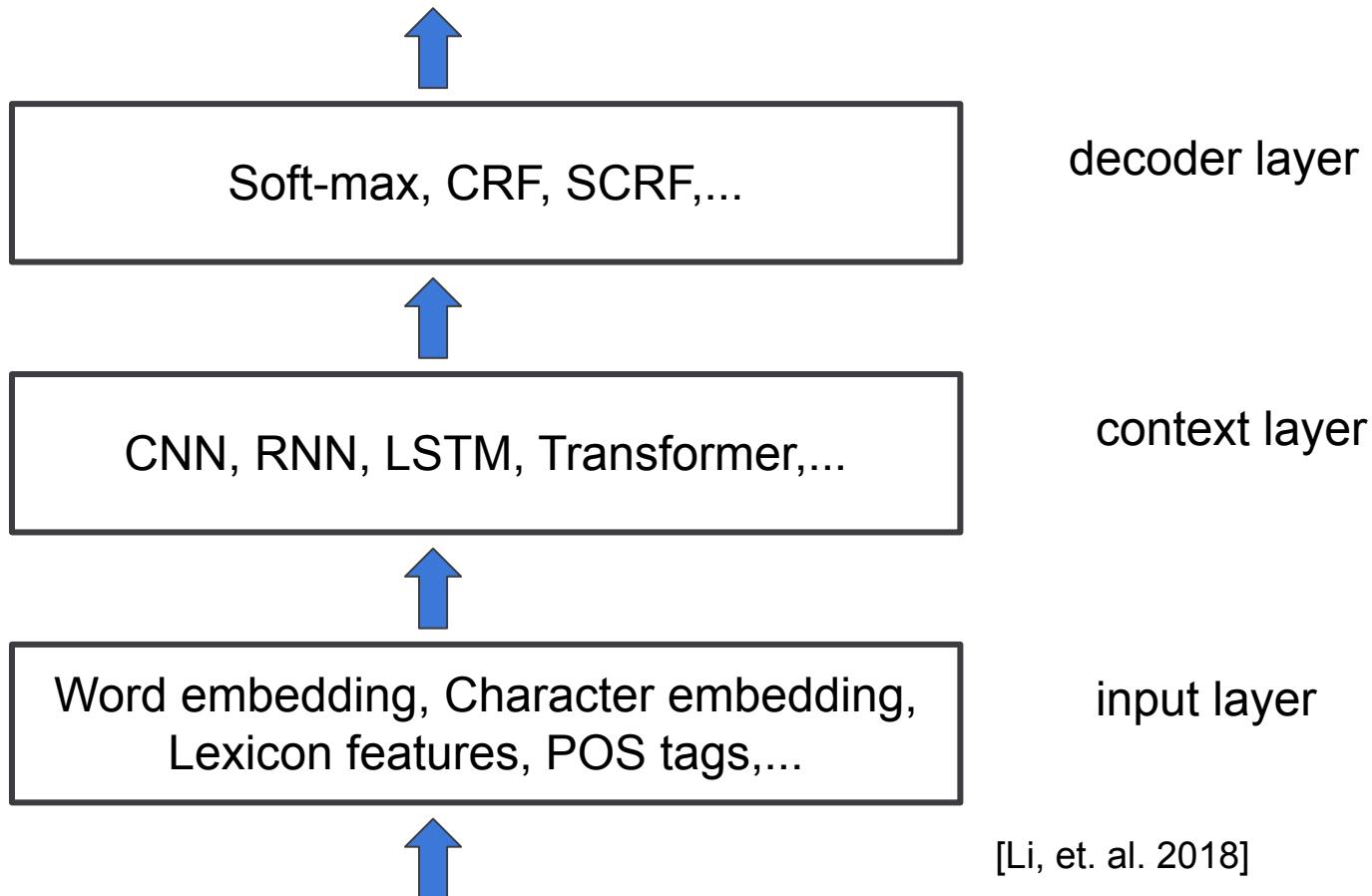


Each tag can correspond to a variable-length phrase.

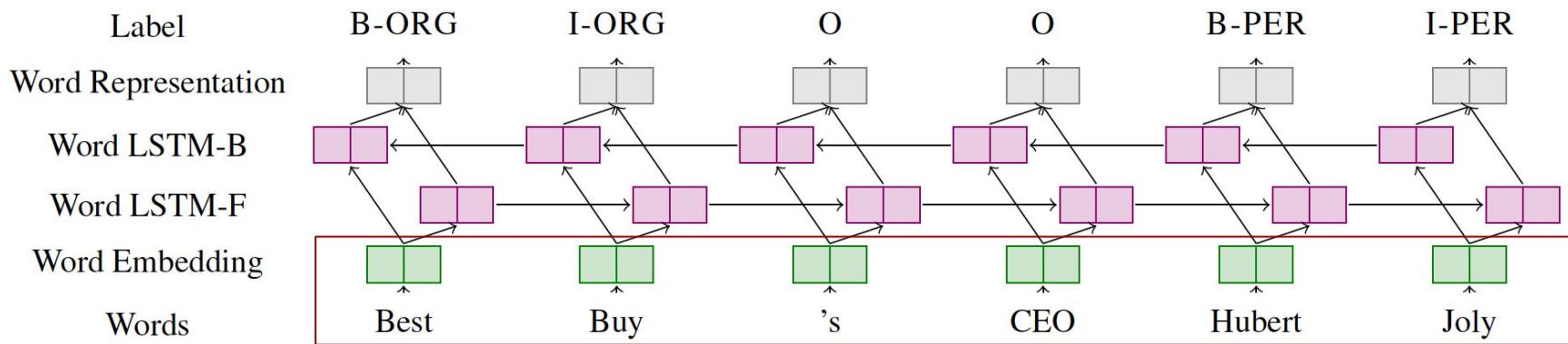
Summary - Traditional Statistical Models

Model	Pros	Cons
Hidden Markov Model	training is simple when states are observed	difficult to include features
Maximum Entropy Markov Model	easy to include features	suffer from “label bias problem” (prefer states with lower number of transitions)
Conditional Random Field	easy to include features, do not suffer from “label bias problem”	training is relatively complex

Deep Learning Tagger Architecture

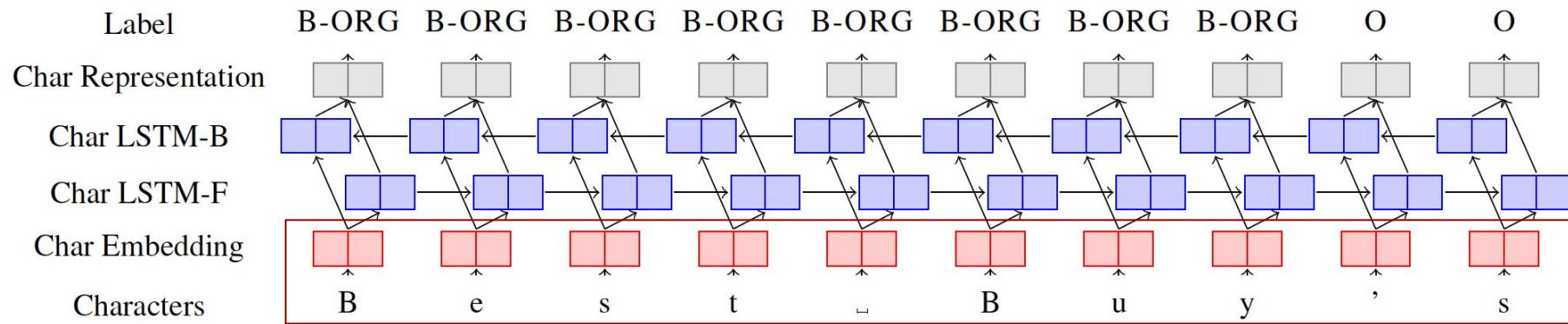


Input Layer - Word Embedding



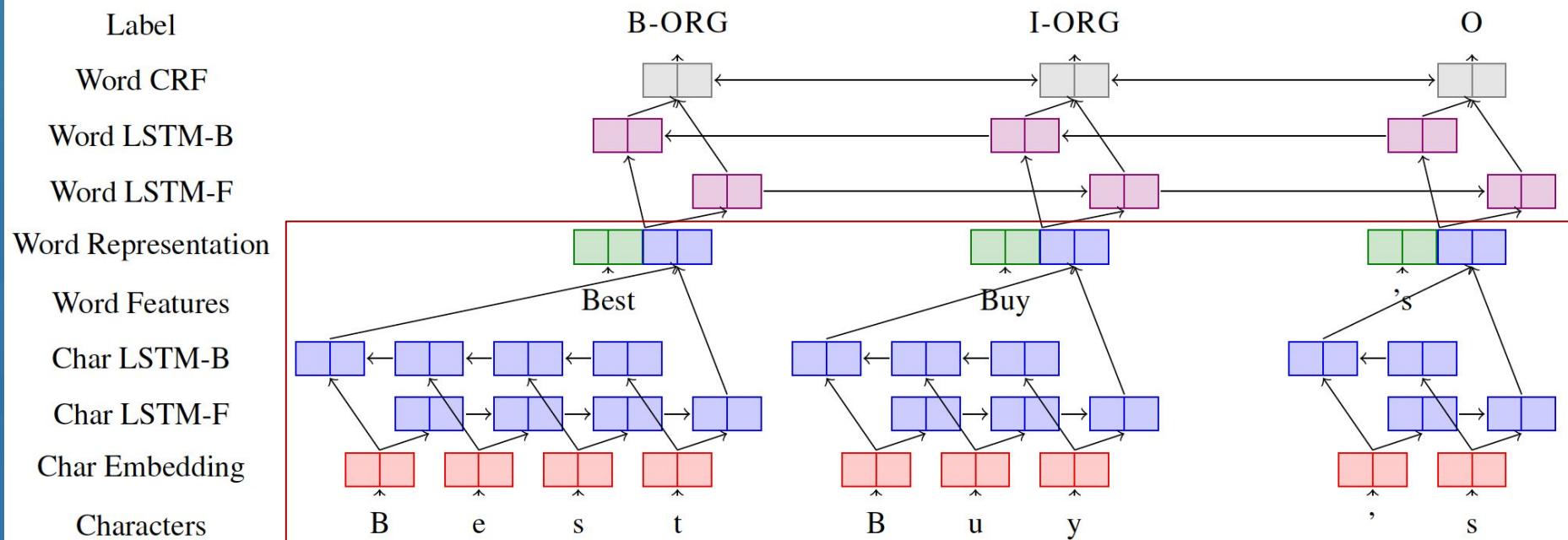
[Yadav, et. al. 2018]

Input Layer - Char Embedding



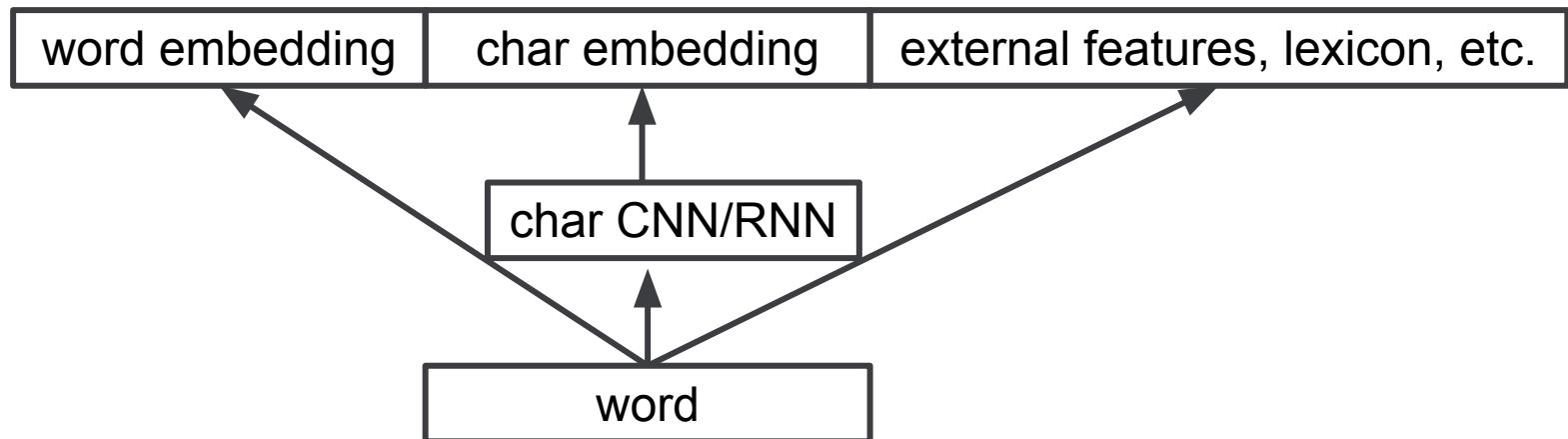
[Yadav, et. al. 2018]

Input Layer - Word and Char Embedding

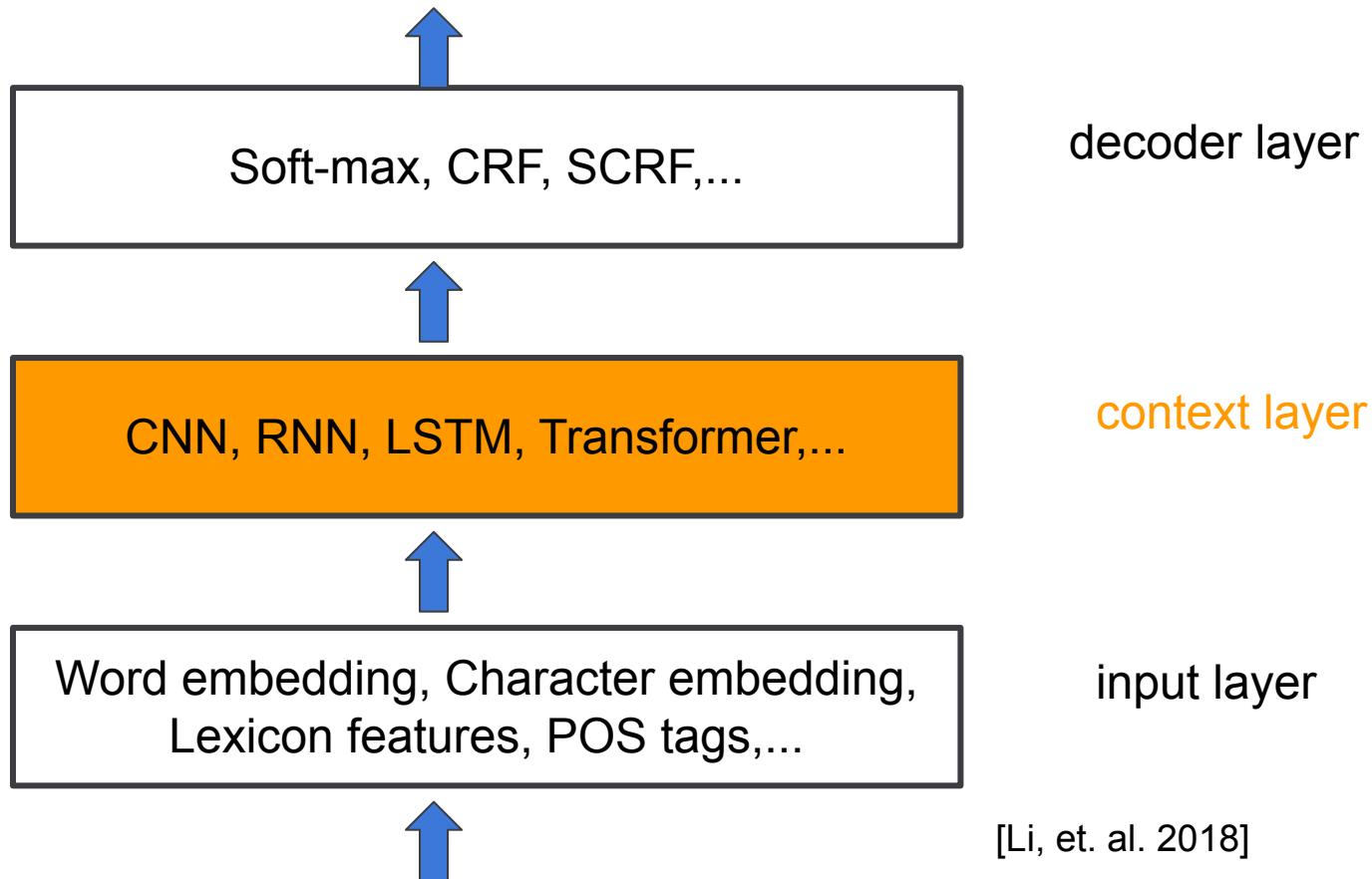


[Yadav, et. al. 2018]

Summary - Input Layer

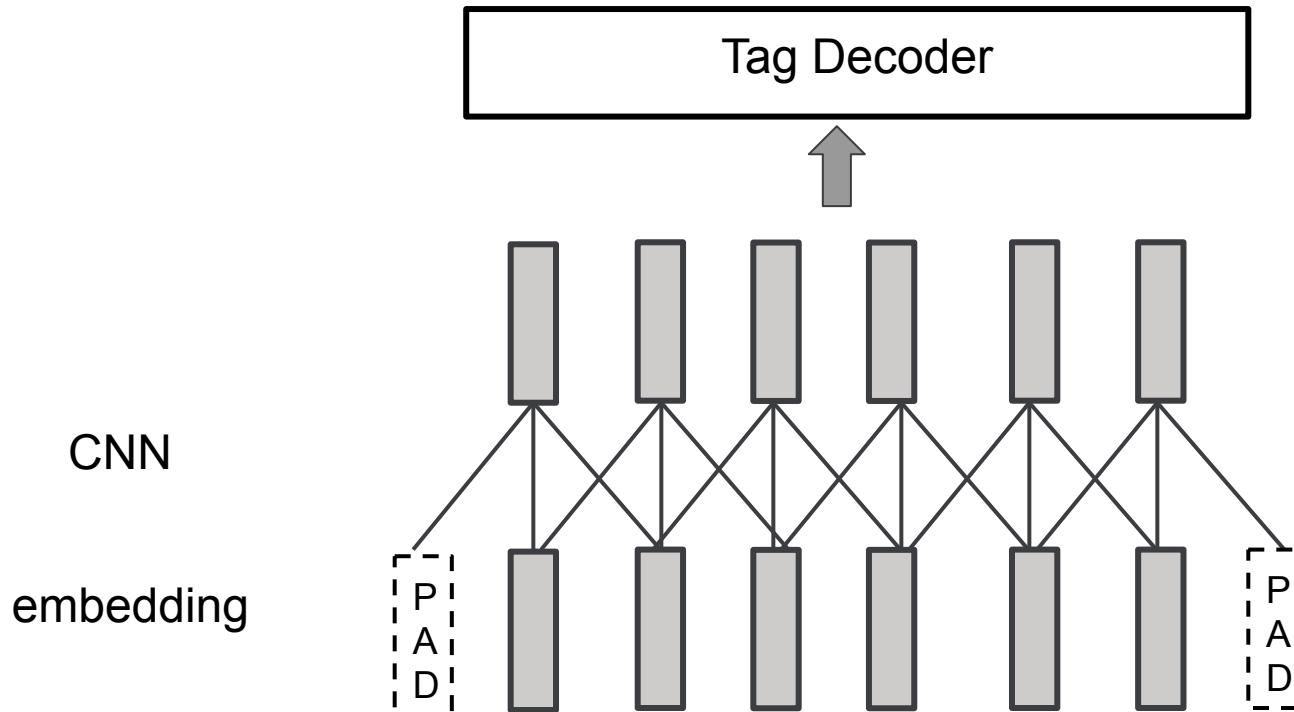


Deep Learning Tagger Architecture



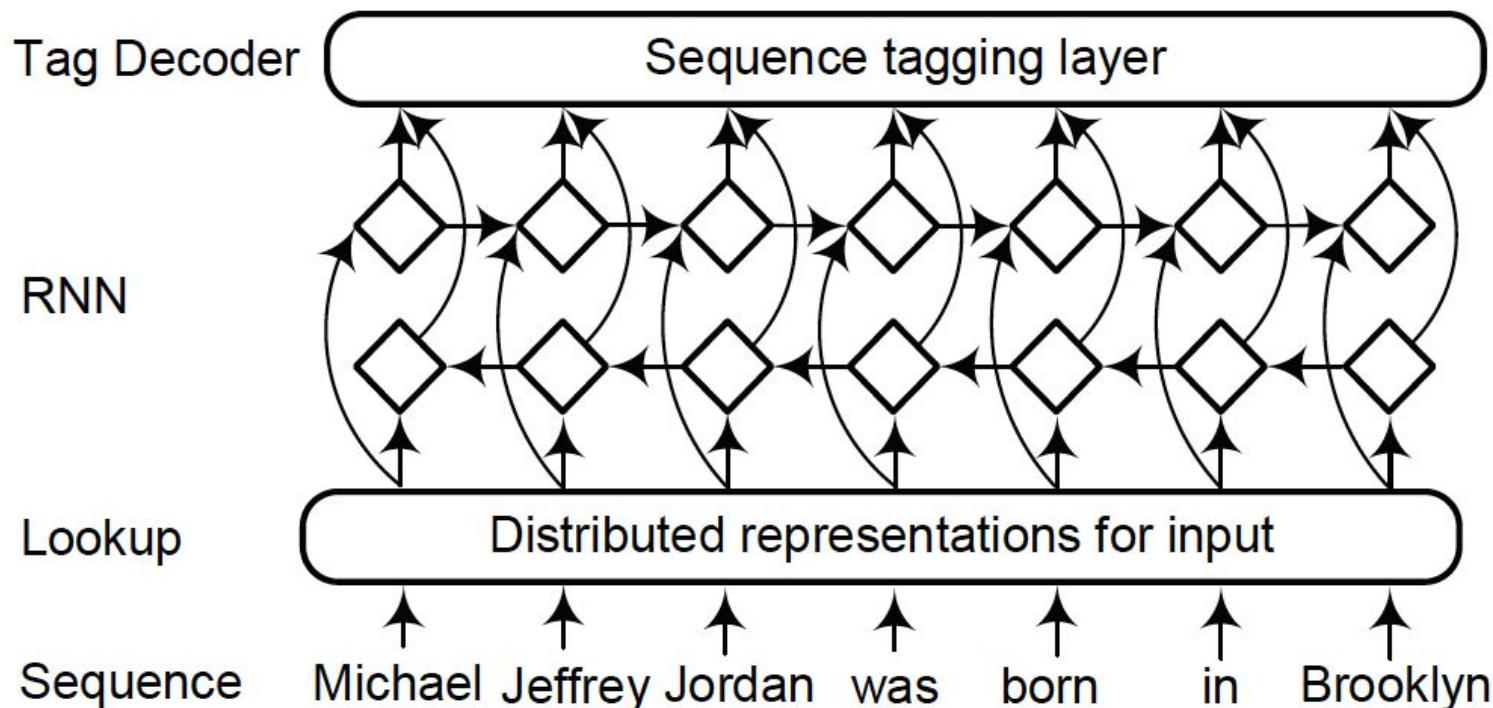
[Li, et. al. 2018]

Context Layer - CNN

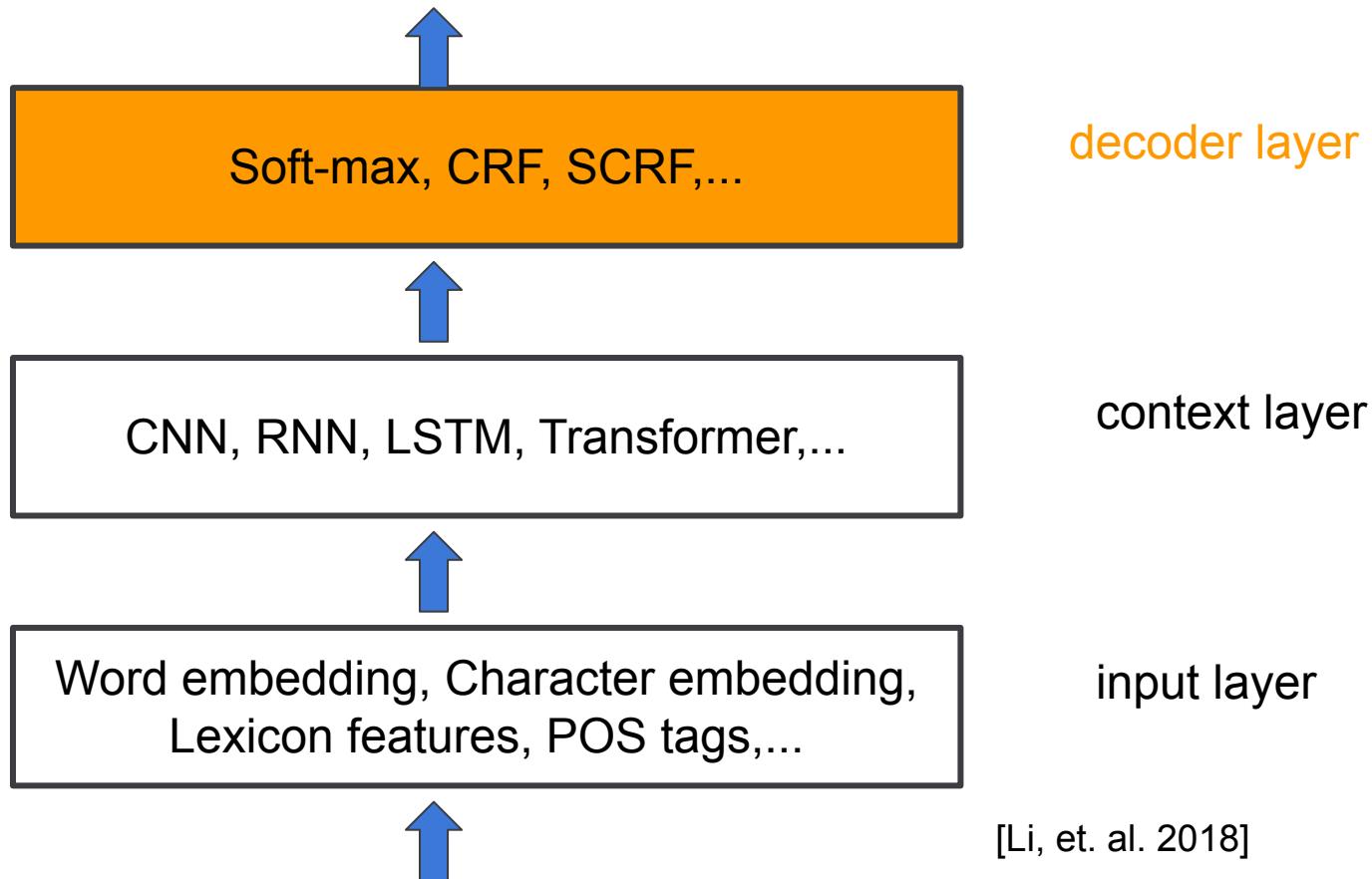


[Li et. al. 2018]

Context Layer - RNN

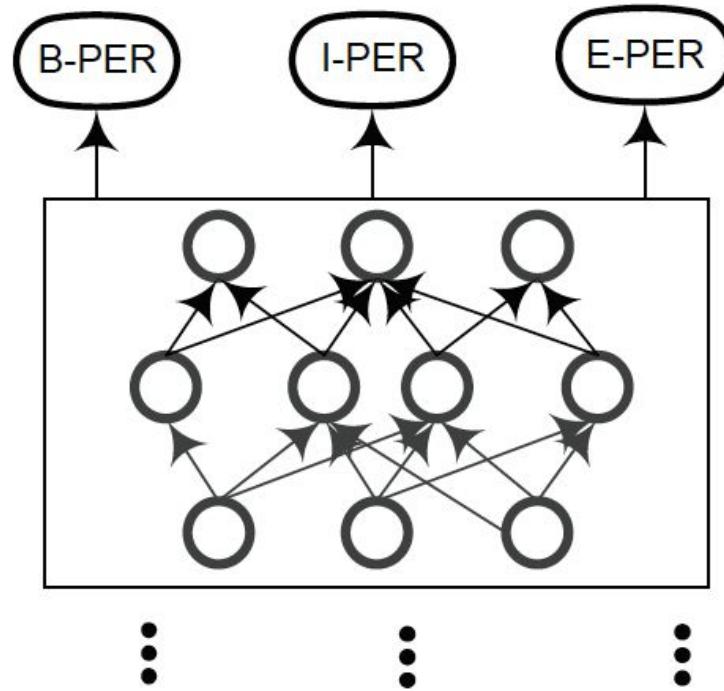


Deep Learning Tagger Architecture



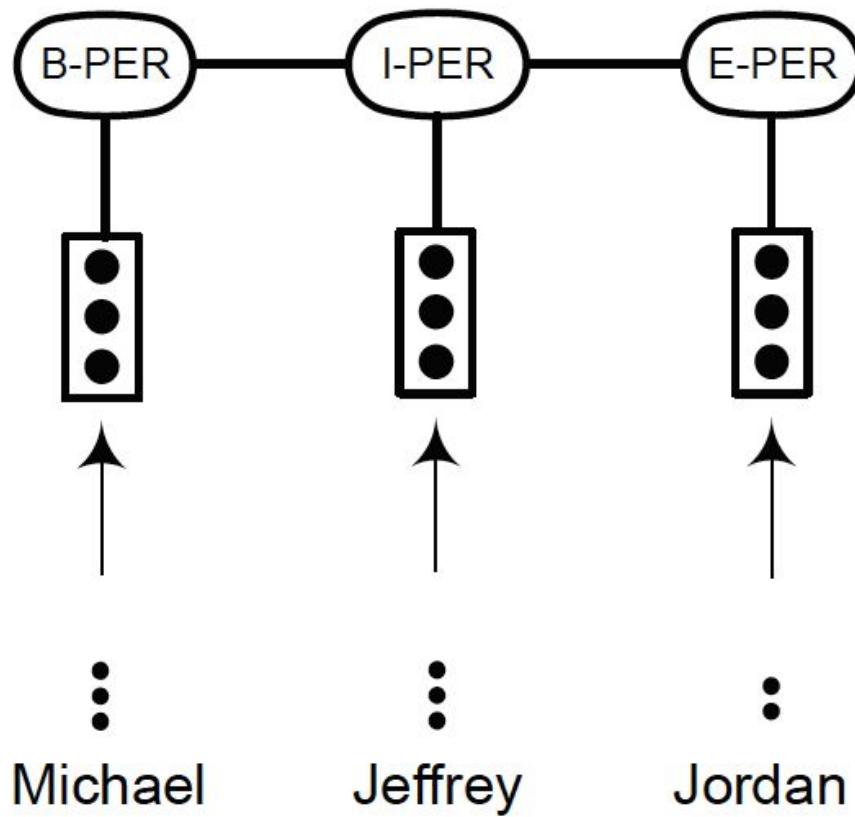
[Li, et. al. 2018]

Tag Decoder - MLP+softmax



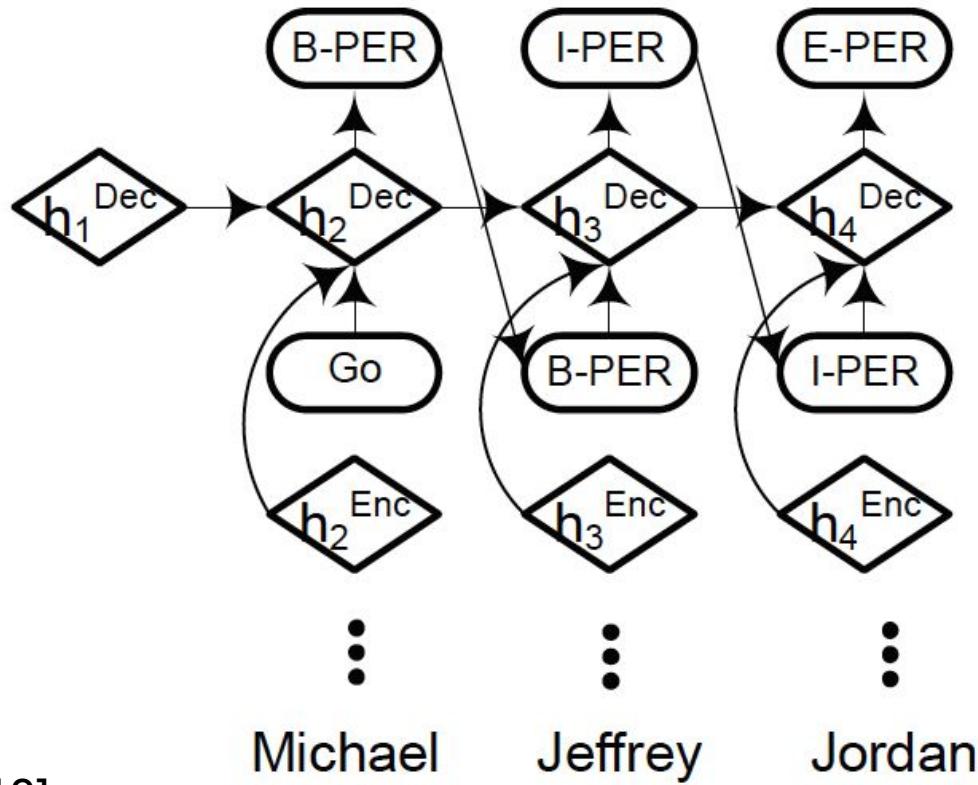
[Li et. al. 2018]

Tag Decoder - CRF



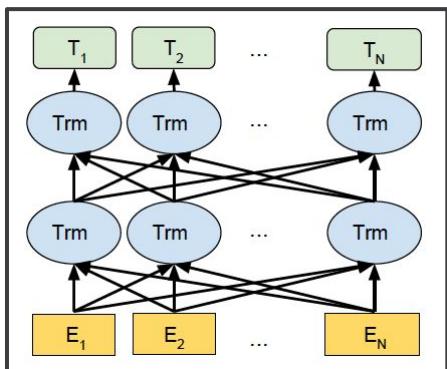
[Li et. al. 2018]

Tag Decoder - RNN



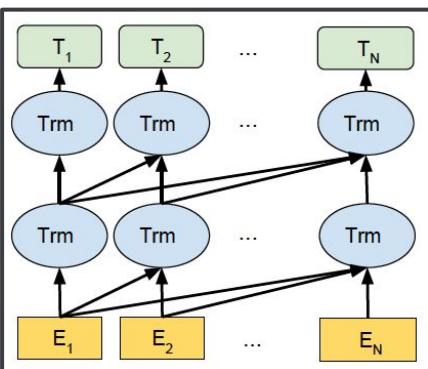
[Li et. al. 2018]

Pre-Training and Fine-Tuning



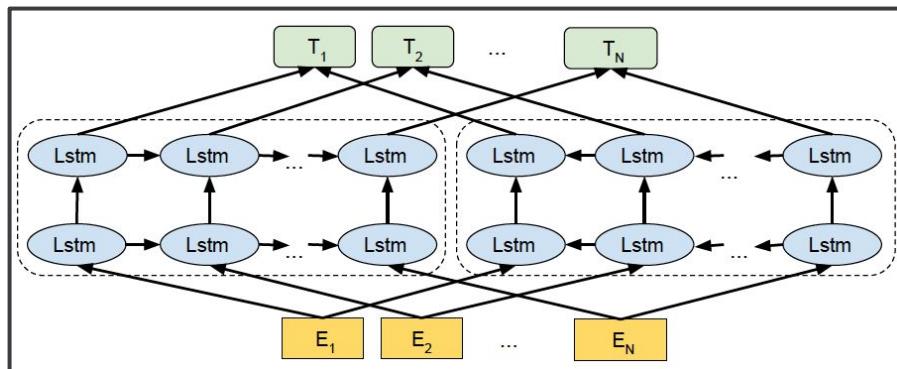
BERT

[Devlin et. al. 2018]



GPT

[Radford et. al. 2018]



ELMo

[Peters et. al. 2018]

Entity Tagging Evaluation

Exact-match evaluation

- segment match
- tag match

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad \text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$\text{F1-score} = \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}$$

Entity Tagging Training Dataset

Corpus	Year	Text Source	#Tags	URL
MUC-6	1995	Wall Street Journal texts	7	https://catalog.ldc.upenn.edu/LDC2003T13
MUC-6 Plus	1995	Additional news to MUC-6	7	https://catalog.ldc.upenn.edu/LDC96T10
MUC-7	1997	New York Times news	7	https://catalog.ldc.upenn.edu/LDC2001T02
CoNLL03	2003	Reuters news	4	https://www.clips.uantwerpen.be/conll2003/ner/
ACE	2000 - 2008	Transcripts, news	7	https://www.ldc.upenn.edu/collaborations/past-projects/ace
OntoNotes	2007 - 2012	Magazine, news, conversation, web	89	https://catalog.ldc.upenn.edu/LDC2013T19
W-NUT	2015 - 2018	User-generated text	18	http://noisy-text.github.io
BBN	2005	Wall Street Journal texts	64	https://catalog.ldc.upenn.edu/ldc2005t33
NYT	2008	New York Times texts	5	https://catalog.ldc.upenn.edu/LDC2008T19
WikiGold	2009	Wikipedia	4	https://figshare.com/articles/Learning_multilingual_named_entity_recognition_from_Wikipedia/5462500
WiNER	2012	Wikipedia	4	http://rali.iro.umontreal.ca/rali/en/winer-wikipedia-for-ner
WikiFiger	2012	Wikipedia	113	https://github.com/xiaoling/figer
N ³	2014	News	3	http://aksw.org/Projects/N3NERNEDNIF.html
GENIA	2004	Biology and clinical texts	36	http://www.geniaproject.org/home
GENETAG	2005	MEDLINE	2	https://sourceforge.net/projects/bioc/files/
FSU-PRGE	2010	PubMed and MEDLINE	5	https://julielab.de/Resources/FSU_PRGE.html
NCBI-Disease	2014	PubMed	790	https://www.ncbi.nlm.nih.gov/CBBresearch/Dogan/DISEASE/
BC5CDR	2015	PubMed	3	http://bioc.sourceforge.net/
DFKI	2018	Business news and social media	7	https://dfki-lt-re-group.bitbucket.io/product-corpus/

[Li et. al. 2018]

Entity Tagging on CoNLL03 English

Source	Method	F1 score
[Passos et al. 2014]	CRF	90.90
[Huang et al. 2015]	Bi-LSTM+CRF	84.26
[Collobert et al. 2011]	Conv-CRF	89.59
[Kuru et al. 2016]	char embedding	84.52
[Chiu and Nichols 2015]	word + char embedding	91.62
[Devlin et. al. 2018]	BERT Large	92.8

References - Entity Tagging

- [Baum et. al. 1966] Baum, L. E.; Petrie, T. (1966). "Statistical Inference for Probabilistic Functions of Finite State Markov Chains". *The Annals of Mathematical Statistics*. 37 (6): 1554–1563.
- [Chiu and Nichols, 2015] Jason PC Chiu and Eric Nichols. 2015. Named entity recognition with bidirectional lstm-cnns. arXiv preprint arXiv:1511.08308
- [Collins and Singer 1999] Michael Collins and Yoram Singer. 1999. Unsupervised models for named entity classification. In 1999 Joint SIGDAT Conference on Empirical Methods in Natural Language Processing and Very Large Corpora.
- [Collobert et. al. 2011] Ronan Collobert, Jason Weston, Leon Bottou, Michael Karlen, Koray Kavukcuoglu, and Pavel Kuksa. 2011. Natural language processing (almost) from scratch. *Journal of Machine Learning Research*, 12(Aug):2493–2537.
- [Collobert and Weston 2008] Ronan Collobert and Jason Weston. 2008. A unified architecture for natural language processing: Deep neural networks with multitask learning. In Proceedings of the 25th international conference on Machine learning, pages 160–167. ACM.
- [Devlin et. al. 2018] Devlin, Jacob and Chang, Ming-Wei and Lee, Kenton and Toutanova, Kristina. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, arXiv preprint arXiv:1810.04805, 2018
- [Huang et. al. 2015] Zhiheng Huang, Wei Xu, and Kai Yu. 2015. Bidirectional lstm-crf models for sequence tagging. arXiv preprint arXiv:1508.01991
- [Kuru et. al. 2016] Onur Kuru, Ozan Arkan Can, and Deniz Yuret. 2016. Charner: Character-level named entity recognition. In Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers, pages 911–921
- [Lafferty et. al. 2001] Lafferty, J., McCallum, A., Pereira, F. (2001). "Conditional random fields: Probabilistic models for segmenting and labeling sequence data". Proc. 18th International Conf. on Machine Learning. Morgan Kaufmann. pp. 282–289
- [Li et. al. 2018] Jing Li, Aixin Sun, Jianglei Han, Chenliang Li, A Survey on Deep Learning for Named Entity Recognition, Dec. 2018, arXiv preprint, <https://arxiv.org/abs/1812.09449>
- [McCallum et. al. 2000] McCallum, Andrew; Freitag, Dayne; Pereira, Fernando (2000). "Maximum Entropy Markov Models for Information Extraction and Segmentation" (PDF). Proc. ICML 2000. pp. 591–598

References (continued)

- [Passos 2014] Alexandre Passos, Vineet Kumar, and Andrew McCallum. 2014. Lexicon infused phrase embeddings for named entity resolution. arXiv preprint arXiv:1404.5367
- [Peters et. al. 2018] Matthew Peters, Mark Neumann, Mohit Iyyer, Matt Gardner, Christopher Clark, Kenton Lee, and Luke Zettlemoyer. 2018a. Deep contextualized word representations. In NAACL.
- [Radford et.al. 2018] Alec Radford, Karthik Narasimhan, Tim Salimans, and Ilya Sutskever. 2018. Improving language understanding with unsupervised learning. Technical report, OpenAI.
- [Yadav et. al. 2018] Vikas Yadav, Steven Bethard, A Survey on Recent Advances in Named Entity Recognition from Deep Learning models, Proceedings of the 27th International Conference on Computational Linguistics, pages 2145–2158 Santa Fe, New Mexico, USA, August 20-26, 2018.

Deep NLP in Search and Recommender Systems

- Language Understanding
 - Entity Tagging: word level prediction
 - **Entity Disambiguation: knowledge base entity prediction**
 - Intent Classification: sentence level prediction
 - Sentiment Analysis and Opinion Mining
- Document Retrieval and Ranking
 - Efficient Candidate Retrieval
 - Deep Ranking Models
- Language Generation for Assistance
 - Auto Completion
 - Query Suggestion
 - Spell Correction
 - Conversational Recommendation

Entity Disambiguation

- Problem statement
- Motivation
- Challenges
- System architecture
 - Candidate Entity Generation
 - Candidate Entity Selection
 - Joint entity tagging and disambiguation
- Evaluation

Entity Disambiguation - Problem Statement

- Also known as entity linking
- Resolves mentions to entities in a given knowledge base
 - Mentions are mostly named entities
 - Example of knowledge base: freebase, wikipedia
- Examples: Jaguar
 - The prey saw the jaguar cross the jungle. *KB:Jaguar*
 - The man saw a jaguar speed on the highway. *KB:Jaguar_Cars*

Motivation

Increase the quality of the retrieved results

Example:

Which Michael Jordan were you looking for?

Michael Jordan (disambiguation)

From Wikipedia, the free encyclopedia

Michael or Mike Jordan may refer to:

[People](#) [edit]

[Sports](#) [edit]

- Michael Jordan (born 1963), American basketball player and businessman
- Michael Jordan (footballer) (born 1986), English goalkeeper
- Mike Jordan (racing driver) (born 1958), English racing driver
- Mike Jordan (baseball, born 1863) (1863–1940), baseball player
- Mike Jordan (cornerback) (born 1992), American football cornerback
- Michael Jordan (offensive lineman), American football offensive lineman
- Michael-Hakim Jordan (born 1977), American professional basketball player
- Michal Jordán (born 1990), Czech ice hockey player

[Other people](#) [edit]

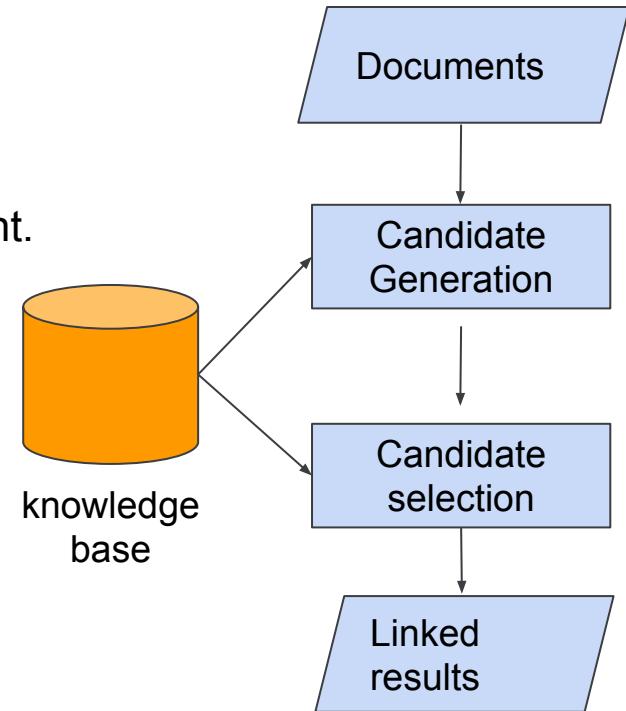
- Michael B. Jordan (born 1987), American actor
- Michael I. Jordan (born 1956), American researcher in machine learning and artificial intelligence
- Michael Jordan (insolvency baron) (born 1931), English businessman
- Michael Jordan (Irish politician), Irish Farmers' Party TD from Wexford, 1927–1932
- Michael H. Jordan (1936–2010), American executive for CBS, PepsiCo, Westinghouse
- Michael Jordan (mycologist), English mycologist

Challenges

- Name variation: *New York* vs *Big Apple*
- Ambiguity: *Michael Jordan*
- Metonymy: *Beijing* (city or Chinese government)
- Absence: no entries in knowledge base
- Evolving information: new company names, e.g. *tic-tok*.

System Architecture

1. Candidate entity generation
 - a. Name dictionary based techniques.
 - b. Surface form expansion from the local document.
 - c. Methods based on search engines.
2. Candidate entity selection
 - a. Graph-based methods
 - b. Text-based methods



Candidate Entity Generation

- Name dictionary based techniques
 - Entity pages
 - Redirect pages
 - Disambiguation pages
 - Bold phrases from the first paragraphs
 - Hyperlinks in Wikipedia articles
- Surface form expansion
 - Heuristic based methods
 - Supervised learning methods
- Methods based on search engines

[Shen et. al. 2014]

Michael Jordan (disambiguation)

From Wikipedia, the free encyclopedia

Michael or Mike Jordan may refer to:

People [\[edit\]](#)

Sports [\[edit\]](#)

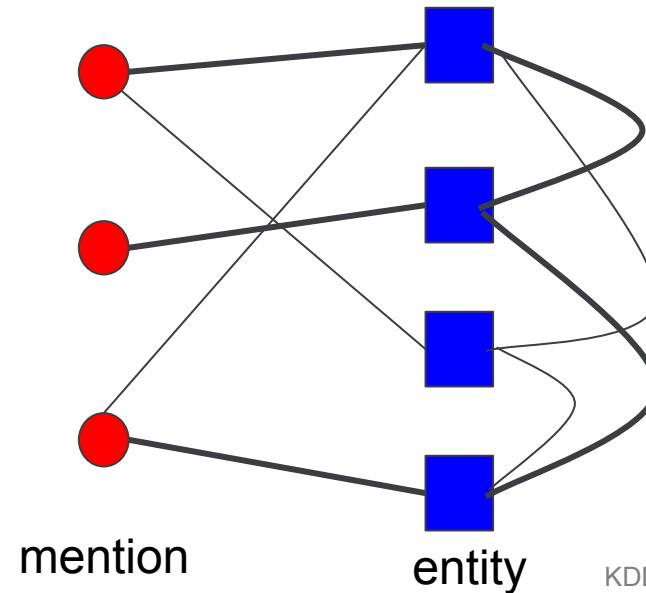
- Michael Jordan (born 1963), American basketball player and businessman
- Michael Jordan (footballer) (born 1986), English goalkeeper
- Mike Jordan (racing driver) (born 1958), English racing driver
- Mike Jordan (baseball, born 1863) (1863–1940), baseball player
- Mike Jordan (cornerback) (born 1992), American football cornerback
- Michael Jordan (offensive lineman), American football offensive lineman
- Michael-Hakim Jordan (born 1977), American professional basketball player
- Michal Jordán (born 1990), Czech ice hockey player

Other people [\[edit\]](#)

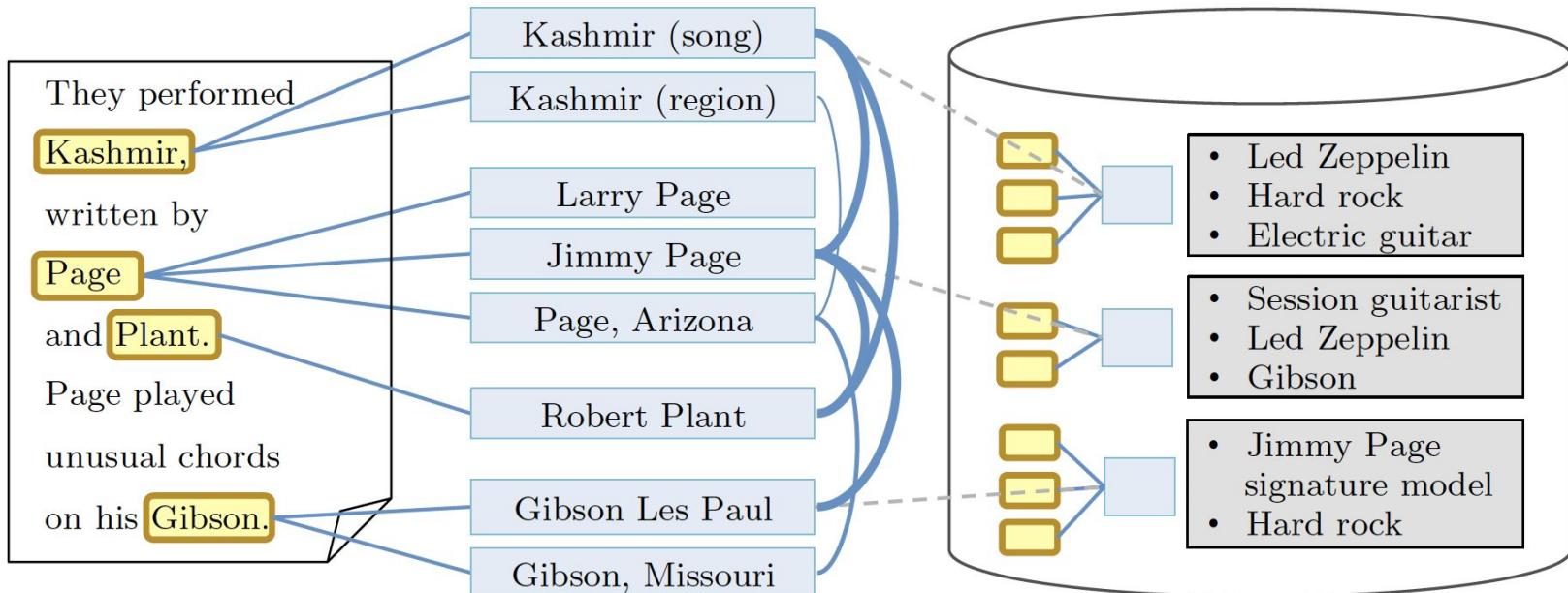
- Michael B. Jordan (born 1987), American actor
- Michael I. Jordan (born 1956), American researcher in machine learning and artificial intelligence
- Michael Jordan (insolvency baron) (born 1931), English businessman
- Michael Jordan (Irish politician), Irish Farmers' Party TD from Wexford, 1927–1932
- Michael H. Jordan (1936–2010), American executive for CBS, PepsiCo, Westinghouse
- Michael Jordan (mycologist), English mycologist

Candidate Entity Selection - Graph based models

- knowledge graph lends itself to graph based methods.
- Build a weighted undirected graph with mentions and candidate entities as nodes
 - weights on mention-entity edges.
 - weights on entity-entity edge.
- Compute an optimal subgraph:
 - optimality is model dependent.
 - contains all mentions
 - one entity per mention



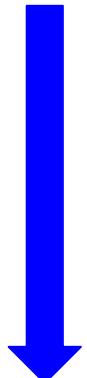
An Example Graph



[Hoffart, et. al. 2011]

Candidate Entity Selection - Text-Based Models

- Idea: find agreement (similarity) between entity and mention.
 - Example: computational intelligence by *Michael Jordan* and Stuart Russell from UC Berkeley
 - wiki/Michael_Jordan
 - wiki/Michael_I._Jordan
- Models according to context.
 - Direct Models.
 - use mention and candidate entities only.
 - Local Models
 - use local context around the mention.
 - Coherence Models.
 - use mentions and entities in a document.
 - Collaborative Models.
 - use mentions and entities across related documents.



increasing
context

Individual vs Joint Approach

Two types of approaches according to how mentions are resolved.

- Individual approach
 - Handle one mention at a time.
 - Rank the candidate entities.
 - Relatively low complexity.
- Joint approach
 - Treat all mentions in a document as a sequence.
 - Jointly optimize the mention-entity pairs.
 - High complexity, usually resorting to approximations.

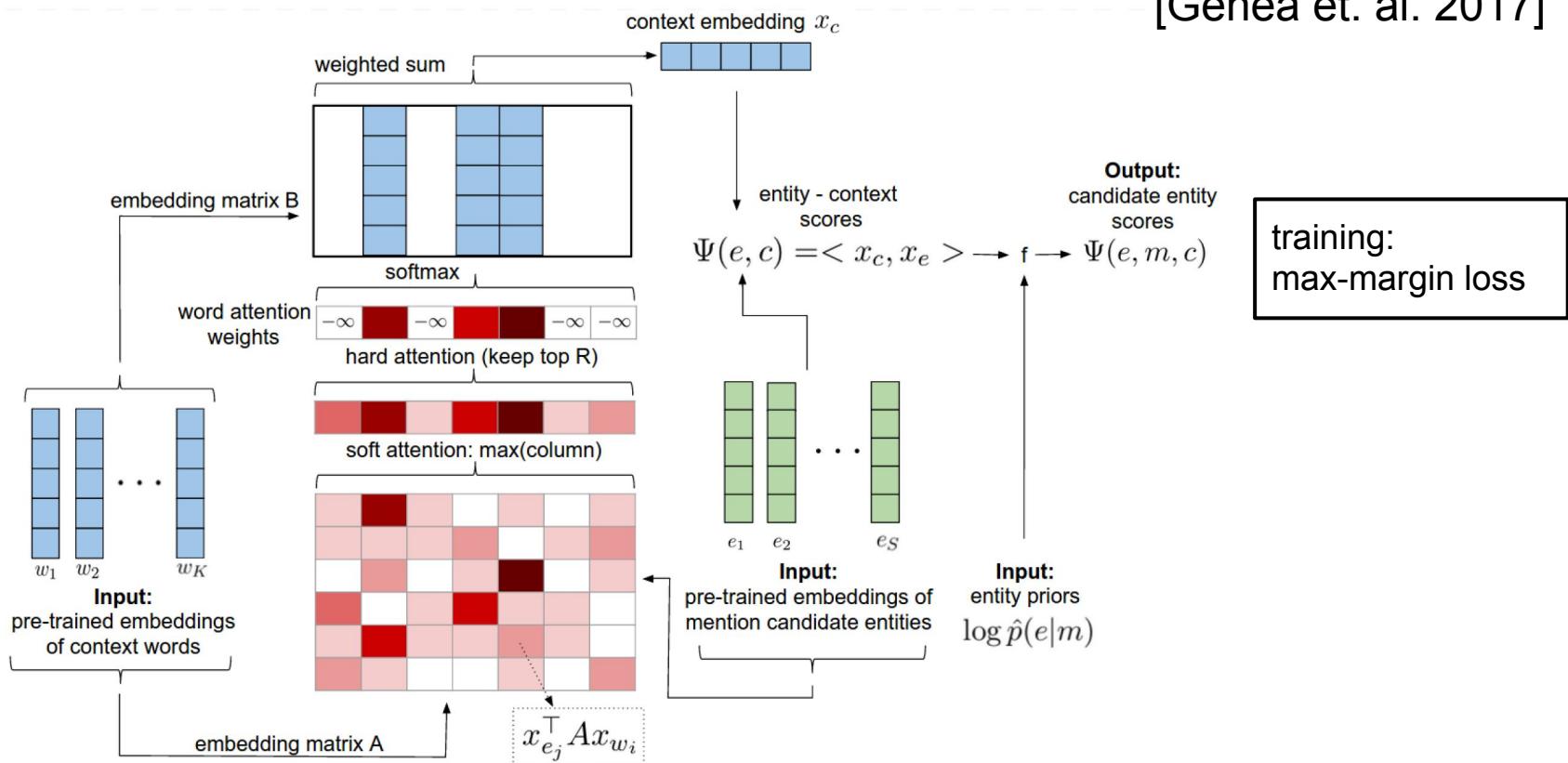
Features

Different features are used in candidate selection modeling.

- Traditional Features
 - entity popularity, entity type.
 - surface form similarity between entity and mention.
 - coherence between mapped entities
- Deep Features [Yamada et. al. 2016]
 - word embedding
 - entity embedding
 - similarity between words, context and candidate entities.

Example: Entity Disambiguation with Attention

[Genea et. al. 2017]

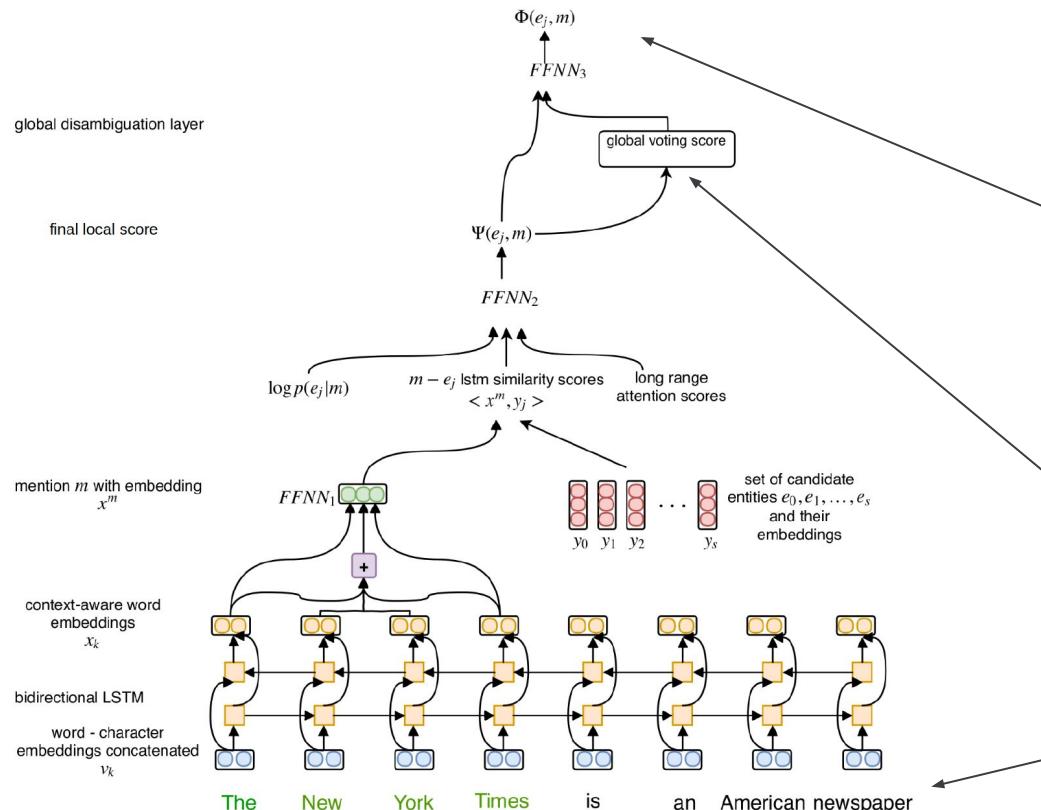


Joint Entity Tagging and Disambiguation

- Previously, we find entities first, then link them. But if entities are identified wrong, disambiguation will likely fail too.
 - Over-split: Romeo and Juliet by Shakespeare
 - Under-split: Baby Romeo and Juliet were born hours apart.
- We could do entity tagging and disambiguation jointly. [Sil, et. al.2013]
 - Over-generate candidate mentions
 - Generate possible entities per mention
 - Score non-overlapping mention-entity pair jointly.

A Joint Neural Entity Tagging and Disambiguation

[Kolitsas, et. al. 2018]



- Trained on golden mention-entity pairs.
- Used max-margin loss.

cosine similarity between current entity and average of other entities in the document

Select mentions that have at least one possible entity

Entity Disambiguation Evaluation

- Entity-tagging-style F1 score
 - A link is considered correct only if the mention matches the gold boundary and the linked entity is also correct
- Accuracy
 - another common metric, simpler than F1 score.
- TAC-KBP B-Cubed+ F1 score
 - not widely used

References - Named Entity Disambiguation

- [Cucerzan 2007] Cucerzan, Silviu. Large-Scale Named Entity Disambiguation Based on Wikipedia Data. Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL): 708–716.
- [Ganea et. al. 2017] Octavian-Eugen Ganea and Thomas Hofmann. 2017. Deep joint entity disambiguation with local neural attention. In Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, pages 2609–2619. Association for Computational Linguistics.
- [Globerson et. al. 2016] Amir Globerson, Nevena Lazic, Soumen Chakrabarti, Amarnag Subramanya, Michael Ringgaard, and Fernando Pereira. 2016. Collective entity resolution with multi-focal attention. In ACL (1).
- [Han et. al. 2011] Han, Xianpei; Sun, Le; Zhao, Jun. Collective Entity Linking in Web Text: A Graph-based Method.
- [Hoffart, et. al. 2011] Johannes Hoffart, Mohamed A. Yosef, Ilaria Bordino, Hagen Fürstenau, Manfred Pinkal, Marc Spaniol, Bilyana Taneva, Stefan Thater, and Gerhard Weikum. 2011. Robust disambiguation of named entities in text. In Proceedings of the Conference on Empirical Method in Natural Language Processing, pages 782–792. Association for Computational Linguistics.
- [Kolitsas et al., 2018] Nikolaos Kolitsas, Octavian Eugen Ganea, and Thomas Hofmann. End-to-end neural entity linking. In CoNLL, 2018.
- [Kulkarni et. al. 2009] S. Kulkarni, A. Singh, G. Ramakrishnan, and S. Chakrabarti, Collective annotation of Wikipedia entities in web text, in SIGKDD, 2009, pp. 457–466
- [Ling et al. 2015] Xiao Ling, Sameer Singh, and Daniel S. Weld. 2015. Design Challenges for Entity Linking. Transactions of the Association for Computational Linguistics, 3:315–328.
- [Le et. al., 2018] Phong Le and Ivan Titov. Improving entity linking by modeling latent relations between mentions. In ACL, 2018.
- [Rao et. al. 2013] Rao, Delip; McNamee, Paul; Dredze, Mark (2013). Entity Linking: Finding Extracted Entities in a Knowledge Base. Multi-source, Multilingual Information Extraction and Summarization. Springer Berlin Heidelberg: 93–115

References (continued)

- [Raiman et. al. 2018] Jonathan Raiman and Olivier Raiman. 2018. DeepType: Multilingual Entity Linking by Neural Type System Evolution. In Proc. of AAAI.
- [Shen et. al. 2014] Wei Shen, Jianyong Wang, and Jiawei Han. 2014. Entity linking with a knowledge base: Issues, techniques, and solutions. TKDE
- [Sil et. al. 2013] Avirup Sil and Alexander Yates. 2013. Re-ranking for joint named-entity recognition and linking. In Proceedings of the 22nd ACM international conference on Conference on information & knowledge management, pages 2369–2374. ACM
- [Sil et. al. 2016] Sil, A., and Florian, R. 2016. One for all: Towards language independent named entity linking. ACL.
- [Yamada et. al. 2016] Ikuya Yamada, Hiroyuki Shindo, Hideaki Takeda, and Yoshiyasu Takefuji. 2016. Joint learning of the embedding of words and entities for named entity disambiguation. CoNLL 2016, page 250.

Deep NLP in Search and Recommender Systems

- Language Understanding
 - Entity Tagging: word level prediction
 - Entity Disambiguation: knowledge base entity prediction
 - **Intent Classification: sentence level prediction**
 - Sentiment Analysis and Opinion Mining
- Document Retrieval and Ranking
 - Efficient Candidate Retrieval
 - Deep Ranking Models
- Language Generation for Assistance
 - Auto Completion
 - Query Reformulation
 - Spell Correction
 - Conversational Recommendation

Intent Classification

- Problem statement
- Deep Learning Models
 - fastText
 - CNN
 - Bi-RNN + Attention

Intent Classification - Problem Statement

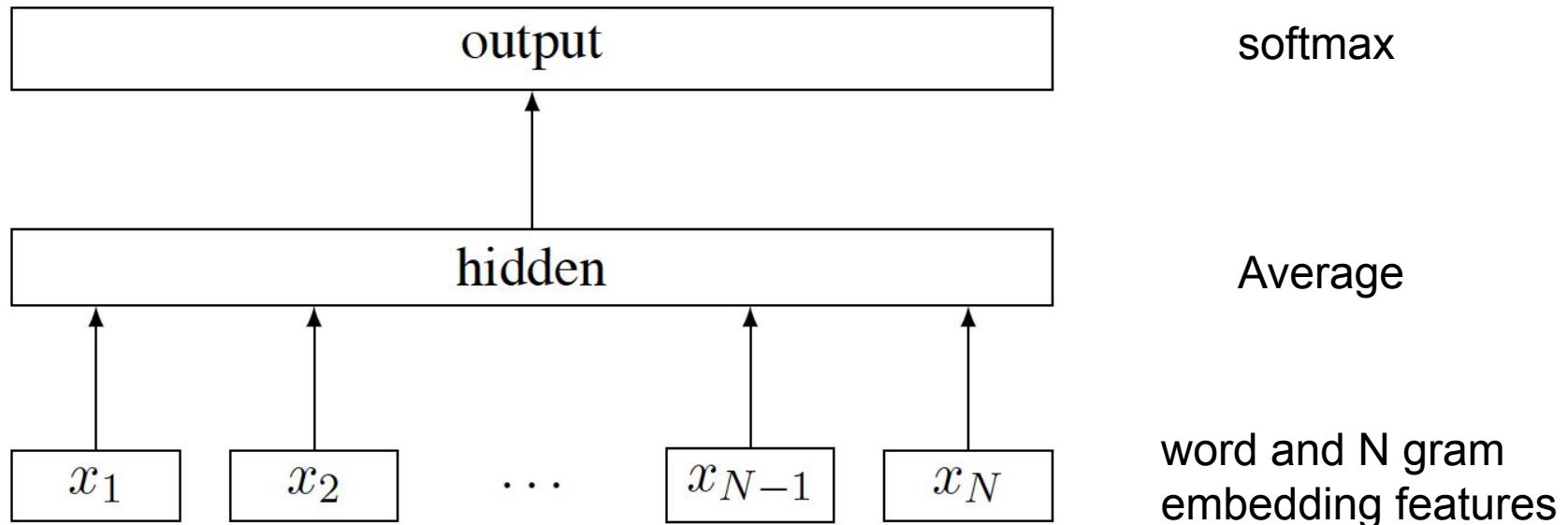
- Web search intent can be classified into 3 class: [Broder 2002]
 - **Navigational.** The immediate intent is to reach a particular site.
 - Greyhound Bus. Probable target <http://www.greyhound.com>
 - **Informational.** The intent is to acquire some information assumed to be present on one or more web pages.
 - San Francisco
 - **Transactional.** The intent is to perform some web-mediated activity.
 - Shopping activities
- In conversational AI, intent is usually task-specific, e.g.
 - I would like to book a flight from SFO to ANC: **FlightBooking**
 - What software can I use to view epub documents:
SoftwareRecommendation.
- We focus on the latter in this tutorial.

Intent Classification - Methods

- Traditional methods
 - Features: bag of words, n-gram, TF-IDF.
 - Models: logistic regression, naive Bayes, SVM, random forest.
- Deep learning methods
 - Word embedding + linear classifier (fastText) [Joulin et. al. 2016].
 - Convolutional neural networks [Hashemi 2016] .
 - Bi-RNN + attention (joint slot filling) [Liu et al.2016] .

Intent Classification - fastText

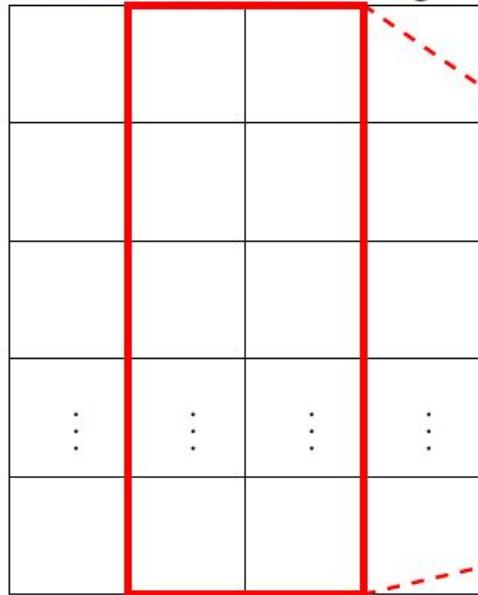
[Joulin et. al. 2016]



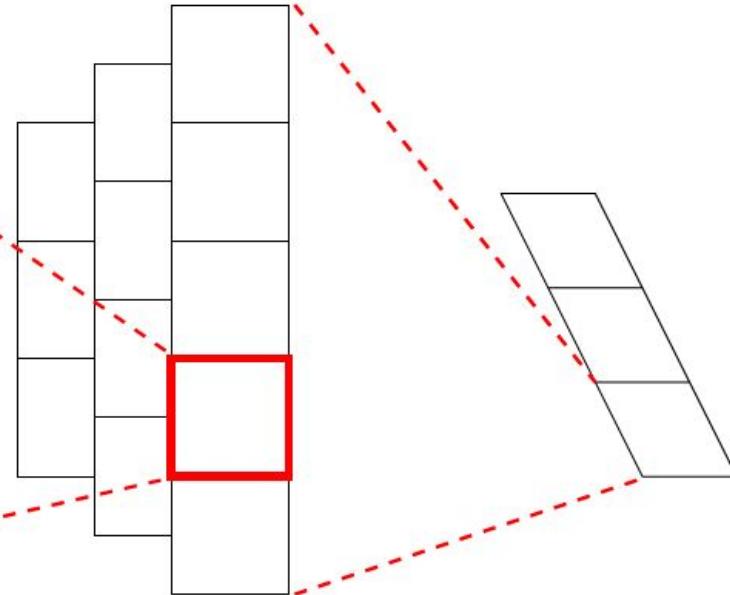
Intent Classification - CNN

[Hashemi 2016]

barack obama wife age



Word vectors in
query matrix

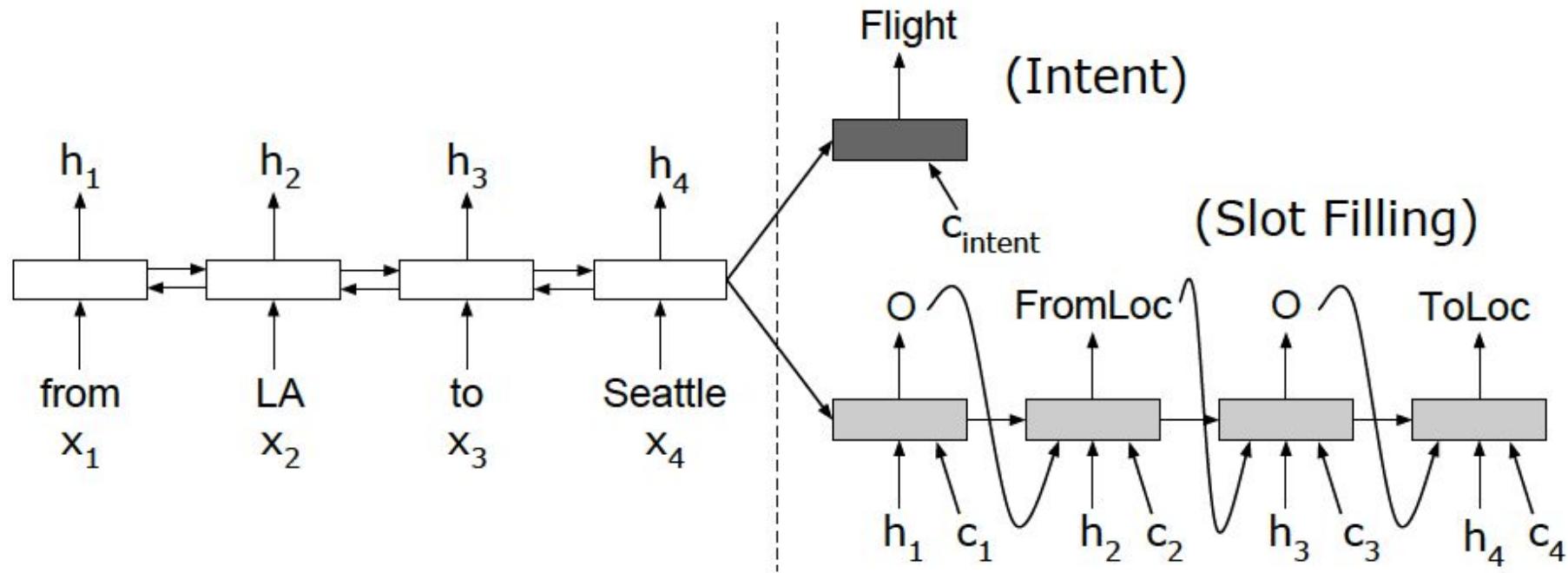


Multiple filters

Max pooling
Query vector

Intent Classification - Bi-RNN+Attention

[Liu et. al. 2016]



References - Intention Classification

- [Broder 2002] A. Broder. A taxonomy of web search. SIGIR Forum, 36(2):3–10, 2002.
- [Kim 2014] Y. Kim. Convolutional neural networks for sentence classification. In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), pages 1746–1751, Doha, Qatar, October 2014. Association for Computational Linguistics.
- [Joulin et. al. 2016] Joulin, A., Grave, E., Bojanowski, P., and Mikolov, T. (2017). Bag of tricks for efficient text classification. In Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics (EACL).
- [Lai et al.2015] Siwei Lai, Liheng Xu, Kang Liu, and Jun Zhao. 2015. Recurrent convolutional neural networks for text classification. In AAAI, pages 2267–2273
- [Liu et al.2016] Pengfei Liu, Xipeng Qiu, and Xuanjing Huang. 2016. Recurrent neural network for text classification with multi-task learning. arXiv preprint arXiv:1605.05101
- [Zhou et al.2016] Peng Zhou, Wei Shi, Jun Tian, Zhenyu Qi, Bingchen Li, Hongwei Hao, and Bo Xu. 2016. Attention-based bidirectional long short-term memory networks for relation classification. In The 54th Annual Meeting of the Association for Computational Linguistics, page 207.
- [Kim et. al. 2016] Joo-Kyung Kim, Gokhan Tur, Asli Celikyilmaz, Bin Cao, and Ye-Yi Wang. 2016. Intent detection using semantically enriched word embeddings. In Proceedings of SLT.
- [Hashemi 2016] Homa B Hashemi, Amir Asiaee, and Reiner Kraft. 2016. Query intent detection using convolutional neural networks. In International Conference on Web Search and Data Mining, Workshop on Query Understanding.
- [Liu et. al. 2016]. Liu and I. Lane, “Attention-based Recurrent Neural Network Models for Joint Intent Detection and Slot Filling,” in Interspeech, 2016.
- [Shi et. al. 2016] Y. Shi, K. Yao, L. Tian, and D. Jiang, “Deep Istm based feature mapping for query classification,” in Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 2016, pp. 1501–1511.

Deep NLP in Search and Recommender Systems

- Language Understanding
 - Entity Tagging: word level prediction
 - Entity Disambiguation: knowledge base entity prediction
 - Intent Classification: sentence level prediction
 - **Sentiment Analysis and Opinion Mining**
- Document Retrieval and Ranking
 - Efficient Candidate Retrieval
 - Deep Ranking Models
- Language Generation for Assistance
 - Auto Completion
 - Query Suggestion
 - Spell Correction
 - Conversational Recommendation

Sentiment Analysis and Opinion Mining

- Problem Statement
- Challenges
- Traditional methods
- Deep learning methods

Sentiment Analysis - Problem Statement

- Interchangeable with Opinion Mining in this tutorial.
- sentiment analysis mainly studies opinions which express or imply positive or negative sentiments [Liu 2012]
- Granularity:
 - Document level
 - Sentence level
 - Aspect level

Posted by: John Smith Date: September 10, 2011

“(1) I bought a Canon G12 camera six months ago. (2) I simply love it.
(3) The **picture quality** is amazing. (4) The **battery life** is also long.
(5) However, my wife thinks it is too heavy for her.”

Sentiment Analysis - Challenges

- Word ambiguity
 - This camera *sucks*.
 - This vacuum cleaner really *sucks*.
- Sentiment-less sentiment words
 - Can you tell me which camera is good?
- Sarcasm
 - What a great car! It stopped working in two days.
- Implicit sentiment
 - This washer uses a lot of water.

Traditional Methods

- Document/Sentence Sentiment Analysis
 - unsupervised: rule based, sentiment orientation
 - supervised: naive Bayes, logistic regression, SVM
- Aspect-based Sentiment Analysis
 - Aspect extraction
 - Frequent nouns and noun phrases
 - Opinion and target relations
 - Sequential tagging: HMM, CRF
 - Topic modeling
 - Aspect sentiment classification
 - Similar methods in document/sentence analysis.
 - limited to aspect dependent features

The **screen** is very **clear** but the
battery life is too **short**.

Deep Learning Approaches

- Document/Sentence Sentiment Analysis
 - Use embedding alone
 - Use CNN/RNN
- Aspect-based Sentiment Analysis
 - Aspect extraction
 - CNN or RNN
 - Aspect sentiment analysis
 - attention-based RNN, memory networks

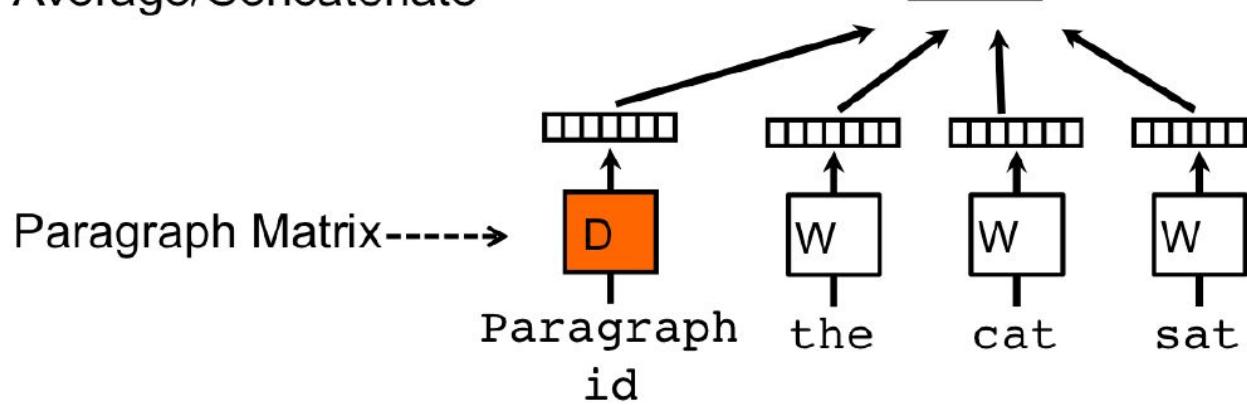
The **screen** is very **clear** but the **battery life** is too **short**.

Document Sentiment Analysis - Paragraph Vector

Classifier

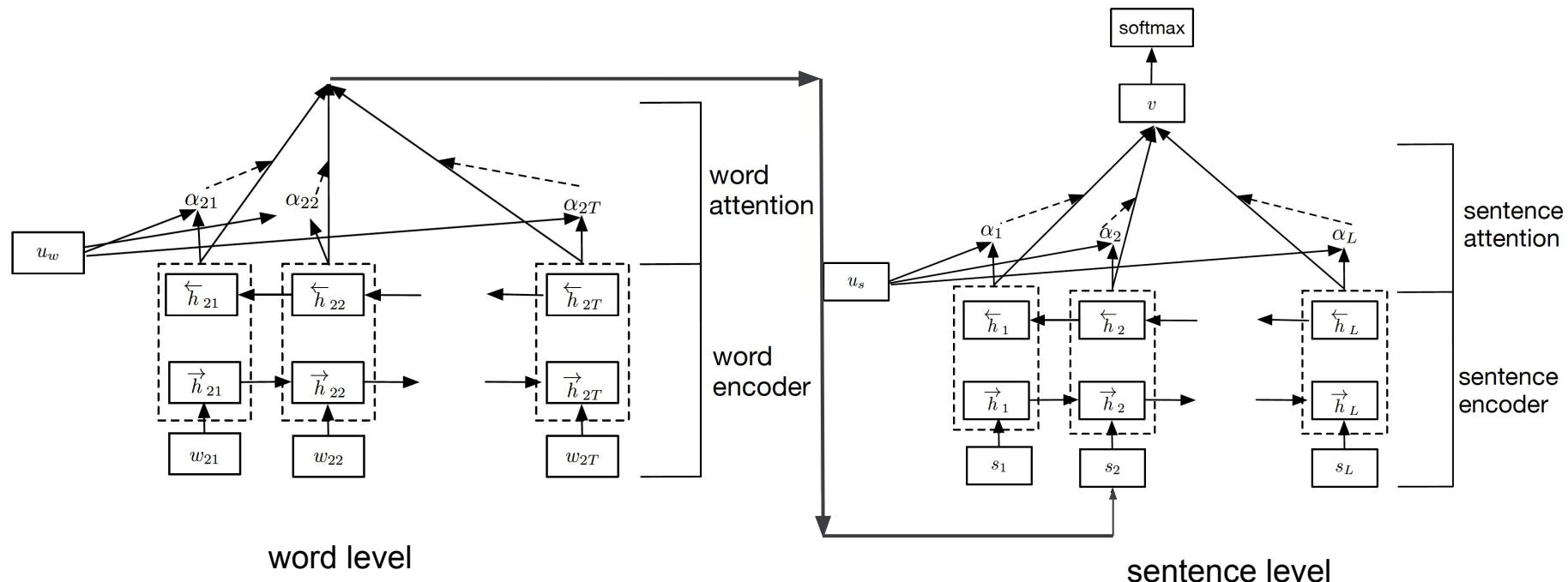
$$p(w_t | w_{t-k}, \dots, w_{t+k}) = \frac{e^{y_{w_t}}}{\sum_i e^{y_i}}$$
$$y = b + U h(w_{t-k}, \dots, w_{t+k}; W)$$

Average/Concatenate



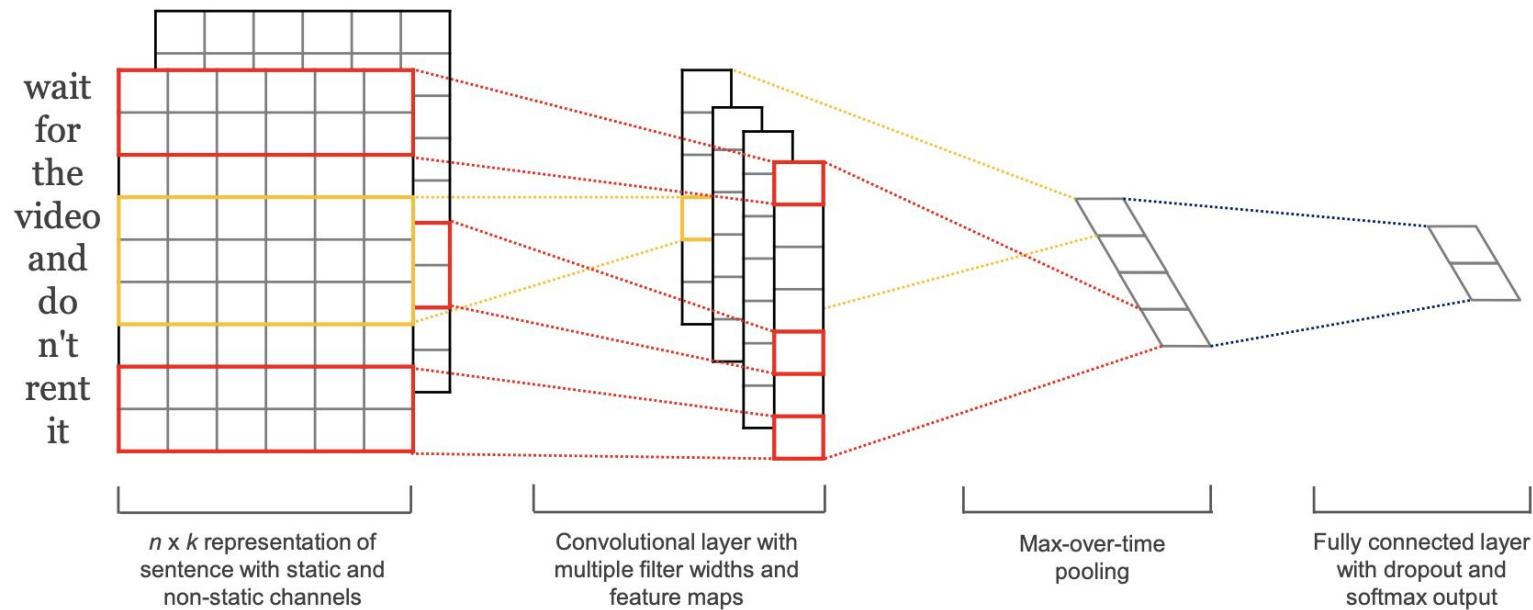
[Le et. al. 2014]

Document Sentiment Analysis - RNN



[Yang et. al. 2014]

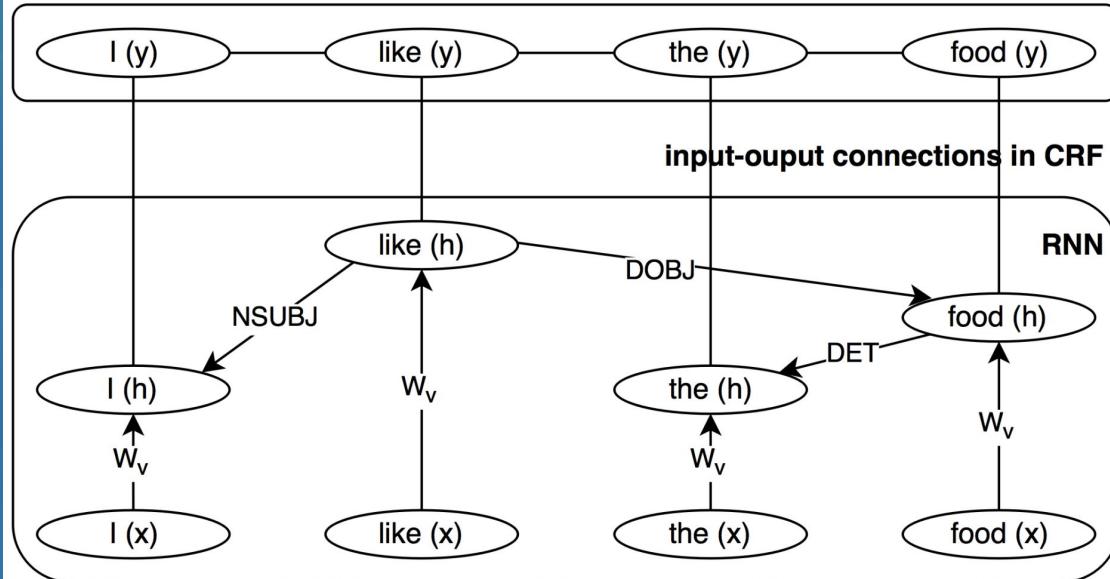
Sentence Sentiment Analysis - CNN



[Kim 2014]

Aspect and Opinion Extraction - RNN

pairwise connections in linear-chain CRF



[Wang et. al. 2016]

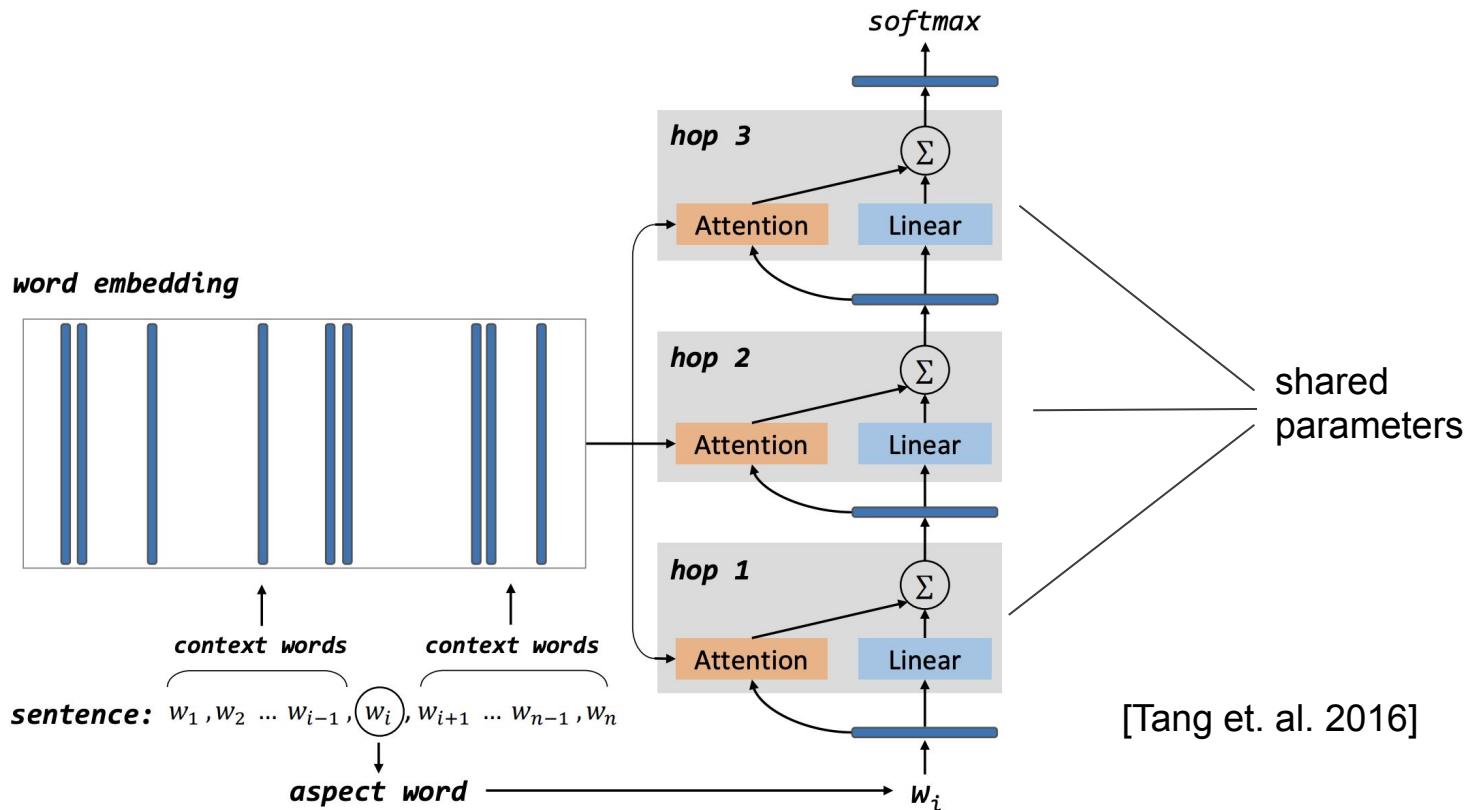
$$\begin{aligned} h_I &= f(W_v \cdot x_I + b), \\ h_{\text{the}} &= f(W_v \cdot x_{\text{the}} + b), \end{aligned}$$

$$\begin{aligned} h_{\text{food}} &= f(W_v \cdot x_{\text{food}} + W_{\text{DET}} \cdot h_{\text{the}} + b), \\ h_{\text{like}} &= f(W_v \cdot x_{\text{like}} + W_{\text{DOBJ}} \cdot h_{\text{food}} \\ &\quad + W_{\text{NSUBJ}} \cdot h_I + b). \end{aligned}$$

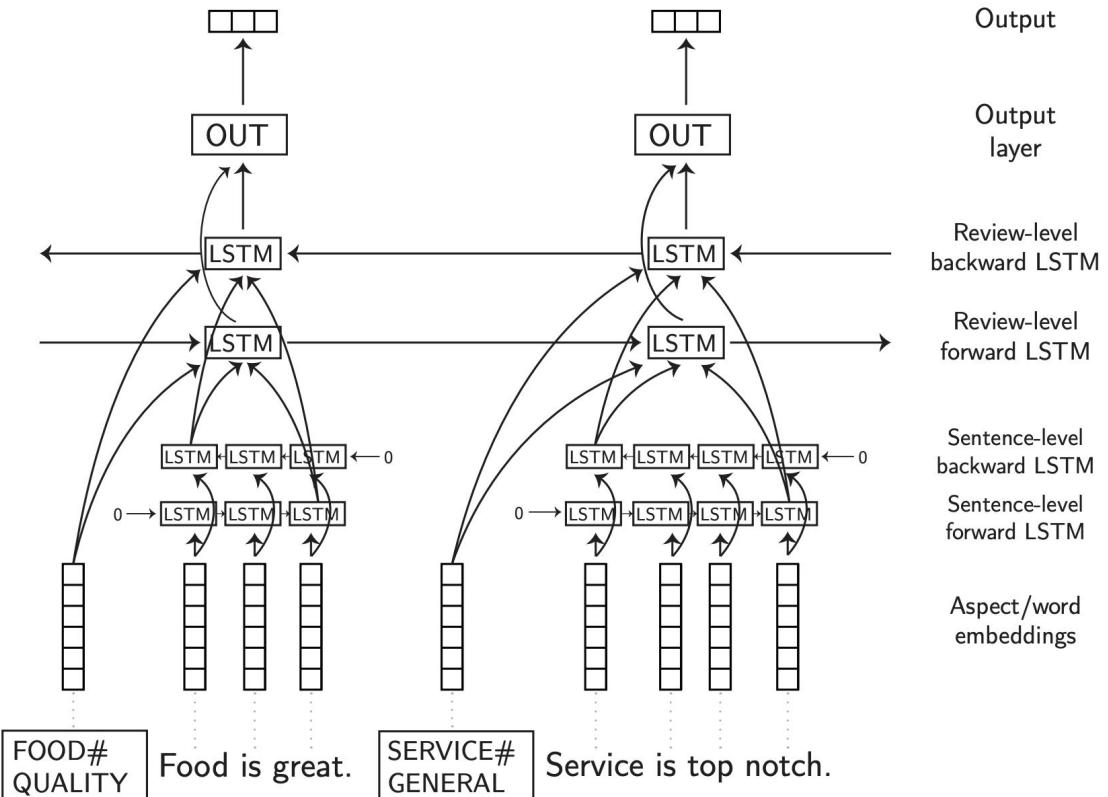
Labels:

- BA (beginning of aspect)
- IA (inside of aspect),
- BO (beginning of opinion)
- IO (inside of opinion)
- O (others)

Aspect Level Sentiment Analysis - Memory Networks



Aspect Level Sentiment Analysis - Hierarchical RNN



- one aspect per sentence
- aspects are given
- Utilize the relationship between aspects

[Ruder et. al. 2016]

References - Sentiment Analysis and Opinion Mining

- [Kim 2014] Y. Kim. Convolutional neural networks for sentence classification. In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), pages 1746–1751, Doha, Qatar, October 2014. Association for Computational Linguistics.
- [Le et. al. 2014] Le Q, Mikolov T. Distributed representations of sentences and documents. In Proceedings of the International Conference on Machine Learning (ICML 2014), 2014.
- [Liu 2015] Liu B. Sentiment analysis: mining opinions, sentiments, and emotions. The Cambridge University Press, 2015.
- [Liu 2012] Liu B. Sentiment analysis and opinion mining (introduction and survey), Morgan & Claypool, May 2012.
- [Ruder et al. 2016] Sebastian Ruder, Parsa Ghaffari, and John G Breslin. A hierarchical model of reviews for aspect-based sentiment analysis. Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, November, 2016.
- [Schouten et. al. 2016] Kim Schouten and Flavius Frasincar. 2016. Survey on Aspect-Level Sentiment Analysis. IEEE Trans. Knowledge and Data Engineering, 28(3):813-830
- [Tang et. al. 2015] Duyu Tang, Bing Qin, and Ting Liu. 2015. Document modeling with gated recurrent neural network for sentiment classification. In Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, pages 1422–1432
- [Tang et. al. 2016] Tang D, Qin B, and Liu T. Aspect-level sentiment classification with deep memory network. arXiv preprint arXiv:1605.08900, 2016.
- [Xu et. al. 2016] Xu J, Chen D, Qiu X, and Huang X. Cached long short-term memory neural networks for document-level sentiment classification. In Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP 2016), 2016.
- [Yang et. al. 2016] Zichao Yang, Diyi Yang, Chris Dyer, Xiaodong He, Alex Smola, and Eduard Hovy. 2016. Hierarchical attention networks for document classification. In Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics
- [Zhang et. al. 2018] Lei Zhang, Shuai Wang, and Bing Liu. 2018. Deep learning for sentiment analysis: A survey. In Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, page e1253. Wiley Online Library

References (Continued)

- [Poria et al.2016] Soujanya Poria, Erik Cambria, and Alexander Gelbukh. 2016a. Aspect extraction for opinion mining with a deep convolutional neural network. *Knowledge-Based Systems*, 108:42–49.
- [Wang et. al. 2016] Wang W, Pan SJ, Dahlmeier D, and Xiao X. Recursive neural conditional random fields for aspect-based sentiment analysis. In Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP 2016), 2016.



Deep NLP in Search and Recommender Systems - Document Retrieval & Ranking



Jun Shi, Weiwei Guo

Deep NLP in Search and Recommender Systems

- Language Understanding
 - Entity Tagging: word level prediction
 - Entity Disambiguation: knowledge base entity prediction
 - Intent Classification: sentence level prediction
 - Sentiment Analysis and Opinion Mining
- Document Retrieval and Ranking
 - Efficient Candidate Retrieval
 - Deep Ranking Models
- Language Generation for Assistance
 - Auto Completion
 - Query Suggestion
 - Spell Correction
 - Conversational Recommendation

Document Retrieval and Ranking

- Efficient Candidate Retrieval
- Deep Neural Ranking

Efficient Candidate Retrieval

- Syntactic retrieval
 - based on string matching
 - use inverted index
 - can include different fields (name, title, etc).
- Semantic retrieval
 - based on vector space representation.
 - approximate nearest neighbor search

Syntactic Retrieval

Query: mike software engineer

Inverted index

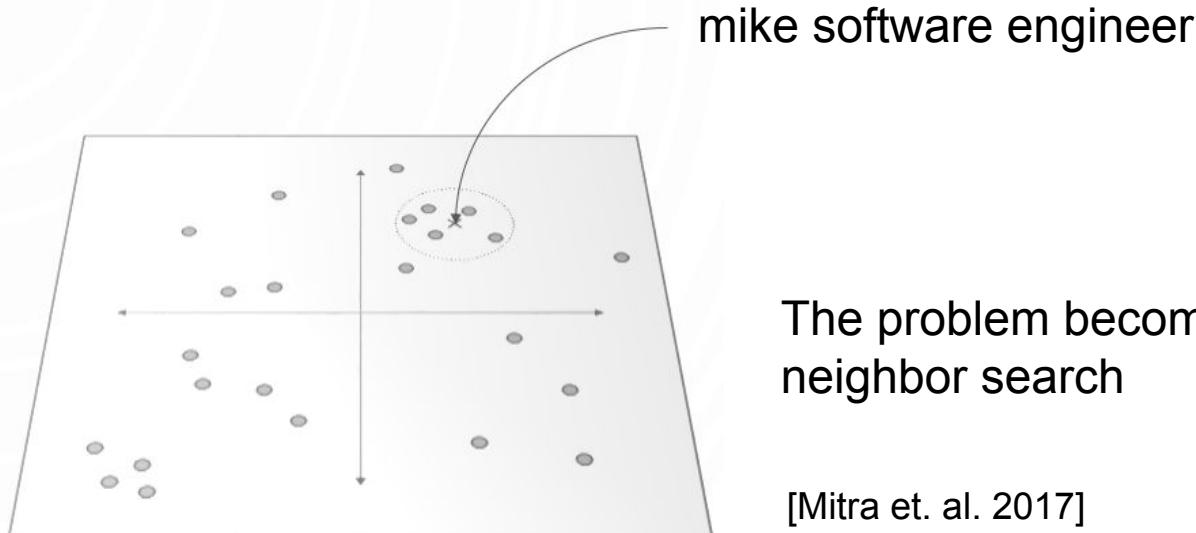
mike	Doc1, Doc2, Doc4 , ...
software	Doc2 , Doc3, Doc4 , Doc7, ...
engineer	Doc1, Doc2, Doc4 , Doc9, ...

Results: **Doc2, Doc4**

It won't retrieve a document contains “mike is a software developer,...”

Semantic Retrieval - Concept

mike software engineer $\xrightarrow{\text{embedding}}$ [0.3, 1.27, -3.4, ...]



[Mitra et. al. 2017]

Semantic Retrieval - Vector Generation

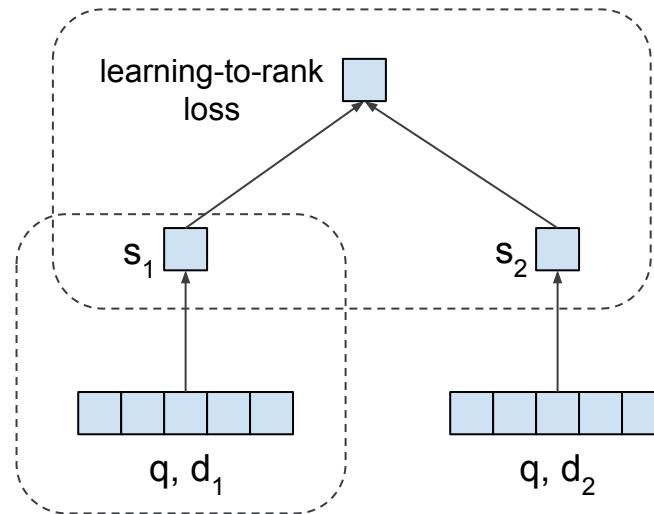
- Bag of words
 - count based, TF-IDF vectors.
- Embedding vectors
 - word embedding: word2vec, Glove, BERT, etc.
 - sentence/document embedding
 - universal sentence encoder [Cer et. al. 2018]
 - Gaussian document representation [Giannis et. al. 2017]
 - power mean concatenation [Rücklé et. al. 2018]

Deep NLP in Search and Recommender Systems

- Language Understanding
 - Entity Tagging: word level prediction
 - Entity Disambiguation: knowledge base entity prediction
 - Intent Classification: sentence level prediction
 - Sentiment Analysis and Opinion Mining
- Document Retrieval and Ranking
 - Efficient Candidate Retrieval
 - **Deep Ranking Models**
- Language Generation for Assistance
 - Auto Completion
 - Query Suggestion
 - Spell Correction
 - Conversational Recommendation

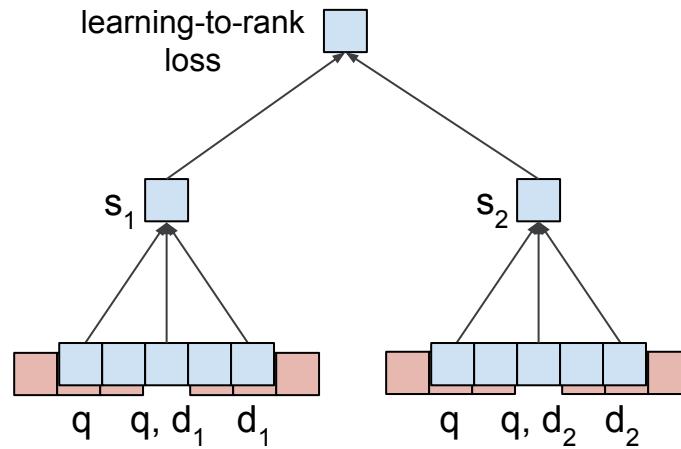
Deep Neural Ranking - Agenda

- Traditional methods
 - Ranking features
 - Learning to rank



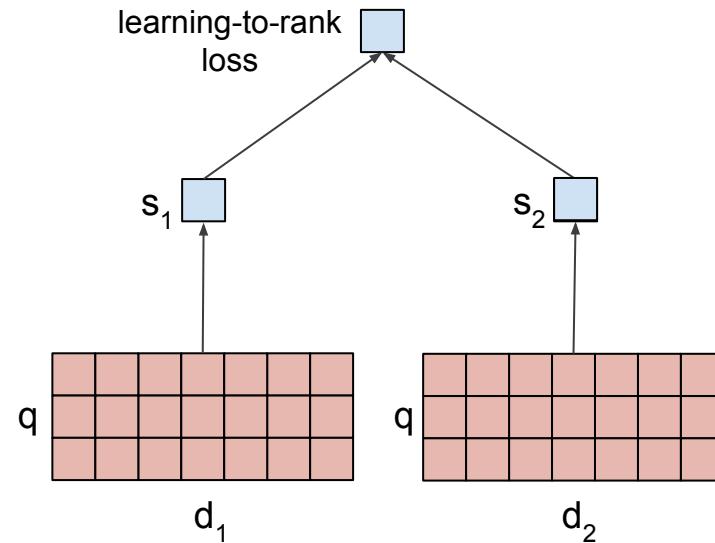
Deep Neural Ranking - Agenda

- Traditional methods
 - Ranking features
 - Learning to rank
- Deep neural ranking
 - Siamese Networks



Deep Neural Ranking - Agenda

- Traditional methods
 - Ranking features
 - Learning to rank
- Deep neural ranking
 - Siamese Networks
 - Interaction-based Networks



Deep Neural Ranking - Agenda

- Traditional methods
 - Ranking features
 - Learning to rank
- Deep neural ranking
 - Siamese Networks
 - Interaction-based Networks
 - Wide & Deep Model

Traditional Ranking Features

- Hand-crafted features
 - Query/document matching features
 - Cosine similarity between query and doc title
 - Entity matching
 - Clickthrough rate from query to this doc based on search log
 -
 - Document alone
 - popularity
 - number of incoming links
 -

Learning to Rank

(Burges, 2010)

- Pointwise ranking
 - Logistic regression

$$\frac{1}{1 + e^{-s}} \text{ for } y = 1$$

- Pairwise ranking

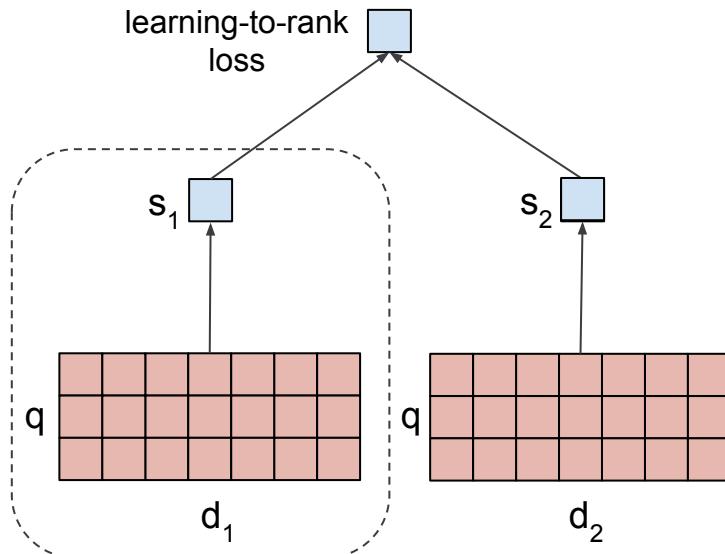
$$\frac{1}{1 + e^{-(s_1 - s_2)}} = \frac{e^{s_1}}{e^{s_1} + e^{s_2}}$$

- Listwise ranking
 - Cross entropy

$$\sum_i y_i \cdot \frac{e^{s_i}}{e^{s_1} + e^{s_2} + \dots + e^{s_n}}$$

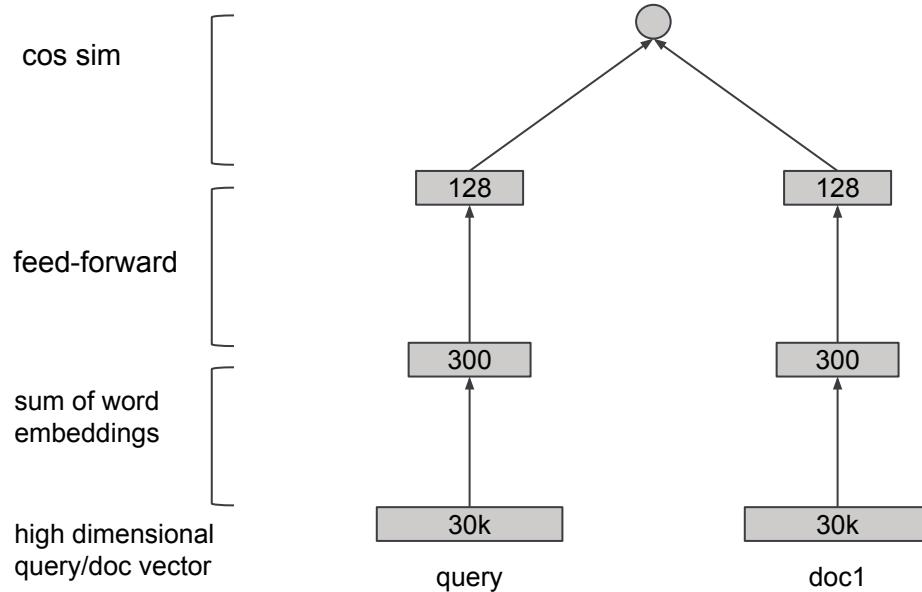
Deep Neural Ranking

- Focus on computing query/document score
- Two categories:
 - Siamese Networks
 - Interaction based Networks



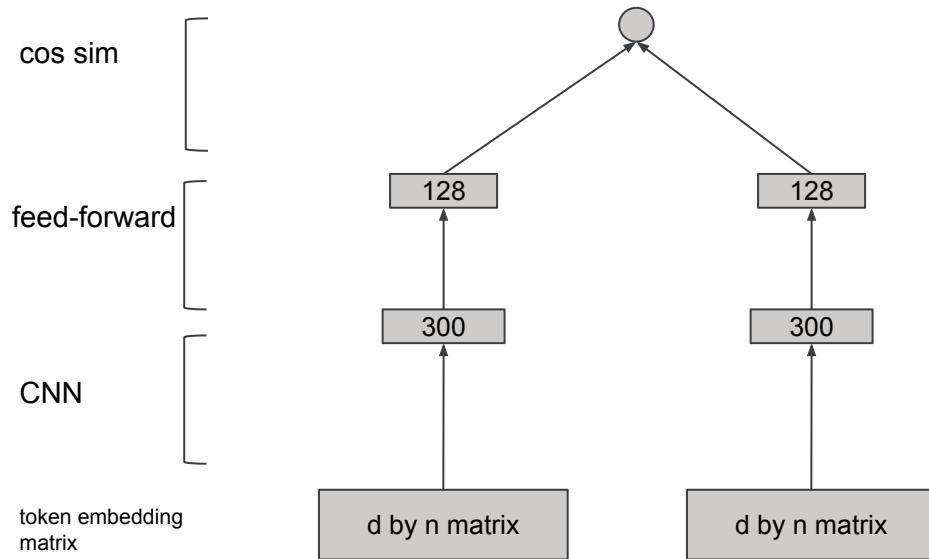
Deep Structured Semantic Model

(Huang et al., 2013)



Modeling Word Sequence by CNN

(Shen et al., 2014)

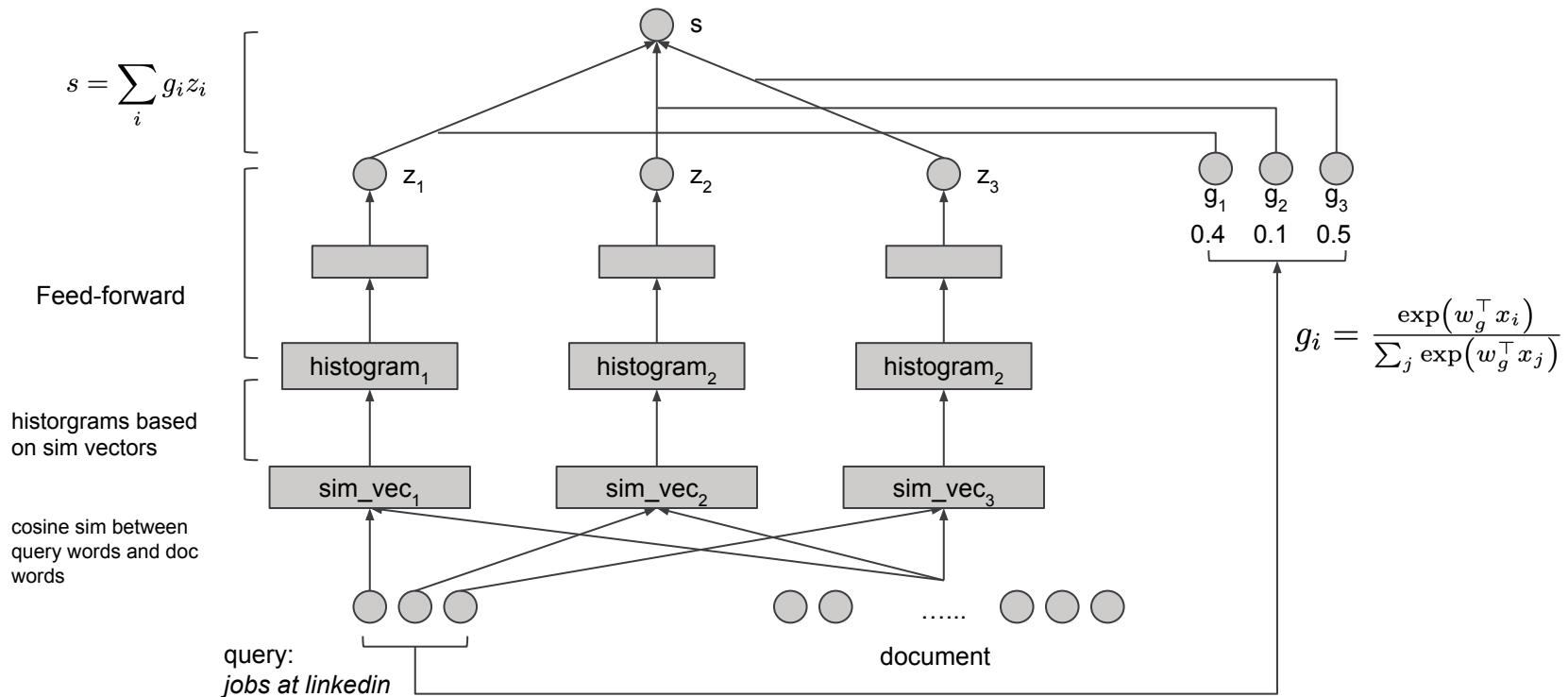


Siamese Networks

- Pros:
 - Generalization; semantic matching
 - Efficient; doc embs can be precomputed
- Cons:
 - Lexical features lost: people/company names, rare words

Interaction-based Networks

(Guo et al., 2016)



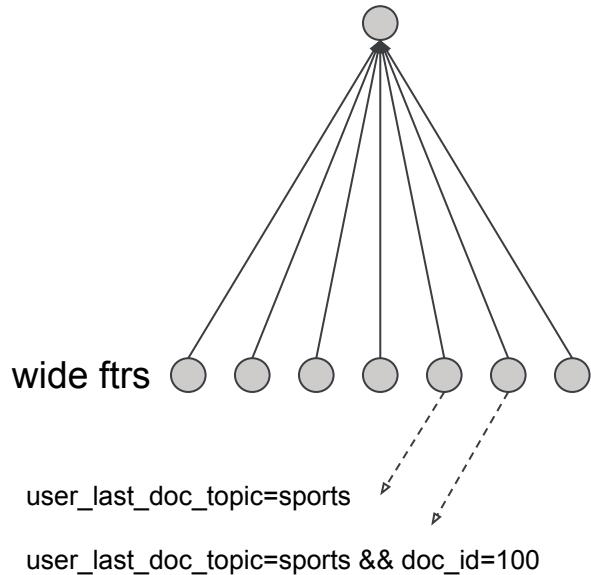
Search Ranking Neural Networks Summary

- Only focus on query/doc scoring
- End-to-end models
- Two popular architectures

	Siamese Network	Interaction Network
Match	topical matches	lexical matches
Latency	small	large

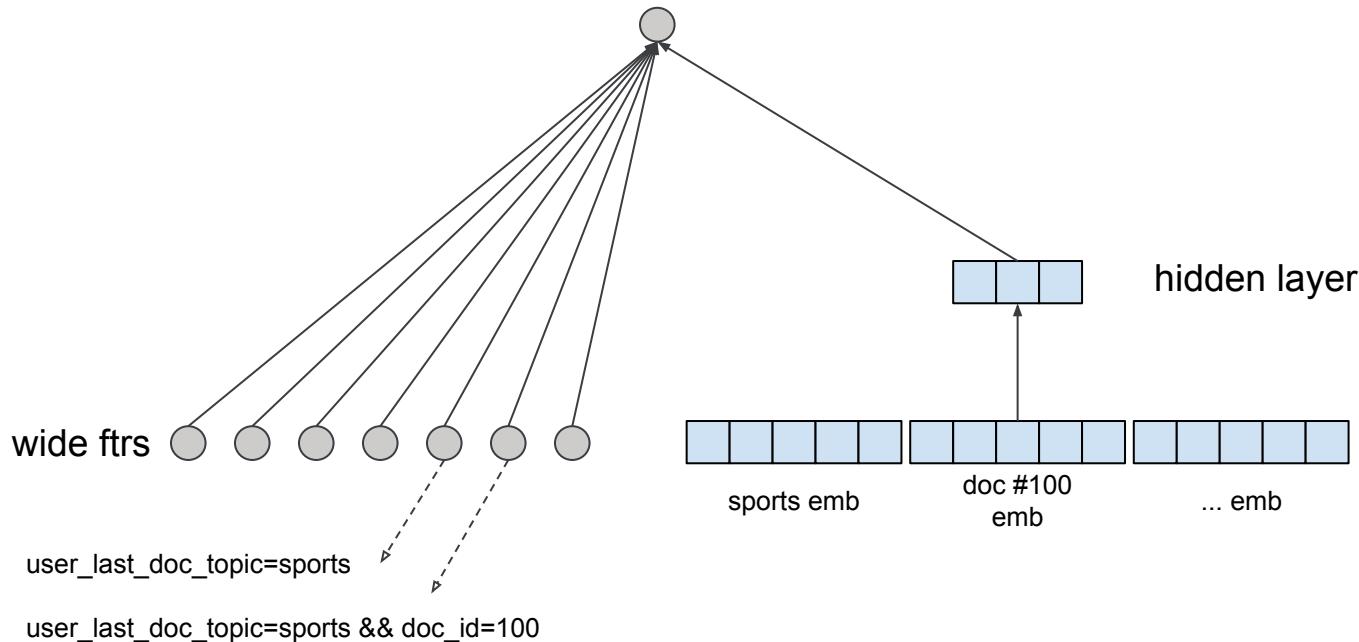
Recommender System: Wide & Deep Model

(Cheng et al., 2016)



Recommender System: Wide & Deep Model

(Cheng et al., 2016)



References - Deep Neural Ranking

- Huang, Po-Sen, Xiaodong He, Jianfeng Gao, Li Deng, Alex Acero, and Larry Heck. "Learning deep structured semantic models for web search using clickthrough data." In Proceedings of the 22nd ACM international conference on Information & Knowledge Management, pp. 2333-2338. ACM, 2013.
- Shen, Yelong, Xiaodong He, Jianfeng Gao, Li Deng, and Grégoire Mesnil. "A latent semantic model with convolutional-pooling structure for information retrieval." In Proceedings of the 23rd ACM international conference on conference on information and knowledge management, pp. 101-110. ACM, 2014.
- Rodrigo Nogueira, Kyunghyun Cho, Passage re-ranking with BERT, 2019
- Burges. "From ranknet to lambdarank to lambdamart: An overview." *Learning*. 2010.
- Guo, Jiafeng, Yixing Fan, Qingyao Ai, and W. Bruce Croft. "A deep relevance matching model for ad-hoc retrieval." In *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management*, pp. 55-64. ACM, 2016.
- Mitra, Bhaskar, Fernando Diaz, and Nick Craswell. "Learning to match using local and distributed representations of text for web search." In Proceedings of the 26th International Conference on World Wide Web, pp. 1291-1299. International World Wide Web Conferences Steering Committee, 2017.
- Cheng, Heng-Tze, Levent Koc, Jeremiah Harmsen, Tal Shaked, Tushar Chandra, Hrishi Aradhye, Glen Anderson et al. "Wide & deep learning for recommender systems." In Proceedings of the 1st workshop on deep learning for recommender systems, pp. 7-10. ACM, 2016.
- Xiong, Chenyan, Zhuyun Dai, Jamie Callan, Zhiyuan Liu, and Russell Power. "End-to-end neural ad-hoc ranking with kernel pooling." In Proceedings of the 40th International ACM SIGIR conference on research and development in information retrieval, pp. 55-64. ACM, 2017.
- Dai, Zhuyun, Chenyan Xiong, Jamie Callan, and Zhiyuan Liu. "Convolutional neural networks for soft-matching n-grams in ad-hoc search." In Proceedings of the eleventh ACM international conference on web search and data mining, pp. 126-134. ACM, 2018.



Deep NLP in Search and Recommender Systems - Language Generation for Assistance

Weiwei Guo

Deep NLP in Search and Recommender Systems

- Language Understanding
 - Entity Tagging: word level prediction
 - Entity Disambiguation: knowledge base entity prediction
 - Intent Classification: sentence level prediction
 - Sentiment Analysis and Opinion Mining
- Document Retrieval and Ranking
 - Efficient Candidate Retrieval
 - Deep Ranking Models
- **Language Generation for Assistance**
 - **Auto Completion**
 - **Query Reformulation**
 - **Spell Correction**
 - **Conversational Recommendation**

Language Generation for Assistance

- Auto Completion
- Query Reformulation
- Spell Correction
- Conversational Recommendation

Common

- Goal: improve user experience
- How: interact with users
- NLP: Language generation

Difference

- Character/word modeling
- Generation models:
 - language modeling, seq2seq

Deep NLP in Search and Recommender Systems

- Language Understanding
 - Entity Tagging: word level prediction
 - Entity Disambiguation: knowledge base entity prediction
 - Intent Classification: sentence level prediction
 - Sentiment Analysis and Opinion Mining
- Document Retrieval and Ranking
 - Efficient Candidate Retrieval
 - Deep Ranking Models
- Language Generation for Assistance
 - **Auto Completion**
 - Query Reformulation
 - Spell Correction
 - Conversational Recommendation

Query Auto-Completion

- Problem statement: save user keystrokes by predicting the entire query

softw|

software engineer salary

software engineer

software

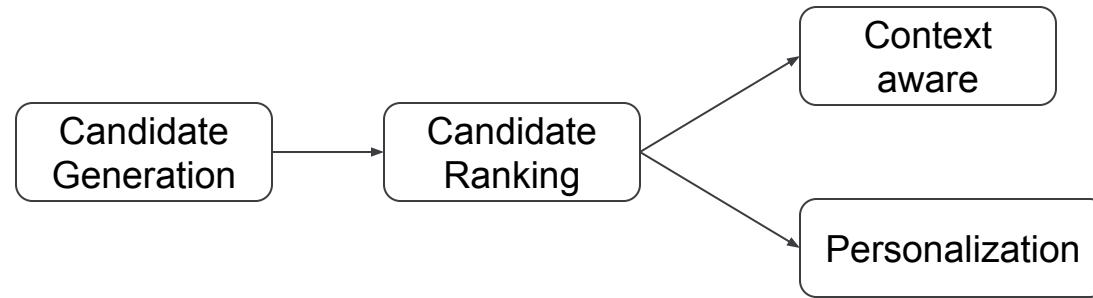
software engineer jobs

software developer

Challenges

- Limited Context
 - Hard to extract features due to limited words in a query
- Latency
 - The search component with most strict requirement on latency

Agenda



Traditional methods

Candidate Generation

- Collect completed queries and associated frequency from search log
- Efficiently retrieve most frequent queries starting with the prefix
 - Using a trie data structure

candidate generation		
<i>soft</i>	→	<i>software engineer</i> 100
		<i>softball</i> 50
		<i>softmax</i> 40
		<i>softbank</i> 35
	

Candidate Generation for Rare Prefix

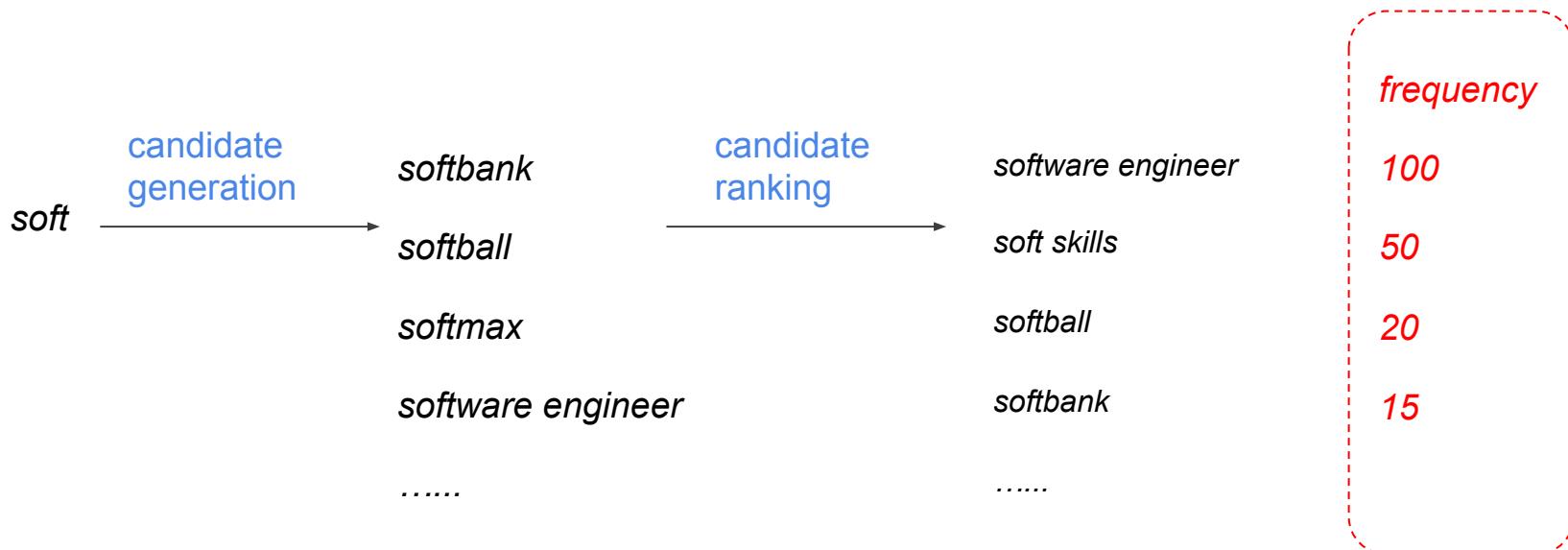
(Mitra & Craswell, 2015)

- No such prefix in search log
 - “*cheapest flights from seattle to*”
- Another trie on suffixes
 - “distance from seattle to dc” →
“*from seattle to dc*”, “*seattle to dc*”, “*to dc*”

“ <i>cheapest flights from seattle to</i> ”	————→	“ <i>to</i> ”	————→	“ <i>to dc</i> ”	100
				“ <i>to sfo</i> ”	50
				“ <i>to airport</i> ”	40
				“ <i>to seattle</i> ”	35
				

Candidate Ranking

- Challenge: very few features can be extracted



Context-Aware

(Bar-Yossef & Kraus, 2011)

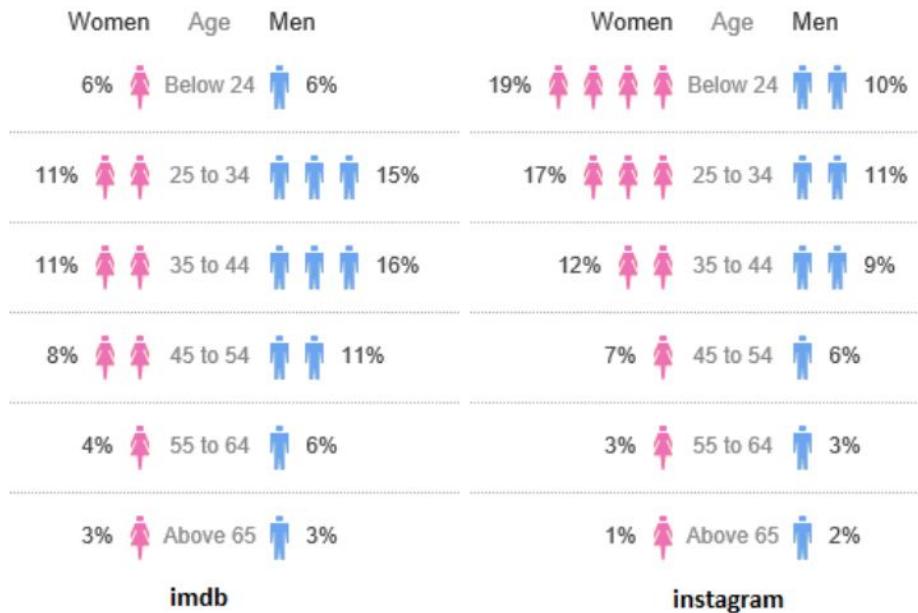
- Key idea: identify the candidates most similar to previous queries

(candidates)		sim score
<i>infant n</i>	→ <i>infant nike shoes</i>	→ <i>infant, nike, shoe, adidas, clothes...</i> 0.1
		<i>infant nutrition</i> → <i>infant, nutrition, baby, eat, food...</i> 0.8

(previous queries)		
	<i>baby eating disorder</i>	→ <i>baby, eating, disorder, nutrition, food...</i>

Personalization

(Shokouhi, 2013)



query prefix is “i”

Feature list

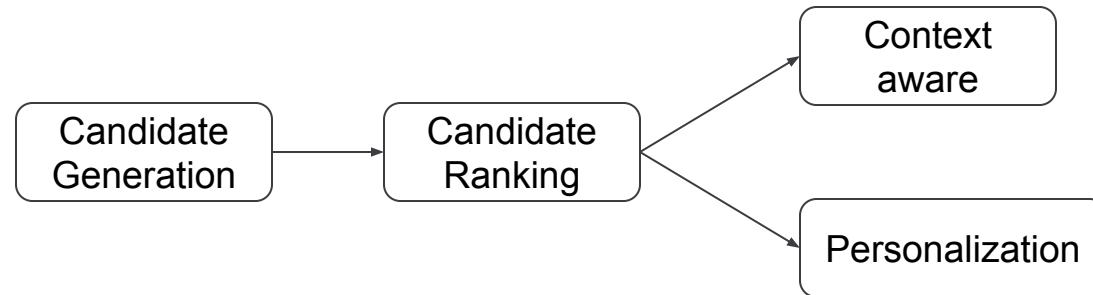
Same age

Same gender

Same region

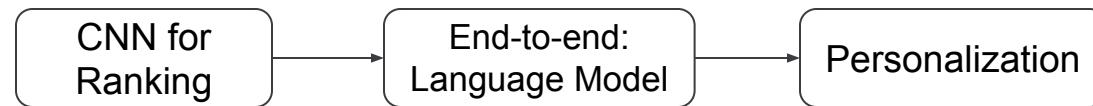
.....

Agenda



Traditional methods

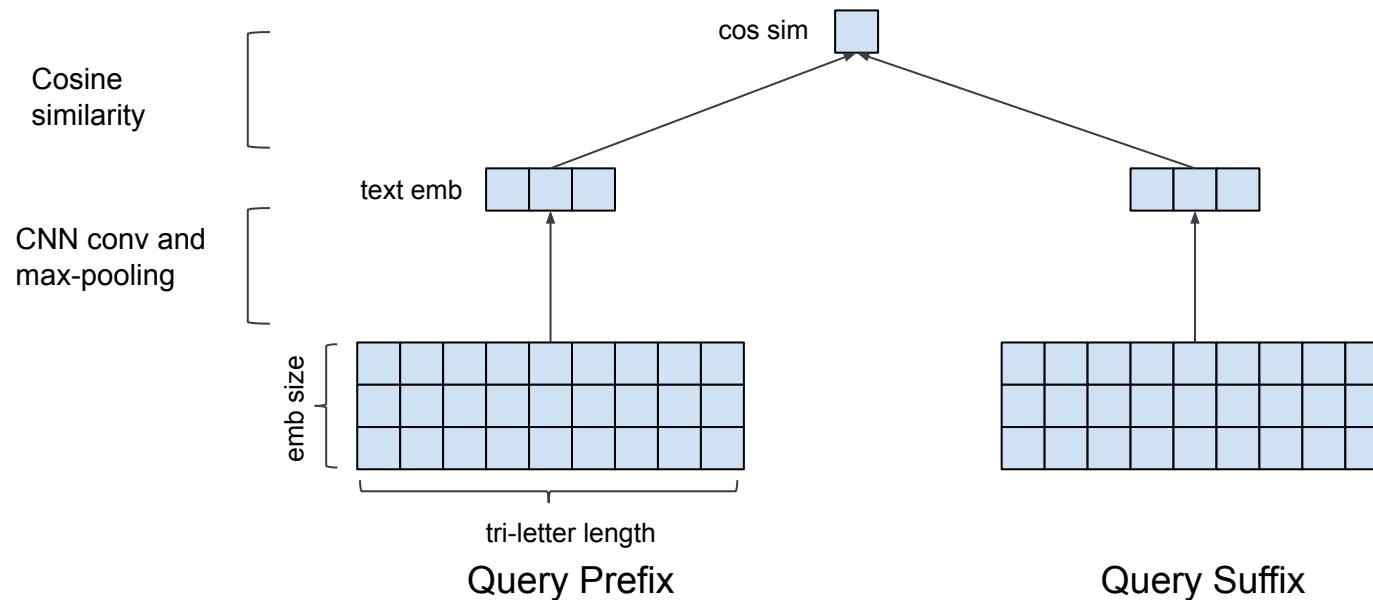
Deep NLP methods



Apply Deep Models in Ranking

(Mitra & Craswell, 2015)

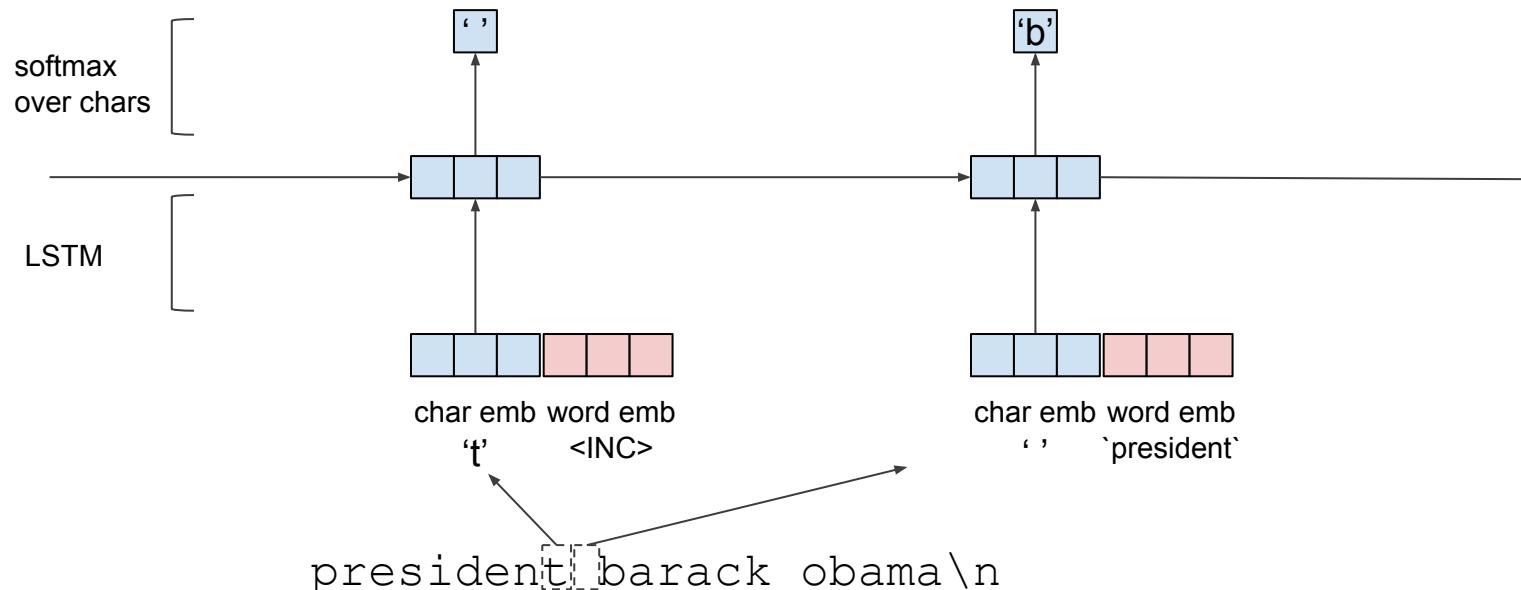
- Measuring the semantic coherence between prefix and suffix



Language Modeling for Auto-Completion

(Park & Chiba, 2017)

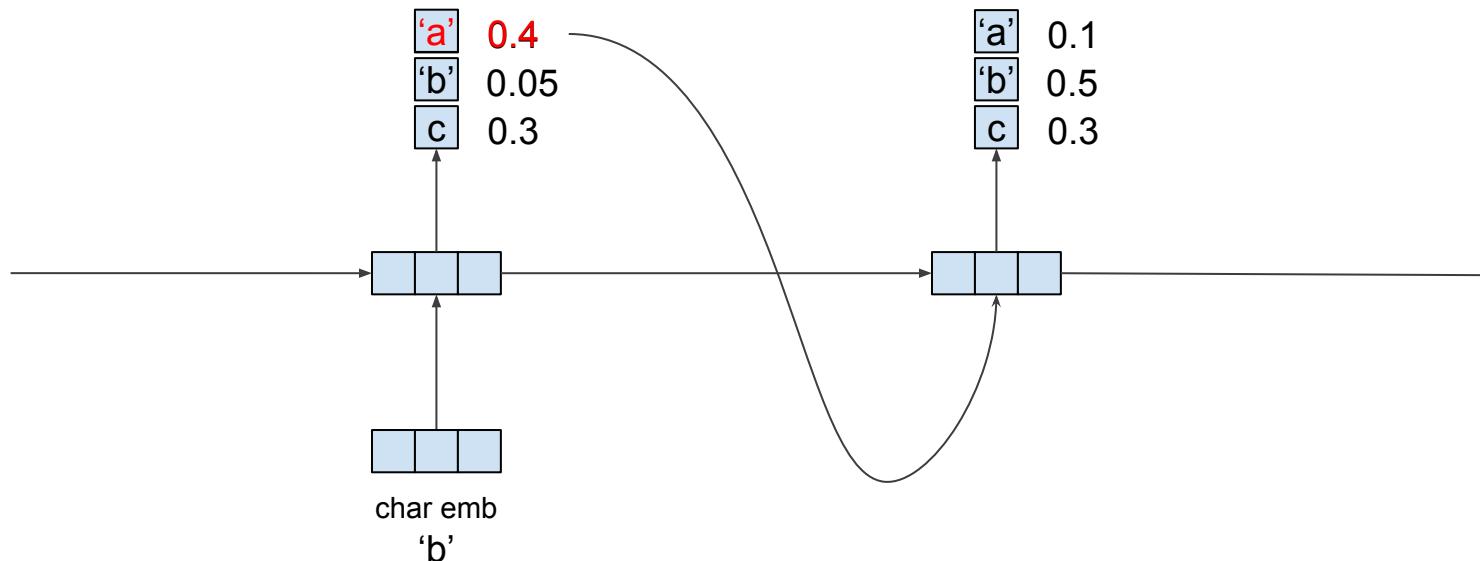
- Training: Character level language model + word embedding



Language Modeling for Auto-Completion

(Park & Chiba, 2017)

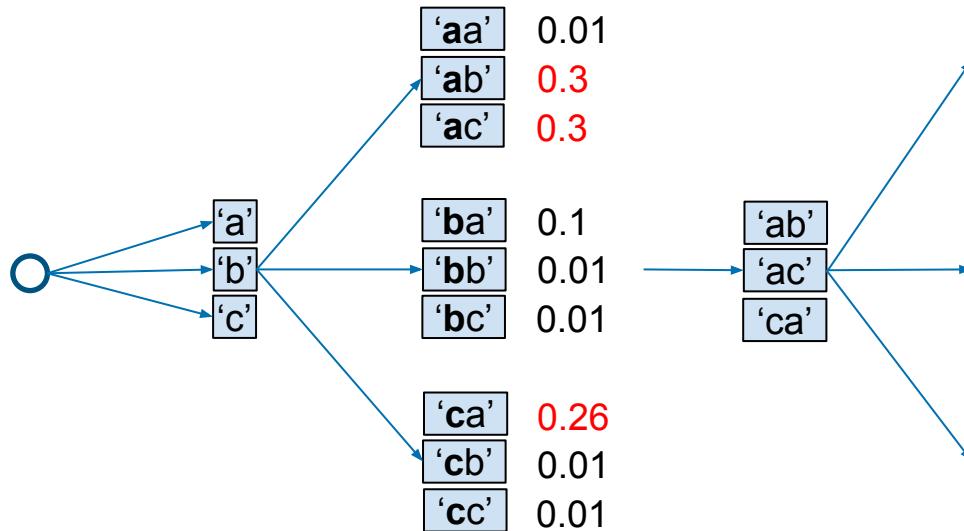
- Testing: Generating and ranking candidates at the same time



Language Modeling for Auto-Completion

(Park & Chiba, 2017)

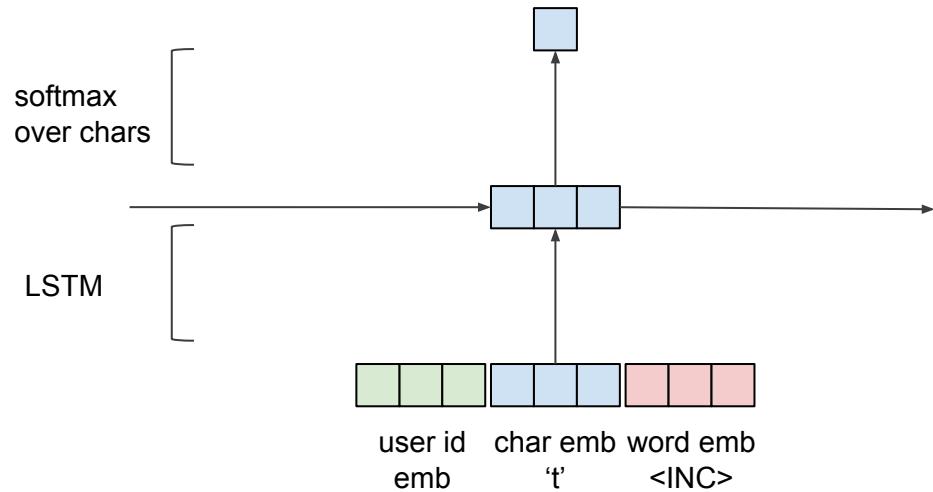
- Testing: Generating and ranking candidates at the same time
- Greedy vs beam search



Personalization

(Fiorin & Lu, 2018)

- Embeddings for User Ids



Query Auto-Completion: Summary

- Traditional methods: hard to extract features
- Deep language model framework:
 - Very flexible to incorporate personalized/contextualized information
 - An end-to-end solution
 - Train: all parameters are optimized together
 - Test: generation and ranking are performed at the same time
 - Cons
 - Time-consuming
 - May generate wrong words

Reference

- Mitra, Bhaskar, and Nick Craswell. "Query auto-completion for rare prefixes." In Proceedings of the 24th ACM international conference on information and knowledge management, pp. 1755-1758. ACM, 2015.
- Park, Dae Hoon, and Rikio Chiba. "A neural language model for query auto-completion." In Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 1189-1192. ACM, 2017.
- Bar-Yossef, Ziv, and Naama Kraus. "Context-sensitive query auto-completion." In Proceedings of the 20th international conference on World wide web, pp. 107-116. ACM, 2011.
- Shokouhi, Milad. "Learning to personalize query auto-completion." In Proceedings of the 36th international ACM SIGIR conference on Research and development in information retrieval, pp. 103-112. ACM, 2013.
- Jaech, Aaron, and Mari Ostendorf. "Personalized language model for query auto-completion." arXiv preprint arXiv:1804.09661 (2018).
- Fiorini, Nicolas, and Zhiyong Lu. "Personalized neural language models for real-world query auto completion." arXiv preprint arXiv:1804.06439 (2018).

Deep NLP in Search and Recommender Systems

- Language Understanding
 - Entity Tagging: word level prediction
 - Entity Disambiguation: knowledge base entity prediction
 - Intent Classification: sentence level prediction
 - Sentiment Analysis and Opinion Mining
- Document Retrieval and Ranking
 - Efficient Candidate Retrieval
 - Deep Ranking Models
- Language Generation for Assistance
 - Auto Completion
 - **Query Reformulation**
 - Spell Correction
 - Conversational Recommendation

Query Reformulation

- Problem Statement: automatically reformulate the previous query

[facebook developers | registration](#)

<https://go.fb.com/become-a-facebook-developer-lp05.html> ▾

Code to connect with 2B+ people with Facebook. Become a Facebook Developer to build and ship your application. Become a Facebook Developer ...

[Developer Tools Archives - Facebook Code](#)

<https://code.fb.com/category/developer-tools/> ▾

Our mission is to increase developer efficiency so that we can continue to ship awesome products quickly. To accomplish this, we have focused on building a ...

[Facebook Developer Conference. April 30 - May 1, 2019.San Jose, CA.](#)

<https://www.f8.com/> ▾

Facebook's annual developer conference spotlights our global community, the latest technology from our family of apps, and the future we are building together.

[Best Facebook App Development Company, Best FB Developers](#)

<https://www.cygnsmedia.com/social-media.../best-facebook-application.html> ▾

Effective branding is possible by attractive Facebook application development. Cygnis Media listed in Best App Development companies for Facebook. Hire best ...

Searches related to fb developer

[fb developers support](#)

[facebook developer support](#)

[facebook for developers products](#)

[facebook developer ecosystem](#)

[facebook developer tutorial](#)

[facebook developer alerts](#)

[facebook for developer tools](#)

[facebook developer forum](#)

Search results for
“*fb developer*”

Reformulated queries for
“*fb developer*”

Agenda

collaborative
filtering

Traditional methods

Traditional Approach: Collaborative Filtering

(Rida et al., 2012)

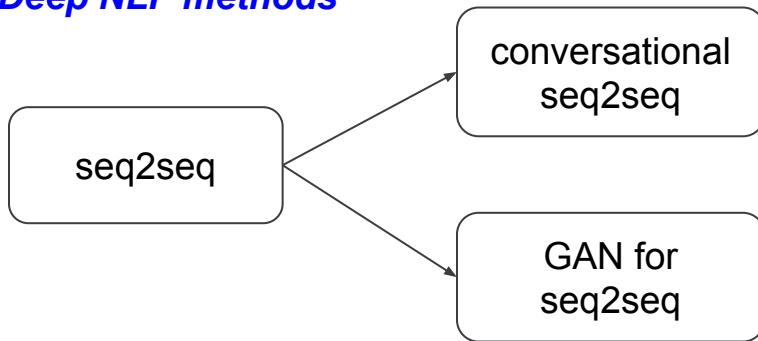
- Collect query pairs issued by the same user from search log
- Treat each query as an ID, and build a query-to-query matrix
 - Value in the matrix is TF-IDF
- Issues:
 - No understanding of words in the query
 - Data sparsity
 - Cannot generate new queries

Agenda

collaborative
filtering

Traditional methods

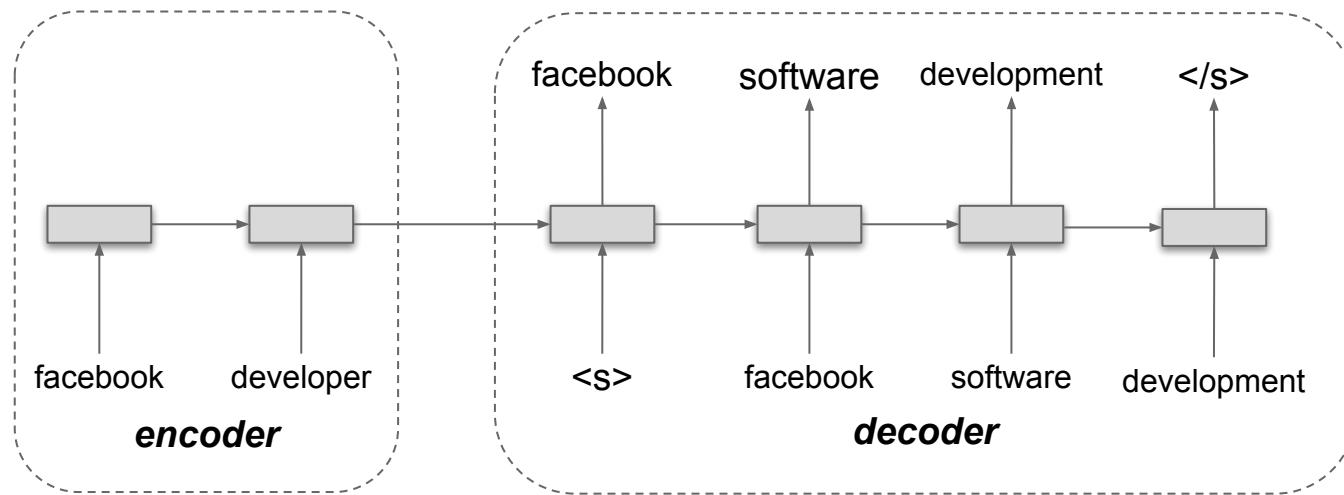
Deep NLP methods



Query Reformulation as a Translation Task

(Sordoni et al. 2015, He et al., 2016)

- Sequence-to-sequence modeling



- Directly modeling the words in a query

Conversational Query Reformulation

(Ren et al, 2018)

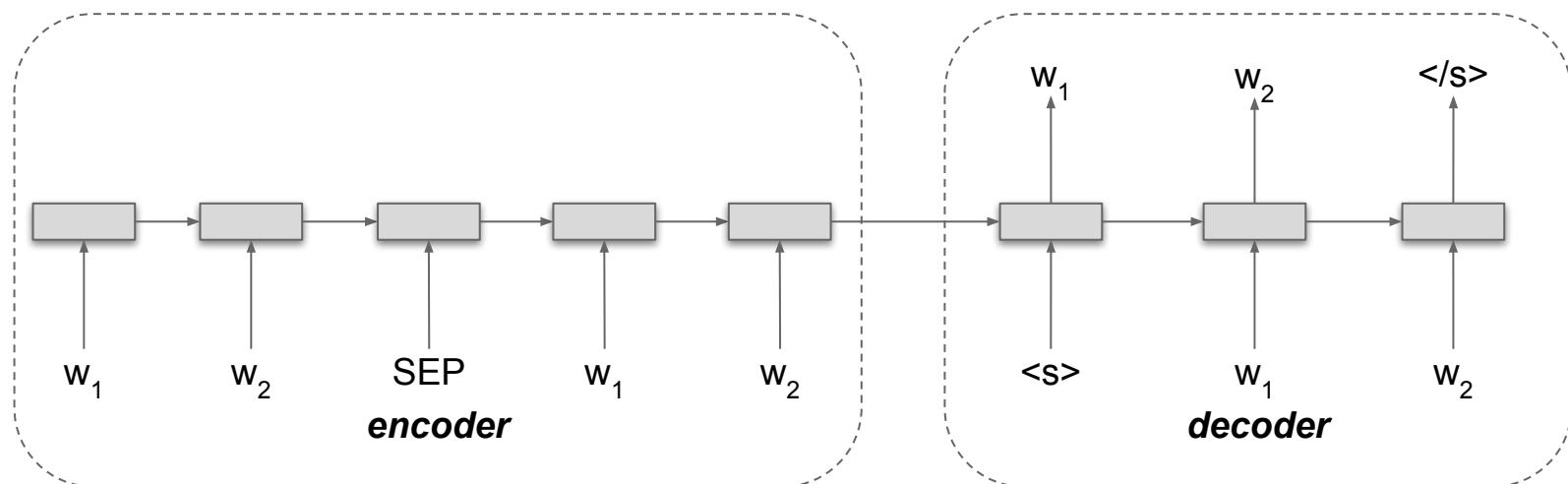
- Conversational queries
- Goal: summarize q1 & q2 into q3

first query (q1)	second query (q2)	summarized query (q3)
when was California founded?	who is its governor?	who is California's governor?
California	population in 1990	population of California in 1990
how tall is kobe bryant?	what about Lebron James?	how tall is Lebron James?
when was the last summer Olympics?	and the winter one?	when was the last winter Olympics?

Conversational Query Reformulation

(Ren et al, 2018)

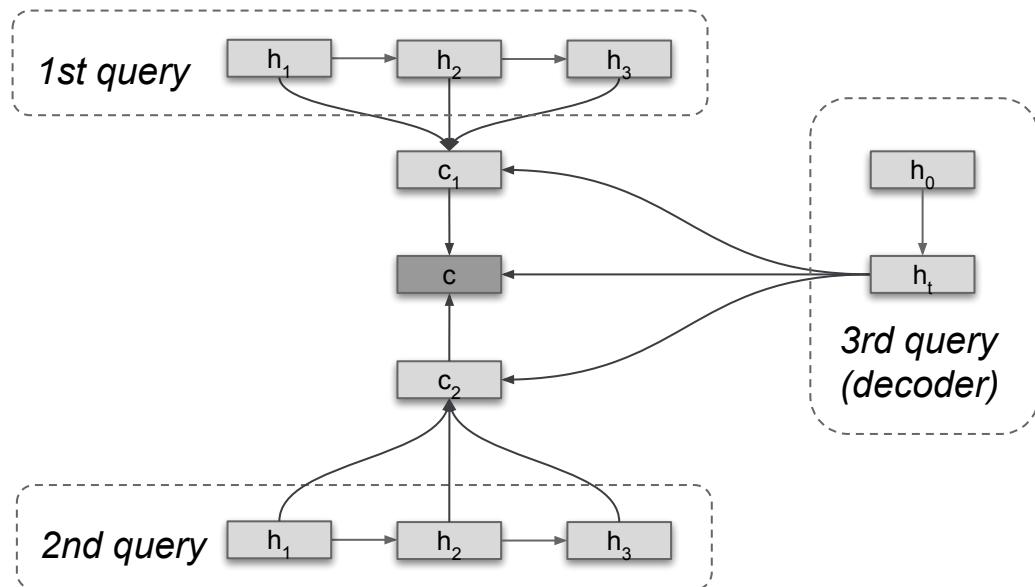
- 1st: concatenated seq2seq



Conversational Query Reformulation

(Ren et al, 2018)

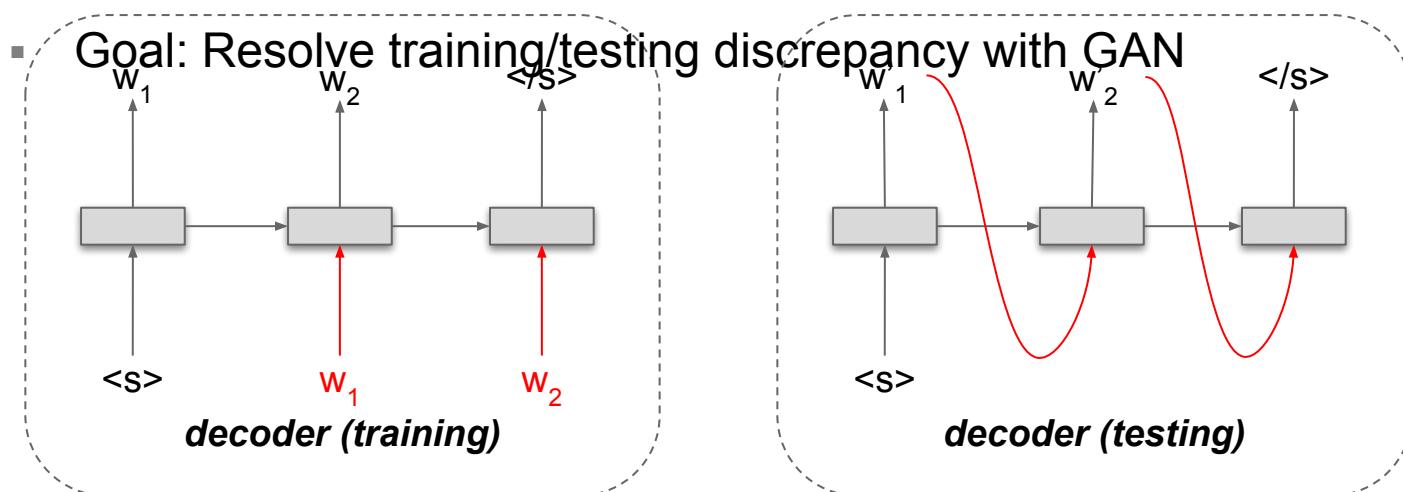
- 2nd: Attention over attention



GAN for seq2seq

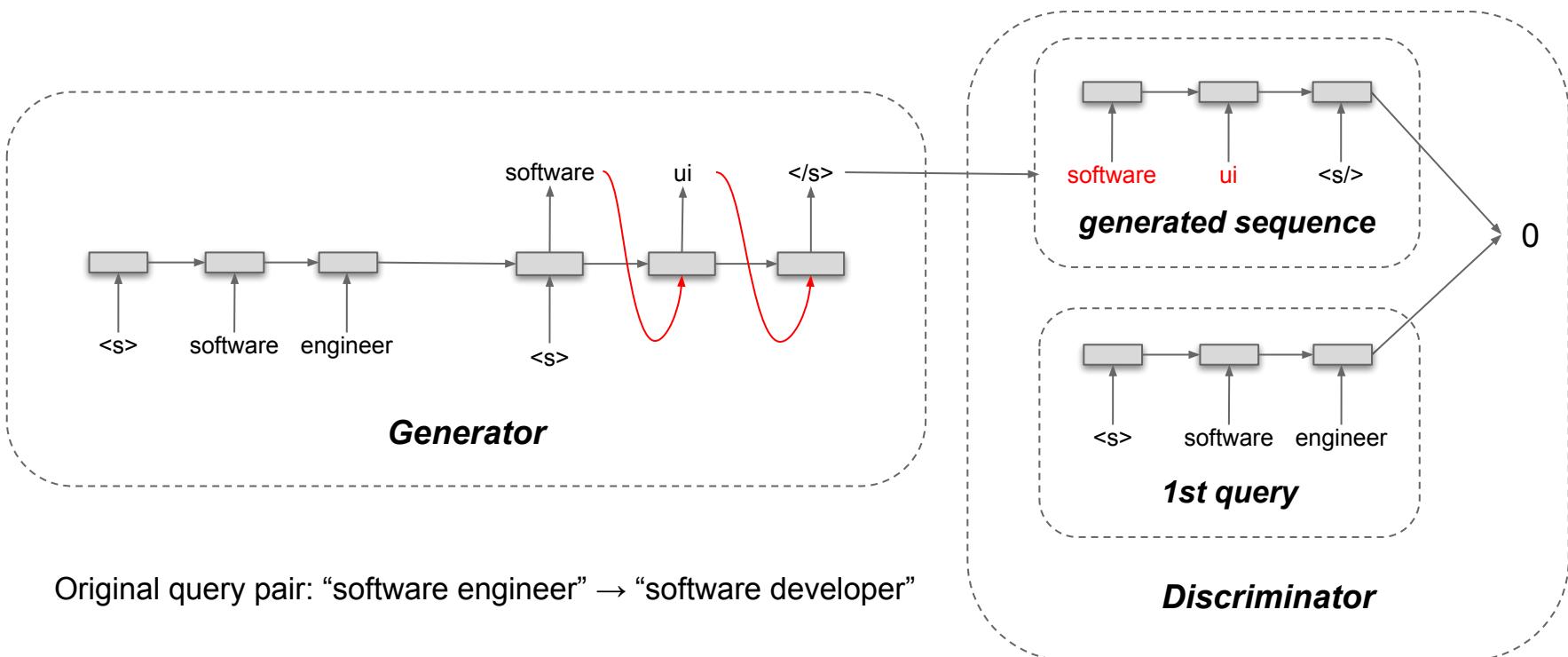
(Lee et al, 2018)

- Motivation: Training/testing discrepancy in decoder of seq2seq
 - Training: inputs are the **groundtruth words**
 - Testing: inputs are the **previous predicted words**



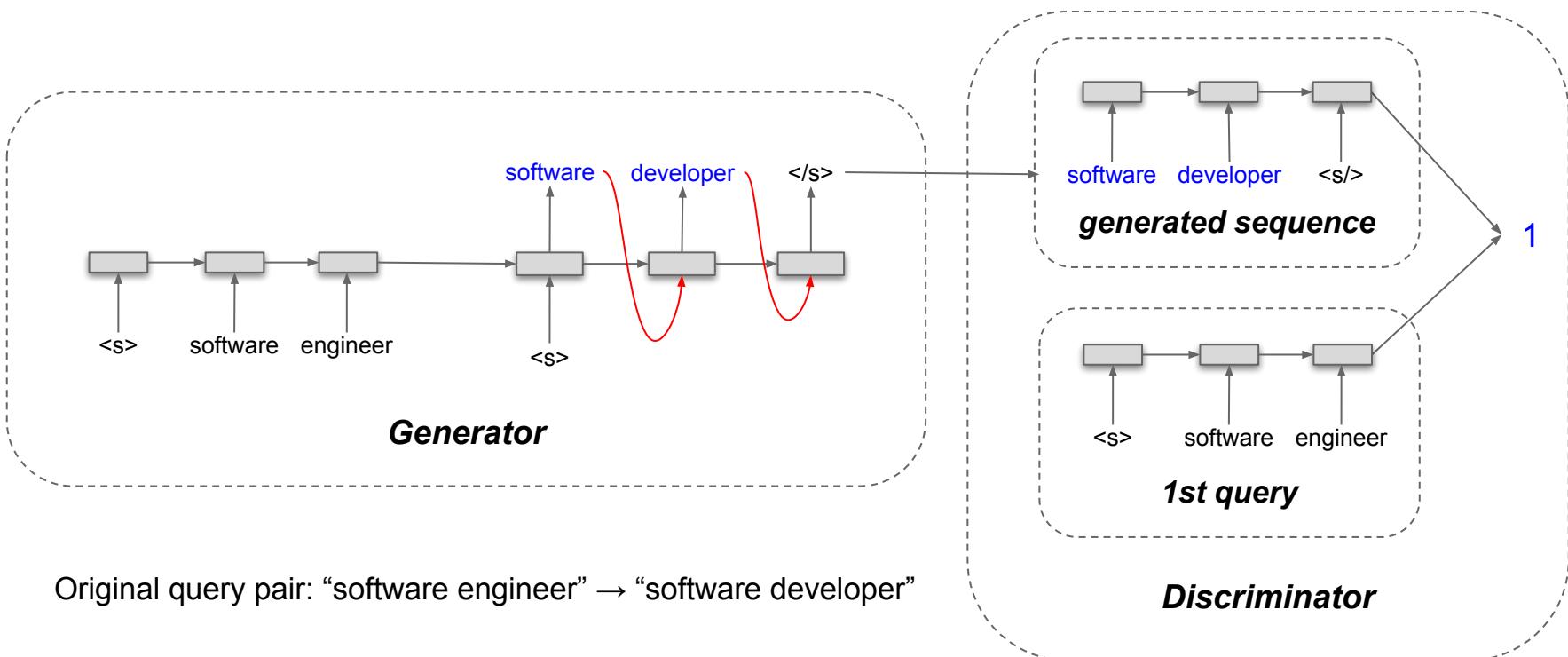
GAN for seq2seq

(Lee et al, 2018)



GAN for seq2seq

(Lee et al, 2018)



Query Reformulation: Summary

- seq2seq framework:
 - Directly modeling the words
 - Very flexible to incorporate session information
 - Achieves great performance (no character modeling)

Reference

- Reda, Azarias, Yubin Park, Mitul Tiwari, Christian Posse, and Sam Shah. "Metaphor: a system for related search recommendations." In Proceedings of the 21st ACM international conference on Information and knowledge management, pp. 664-673. ACM, 2012.
- He, Yunlong, Jiliang Tang, Hua Ouyang, Changsung Kang, Dawei Yin, and Yi Chang. "Learning to rewrite queries." In Proceedings of the 25th ACM International Conference on Information and Knowledge Management, pp. 1443-1452. ACM, 2016.
- Ren, Gary, Xiaochuan Ni, Manish Malik, and Qifa Ke. "Conversational query understanding using sequence to sequence modeling." In Proceedings of the 2018 World Wide Web Conference, pp. 1715-1724. International World Wide Web Conferences Steering Committee, 2018.
- Lee, Mu-Chu, Bin Gao, and Ruofei Zhang. "Rare query expansion through generative adversarial networks in search advertising." In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, pp. 500-508. ACM, 2018.
- Cao, Huanhuan, Dixin Jiang, Jian Pei, Qi He, Zhen Liao, Enhong Chen, and Hang Li. "Context-aware query suggestion by mining click-through and session data." In Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining, pp. 875-883. ACM, 2008.
- Sordoni, Alessandro, Yoshua Bengio, Hossein Vahabi, Christina Lioma, Jakob Grue Simonsen, and Jian-Yun Nie. "A hierarchical recurrent encoder-decoder for generative context-aware query suggestion." In Proceedings of the 24th ACM International Conference on Information and Knowledge Management, pp. 553-562. ACM, 2015.
- Mei, Qiaozhu, Dengyong Zhou, and Kenneth Church. "Query suggestion using hitting time." In Proceedings of the 17th ACM conference on Information and knowledge management, pp. 469-478. ACM, 2008.

Deep NLP in Search and Recommender Systems

- Language Understanding
 - Entity Tagging: word level prediction
 - Entity Disambiguation: knowledge base entity prediction
 - Intent Classification: sentence level prediction
 - Sentiment Analysis and Opinion Mining
- Document Retrieval and Ranking
 - Efficient Candidate Retrieval
 - Deep Ranking Models
- Language Generation for Assistance
 - Auto Completion
 - Query Reformulation
 - **Spell Correction**
 - Conversational Recommendation

Spell Correction

micrsoft



All



News



Maps



Shopping

About 1,780,000,000 results (0.58 seconds)

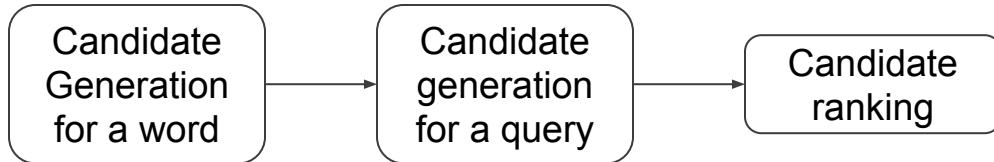
Showing results for ***microsoft***
Search instead for micrsoft

Spell Correction

- Why spell correction:
 - Reduce the no results
- Challenges:
 - Many rare words (people/company names) look like misspelled words
 - Modeling characters and words at the same time

query	similar query	has error?
<i>tumblr</i>	<i>tumble</i>	No
<i>tumblw</i>	<i>tumble</i>	Yes
<i>galaxy s10e</i>	<i>galaxy s10</i>	No
<i>galaxy s10d</i>	<i>galaxy s10</i>	Yes

Agenda



Traditional methods

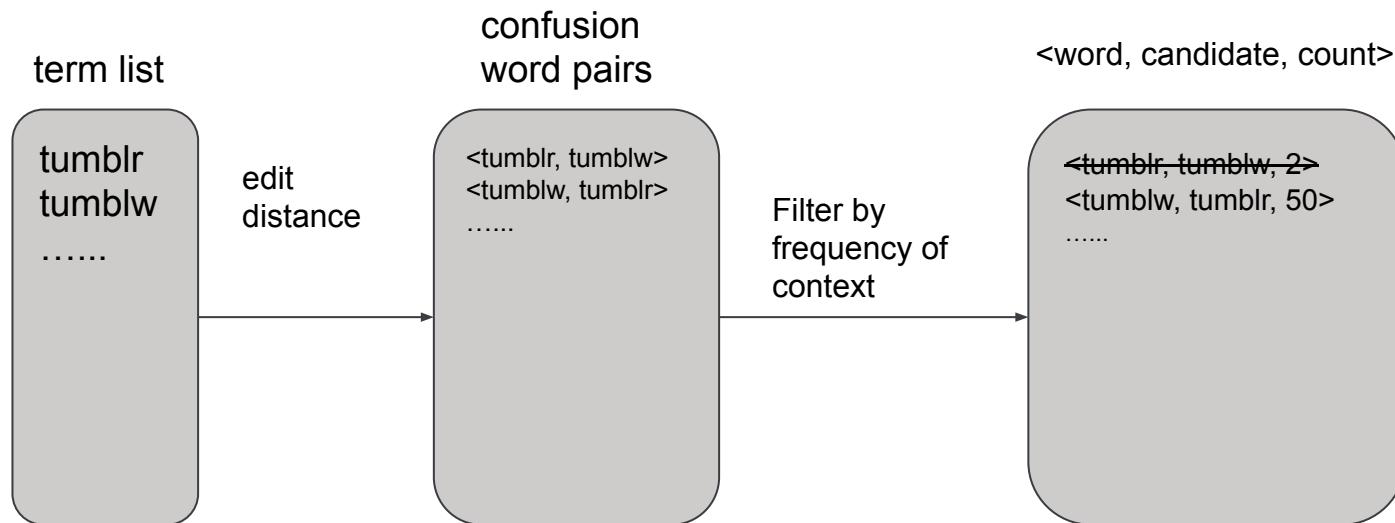
Candidate Generation for a Word

(Whitelaw et al 2009)

- Goal:
 - given “tumblw”, suggest “tumblr”, “tumble”...
 - given “tumblr”, suggest “tumble”, but not “tumblw”
 - Key challenge: what is a legit word? (“tumblr”  “tumblw”  - Coverage of a dictionary is not enough
- Solution: use statistics in web noisy data
 - Correct words appear more frequent than incorrect words

Candidate Generation for a Word

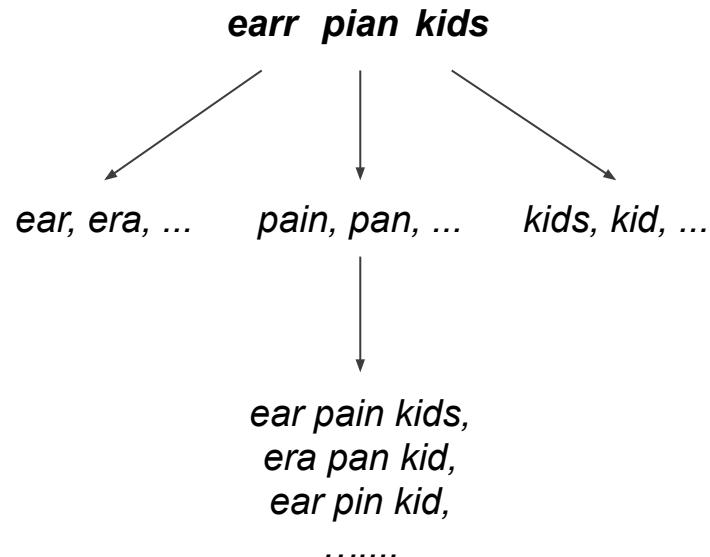
(Whitelaw et al 2009)



- For the context “social media X”:
 - $\text{freq}(\text{"social media tumblw"}) < \text{freq}(\text{"social media tumblr"})$

Candidate Generation for a Query

(Chen et al 2007)



- Problem: query candidate size grows exponentially with # of words
- Solution: prune with language model
 - **ear** pian kids: 0.8
 - ~~**era** pian kids: 0.1~~
 - **earr** **pain** kids: 0.9
 - **earr** **pan** kids: 0.7
 - ~~**earr** **pien** kids: 0.2~~
 - **earr** pian kid: 0.8

Candidate Ranking

(Li et al 2006, Chen et al 2007)

$$\operatorname{argmax}_c P(c|q) = \operatorname{argmax}_c P(q|c)P(c)$$

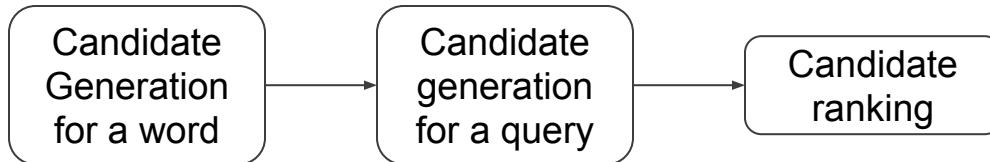
candidate query



Feature Types	Examples
similarity $P(q c)$	Edit distance
	Frequency of user reformulation

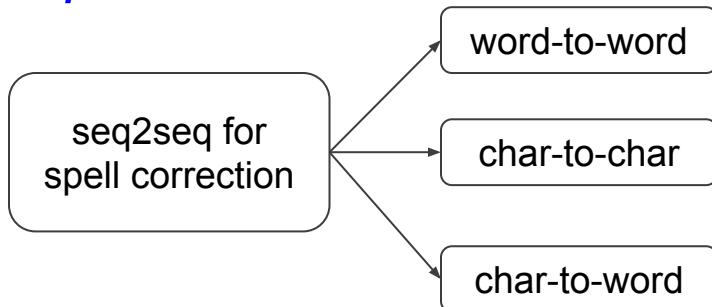
Likelihood $P(c)$	Language model score of the candidate
	Frequency of candidate terms appearing in the page titles

Agenda



Traditional methods

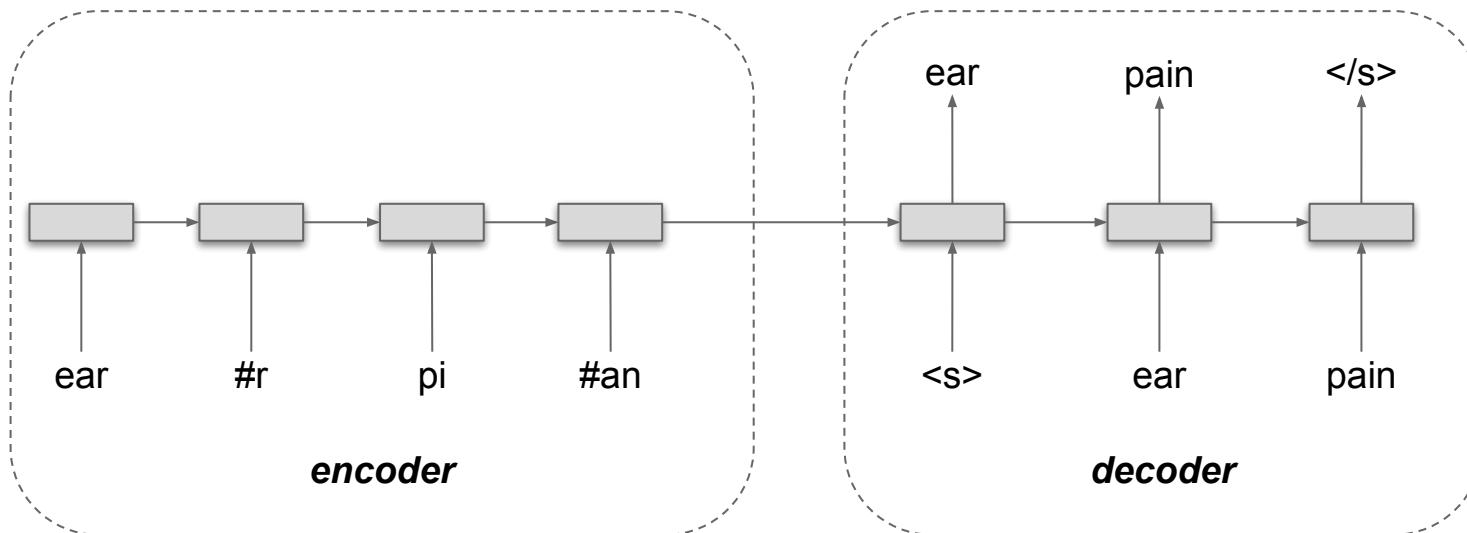
Deep NLP methods



Seq2seq for Spell Correction

(Ghosh and Kristensson, 2017, Zhou et al., 2017)

- From subwords to subwords



Seq2seq for Spell Correction

(Ghosh and Kristensson, 2017, Zhou et al., 2017)

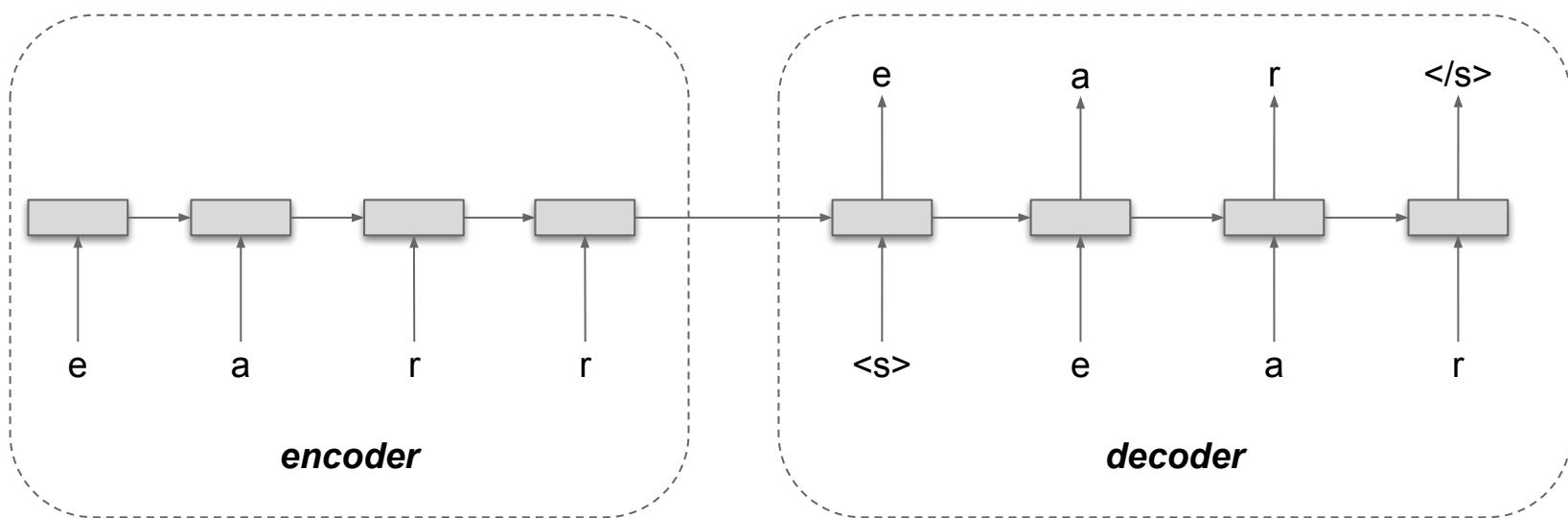
- From subwords to subwords
- Issue: subword is designed for texts without errors

	Normal texts	Spell errors
Example	$hunter \rightarrow hunt \#er$	$hunetr \rightarrow hu \#net \#r$
Subword semantics	<i>relevant</i>	<i>irrelevant</i>

Seq2seq for Spell Correction

(Ghosh and Kristensson, 2017, Zhou et al., 2017)

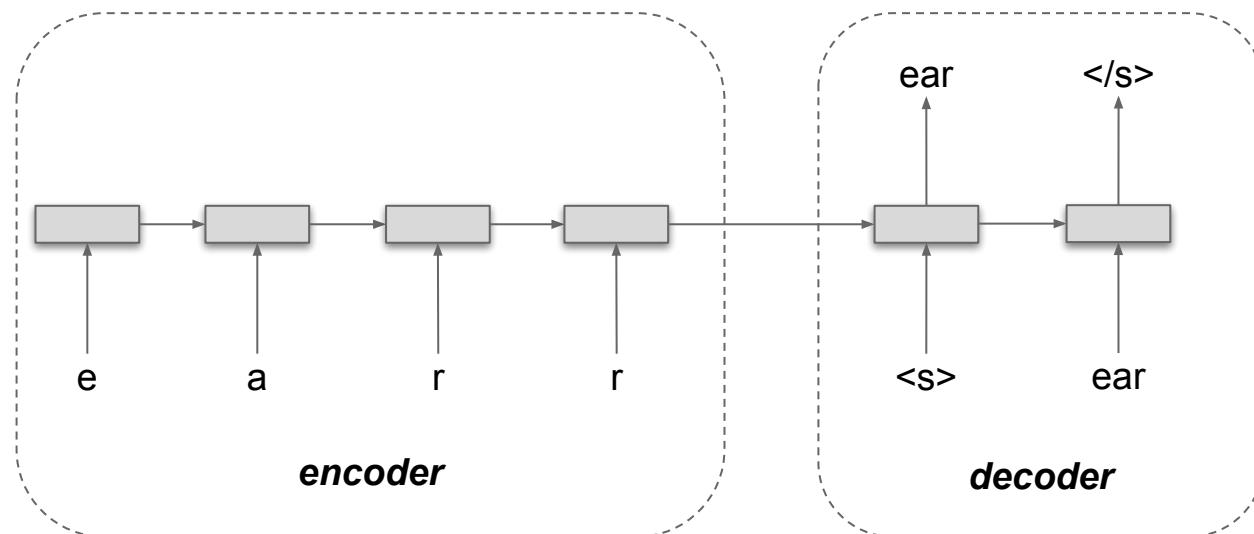
- From characters to characters
- Issue: on the decoder, no word information
 - Might produce words with wrong spelling



Seq2seq for Spell Correction

(Ghosh and Kristensson, 2017, Zhou et al., 2017)

- From characters to words
 - Most popular structure
 - Can leverage pretrained language model



Reference

- Hasan, Saša, Carmen Heger, and Saab Mansour. "Spelling correction of user search queries through statistical machine translation." In Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, pp. 451-460. 2015.
- Yingbo Zhou, Utkarsh Porwal, Roberto Konow. "Spelling Correction as a Foreign Language." arXiv. 2017.
- Ghosh, Shaona, and Per Ola Kristensson. "Neural networks for text correction and completion in keyboard decoding." arXiv preprint arXiv:1709.06429 (2017).
- Chen, Qing, Mu Li, and Ming Zhou. "Improving query spelling correction using web search results." In Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL), pp. 181-189. 2007.
- Li, Mu, Yang Zhang, Muhua Zhu, and Ming Zhou. "Exploring distributional similarity based models for query spelling correction." In Proceedings of the 21st International Conference on Computational Linguistics and the 44th annual meeting of the Association for Computational Linguistics, pp. 1025-1032. Association for Computational Linguistics, 2006.
- Whitelaw, Casey, Ben Hutchinson, Grace Y. Chung, and Gerard Ellis. "Using the web for language independent spellchecking and autocorrection." In Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing: Volume 2-Volume 2, pp. 890-899. Association for Computational Linguistics, 2009.

Deep NLP in Search and Recommender Systems

- Language Understanding
 - Entity Tagging: word level prediction
 - Entity Disambiguation: knowledge base entity prediction
 - Intent Classification: sentence level prediction
 - Sentiment Analysis and Opinion Mining
- Document Retrieval and Ranking
 - Efficient Candidate Retrieval
 - Deep Ranking Models
- Language Generation for Assistance
 - Auto Completion
 - Query Reformulation
 - Spell Correction
 - **Conversational Recommendation**

Conversational recommender systems

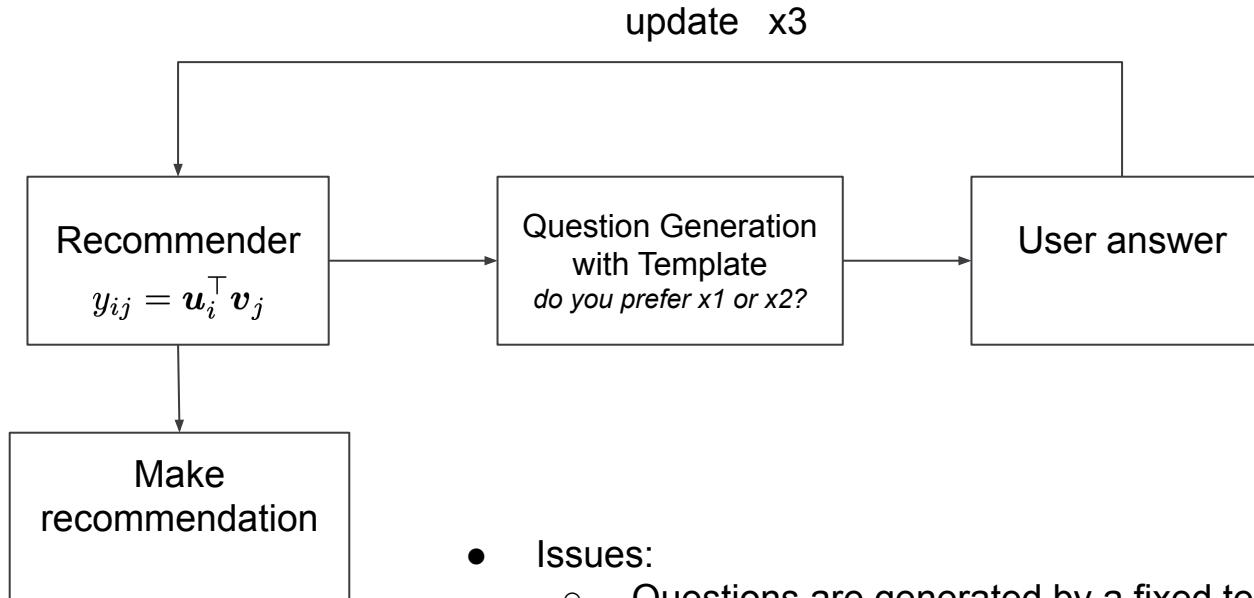


- Motivation:
 - Actively gather more information or resolve ambiguity
 - Better user experience

- Challenges:
 - What to ask to the user
 - Ask vs Rec

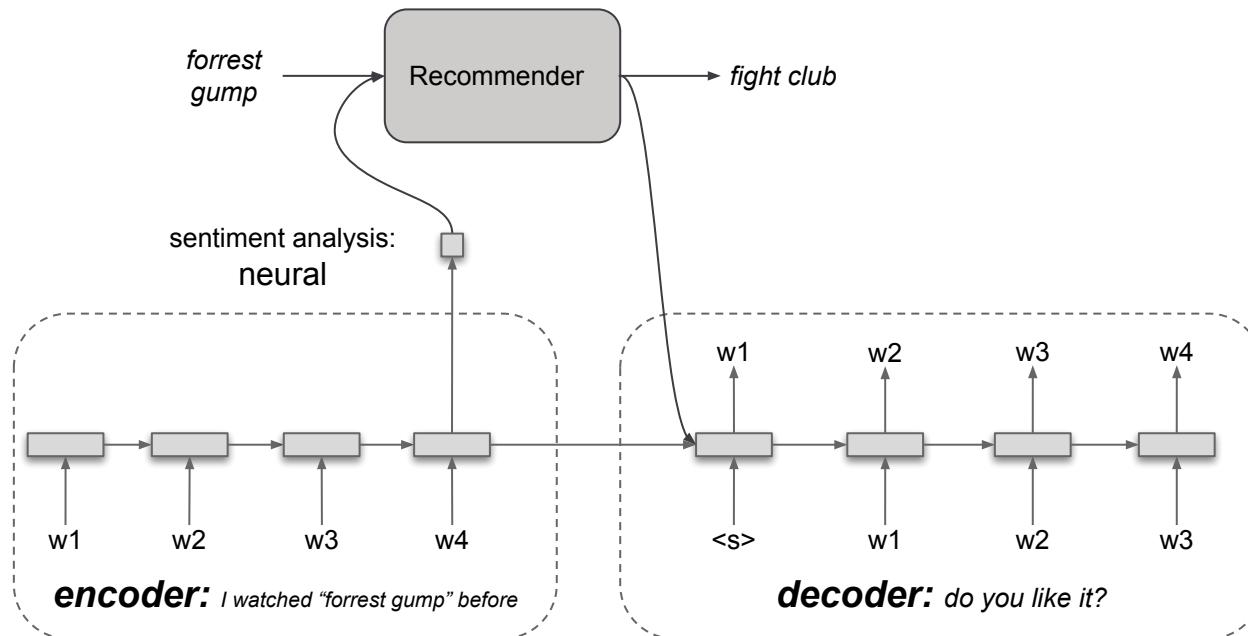
Traditional Methods: Collaborative Filtering

(Christakopoulou et al., 2016)



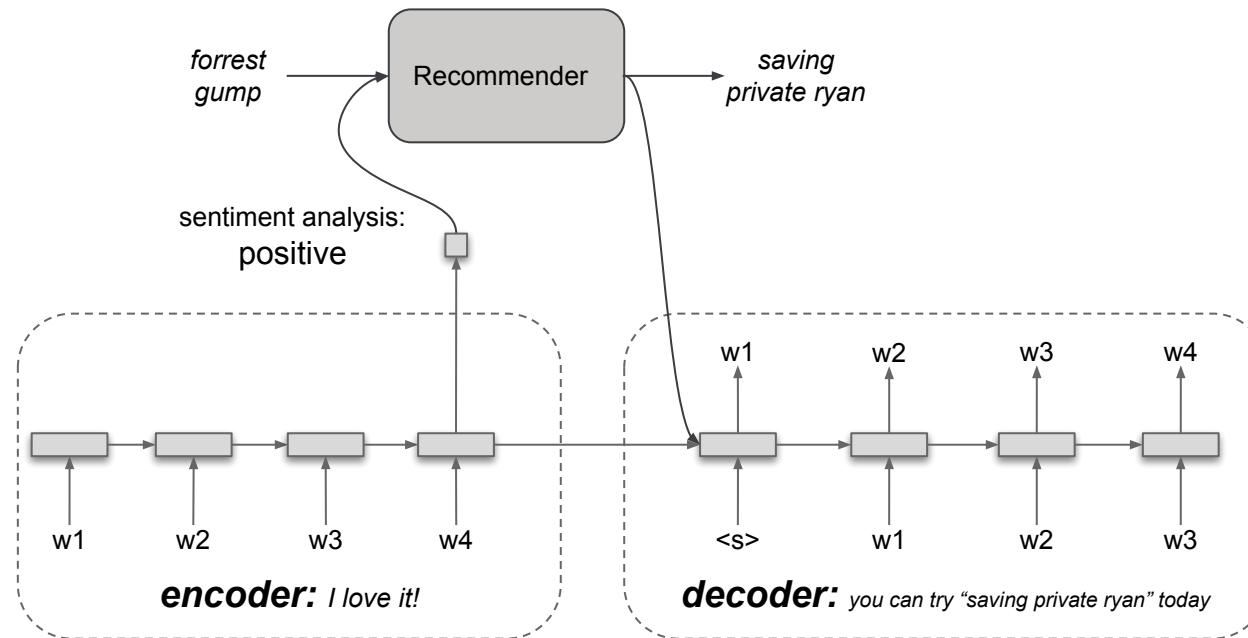
Seq2seq with Recommender Systems

(Li et al., 2018)



Seq2seq with Recommender Systems

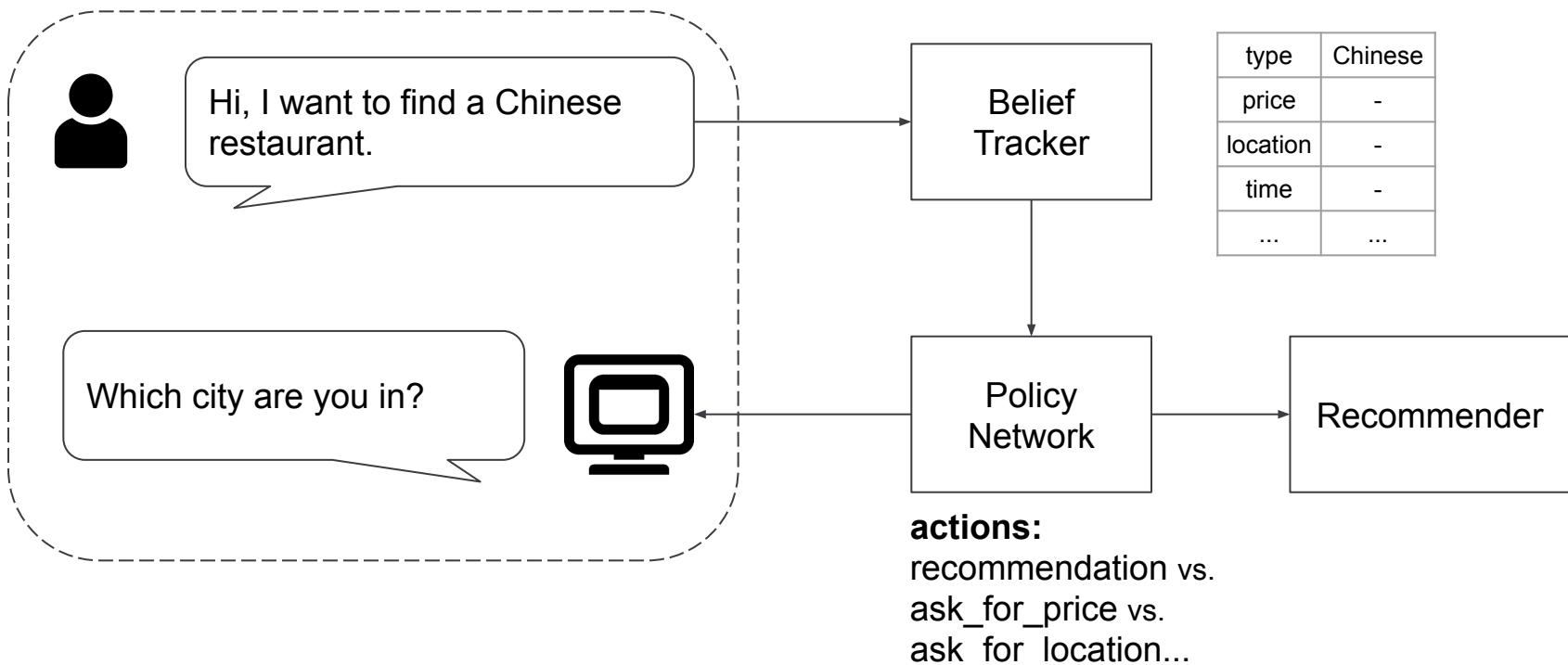
(Li et al., 2018)



- Issue: only sentiment is encoded; not applicable for facets, such as directed-by, release date

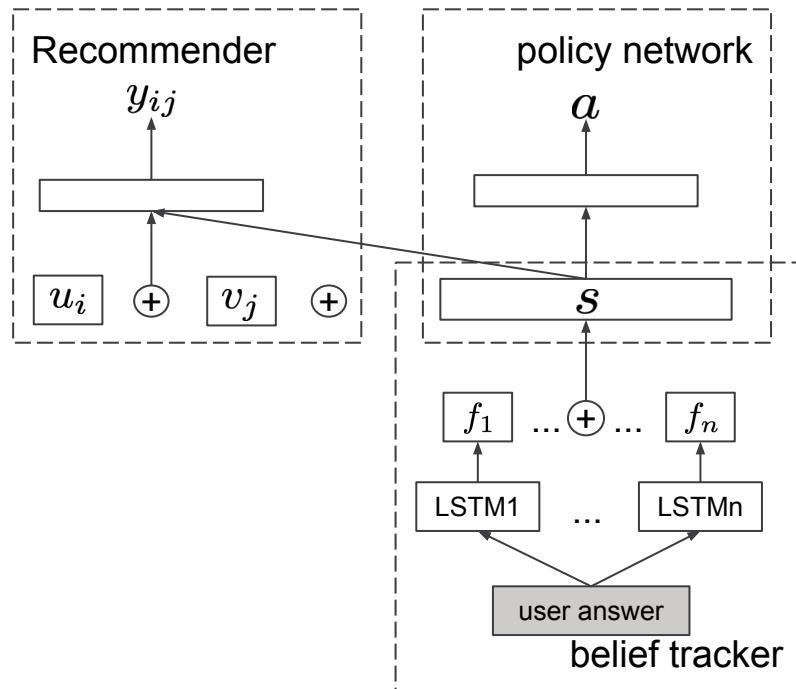
Recommendation with Belief Tracker

(Dhingra et al., 2016, Sun & Zhang, 2018)



Recommendation with Belief Tracker

(Dhingra et al., 2016, Sun & Zhang, 2018)



- Reinforcement learning to consider future actions

rec

location → rec

food_type → location → rec

$$r_0 + \gamma r_1 + \gamma^2 r_2 + \dots$$

Reference

- Sun, Yueming, and Yi Zhang. "Conversational recommender system." In The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval, pp. 235-244. ACM, 2018.
- Li, Raymond, Samira Ebrahimi Kahou, Hannes Schulz, Vincent Michalski, Laurent Charlin, and Chris Pal. "Towards deep conversational recommendations." In Advances in Neural Information Processing Systems, pp. 9725-9735. 2018.
- Christakopoulou, Konstantina, Filip Radlinski, and Katja Hofmann. "Towards conversational recommender systems." In Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining, pp. 815-824. ACM, 2016.
- Dhingra, Bhuwan, Lihong Li, Xiujun Li, Jianfeng Gao, Yun-Nung Chen, Faisal Ahmed, and Li Deng. "Towards end-to-end reinforcement learning of dialogue agents for information access." arXiv preprint arXiv:1609.00777 (2016).
- Christakopoulou, Konstantina, Filip Radlinski, and Katja Hofmann. "Towards conversational recommender systems." In Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining, pp. 815-824. ACM, 2016.

Language Generation for Search: Summary

	Traditional methods	deep NLP methods
Candidate generation	Rule based	End-to-end solution Language modeling
Candidate ranking	Few features	
Latency	Low	High

Agenda

- 1** Introduction
- 2** Deep Learning for Natural Language Processing
- 3** Deep NLP in Search and Recommender Systems
- 4** Real World Examples



Deep NLP in Search and Recommender Systems - Real World Examples at LinkedIn

Huiji Gao

LinkedIn Search and Recommender Systems

Job Search

in machine learning engineer United States Search

Jobs Date Posted LinkedIn Features Company Benefits All filters

Sort by: Relevance

Machine learning engineer in United States 1,454 results Job Alert Off

Sr. Machine Learning Engineer - Home Timeline Promoted Twitter San Francisco, CA, US \$144K – \$314K 3 days ago · 2 applicants

Machine Learning Field Engineer Promoted Cloudera Palo Alto, CA, US 5 days ago

Machine Learning Engineer Apple Cupertino, CA, US \$147K – \$400K 1 day ago · 7 applicants

Senior Machine Learning Device Software Engineer Amazon Web Services (AWS) East Palo Alto, CA, US 5 days ago · 2 applicants

Machine Learning Engineer PayPal San Jose, CA, US 1 day ago · 1 applicant

Sr. Machine Learning Engineer - Home Timeline Twitter · San Francisco, CA, US Posted 3 days ago · 19 views Save Apply

Job Company Connections

- 2 applicants 1,001–5,000 employees
- Associate San Francisco, CA 4 connections

42 company alumni

Job description

Twitter's Consumer Product Teams are responsible for core features of twitter.com, which includes Timelines, Tweets, Search, Trends, Recommendations, Notifications, Tweet details/permalink, and more! Our code operates at massive scale and speed, serving billions of requests per day, connecting hundreds of millions of active Twitter users to real-time information about their lives and the world we live in.

Who We Are

At Twitter, our mission is to instantly connect users to the information most meaningful to them. Realizing this involves work in areas such as machine learning, applied data science, recommendation systems, information retrieval systems, natural language

Job Recommendation

Jobs you may be interested in	
 Senior Manager - Machine Learning Infrastructure Tinder, Inc. San Francisco Bay Area 1 week ago ·  Easy Apply	 Senior Manager - Machine Learning Infrastructure Tinder, Inc. · San Francisco Bay Area Posted 1 week ago · 803 views  
 Director of Applied Machine Learning Arm San Jose, CA, US 2 weeks ago	
 Head of Data Science Govzilla Pleasanton, CA, US 1 day ago ·  Easy Apply	
 AI Research Engineering Manager Facebook Menlo Park, CA, US 1 week ago	Job <ul style="list-style-type: none">• 6/10 skills match• 50 applicants Company <ul style="list-style-type: none">• 201-500 employees• Los Angeles, California Connections  1 connection  4 company alumni
 Vice President, AI Intuit Mountain View, CA, US 6 days ago	Job description <p>Tinder brings people together. With tens of millions of users, hundreds of millions of downloads, 2+ billion swipes per day, 20+ million matches per day and a presence in 190+ countries, our reach is expansive—and rapidly growing. Machine learning plays a critical role at Tinder, we have dozens of Machine learning models in production that power product features like user recommendations, photo moderation, anti-spam, etc. with a variety of algorithms - from linear models to decision trees to deep neural networks and these models operate at a large scale. We are looking for a Senior Manager that will build and manage our machine learning infra team to enable and empower data science and product teams on ML projects at Tinder.</p>
 Director of Machine Learning Tubi San Francisco, CA, US	In this Senior Manager role, you will:

Rank job posts for LinkedIn member to help the user to find jobs that he/she would like to apply for (then, click the “apply” button)

LinkedIn Search and Recommender Systems

People Search

The screenshot shows the LinkedIn search interface for 'bo long artificial intelligence'. The search bar at the top contains the query. Below it, there are dropdown menus for 'People', 'Connections', 'Locations', and 'Current companies', followed by a 'All Filters' button. A search summary indicates 'The Berkeley MBA - Same degree, different schedules. Evening. Weekend. Executive'. Below this, a 'Continue Search in Recruiter' button is shown, along with a note about 8+ additional advanced filters. The main results section displays 5,164 results for 'Bo Long'. Each result includes a profile picture, name, title ('Machine Learning/Artificial Intelligent, Recommendation and Search'), location ('San Francisco Bay Area'), skill ('Artificial Intelligence'), and a 'Message' button. A note below the results states 'Showing 5,164 results'.

Recruiter Search

The screenshot shows the LinkedIn Recruiter search interface for 'machine learning engineer'. The search bar at the top contains the query. Below it, there are sections for 'Search history and alerts', 'Showing results for', 'Custom filters', 'Job titles', 'Locations', and 'Candidate geographic locations'. The main results section displays 619,529 candidates. Key statistics shown include 187,915 candidates more likely to respond, 109,083 open to new opportunities, and 79,232 with company connections. A 'View search insights' button is present. Below the statistics, a card for 'Machine Learning Engineer' at 'NANO Web Group' is shown, along with a 'Save to a project' button and three dots. A note at the bottom indicates 'Past Co-op at IBM 2018 - 2018' and 'Research Assistant at Florida Institute for Cybersecurity Research 2016 - 2018'.

Lead Search

The screenshot shows the LinkedIn Sales Navigator Lead Search interface for 'vp of marketing'. The search bar at the top contains the query. Below it, there are tabs for 'Lead results' and 'Account results'. The 'Lead results' tab is selected, showing 3,242,889 total results. A summary table provides details: 115,241 changed jobs in past 90 days, 69,508 leads with TeamLink intro, 9,505 mentioned in news in past 30 days, and 514,2 posted or past 30 days. The main results section displays two leads, each with a profile picture, name, title ('VP Marketing at Wix.com' and 'VP Marketing at Apple'), location ('Israel' and 'Greater Los Angeles Area'), and a 'Save' button. A note below the results states 'Showing 3,242,889 results'.

People Recommendation

The screenshot shows the LinkedIn People Recommendation interface. It features a heading 'People you may know with similar roles' and a 'See all' link. Below this, four recommendation cards are displayed, each with a circular profile picture and a 'Connect' button. The cards are: 1) 'Engineering Leader @Uber (We are hiring!)' with 44 mutual connections; 2) 'Sr Manager, Software Engineering' with 36 mutual connections; 3) 'Manager Machine Learning and Data' with 49 mutual connections; and 4) 'Engineering Manager at Change.org' with 19 mutual connections.

LinkedIn Search and Recommender Systems

Content Search

The screenshot shows a LinkedIn search results page for the query "KDD 2018 LINKEDIN". At the top, there's a search bar with the query, followed by navigation links for Home, My Network (with 2 notifications), Jobs, and Messaging. Below the search bar, a post from Steve Gebrezgier is displayed, featuring his profile picture, name, title, and a brief bio. The post content discusses attending KDD 2018 and mentions machine learning and artificial intelligence. Below the post, there's a summary of the LinkedIn company page for "LinkedIn", which has 8,487,084 followers and was updated 1 year ago. A call-to-action encourages users to visit the LinkedIn booth at KDD 2018. At the bottom of the page, a large blue LinkedIn logo is displayed, and the text "LinkedIn @ KDD 2018" and the URL "engineering.linkedin.com" are shown.

Content Recommendation

The screenshot shows a LinkedIn content recommendation feed. At the top, there's a search bar with the word "Search", followed by navigation links for Home, My Network (with 14 notifications), Jobs, and Messaging. The main feed features a post from Nadiya Hayes, Chief of Staff at LinkedIn, with a "PREMIUM" badge. The post discusses attending KDD 2019 and includes a link to a preview. Below the post, there are like, comment, and share buttons. To the right of the post, there's a sidebar with recent activity and a "Ads" section for Microsoft. The sidebar shows posts from groups like ACM SIGKDD & Annual KD... and LinkedIn Company Group, along with hashtags for finance, healthcare, and entrepreneurship. The "Ads" section displays an advertisement for Microsoft 365 Training Day: Desktop Deployment in Sunnyvale, CA on August 21, 2019. The bottom of the page shows a comment from Brendan McWeeney and various footer links and icons.

Natural Language Data at LinkedIn

The screenshot shows a LinkedIn search interface with a red box highlighting the search bar containing 'machine learning' and the word 'Query'. A blue arrow points from the 'Query' text in the search bar to the 'Huiji Gao' profile on the right. Another blue circle highlights the user profile picture of Huiji Gao.

Search Bar: machine learning **Query** United States

Filters: Jobs ▾ Date Posted ▾ LinkedIn Features ▾ Company ▾ Experience Level ▾ All filters

Sort by: Relevance

Course details: Learning Courses

6h 32m · Beginner + Intermediate · Released: April 10, 2017 · 11 chapter quizzes

Exercise Files: See all

By using Python to glean value from your raw data, you can simplify the often complex journey from data to value. In this practical, hands-on course, learn how to use Python for data preparation, data munging, data visualization, and predictive analytics. Instructor Lillian Pierson, P.E. covers the essential Python methods for preparing, cleaning, reformatting, and visualizing your data for use in analytics and data science. She helps to provide you with a working understanding of machine learning, as well as outlier analysis, cluster analysis, and network analysis. Plus, Lillian explains how to create web-based data visualizations with Plot.ly, and how to use Python to scrape the web and capture your own data sets.

Learning Objectives:

- Getting started with Jupyter Notebooks
- Visualizing data: basic charts, time series, and statistical plots
- Preparing for analysis: treating missing values and data transformation
- Data analysis basics: arithmetic, summary statistics, and correlation analysis
- Outlier analysis: univariate, multivariate, and linear projection methods
- Introduction to machine learning
- Basic machine learning methods: linear and logistic regression, Naïve Bayes

Huiji Gao
Engineering Manager - Machine Learning and AI at LinkedIn

Member Profiles

Experience

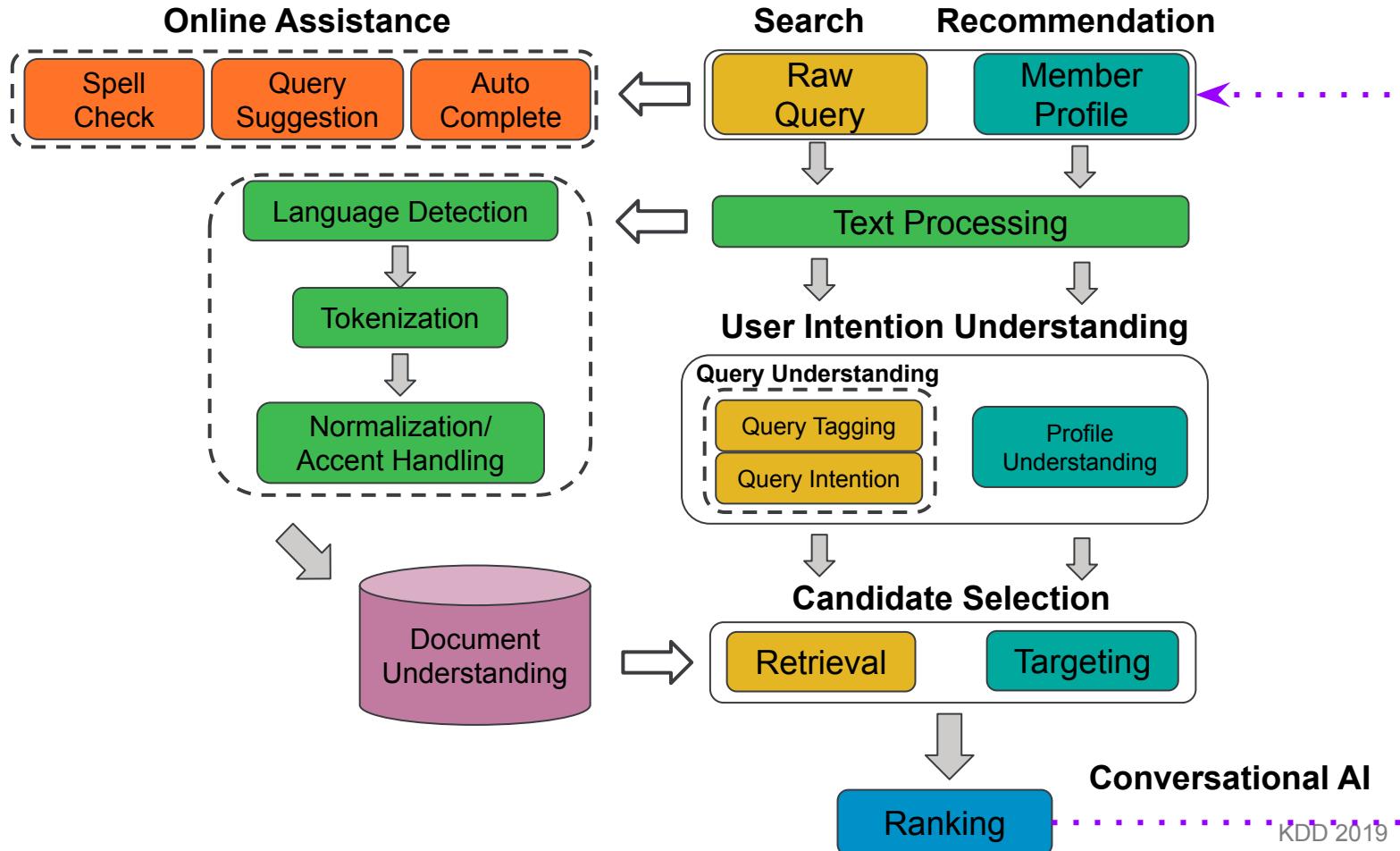
- LinkedIn 4 yrs 3 mos
Engineering Manager - Machine Learning and AI
Aug 2018 – Present · 9 mos
San Francisco Bay Area
Lead LinkedIn Personalization and Search AI Foundation team - Provide LinkedIn users with intelligent experience through natural language understanding (across multiple languages) and personalization powered by Machine Learning and Artificial Intelligence.
- Staff Machine Learning Engineer
Mar 2018 – Jul 2018 · 5 mos
San Francisco Bay Area
Promote LinkedIn's search relevance foundation with high-quality search results and satisfactory searcher experience powered by Machine Learning and Artificial Intelligence
- Senior Machine Learning Engineer - Computational Advertising and Information Retrieval
Jun 2016 – Feb 2018 · 1 yr 9 mos
San Francisco Bay Area
Ads Relevance:
Worked on a variety of ads relevance products, including audience expansion behavior modeling, campaign performance optimization, and CTR prediction. Developed several important classes of machine learning models that have generated double-digit % increases in ad relevance.
- Applied Research Engineer
Feb 2015 – May 2016 · 1 yr 4 mos
San Francisco Bay Area
Computational Advertising

Role

- Research Assistant
Arizona State University
Aug 2009 – Dec 2014 · 5 yrs 5 mos
Design and implement a disaster relief system ACT (ASU Coordination Tracker) to enhance the coordination among relief organizations.
- Mining large-scale location-based social network data to study human mobile behavior

Buttons: Save Apply

LinkedIn Search and Recommendation Ecosystem



NLP in LinkedIn: Challenges

- **Data Ambiguity**
 - Short Query Text
 - “abc”
ABC News? ABC Stores?
 - No Strict Syntax
 - “bing search engineer”
“Bing Search, Engineer” “Bing, Search Engineer”
 - Personalization
 - “looking for new jobs”
Job Seeker looks for jobs
Recruiter looks for candidates
- **Deep Semantics**
 - Representations of documents w.r.t. member intent
 - “developer proficient at cloud computing” -> “Azure, AWS ...”
 - “engineering openings” -> Job Posts

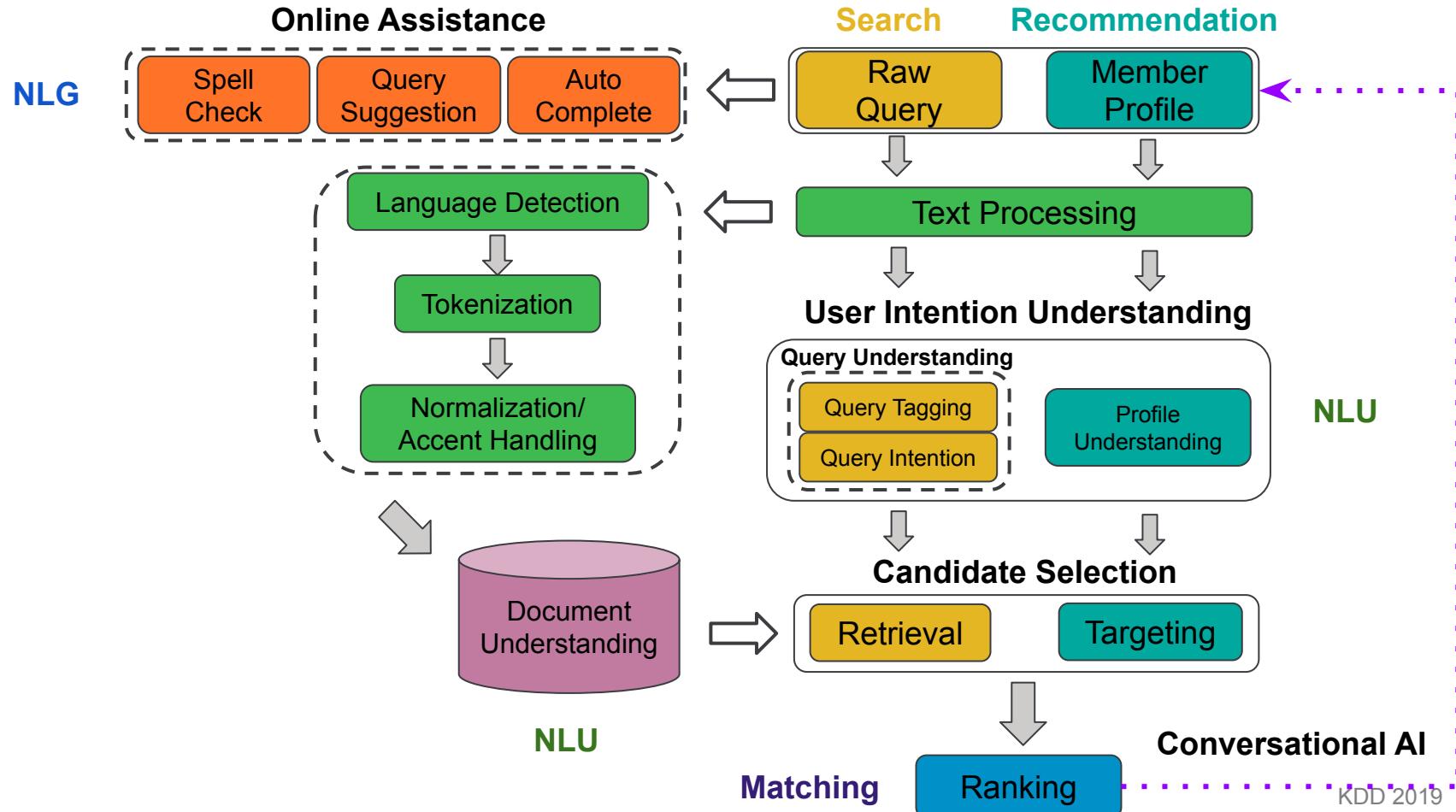
Deep NLP in LinkedIn Search & Recommendation: Challenges

- **Product Oriented Model Design**
 - Design deep NLP algorithms for specific search & recommendation components
 - Consider business rules, post filters, results blender, user experience, etc
- **Model Debugging**
 - Understand model for online performance investigation
- **Online Latency**
 - Serving deep NLP models with product latency restriction

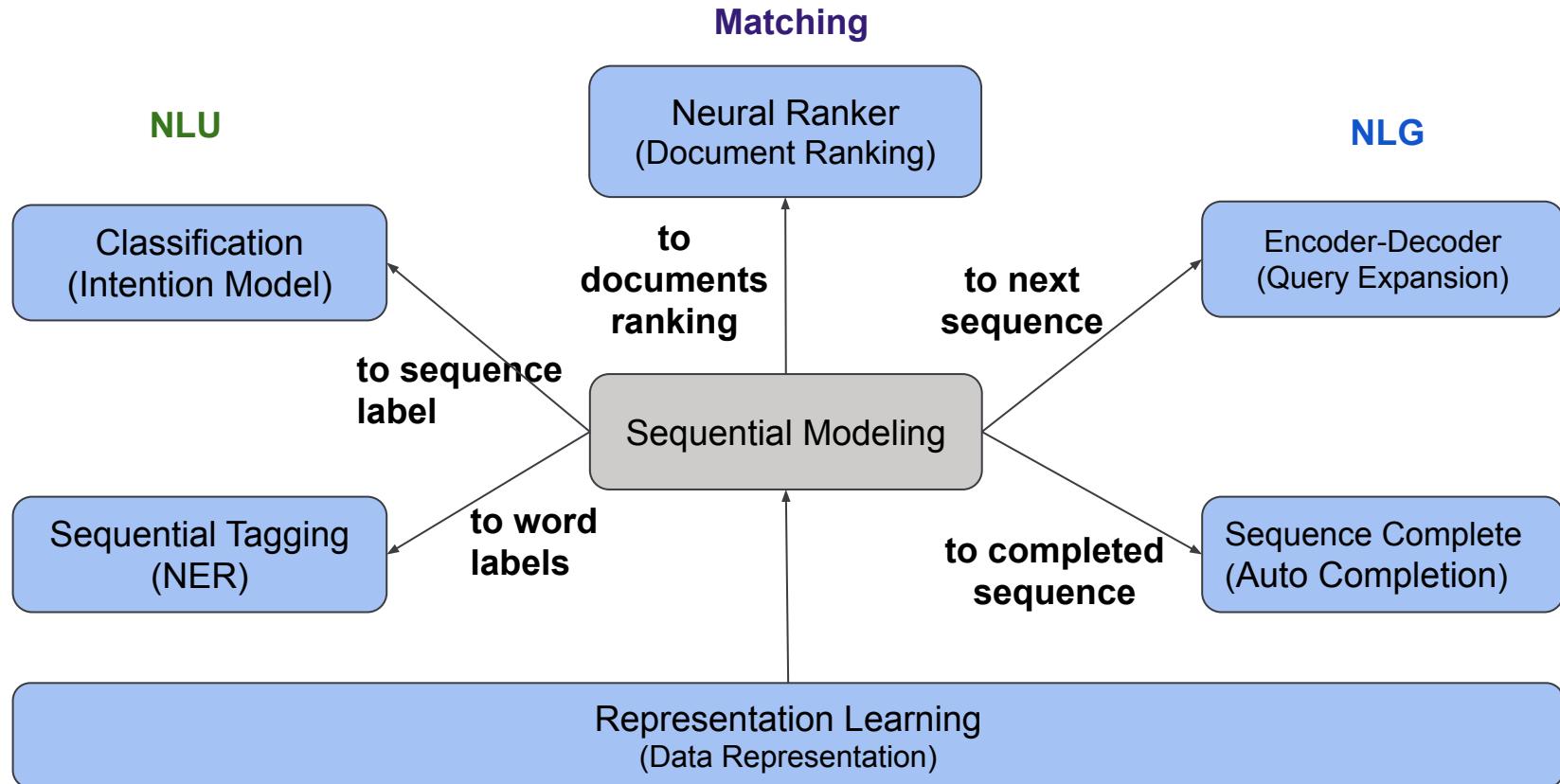
Applying Deep NLP in LinkedIn Search & Recommendation

- **Feature Driven**
 - Representation Learning
 - Using features generated from deep learning models
e.g., word embedding
- **Model Driven**
 - Power product features directly with deep learning models
 - CNN/LSTM/Seq2seq/GAN/BERT based deep NLP models

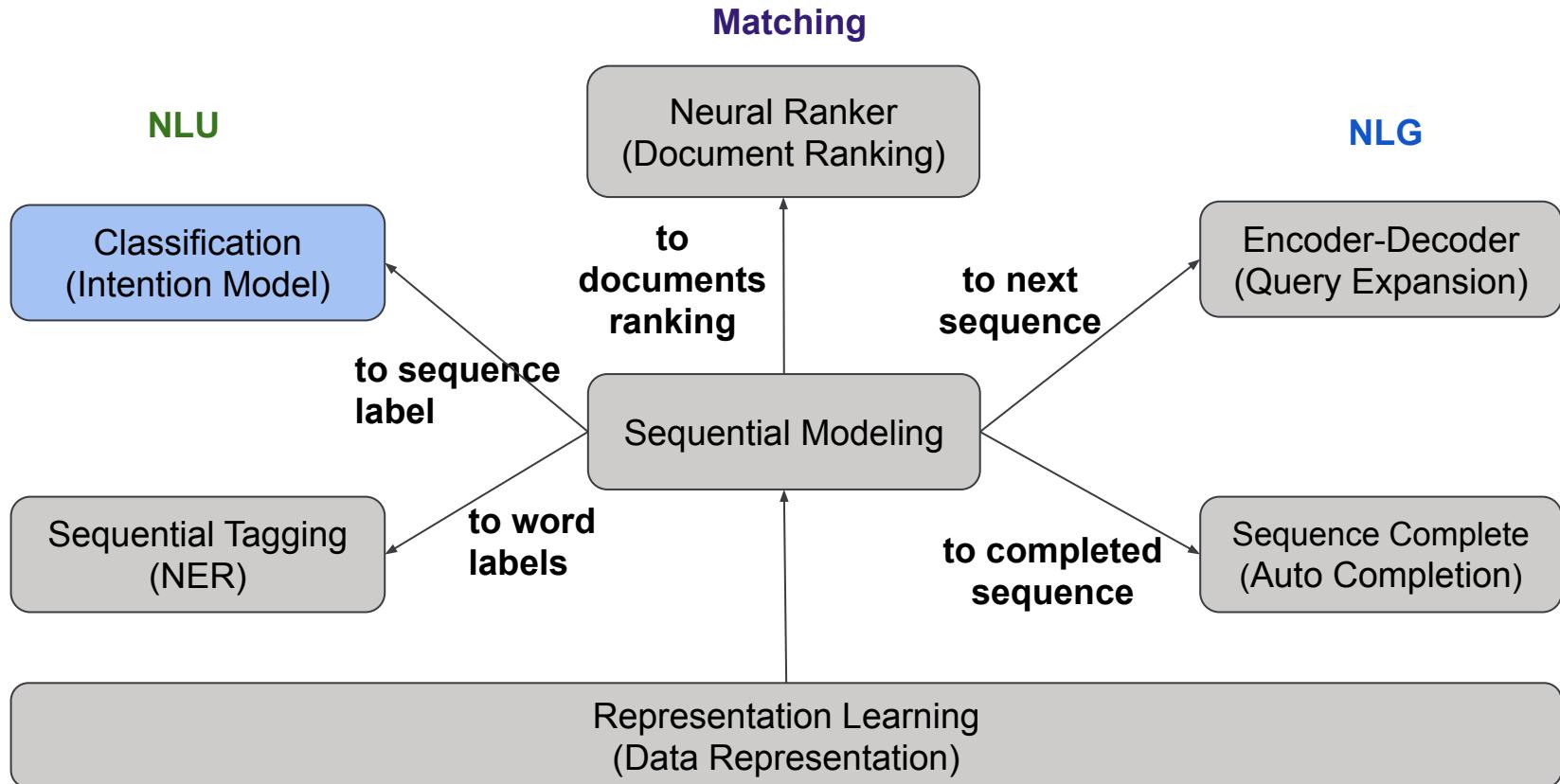
LinkedIn Search and Recommendation Ecosystem



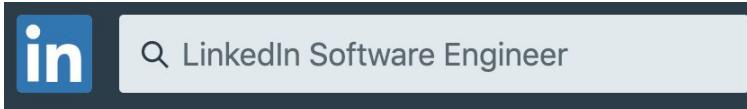
Deep Learning for Natural Language Processing



Deep Learning for Natural Language Processing



Query Intention Model: Goal



- **Output of Query Intention Model**

- Search Vertical Prediction
 - **People, Job Posts, Articles ...**
- Properties Associated with the Vertical
 - **Extracted** from raw query **OR Inferred** from other information

The searcher is looking for:

0.99

People
from “LinkedIn” as a
“Software Engineer”

0.06

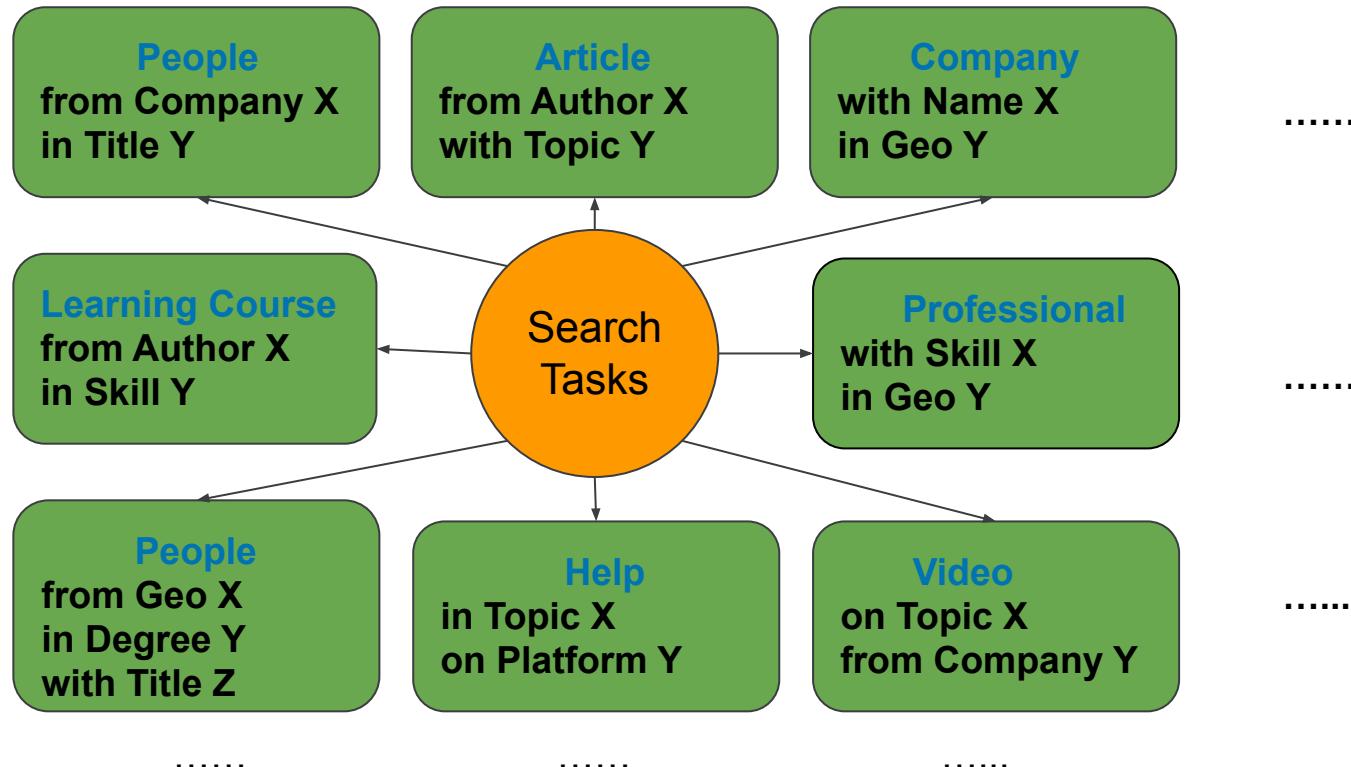
Job Post
from “LinkedIn” on
“Software Engineer”
position

0.03

Article
from “LinkedIn” on
“Software Engineer”
topic

.....

Query Intention Model: Task Oriented Intention Prediction



Query Intention: Member Search Footprint

Bridging the Gap between Search and Recommendation



Query Intention Model: Goal

Query: LinkedIn Software Engineer

- **Output of Query Intention Model**

- Vertical Prediction
 - **People, Job Posts, Articles ...**
- Properties Associated with the Vertical
 - Extracted from raw query **OR** Inferred from other information

The searcher is looking for:

0.99

People
from “LinkedIn” as a
“Software Engineer”

0.06

Job Post
from “LinkedIn” on
“Software Engineer”
position

0.03

Article
from “LinkedIn” on
“Software Engineer”
topic

.....

Query Intention on Search Blending

The screenshot shows the LinkedIn search interface for the query "google machine learning engineers". The search bar at the top contains the query. Below it, the navigation bar includes Home, My Network (with 3 notifications), Jobs (with 4 notifications), and Messaging. The "Connections" and "Locations" dropdowns are highlighted with orange arrows pointing to them from the right side of the slide.

The main search results section displays 5 people profiles:

- Prashant Malani** • 3rd
Software Engineer at Google
San Francisco Bay Area
Past: Teaching Assistant - 18-603 : Leadership for Engineers at Carnegie Mellon University
- Tyler Beneke** • 3rd
Network Operations - Data Center Networking Program Manager
San Francisco Bay Area
Current: Network Operations - Data Center Networking Program Manager at Google
- Juhil Somaiya** • 3rd
Google Cloud Facilitator || Python developer || Tech Speaker || Machine learning fan...
Ahmedabad Area, India
Current: Google Assistant Developer Community member at Google

Below these results, a section titled "Job results for google machine learning engineers" shows 748 results. It includes a "See all" link and three job cards:

- Software Engineer, Machine Learning
Google
- Software Engineer, Chrome Networking
Google
- Partner Engineer, Machine Learning, Cloud Global...
Google

Annotations on the right side of the slide identify specific elements:

- A large orange arrow points to the "Connections" and "Locations" dropdowns with the label "Search verticals".
- Two green arrows point to the first two people profiles (Prashant Malani and Tyler Beneke) with the label "Primary result".
- A yellow arrow points to the job results section with the label "Secondary cluster".

Goal:

- To understand the vertical preference of user on LinkedIn search

Query Intention Model: Challenges

- Complicated Semantics



Query Intention Model: Challenges

- Personalization
 - **Query:** Blockchain Tech
 - **Job seeker** looks for Blockchain Technology job
 - **Company** named Blockchain Tech
 - **Learning course** on Blockchain Tech
 - **Content** on Blockchain Technology
 - **Video** about blockchain technology
 - **Recruiter** looks for candidates with Blockchain tech skill
 -

The screenshot shows the LinkedIn search interface with the query 'Blockchain Tech'. It displays three company profiles under the heading 'Showing 2,434 results':

- Blockchain Tech** (Information Technology and Services)
- Blockchain Tech News** (Online Media)
- Blockchain Tech** (Information Technology and Services)

The screenshot shows the LinkedIn search interface with the query 'Blockchain'. It displays several groups and hubs:

- Blockchain (Software & Technology Professionals Managers | HR | Recruiters | Blockchain & Investors Group) - 1,796,156 members
- Blockchain Industry Group (Blockchain Industry Group) - 1,796,156 members
- Hub-Centres - Startups, Financing, Blockchain, Bitcoin & Jobs (Hub-Centres - Startups, Financing, Blockchain, Bitcoin & Jobs) - 40,276 members
- Information Technology, Fintech, Blockchain and Bitcoin Innovation (Information Technology, Fintech, Blockchain and Bitcoin Innovation) - 2,024,055 members
- Developers - Android, iOS developer , Blockchain, Ethereum, Java, Ruby, .net, PHP, django, etc (Developers - Android, iOS developer , Blockchain, Ethereum, Java, Ruby, .net, PHP, django, etc) - 10,139 members

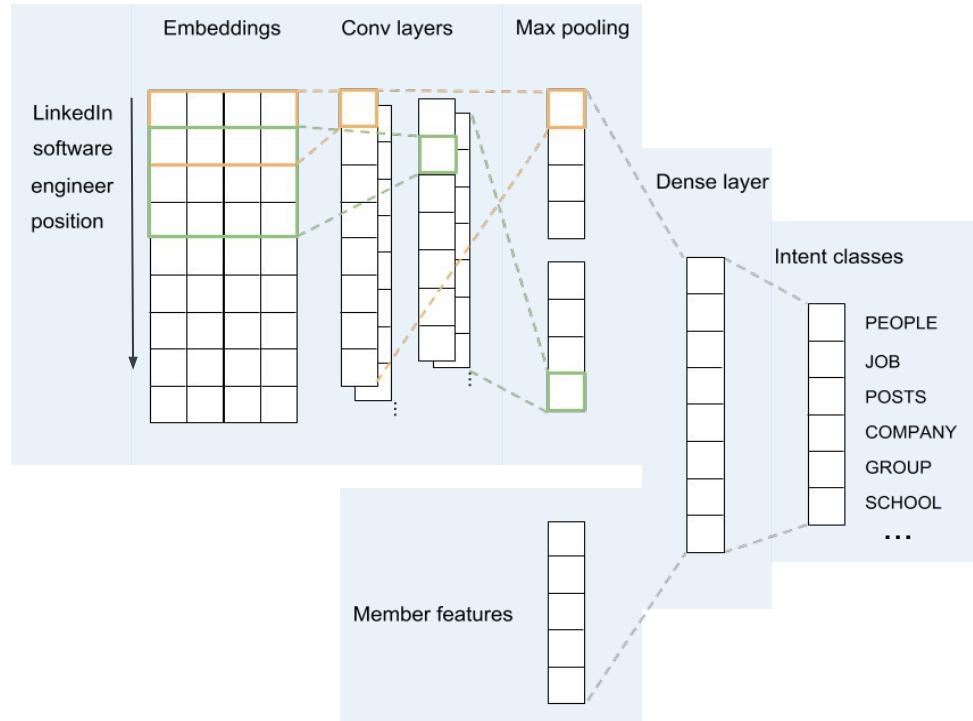
CNN Based Query Intention Model

CNN for Semantic Feature Extraction

- Word/query representations
- Generalization power
- Word n-gram patterns

Personalization

- Member-level Features



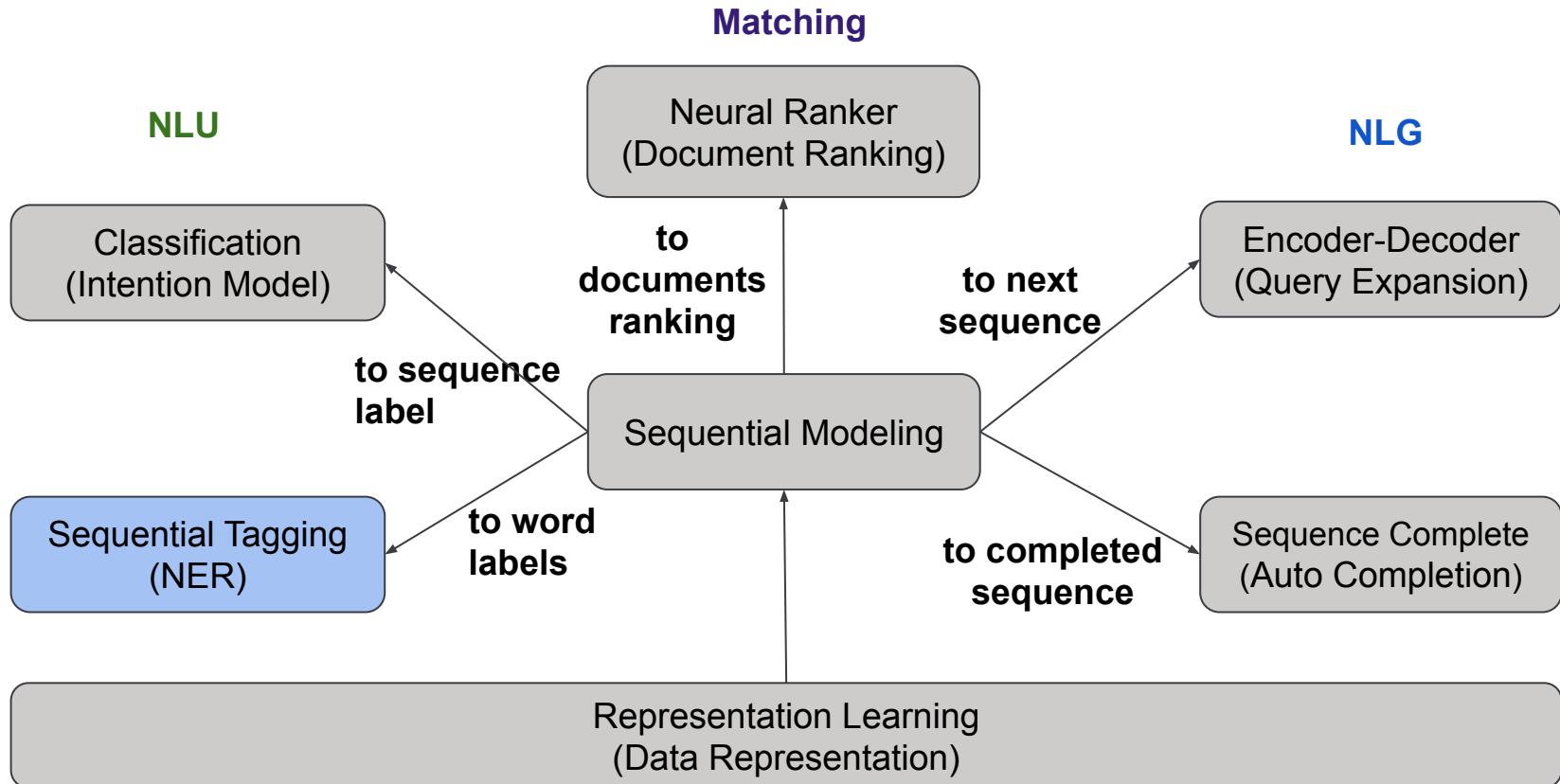
Query Intent - Experiment Results

- Offline Results

	Overall Accuracy	F1 on PEOPLE	F1 on JOB
Baseline (ML Model)	-	-	-
CNN	+2.9%	+11.9%	+1.7%

- Online Results
 - +0.65% Job CTR@1
 - +0.90% Overall Cluster CTR, +4.03% Cluster CTR via Entity Click

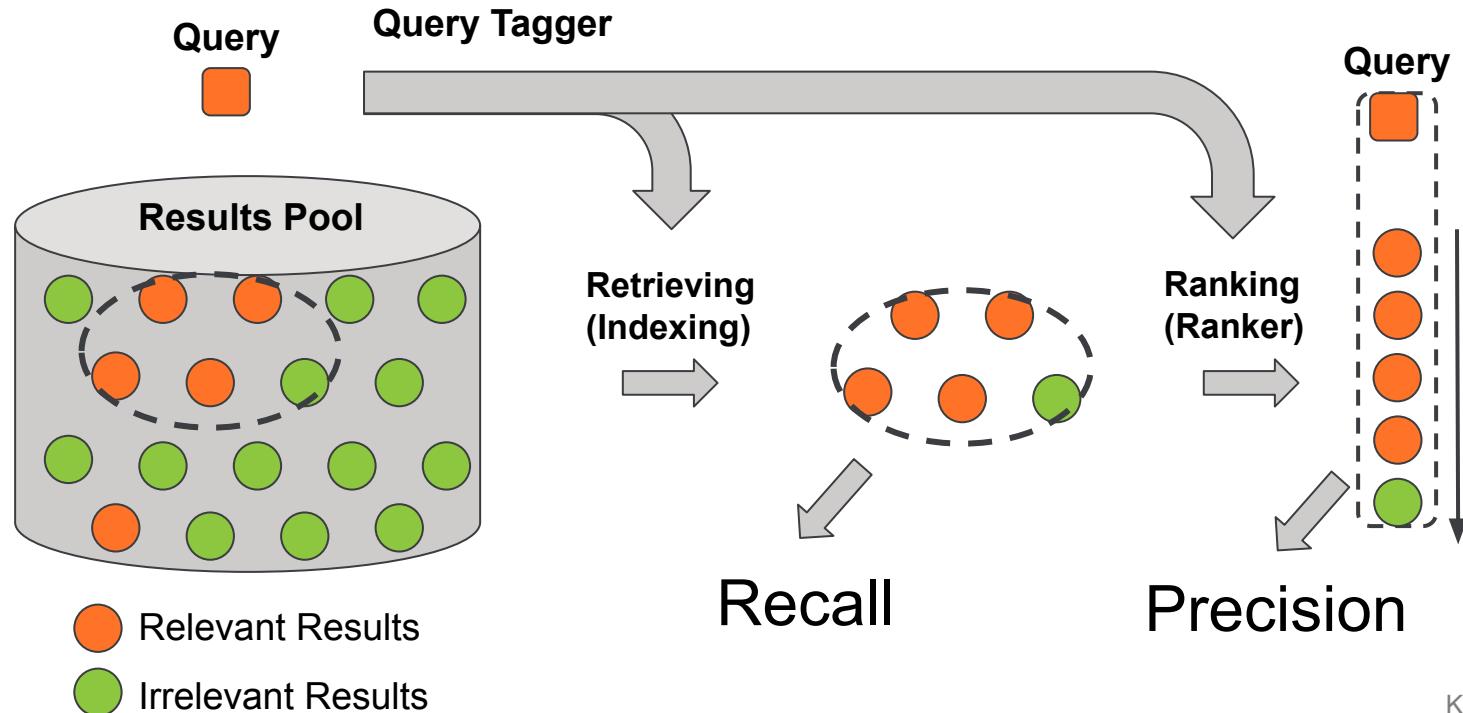
Deep Learning for Natural Language Processing



Entity Tagging at LinkedIn

- **Search Systems**

Query: Mike LinkedIn Software Engineer



Entity Tagging at LinkedIn

- **Understanding Queries & Documents with Entity Tagger**

Query: Mike LinkedIn Software Engineer

- **Query Tagger for Retrieval**

CN: company name

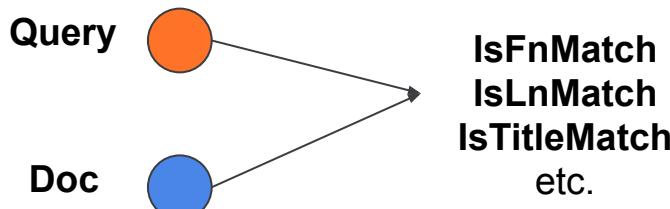
FN: first name

T: title

Mike LinkedIn Software Engineer
FN CN T T

- **Ranking Features**

Index	
FN:{mike}	Doc1, Doc2 , Doc4, ...
CN:{linkedin}	Doc2 , Doc4, Doc5, Doc6, ...
T:{software}	Doc2, Doc3, Doc4 , Doc7, ...
T:{engineer}	Doc1, Doc2 , Doc4, Doc9, ...



Entity Tagging at LinkedIn

● Recommender Systems

Ads Recommendation

Who is your target audience?

Advertiser Targeting

INCLUDE people who have ANY of the following attributes

Member Skills

- Customer Satisfaction X Customer Experience Management X Customer Experience X
- Customer Acquisition X Customer Retention X Strategic Sales Plans X
- Marketing Strategy X Key Client Relationships X Customer Loyalty Management X
- High Performance Sales Teams X Strategic Account Development X
- Key Account Acquisition & Retention X New Customer Acquisitions X Customer Loyalty X
- Loyalty Marketing X Direct Marketing X + Add Member Skills

AND also have ANY of the following attributes

Home > Interests > Member Interests Search

Member Groups

Member Interests

- Arts and Entertainment
- Business and Management
- Finance and Economy
- Marketing and Advertising
- Politics and Law
- Sales and Retail



Member Profile

Natural Language Understanding: Entity Tagging

LinkedIn	software	engineer	data	scientist	jobs
CN	T	T	T	T	O
B-CN	B-T	I-T	B-T	I-T	O

B-CN: beginning of a company name

I-CN: Inside of a company name

B-T: beginning of a job title

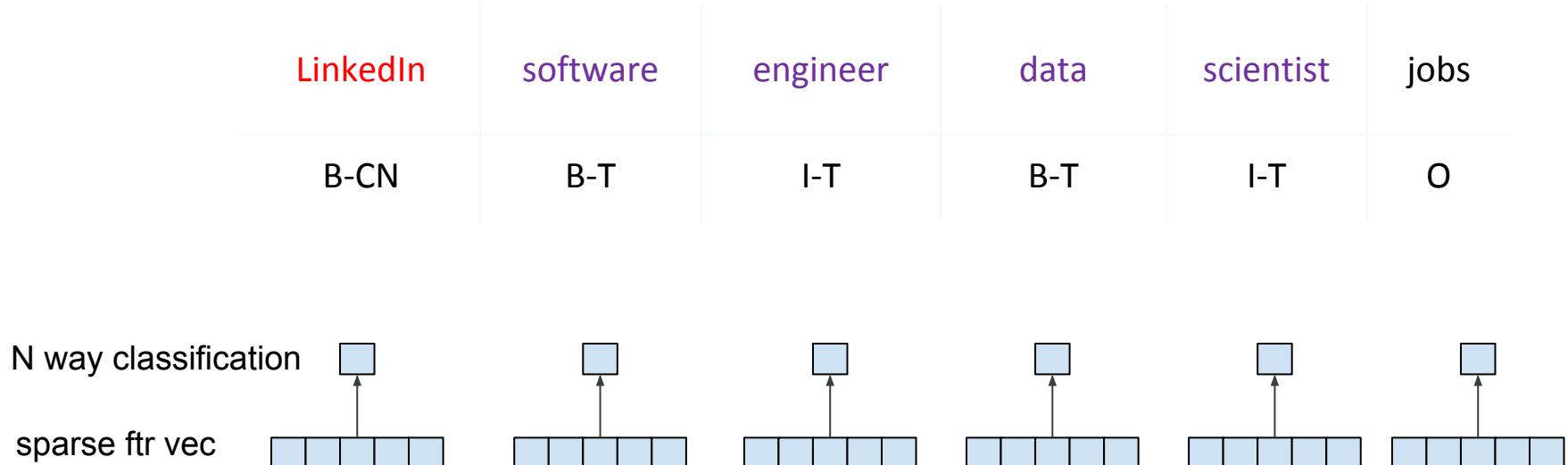
I-T: Inside of a job title

O: Not an entity

B-PN: beginning of person name

...

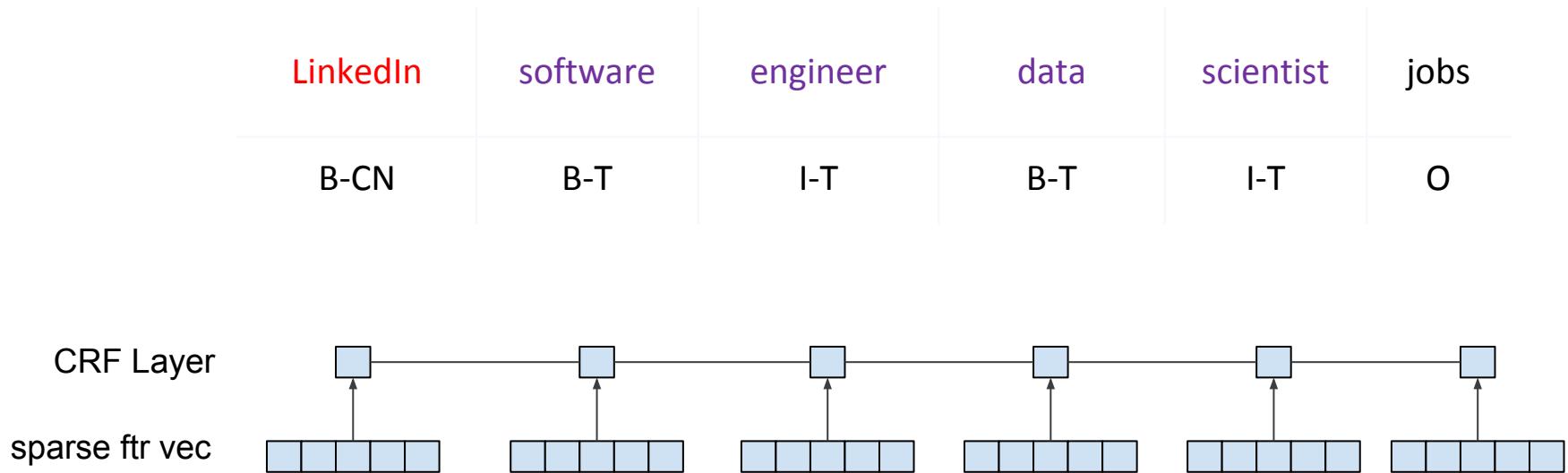
Query Tagger: Logistic Regression



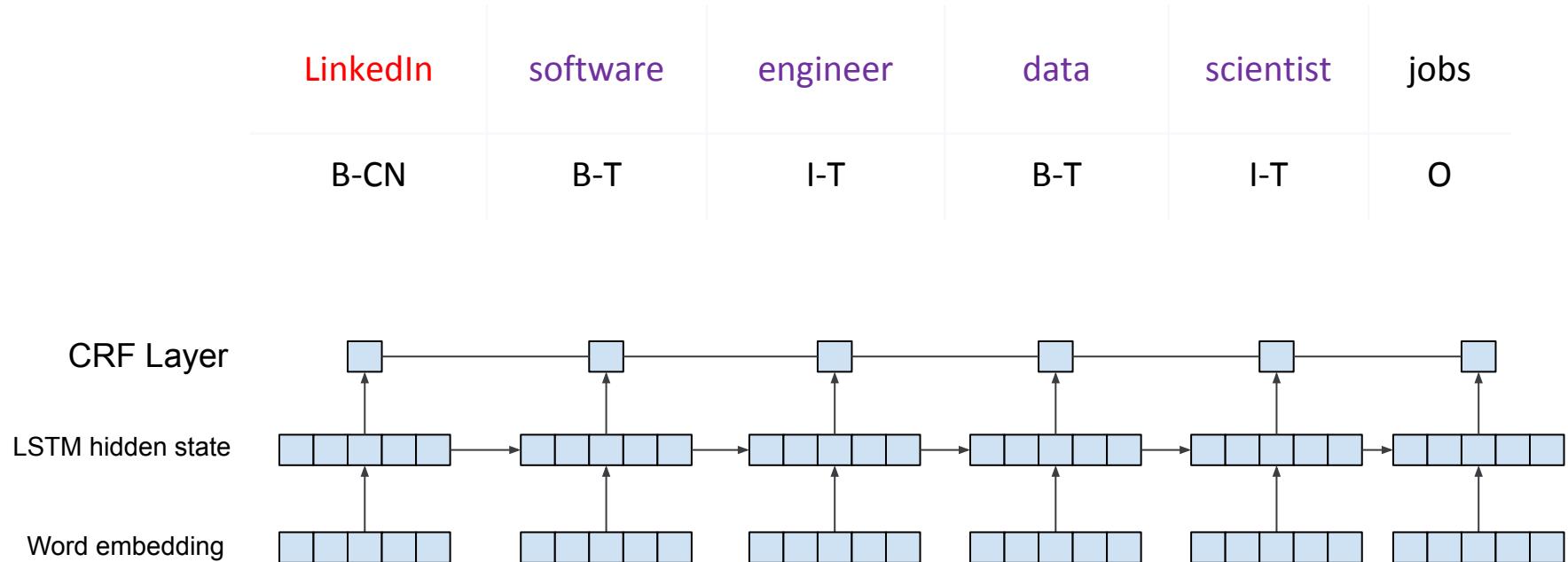
ftr 0: whether the current word is “linkedin”
ftr 1: whether the current word is “facebook”
...

ftr n: whether the next word is “software”
ftr n+1: whether the next word is “linkedin”

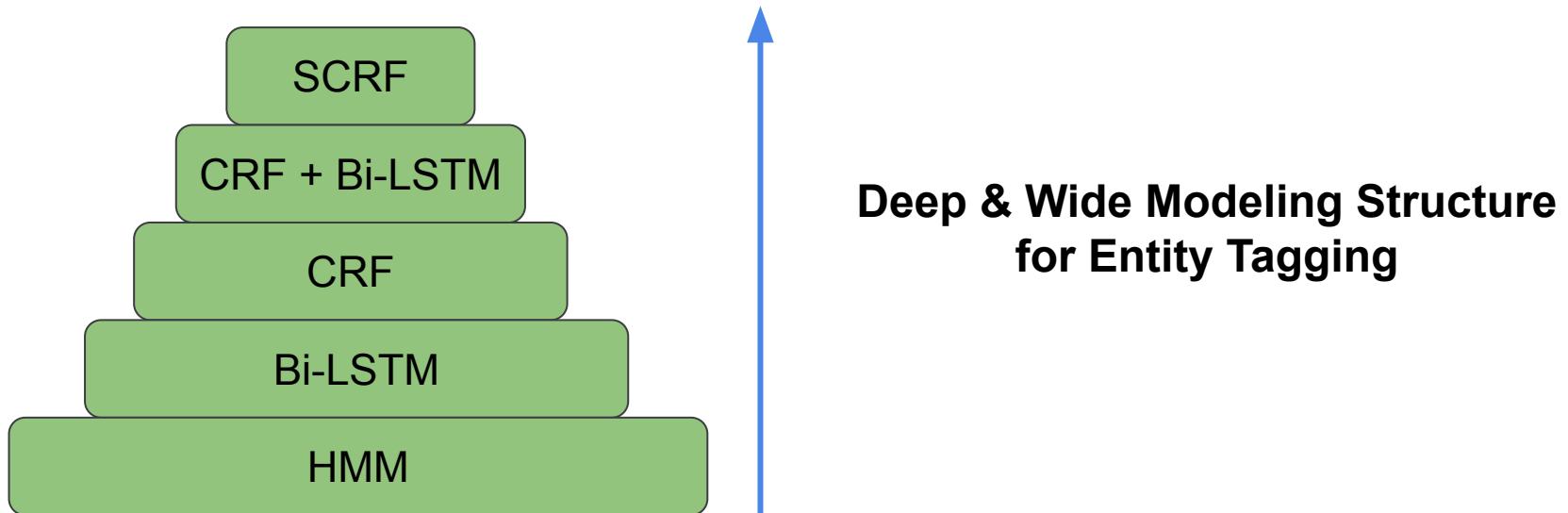
Query Tagger: CRF



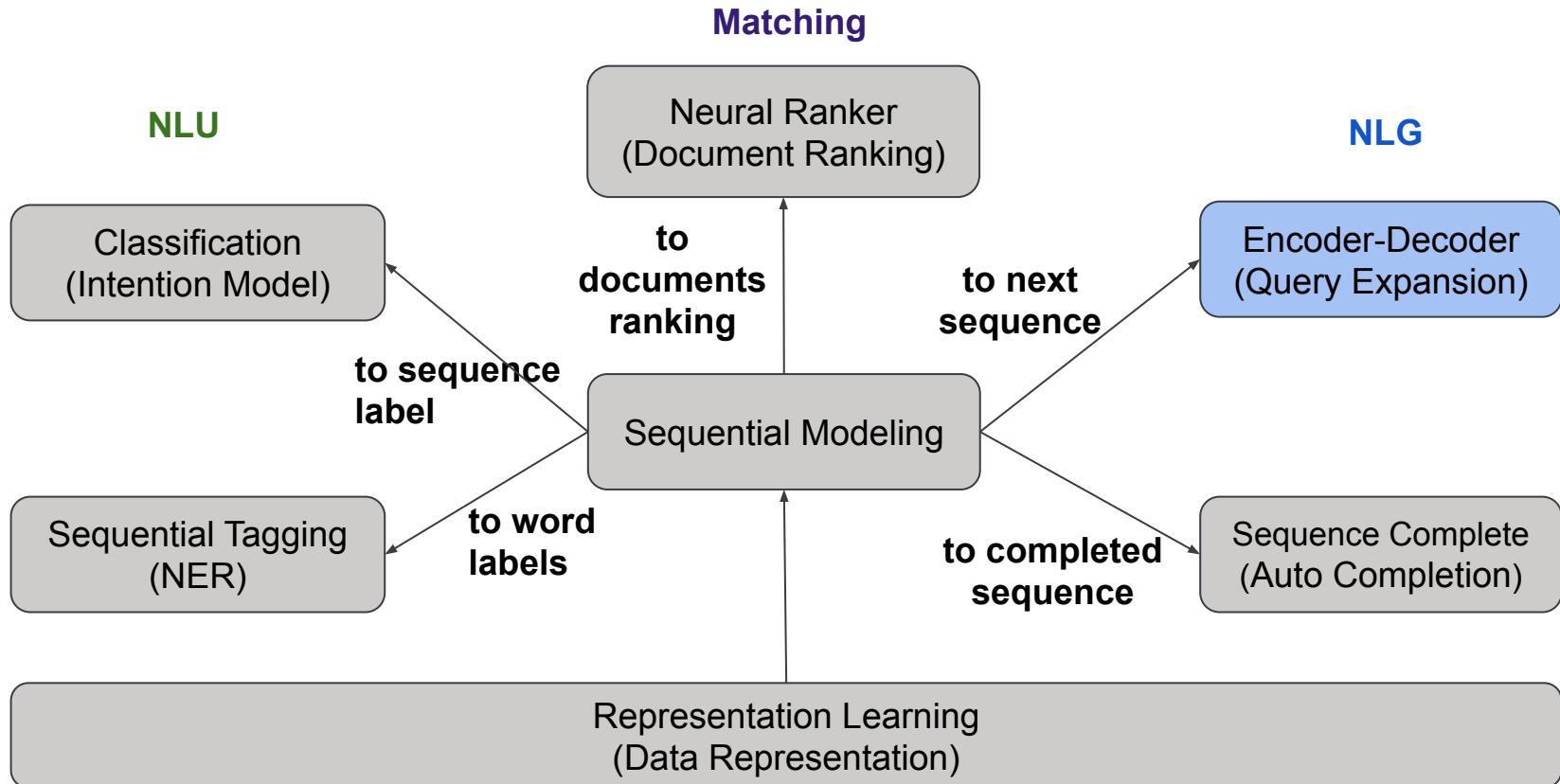
Query Tagger: CRF + LSTM



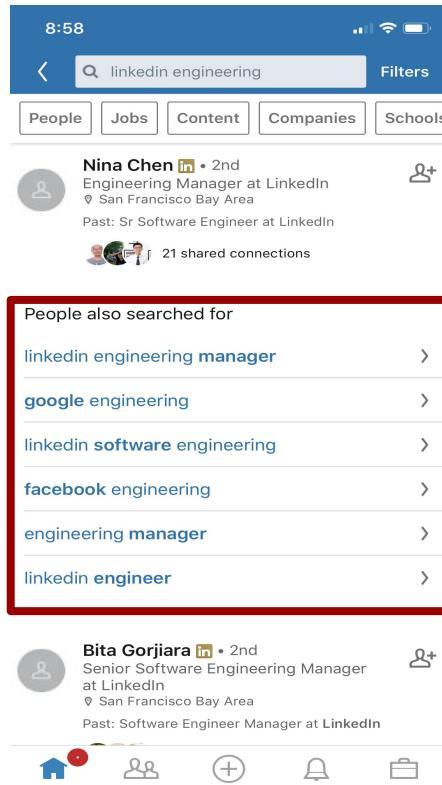
Query Tagger Performance



Deep Learning for Natural Language Processing

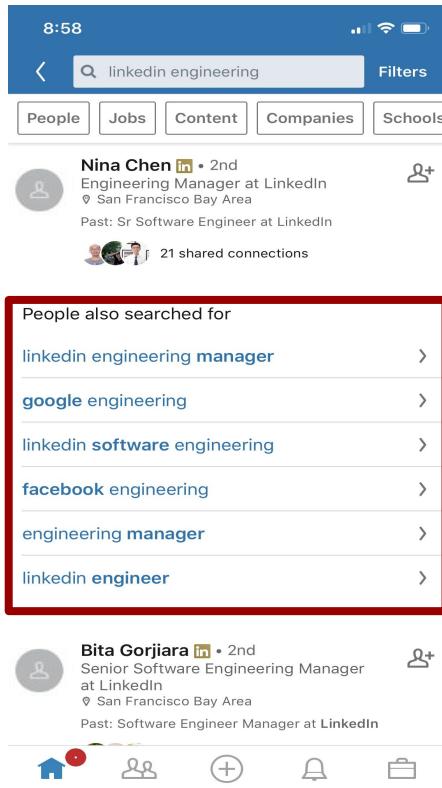


Natural Language Generation: Query Suggestion



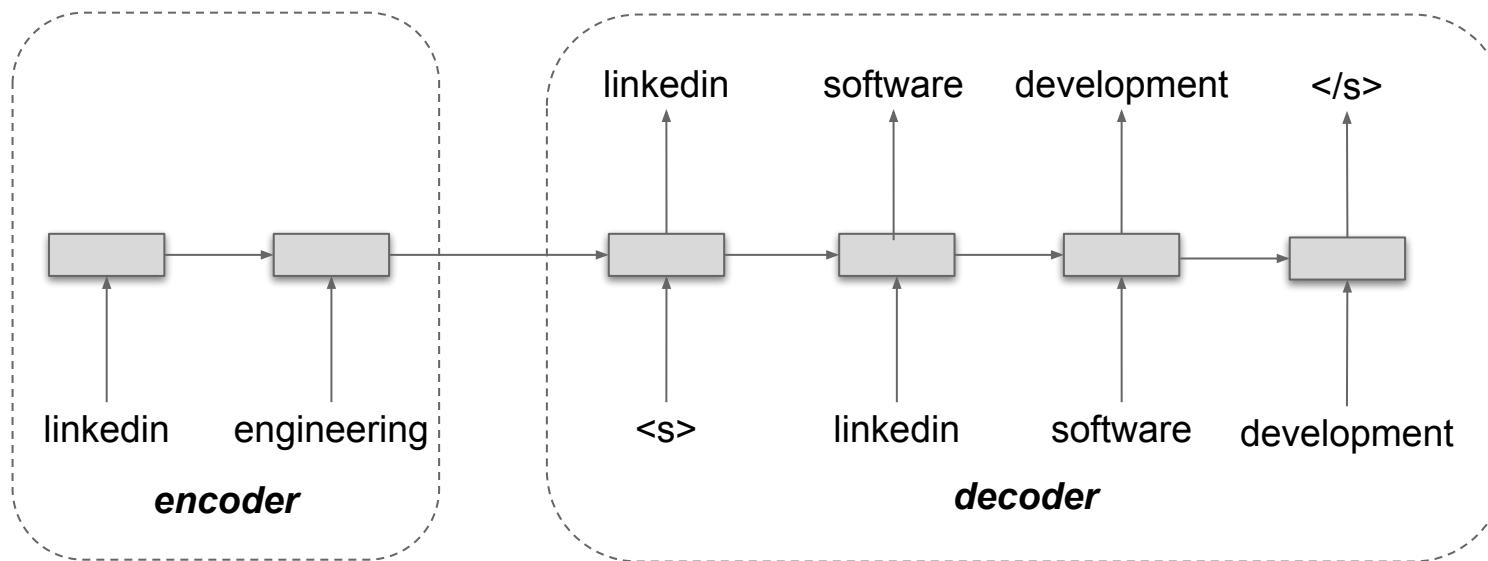
- Save User Effort of Rephrasing Informative Queries
- Capture Users' Search Intention
- Personalization

Natural Language Generation: Query Suggestion



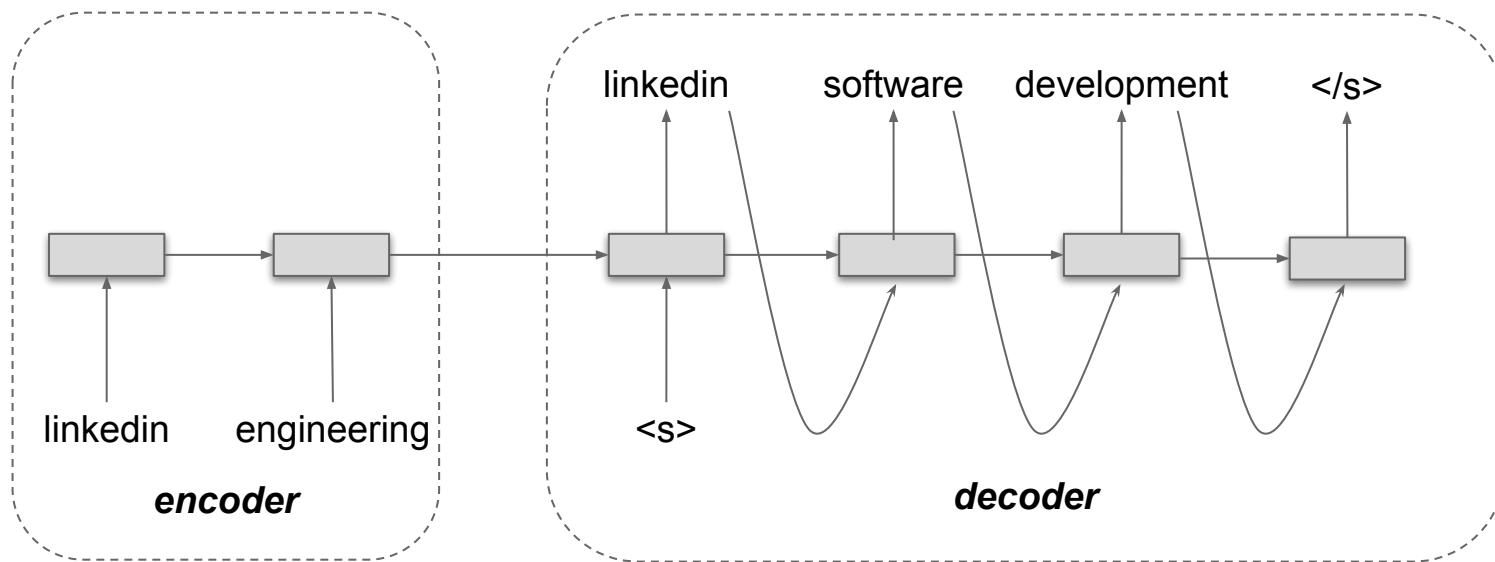
- Traditional Frequency Based Methods
 - Collect $\langle q_1, q_2 \rangle$ pairs from search log
 - Save the frequent pairs in a key-value store
- Lack of Generalization
 - Purely string matching
 - Cannot handle unseen queries, rare words
- Seq2seq: Model Query Reformulation Behavior

Query Suggestion: Reformulate to Related Queries



- Training: the 2nd query is given
- Maximize $P(\mathbf{y}|\mathbf{x}) = \prod P(y_i|h_i)$

Query Suggestion: Reformulate to Related Queries



- Inference: the 2nd query is unknown
- Beam search instead of greedy search

Query Suggestion: Personalization



Software engineer



Accountant

A screenshot of a LinkedIn search results page. The search bar at the top contains the query "microsoft". Below the search bar, there are navigation links for Home, My Network, Jobs, and Messaging. A "People also searched for" section is displayed, containing six suggestions arranged in two rows of three: "microsoft recruiter", "microsoft marketing", "microsoft azure", "microsoft director", "microsoft sales", and "microsoft intern".

in

Q microsoft

Home My Network Jobs Messaging

People also searched for

microsoft recruiter microsoft marketing microsoft azure

microsoft director microsoft sales microsoft intern

A screenshot of a LinkedIn search results page, identical in layout to the first one. The search bar contains "microsoft". The "People also searched for" section shows the same six suggestions: "microsoft recruiter", "microsoft marketing", "microsoft azure", "microsoft director", "microsoft sales", and "microsoft intern".

in

Q microsoft

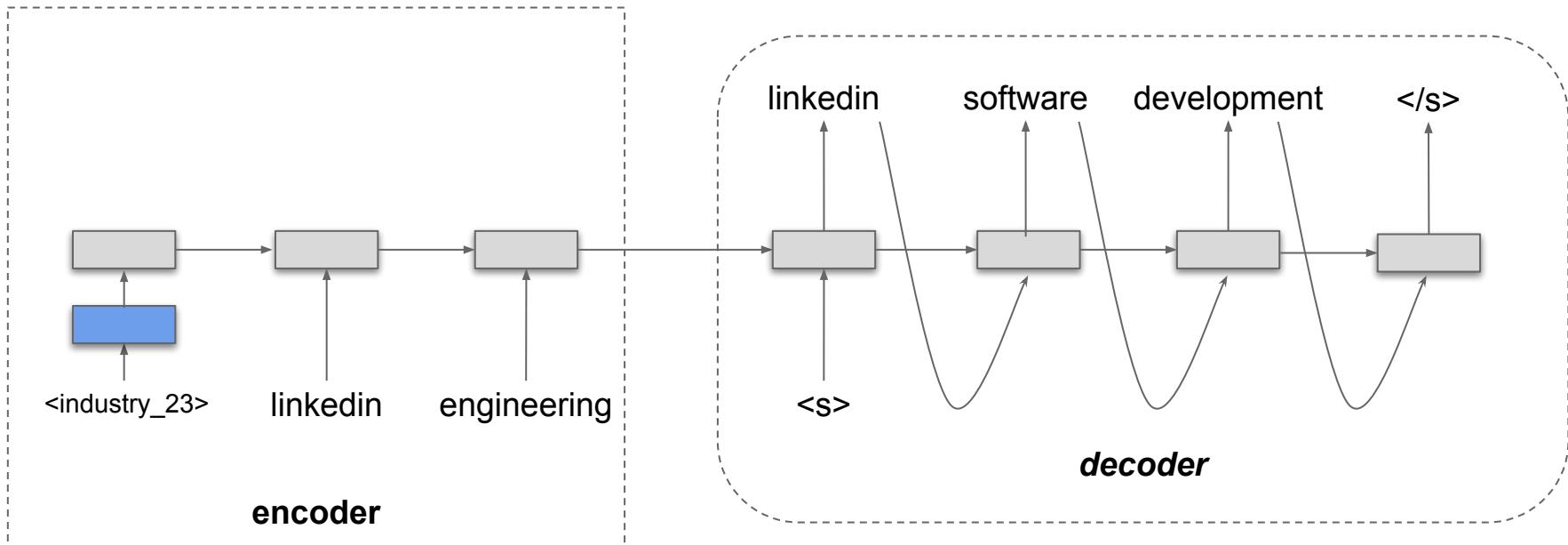
Home My Network Jobs Messaging

People also searched for

microsoft recruiter microsoft marketing microsoft azure

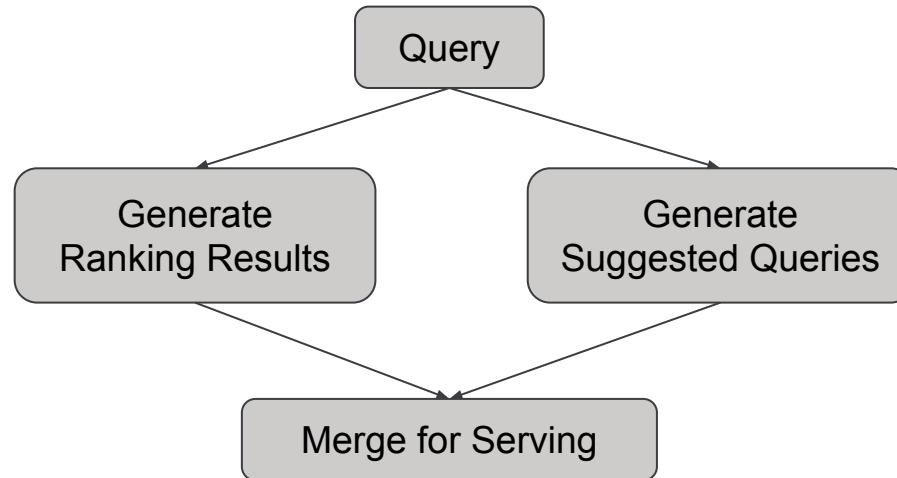
microsoft director microsoft sales microsoft intern

Query Suggestion: Personalization



Query Suggestion: How to Handle Online Latency

- Latency is strictly constrained for one query
 - Make it parallel with search ranking

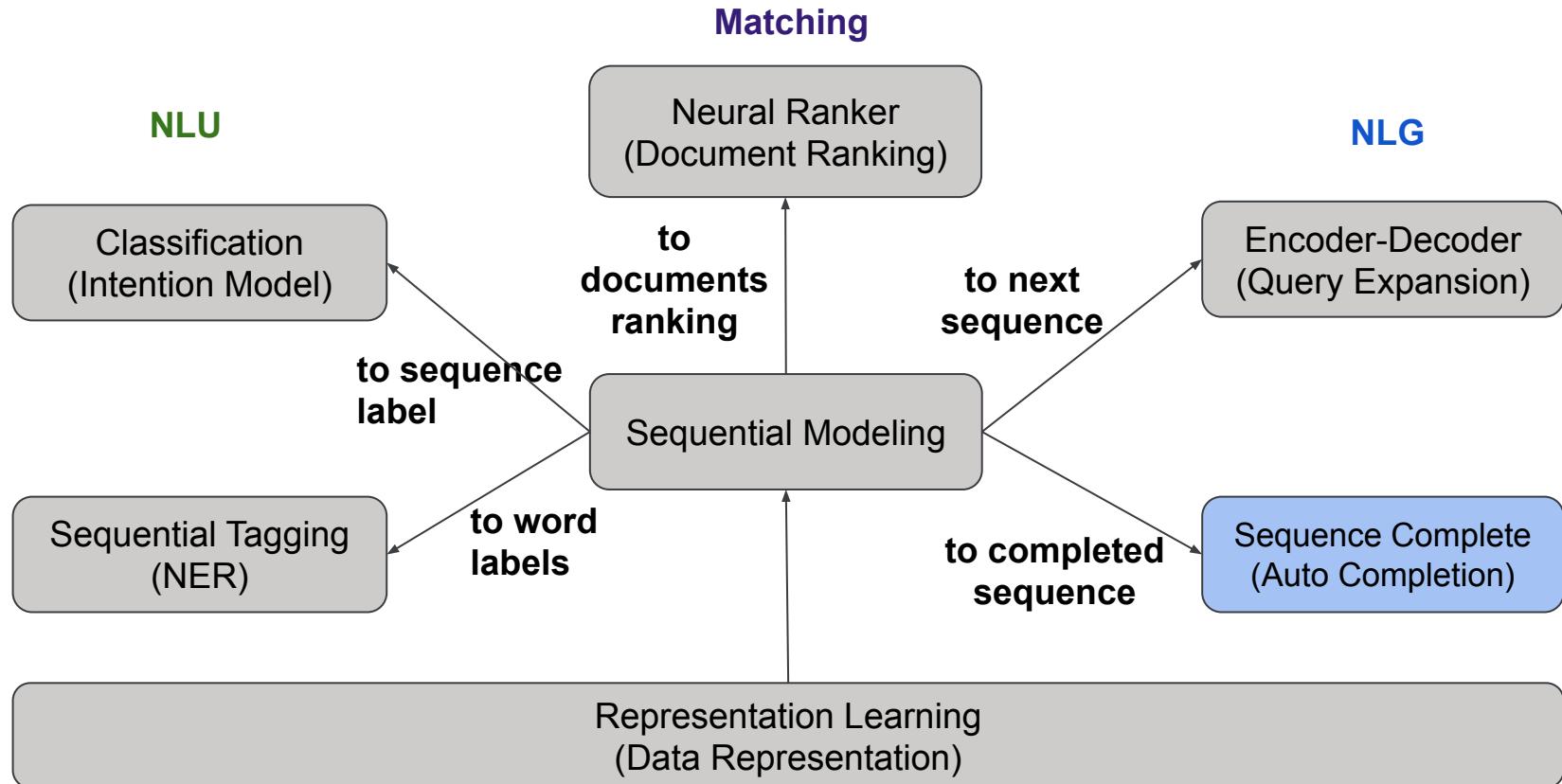


Online Performance

English Market

- No Personalization
 - Coverage: +80% Impressions, +26% CTR
 - +1% Total job application
- With Personalization
 - +1.18% Session Success for passive job seekers

Deep Learning for Natural Language Processing



Natural Language Generation: Auto-Completion

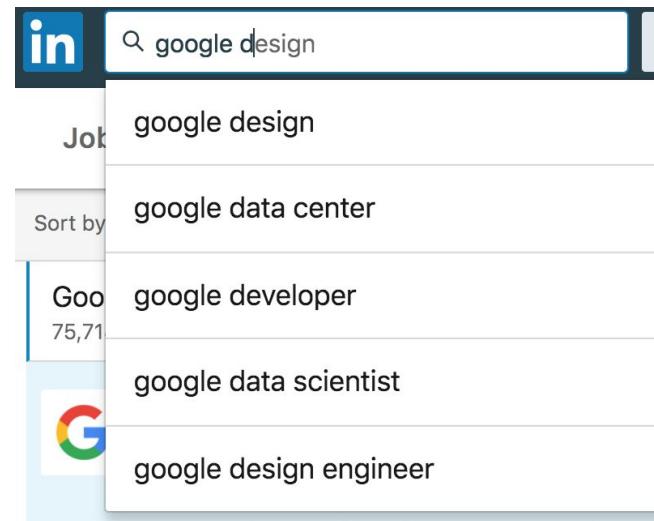
softw

software engineer salary
software engineer
software
software engineer jobs
software developer

- Given a prefix, predict the **completed query**, rather than the **completed word**

Auto-completion Challenges

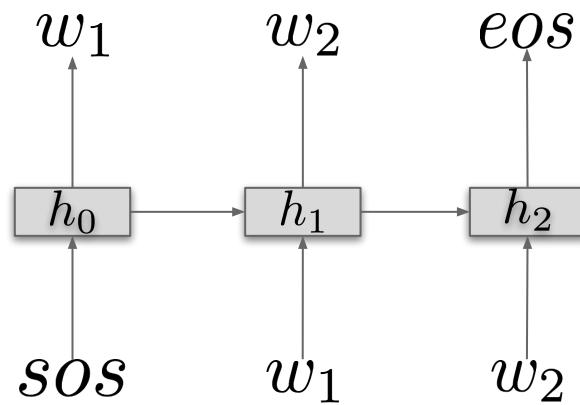
- **Short Context**
 - How to enrich semantic features
- **Personalization**
- **Latency Restriction**
 - Have to adopt simple and fast models



A Two-step Approach: Generation and Ranking

- **Candidate Generation**
 - Collect query frequency from search log
- **Candidate Ranking**
 - Neural Language Model serves as a scoring function

Auto-Completion: Neural Language Model as Scoring/Ranking



$$s(q) = \sum_i \log P(w_{i+1}|h_i)$$

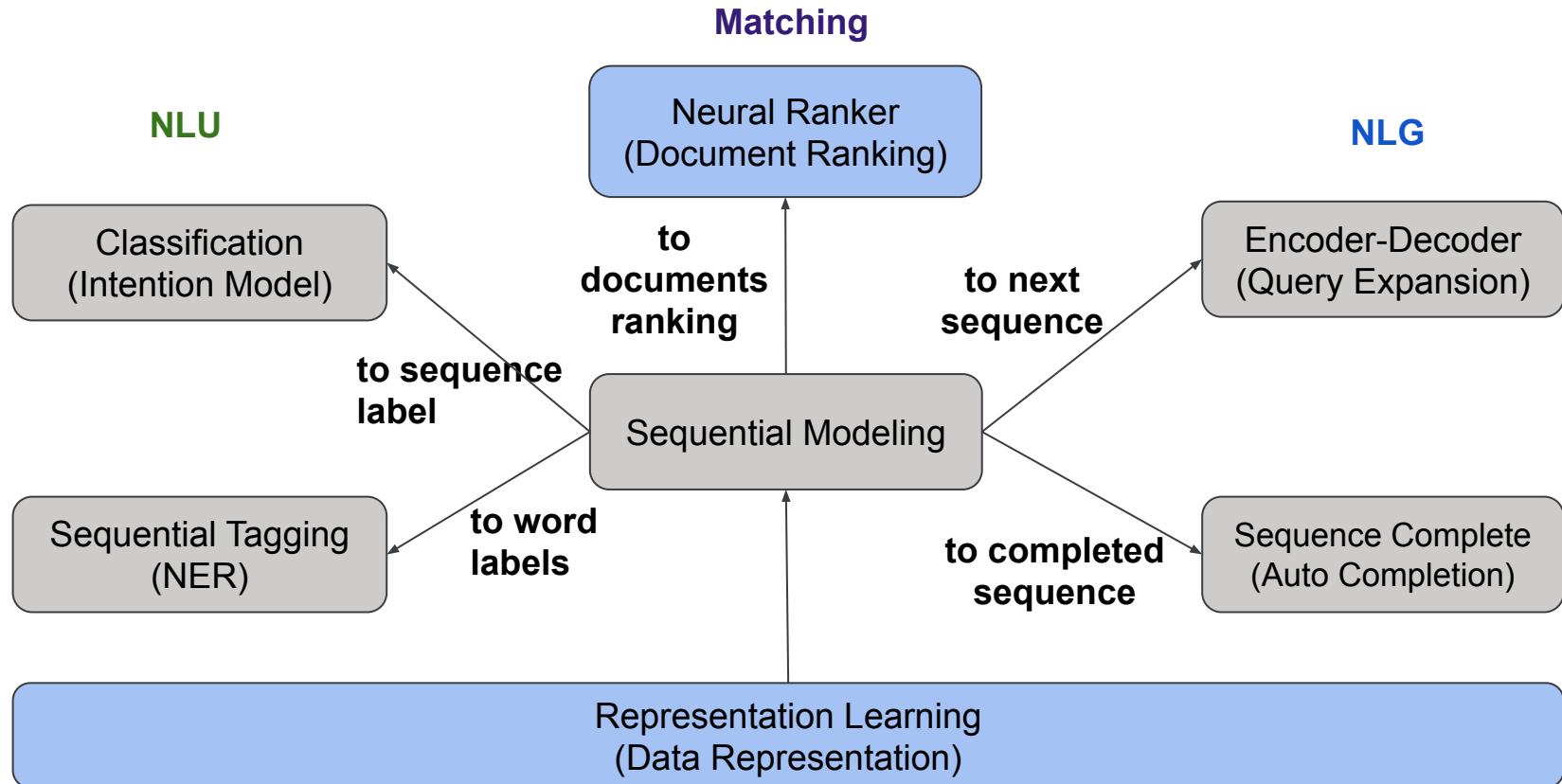
- Achieved 6x speedup with optimization

Auto-Completion Online Experiments

Neural Auto-completion vs Frequency Based Methods

- People Search English Market
 - +88.75% autocomplete clicks
 - +0.81% CTR@1 in SERP
- Job Search German Market
 - +3.24% Results clicked
 - -6.24% Raw searches w/o results
 - +3.24% Entity view

Deep Learning for Natural Language Processing



Natural Language Processing: Document Ranking

Job Search

LinkedIn search results for "machine learning engineer" in United States:

- Sr. Machine Learning Engineer - Home Timeline** (Promoted)
Twitter · San Francisco, CA, US
\$144K – \$314K
3 days ago · 2 applicants
- Machine Learning Field Engineer** (Promoted)
Cloudera · Palo Alto, CA, US
5 days ago
- Machine Learning Engineer**
Apple · Cupertino, CA, US
\$147K – \$400K
1 day ago · 7 applicants
- Senior Machine Learning Device Software Engineer**
Amazon Web Services (AWS) · East Palo Alto, CA, US
5 days ago · 2 applicants
- Machine Learning Engineer**
PayPal · San Jose, CA, US

Job description (from Twitter timeline):

Twitter's Consumer Product Teams are responsible for core features of twitter.com, which includes Timelines, Tweets, Search, Trends, Recommendations, Notifications, Tweet details/permalink, and more! Our code operates at massive scale and speed, serving billions of requests per day, connecting hundreds of millions of active Twitter users to real-time information about their lives and the world we live in.

Who We Are:

At Twitter, our mission is to instantly connect users to the information most meaningful to them. Realizing this involves work in areas such as machine learning, applied data science, recommendation systems, information retrieval systems, natural language

Ads Recommendation

LinkedIn feed recommendations:

- Xi Fang, Jiayu Zhou and 408 others follow Microsoft**
Microsoft · 7954,475 followers
Promoted
Attention IT pros, business and operations managers, and security analysts — it's time to experience how the cloud is transforming the way we work.
- Microsoft 365 Training Day: Desktop Deployment**
Sunnyvale, CA · August 21, 2019
Join us in Sunnyvale for a free technical training day
microsoftevents.com
- Huiji, picture yourself at Google**
Technical Program Manager, G... · Palo Alto, CA, US
Technical Program Manager, B... · Sunnyvale, CA, US
Technical Program Manager, C... · New York City, NY, US

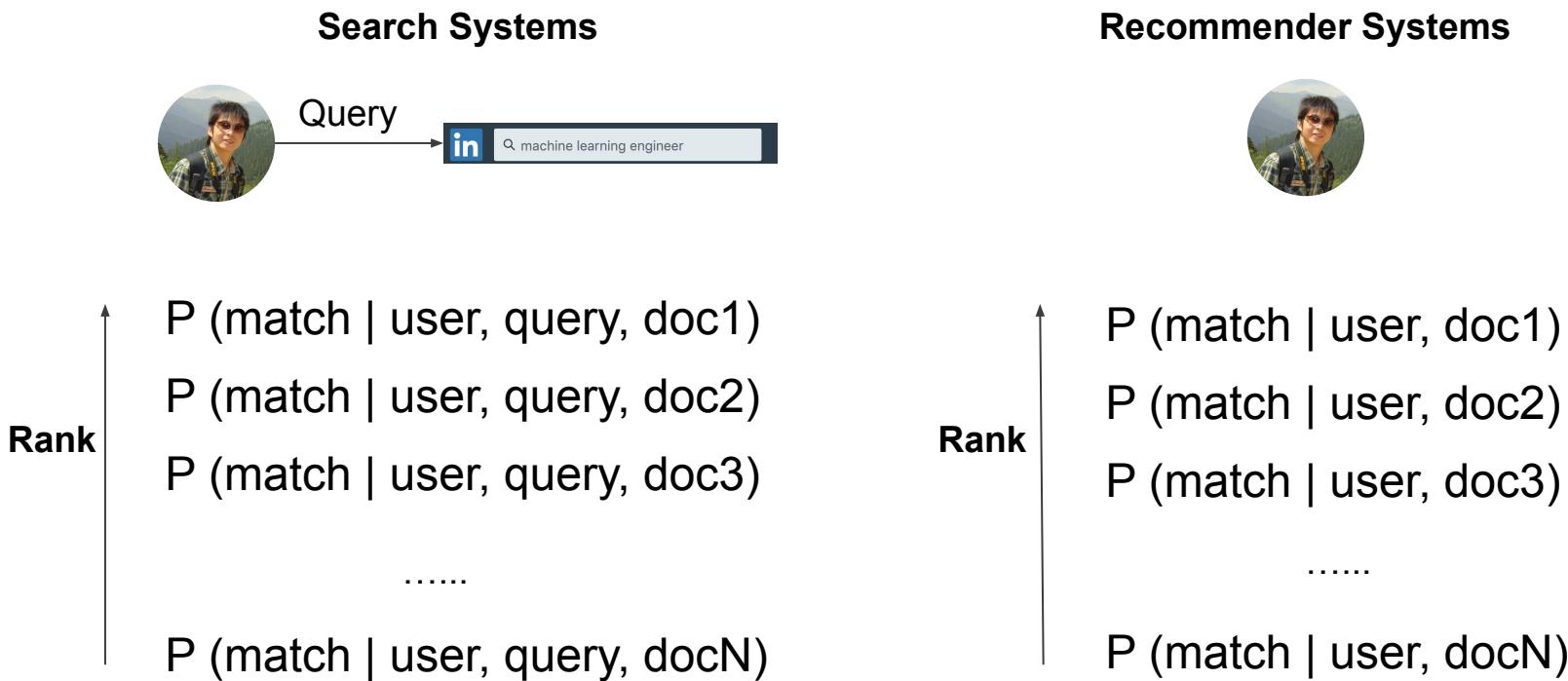
Learning section:

- Similar to videos you've watched**
From Project Management Foundations: Teams · 4:21
Making team changes to benefit the group
- 5 coworkers like this**

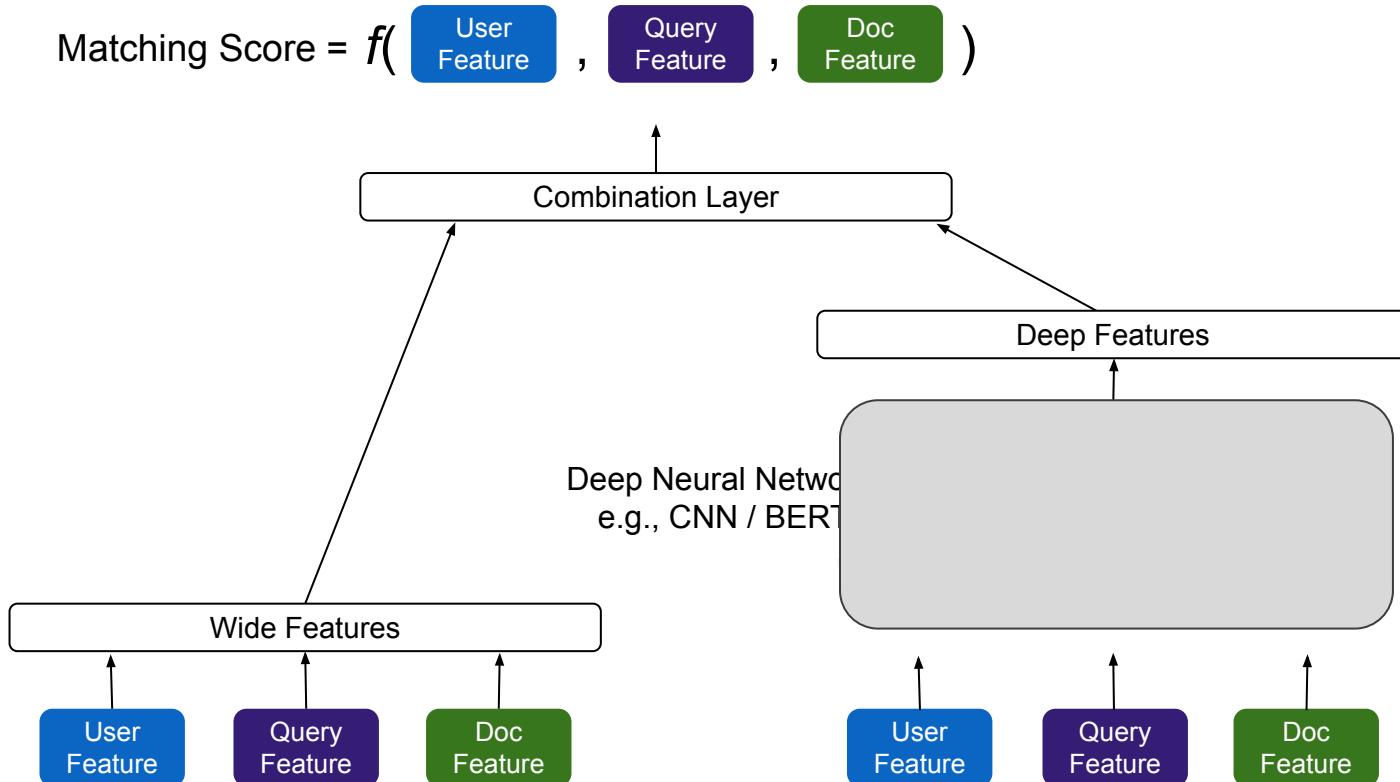
Bottom navigation:

- About · Help Center · Privacy & Terms · Advertising · Business Services · Get the LinkedIn app · More
- LinkedIn · LinkedIn Corporation © 2019

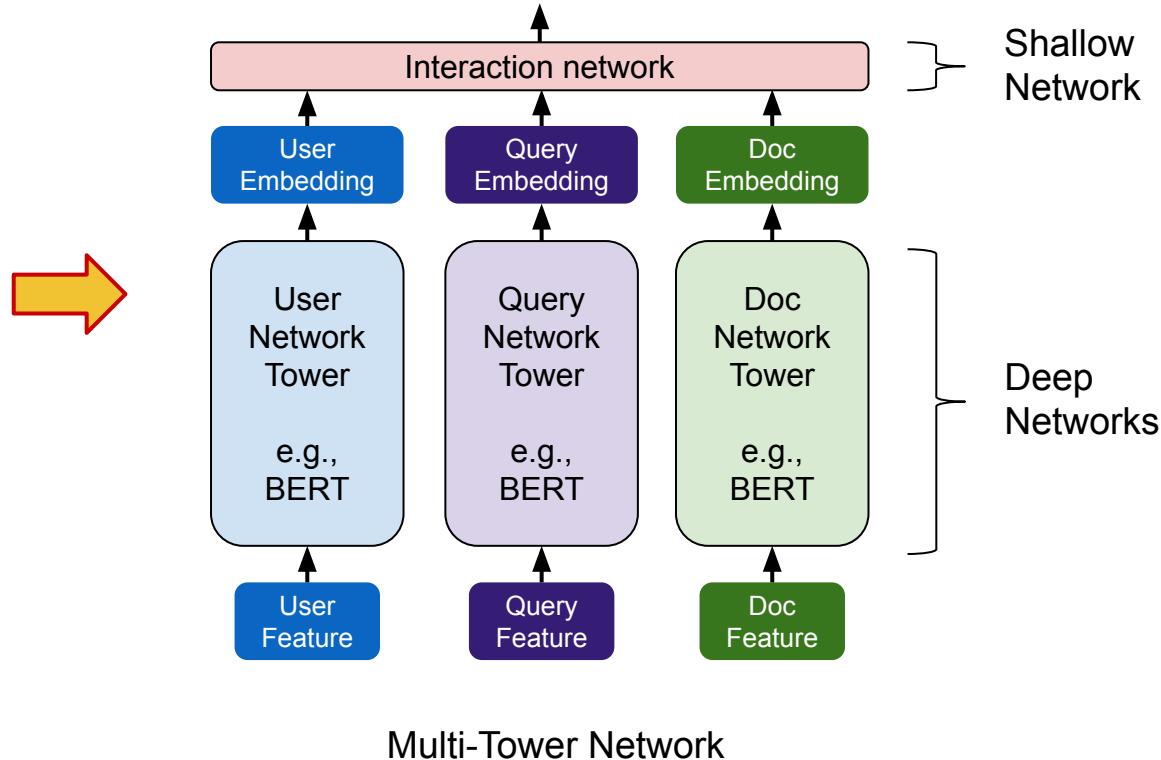
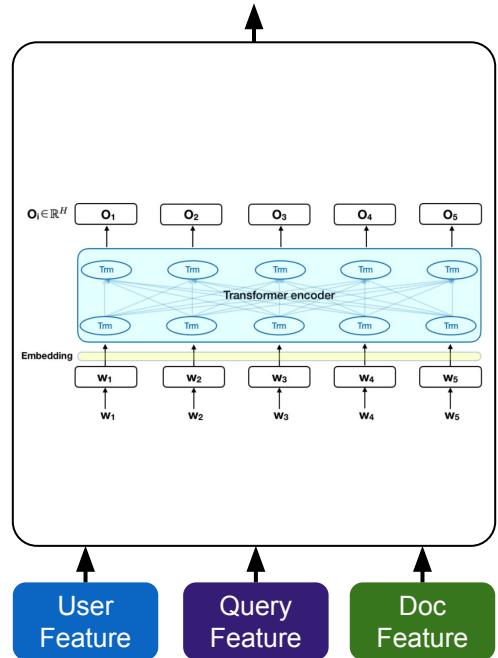
Natural Language Processing: Document Ranking



Natural Language Processing: Document Ranking



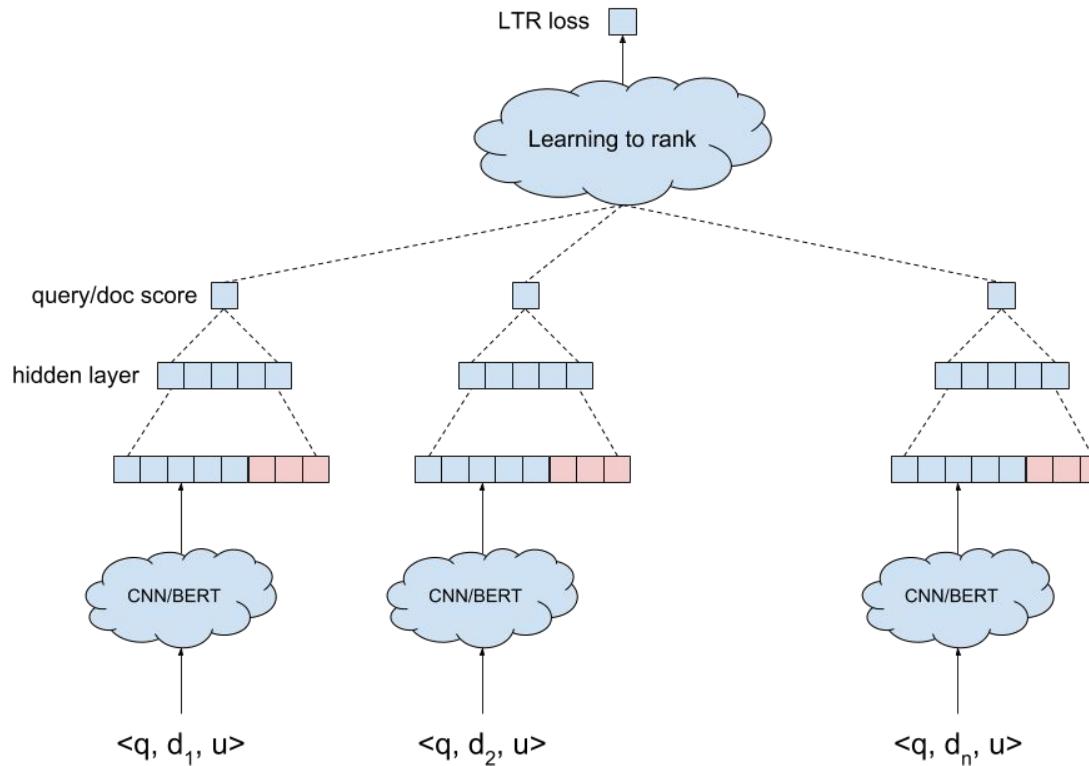
Natural Language Processing: Document Ranking



One Big Network

Multi-Tower Network

Neural Ranking



Experiments

- Offline

People Search (NDCG@10)		Job Search (NDCG@10)	Ads Click-through Rate (AUC)	Ads Conversion (AUC)
Wide Features	-	-	-	-
CNN	+1.32%	+2.36%	+3.43%	+1.37%
LinkedIn BERT	+1.98%	+5.14%	-	-

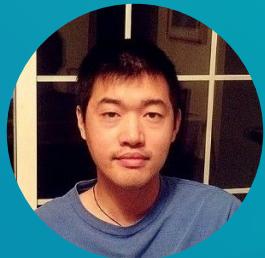
- Online (CNN)

- Job Search

- Job apply clicks: +1.62%
 - Onsite paid applicants: +1.38%

- People Search

- Satisfied clicks: +0.78%
 - Successful sessions: +0.37%



Deep NLP in Search and Recommender Systems - Question Answering at LinkedIn

Liang Zhang

Traditional Web Search



question answering deep learning



All Images Videos Maps News Shop I My saves

9,230,000 Results Any time ▾

[PDF] [Question Answering Using Deep Learning - ...](#)

<https://cs224d.stanford.edu/reports/StrohMathur.pdf>

[Question Answering Using Deep Learning](#) Eylon Stroh SCPD Student maestroh@stanford.edu Priyank Mathur SCPD Student priyankm@stanford.edu Abstract With advances in **deep learning**, neural network variants are becoming the dom-

State of the art deep learning model for question answering

<https://einstein.ai/research/state-of-the-art-deep-learning-model...> ▾

The cloze-form **question answering** task is not as natural as open-domain **question answering**, but the ease with which cloze-form datasets can be created has led to dramatic progress in the development of expressive models such as **deep** neural networks for **question answering**.

Deep Learning for Visual Question Answering - ...

<https://www.kdnuggets.com/2015/11/deep-learning-visual-question...> ▾

An year or so ago, a chatbot named Eugene Goostman made it to the mainstream news, after having been reported as the first computer program to have passed the famed Turing Test in an event organized at the University of Reading. While the organizers hailed it as a historical achievement most of the

- User Input: Keyword query
- Output: A list of documents
- A classic way of obtaining information from machines

What can be done better:

- *Understand user intent and semantic meaning of the query*
- *Directly show the answer*
- *Interact with the users when the results do not satisfy the user's need*

Modern Web Search with Question Answering

The screenshot shows a search interface with the following elements:

- Search Bar:** Displays the query "what is question answering".
- Search Buttons:** Includes a blue magnifying glass icon and a blue 'Share' icon.
- Filter Buttons:** Labeled "All", "Images", "Videos", "Maps", "News", "Shop", and "My saves".
- Search Results Summary:** Shows "221,000,000 Results" and a date range "Any time".
- Result Preview:** A box for "Question answering" with a snippet: "Question answering (QA) is a computer science discipline within the fields of information retrieval and natural language processing (NLP), which is concerned with building systems that automatically answer questions posed by humans in a natural language." It includes a "Share" button.
- Result Link:** A link to "Question answering - Wikipedia" with the URL "https://en.wikipedia.org/wiki/Question_answering".
- More Information:** A button labeled "See more about Question answering" with a dropdown arrow icon.
- Page Navigation:** A horizontal navigation bar with tabs: <, Overview (highlighted in red), Contents, History, Architecture, Question ↗, and >.

Question answering is a computer science discipline within the fields of information

Modern Web Search with Question Answering

Q who is the ceo of LinkedIn

Home My Network Jobs Messaging Notifications Me Work Recruiter

People Jobs Content More ▾ People filters Connections Locations Current companies All Filters

Python Developer Jobs - Let companies apply to you. No resume needed. Salaries: £35K-£100K Ad ...

Showing 36,690 results

Jeff Weiner in CEO at LinkedIn San Francisco Bay Area
Current: Member, Board of Directors at Intuit
93 shared connections

Brian Rumao • 2nd in Chief of Staff to the CEO at LinkedIn San Francisco Bay Area
Current: Sr. Director, Chief of Staff to the CEO at LinkedIn
91 shared connections

Follow

Knowledge Card

Jeff Weiner
CEO at LinkedIn
San Francisco Bay Area
Internet executive with over 20 years of experience, including general management of mid to large si... See more

People Also Viewed

Kevin Systrom
CEO, Instagram

Jess Whitsen

Job results for who is the ceo of LinkedIn 763 results See all

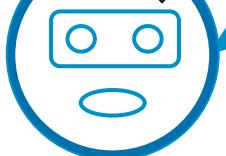
How Question Answering Can Help LinkedIn

Recruiter Assistant

How should I hire
for this AI engineer
position?



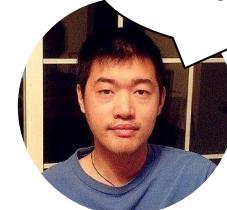
These are good
candidates ...



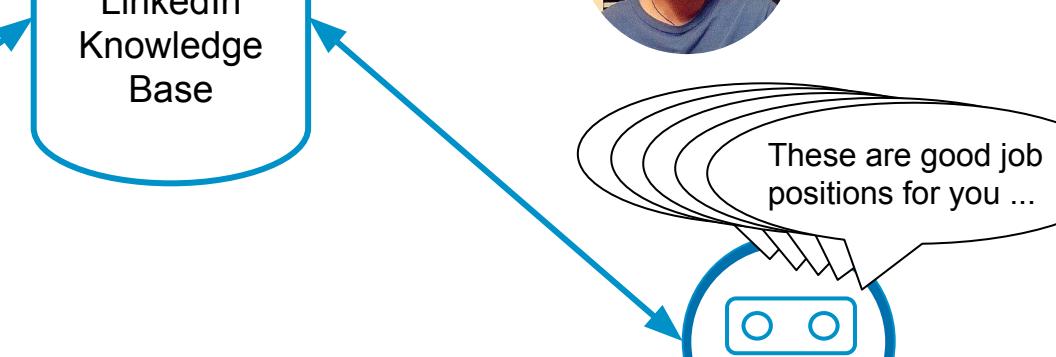
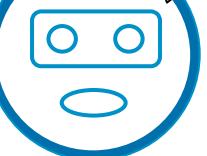
LinkedIn
Knowledge
Base

Job Seeker Assistant

Find good jobs
for me.



These are good job
positions for you ...



LinkedIn Help Center



Searching for Jobs on the Job Search Mobile App ARTICLE

The LinkedIn Job Search mobile app gives you the ability to search for jobs based on job title, related keywords, company, and location. iOS To search for a job on the Job Search mobile app: Tap the Search icon at the...

Searching for Jobs on LinkedIn ARTICLE

You can search and apply for job opportunities based on keyword, title, or location. You can also save your search and access it later. To search for a job: Click the Jobs icon at the top of your LinkedIn homepage. In the search box, type a...

- Input: Question in Natural Language
- Traditional keyword search does not work well
- Our solution:
 - Question-answering through Search
 - Chatbot (Demo)

Question Answering Through Search

LinkedIn Help Center Search

How do I keep my job search a secret?

x



Searching for Jobs on the Job Search Mobile App

ARTICLE

The LinkedIn Job Search mobile app gives you the ability to search for jobs based on job title, related keywords, company, and location. iOS To search for a job on the Job Search mobile app: Tap the Search icon at the...

Searching for Jobs on LinkedIn

ARTICLE

You can search and apply for job opportunities based on keyword, title, or location. You can also save your search and access it later. To search for a job: Click the Jobs icon at the top of your LinkedIn homepage. In the search box, type a...

- Input: Question as a Query
- Output: Article / Section of the Article that contains the answer
 - Few thousands candidates
 - New candidates get added regularly

Our Data

- A random sample of user click-through data in 01/2017 - 04/2018
- 200K \langle query, list of articles \rangle samples, click / non-click as responses for each \langle query, article \rangle
- Only contains queries with > 4 words (*where traditional keyword search generally fails!*)

Our Modeling Approach

 x 🔍

Privately Looking for a Job ARTICLE

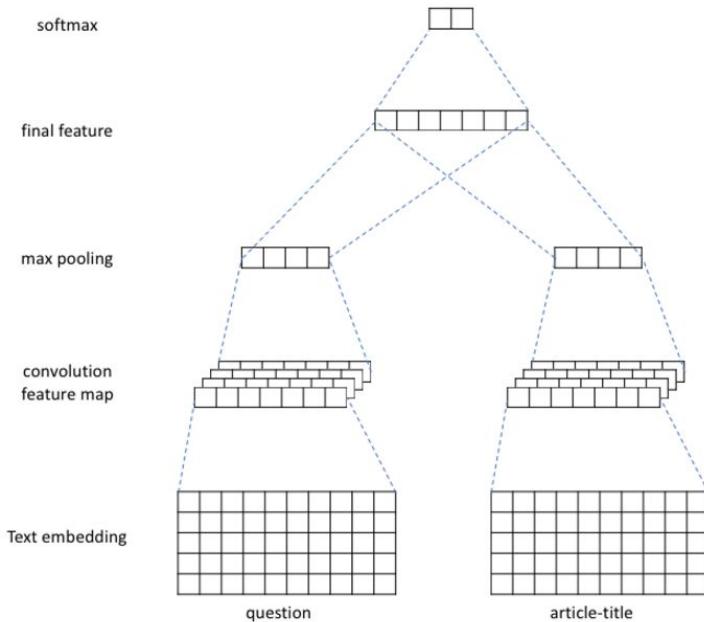
If you're embarking on a job search, consider these tips: Just because you use LinkedIn doesn't mean you're looking for a job. Many people use LinkedIn to keep in contact with others and to help them succeed in their current position. No...

Searching for Jobs on LinkedIn ARTICLE

You can search and apply for job opportunities based on keyword, title, or location. You can also save your search and access it later. To search for a job: Click the Jobs icon at the top of your LinkedIn homepage. In the search box, type a...

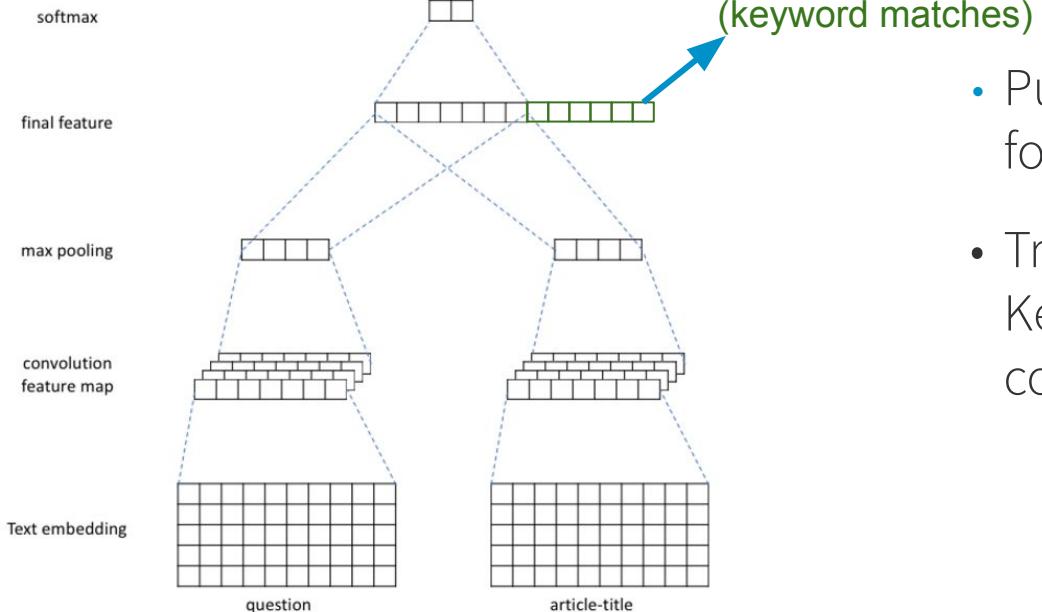
- Consider each candidate article as an intent
- Problem: Natural language Query => Intent
- Need to handle regularly added new articles without too-frequently retraining the model

Our Modeling Approach -- CNN



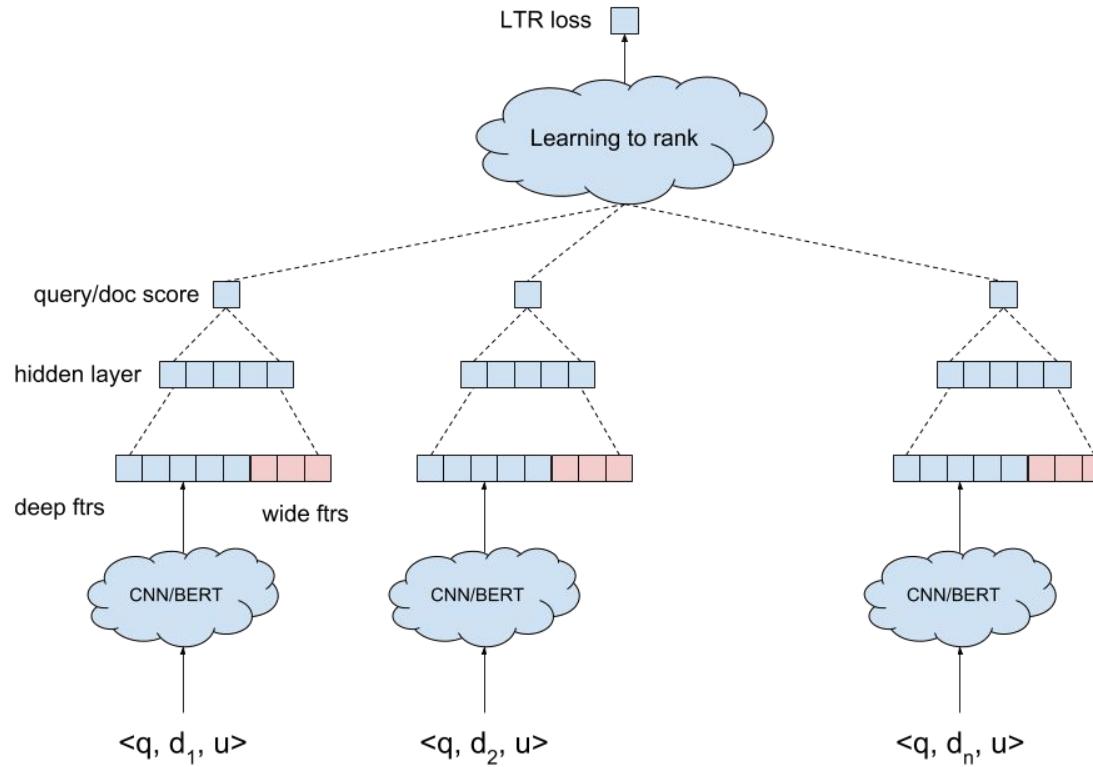
- Independent CNN layers for Question & article title, keywords, content
 - Word2Vec Embedding to initialize
- Final layers: $\text{Sim} < \text{Question Embedding}, \text{Article Embedding} >$ \Rightarrow Softmax over all candidate questions
- For new articles in test data: Use embedding similarity instead of retraining the model

Deep + Wide Model with Better Performance



- Pure CNN Model did not work well for out-of-vocabulary queries
- Traditional IR-based Keyword-match features added to complement (wide component)

Our Modeling Approach -- BERT + Wide



Offline Evaluation

Model	Lift % in Precision@1 vs Control (IR-based)
Deep Model (CNN)	+26%
CNN + Wide Model	+32%
BERT + Wide Model	+36%

Online A/B Test Performance

Metric	CNN vs Control	BERT vs CNN
Happy Path Rate	+15%	+11.3%
Clicked Session	+5.7%	+5.2%

- Happy Path: Users who clicked only one search result and then left help center without creating a case

More Examples...



how to download my resume



Saving a Profile in a PDF Format

ARTICLE

You can save a copy of your profile or someone else's profile in a PDF format from the introduction card on your profile. Note: We currently don't offer the option to choose which sections will be included in saved versions of your profile. To...

Uploading Your Resume when Applying for Job on LinkedIn

ARTICLE

If you're applying for a job through LinkedIn and not through a company's web site, you can attach your resume to the application in addition to applying with your LinkedIn profile. To upload a new resume: Search for a job. Click on a job...

Using LinkedIn Resume Builder (currently available in India only)

ARTICLE

Important: Resume Builder is currently available to only LinkedIn members in India who've subscribed to Premium Career. Please note that this new feature will be gradually rolled out to our members and may not be available to you at this time...

Troubleshooting Learning Video Downloads and Mobile Viewing

ARTICLE

If the video you're streaming is pausing, or you're unable to download a video on your mobile device, you may be experiencing a problem with your network connection. While you can stream or



I am a recruiter, how to find good candidates?



Finding Additional Job Candidates

ARTICLE

You can find additional job candidates using various features of LinkedIn. Starting from basic search to Recommended Matches and LinkedIn Recruiter Lite, LinkedIn offers many ways to help you find the best candidates for your open job...

Differences Between Recruiter, Recruiter Premium Subscription (RPS), and Recruiter Lite

ARTICLE

While Recruiter Lite lets you search for, filter, and contact LinkedIn members who may be good candidates for your role and comes with InMail, a Recruiter account gives you unrivaled access to the entire LinkedIn network. Designed specifically for...

Best Practices for Posting Jobs

ARTICLE

Recruiters can find the best candidates for the open job role faster and easier with effective use of all the LinkedIn features. We recommend the below best practices to enhance your job posting experience: Creating and maintaining the...

Targeting Options and Best Practices for LinkedIn Advertisements

ARTICLE

When advertising on LinkedIn, you can get your message in front of the right people when they are most engaged by using targeting. You can deliver your content to those who matter most to your

Interactive Search (Demo)

account



Let's narrow down your search:

[Restrict account](#)

[Merge account](#)

[Close account](#)

Change your account and privacy settings

The Privacy & Settings page allows you to manage your account settings, update your privacy and security settings, and set your preferences for how frequently you're contacted by and through LinkedIn.

[View and change your settings](#)

ChatBot for LinkedIn Help Center



Problems to Solve

- Out-of-scope questions
 - “How are you?”
 - “Will you marry me?”
- Ambiguity
 - Incomplete questions (e.g. verb / object missing)
 - e.g. “premium” => “cancel premium”? “be premium”?
 - Question is complete (verb + object) but ambiguous
 - e.g. “Cancel account” => “cancel premium subscription”? “closing LinkedIn account”?
- Actions to take instead of just article link



Editing Your Profile ARTICLE

You can edit sections of your LinkedIn profile to best reflect your own, personal brand. Use the introduction card on your profile to display details of your current personal and professional status, just like a business card. You can manage...

To edit sections on your profile:

1. C ▾ iOS

2. C ▾ To ▾ Android

3. S

To edit sections on your profile:

4. C

1. Tap your profile picture.

5. M

2. Tap the Edit icon in the top right of the section you wish to edit.

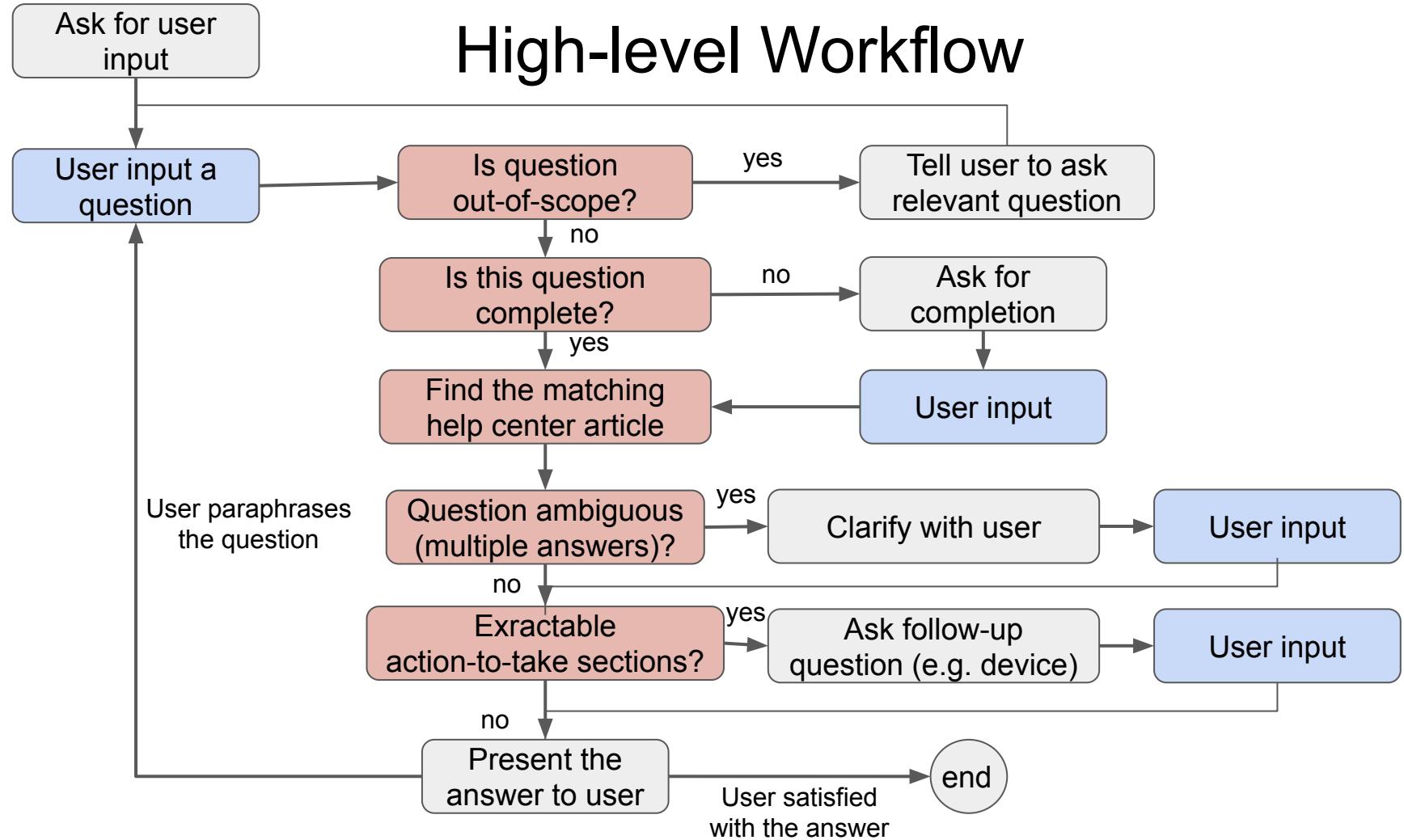
6. C

3. Tap the **Add button** to add new items to that section or tap the Edit icon to the right of that entry.

4. Make changes in the fields provided.

5. Tap **Save**.

High-level Workflow



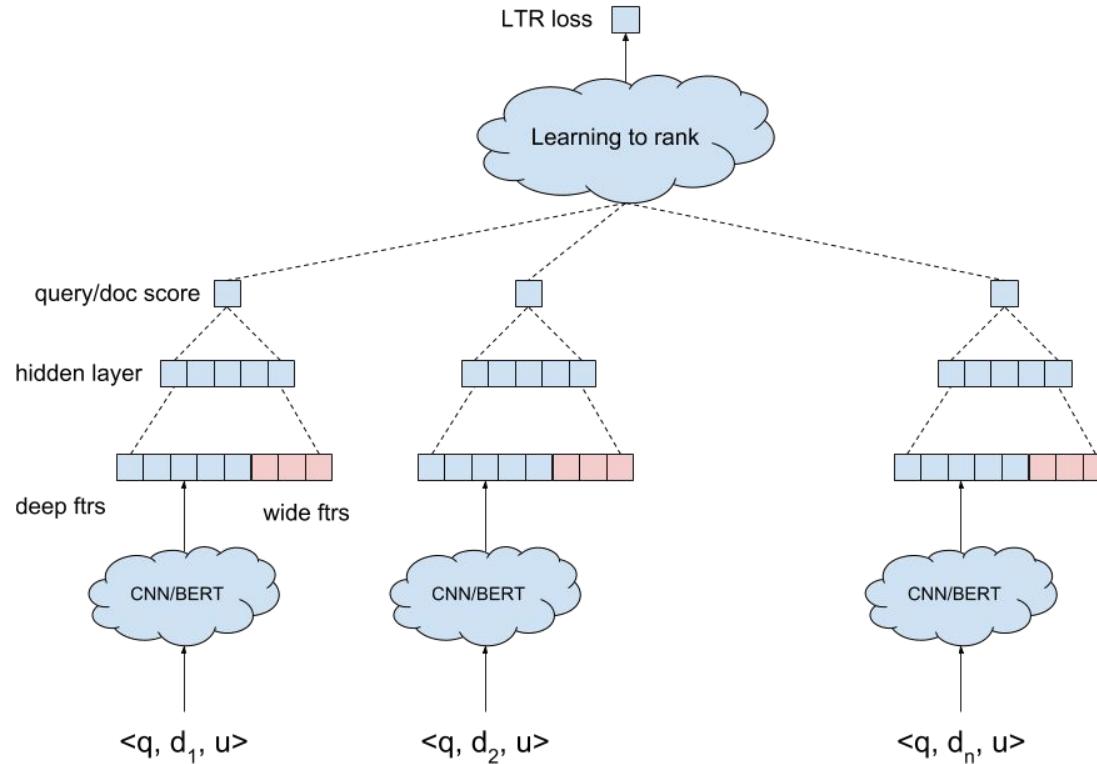
Out-of-Scope Detection Model

- “*Who is abraham lincoln?*” → Not help center query
- Current: A binary classification approach
 - Positive: help center queries
 - Negative: LinkedIn search queries, movie dialogue, quora questions
- Future:
 - Being able to respond to other intents, e.g.
 - Greetings
 - Ask for speaking to an agent
 - Looking up personal account info (has my premium account expired?)

Incomplete Question Handling

- Incomplete Questions:
 - Object (noun) missing
 - e.g. “cancel” -> what to cancel?
 - Verb missing
 - e.g. “Profile” → what to do with profile?
- We assume the question must contain nouns (objects) and verbs
 - If not, explicitly ask for it
- Standard POS Tagging model from openNLP
- If incomplete: Suggest most related & frequent complete queries from query logs
- Future work
 - Use Entity Tagging in help center context to detect question completeness

Question-Answering Model (BERT + Wide)



Actions-to-take Section Extraction & Ambiguity Handling

- Ambiguity Handling
 - Clarify with user if scores among top 3 articles are too close
- Extract Action-to-take Sections from Top-ranked Article
 - Pattern mining (e.g. ios / android section heads) + Human labeling
 - Follow up with user to pick the right section with actions to take (e.g. desktop / ios / android)

Demo



Conclusion

Lessons & Future Trends

- **Ground Truth Data**
 - Human Labelled Data (Crowdsourcing)
 - User Behavior Data
 - Automatic Data Generation with Generative Models (e.g., GANs)
- **Efficient Modeling Training and Serving**
 - Pre-trained Model for Multiple Products
 - Automatic Hyperparameter Tuning & Structure Learning
 - Memory and Latency Optimization for Efficient Online Serving

Lessons & Future Trends (Cont'd)

- **Internationalization**
 - Transfer Learning
 - Multilingual Learning
- **Reinforcement Learning**
 - Long-term reward optimization in near-real-time user feedback loop
 - Multi-turn question answering
- **Model Debugging**
 - Model Interpretation
 - Model Fairness & Privacy

Thank You!

