Using Historical Context to improve Dialog State Tracking

By Sanuj Sharma
With Prafulla Choubey and Prof Ruihong Huang

Dialog Turn

- User Utterance: **I'm looking for a Chinese restaurant in east part of town.**
- System Utterance: **What's your budget?**
- System Actions: **price range**
- Dialog State
  - **inform(food = Chinese)**
  - **inform(area = east)**

Dialog State

- **Requests:** information requested by user.
  - request = address, request = phone
- **Joint Goals:** the set of accumulated turn goals.
  - food = french, price range = cheap

- **REQUEST: price range**
- **INFORM: food = chinese, area = east**

WoZ 2.0 Dataset

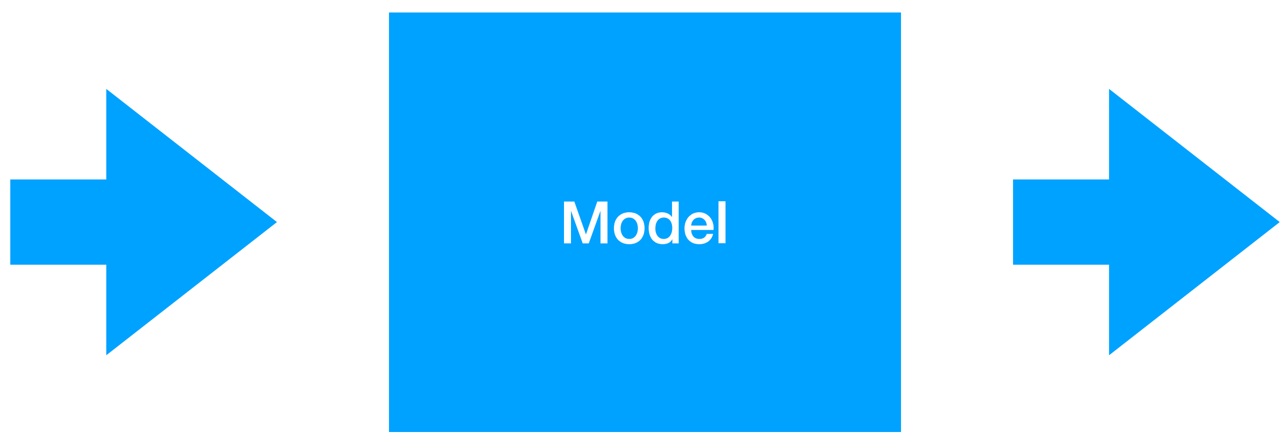
1400 dialogs, 10-15 turns / dialog

- **Area:** north, south, east, west, central.
- **Price Range:** cheap, expensive, moderate, don’t care.
- **Food:** French, Chinese, Mexican, Indian.
- **Request:** address, phone, pinched, area, price.

Binary Classification

- For **each slot-value pair**:

- User Utterance
- Previous turn System Acts
- Slot-value

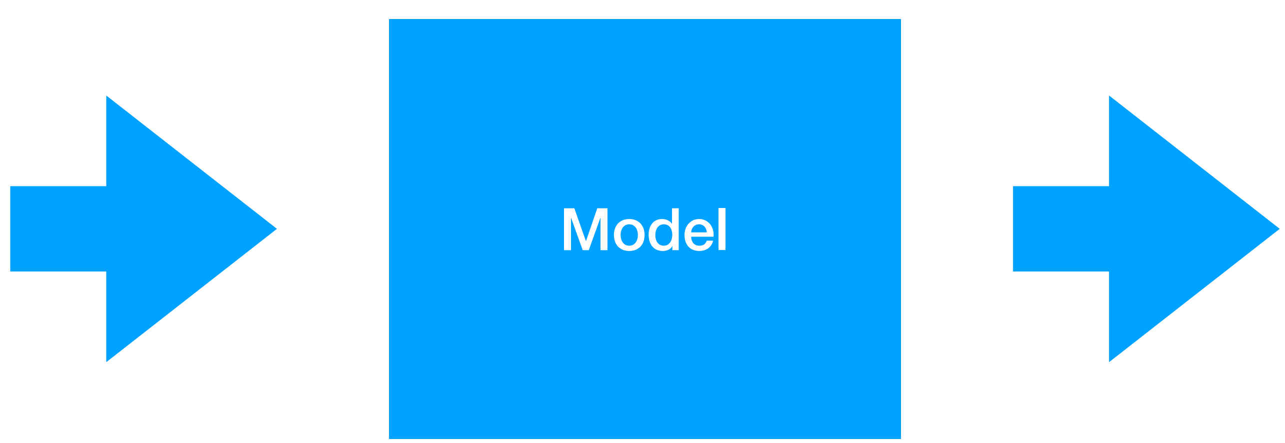


- Probability of the slot-value in dialog state

Binary Classification

- For **each slot-value pair**:

- I would like to eat Chinese food.
- food
- Food = Indian

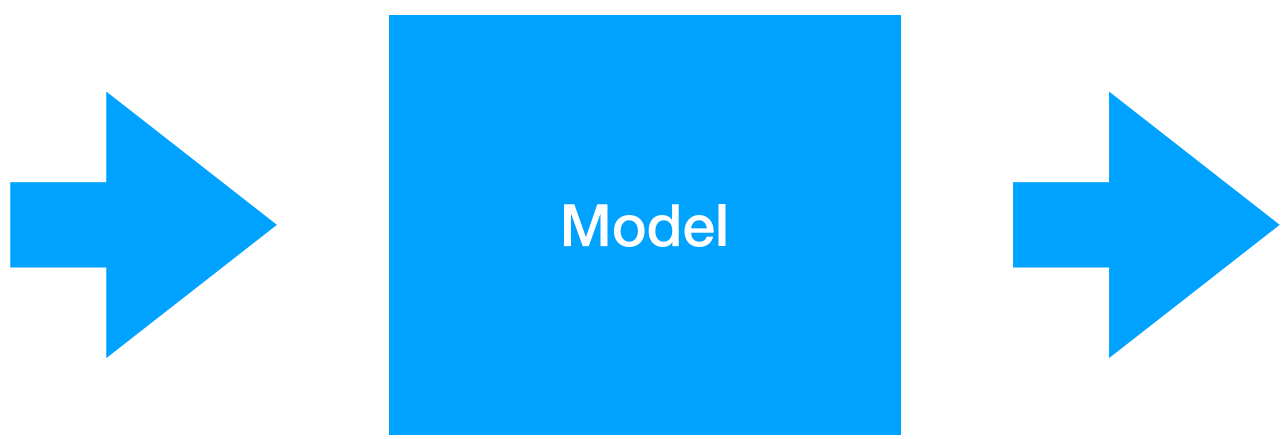


- > 0.5 then yes
- Else no

My Contribution

- For **each slot-value pair**:

- **User + previous system utterance**
- Previous turn System Acts
- Slot type-value



- Probability of the slot-value in dialog state

- **Historical Context: Previous utterances and slot value where current slot type was last modified.**

An Example

Turn 5 prediction for **slot type food** will use **sys utterance** from **turn 2** and **user utterance** from **turn 3**

Turn 1

```
hello, i'm looking for a restaurant with fair prices

Price range: moderate

There are 31 places with moderate price range.
Can you please tell me what kind of food you
would like?

Sys act: food
```

Turn 2

```
well I want to eat in the North, what's up that way?

Area: north

I have two options that fit that description, Golden
Wok chinese restaurant and The Taj which serves
Indian food. Do you have a preference?

Sys act: food
```

Turn 3

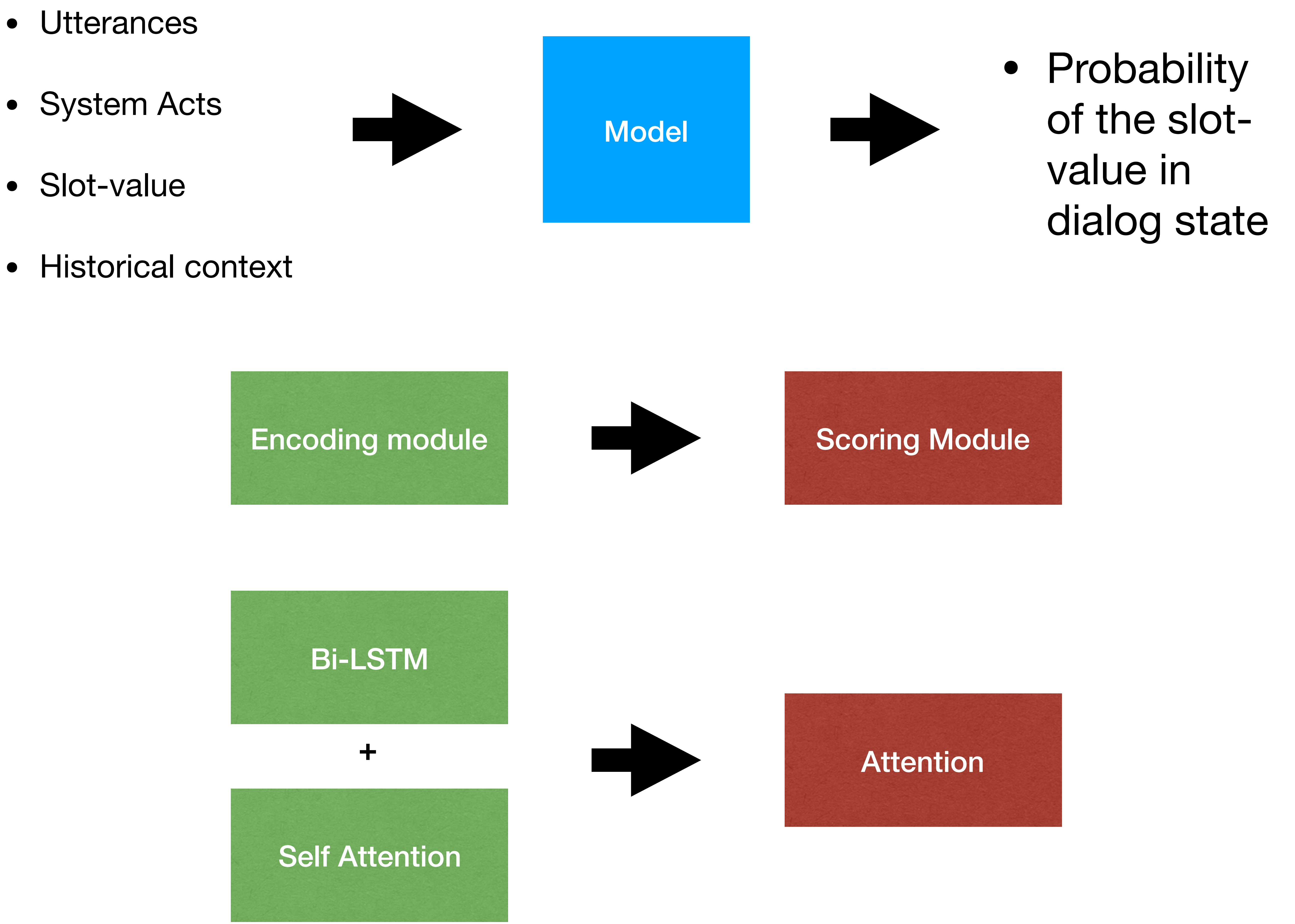
```
Can I have the address and phone number
for the Golden Wok chinese restaurant?

Request: address
Request: phone number
Food: Chinese

The phone number is 01223 350688.
```

Turn 4

```
thank you. what is the address?

Request: address

The address is 191 Histon Road Chesterton.
```

Turn 5

```
Okay, what about Taj, what's the address and
phone of that?

Request: address
Request: phone number
Food: Indian
```



Utterances
System Acts
Slot-value
Historical context
Model
Probability of the slot-value in dialog state
Encoding module
Scoring Module
Bi-LSTM
+
Self Attention
Attention

Results

Model	Joint Goal Accuracy	Turn Request Accuracy	Approximate # of parameters (in million)
Baseline	84.5%	95%	1.2
Baseline + Historical Context	**88.4%**	**96%**	**6**
GLAD	86.4%	97%	17
GLAD + Historical Context	**89%**	**97%**	**28**

Baseline: Bi-lstm + self-attention encoder with attention scorer.
GLAD: Global-locally self-attentive dialog state tracker

Thank you

Personality-Based Chatbot

Sameer Kumar Behera
Srishti Agarwal
Shubham Bhargava

About Chatbots

- **Rule-Based** : Answers using set of hand-crafted rules.
- **Retrieval** : Answer selected based on set of answers to the question.
- **Generative** : Generate proper responses, Seq2Seq, Encoder-Decoder, can generate new, complex responses.
- **Speaker-Addressee** : Predict how Speaker A would respond to message by Speaker B.

Personality Chatbots!

- Chatbots with Persona

- Persona - elements of identity (background facts or user profile), language behavior, and interaction style.

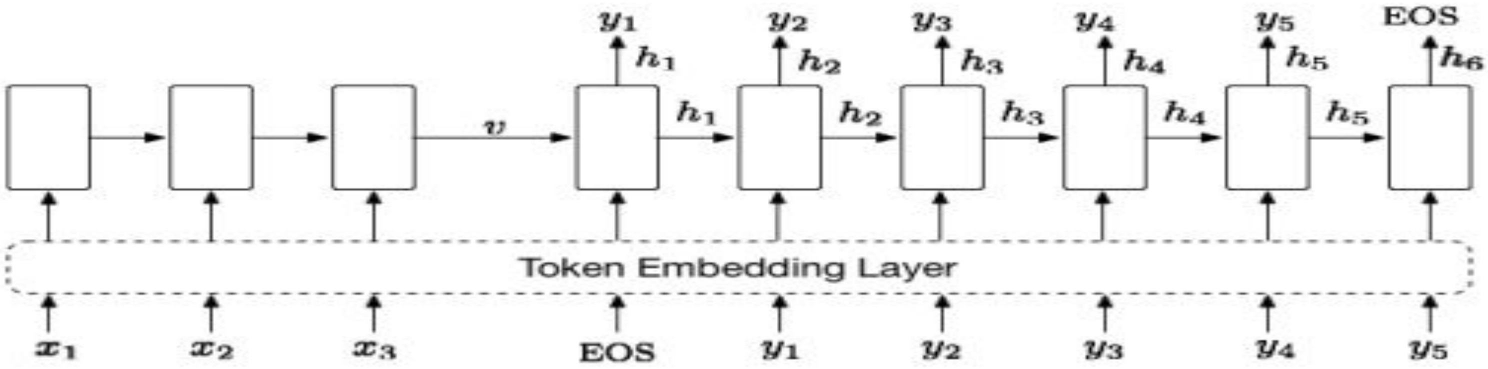
- Wide applications
  - Domain Specific Assistance like IT Helpdesk, Customer Care Representatives..
  - Entertainment

Dataset

- TV Series Corpus - Friends(Joey) , and Big Bang Theory (Sheldon) scripts
- Processed Scripts to have Q & A like format.
- Replaced the change in Scene by a separator as to differentiate the contexts.

Approach

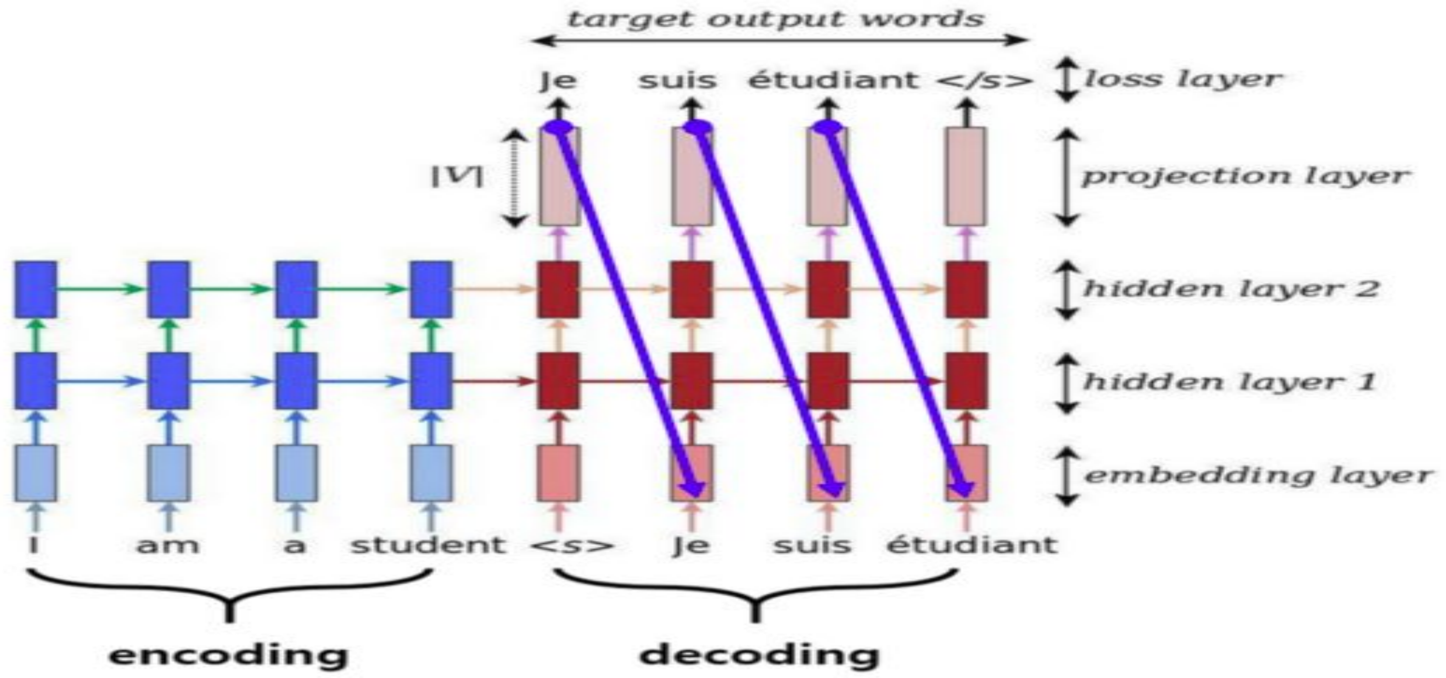
Sequence-to-sequence encoder-decoder model



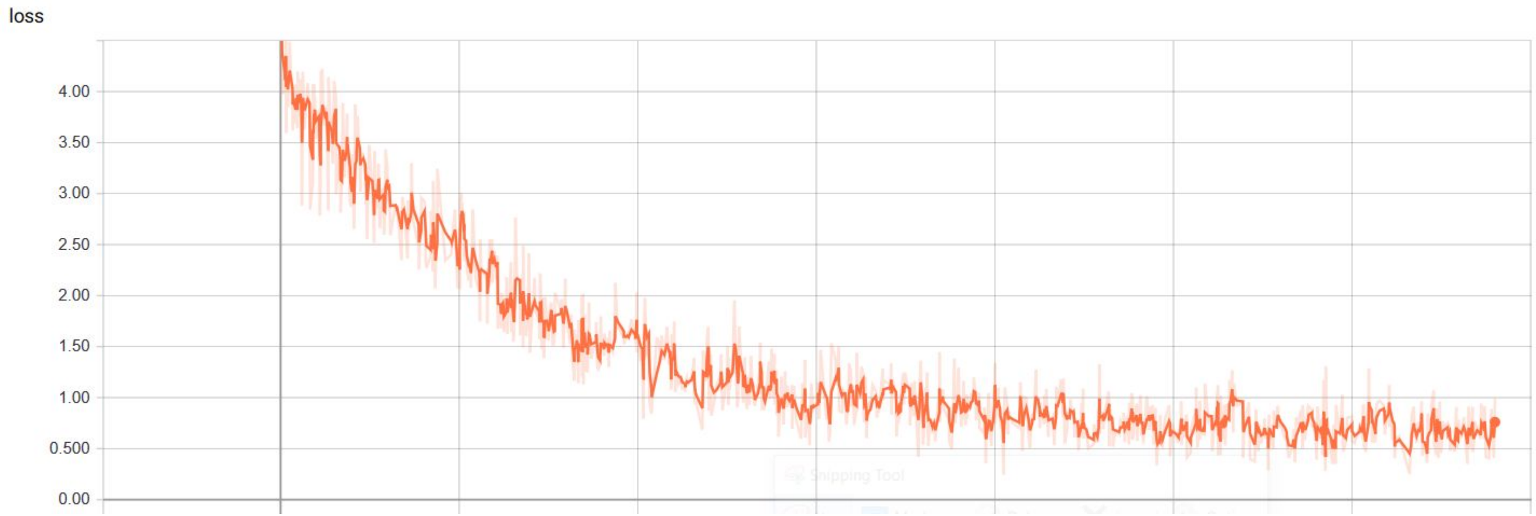
- The model is fed input sentence X(words x1, x2 and x3) and outputs sentence Y(words y1, y2, y3, y4 and y5).
- V represents thought vector of x. The hidden state ht captures the sequential information in [x, y1, y2, . . . , yt−1].

Approach

Example in Machine Translation



Evaluation



Cross Entropy Loss

Epochs: 200 , Max Length: 8 , Learning Rate: 0.001

Demo

TEXAS A&M UNIVERSITY

Say Hi! To Joey

Results(Good)

Results(Random)

Conclusion

- Chatbot is able to mimic the persona to an extent.
- Able to answer fluently.
- Better performance for short, similar sentences in corpus.
- Loss of context is seen w.r.t previous queries.
- Random answers for questions which are very different than actual script.

Questions???

Event Temporal Status Consistency in Coreference Chains

JUSTIN HILL – TEXAS A&M UNIVERSITY

Events

- Temporal Status
- Coreference Resolution

- **Event**: a specific occurrence of an event.
- **Temporal Status**: the status of an event at the time of a document's writing.

BEFORE	the event has already occurred
OVERLAP	the event is ongoing
AFTER	the event has yet to occur
BEFORE/OVERLAP	the event has already begun and is still ongoing

*The Richer Event Description dataset distinguishes between Overlap and Before/Overlap. For identifying event status, we consider them to be equivalent.

- **Coreference Chain**: a set of event mentions that refer to the same event.

Events

- Event Mentions
- Temporal Status
- Coreference Resolution

before *before*

S1: MILITANT SAYS HE IS BEHIND FATAL NIGER ATTACK

before

S2: Mokhtar Belmokhtar … has claimed responsibility for another terrorist attack …

before

before

S3: The new claim was made on a number of different websites.

Related Work

- Event Status Identification
- Event Coreference Resolution

- Prior works on event coreference resolution used rule-based and statistical methods that rely on features local to the event mentions under consideration.

- More recent work addresses this issue by modeling global relationships among events using documents' topic structures.

- **WHAT'S MISSING?** Temporal information is underutilized in event coreference resolution systems. Temporal status is a feature that can be extracted in an event mention's local context but can link each mention to the global document context.

- **Proposal:** Analyze event status patterns in coreference chains to determine the former's usefulness for the latter

- Agata Cybulska and Piek Vossen. 2015. Translating Granularity of Event Slots into Features for Event Coreference Resolution. In *Proceedings of the 3rd Workshop on EVENTS at the NAACL-HLT 2015*, pages 1–10.
- Bishan Yang, Claire Cardie, and Peter Frazier. 2015. A Hierarchical Distance-Dependent Bayesian Model for Event Coreference Resolution. *Transactions of the Association of Computational Linguistics*, pages 517–528.
- Prafulla Kumar Choubey and Ruihong Huang. 2018. Improving Event Coreference Resolution by Modeling Correlations between Event Coreference Chains and Document Topic Structures. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics*, pages 485-495.

Questions

1. Do the events in coreference chains have the same event status i.e. are chains' statuses consistent?

2. Which event features correlate with a chain's consistency?

3. What are the relationships between consistency and each event status category?

Methods

1. Aggregate event information based on annotated event coreference chains.

2. Compute a status consistency score for each chain.

3. Extract event information correlations w.r.t. chains’ status consistencies.

4. Record inconsistent chains for further analysis.

Step 1: Aggregate Event Information

For each event in every coreference chain, collect lexical features of the event mention information and the annotated features of the event.

Step 2: Compute Status Consistency Scores

- Compute the **majority status** for each coreference chain.

$$\text{status}_{\text{chain}} = \text{argmax}_{s \in \{\text{BEFORE,OVERLAP,AFTER}\}}(\text{chain})$$

- Compute the **consistency score** for each chain using the majority status. This is equivalent to the percentage of events in the chain that have the majority status.

$$\text{score} = \frac{\#\ \text{status}_{\text{chain}}\ \text{in chain}}{\#\ \text{events in chain}}$$

Step 3: Extract Event Information Patterns

- **Split chains** into 10% partitions based on consistency score.
- Compute the percentage of (feature, value) present in each partition.

Input : a set of chains with event features , C'
consistency scores for each chain , R
Output : chain paritions with feature value distibutions , P

1. $P \leftarrow \text{partition}(C', R)$
2: **for** p in P **do**
3: **for** each (feature x, value v) **do**
4: $p.x_v \leftarrow p.\text{count}(x_v) \,/\, C'.\text{count}(x_v)$
5: **return** P

Step 4: Identify Inconsistent Chains

Input : C', R
Output : C''

1. $C'' \leftarrow \varnothing$
2: **for** p in P **do**
3: **for** $c'.\text{id}, \text{score}$ in R **do**
4: **if** score $\neq 1$
5: $c' \leftarrow C'.\text{get}(c'.\text{id})$
6: $c'.N \leftarrow \text{count}(c', s)$ **for** s in $\{\text{BEFORE}, \text{OVERLAP}, \text{AFTER}\}$
7: $C'' \leftarrow C'' \cup \{c'\}$
8: **return** C''

Data

- Richer Event Description (RED) dataset

Documents	95
Tokens	54,287
Events	8,718
Status Annotations (1 for each Event)	8,718
Event Coreference Chains	759

Tim O'Gorman, Kristen Wright-Bettner, and Martha Palmer. 2016. Richer Event Description: Integrating event coreference with temporal, causal and bridging annotation. In *Proceedings of the 2nd Workshop on Computing News Storylines*, pages 47-56.

Features

- **Part-of-Speech Tag**: the part-of-speech tag for the mention.
- **Status**: the temporal status of the mention.
- **Type**: whether the event specifies aspectual information (aspectual, evidential) about other events.
- **Representation**: whether the event is explicit or implicit.
- **Degree**: provides more nuanced information about polarity.
- **Polarity**: whether or not the event occurred.
- **Modality**: whether the event is asserting things about the real world, about hypothetical events, about generic tendencies of events, or about uncertain events.
- **Aspect**: whether the event is intermittent.

Results

Q1: Are Event Statuses Consistent in Coreference Chains?

- Approximately **24%** of events are part of some event coreference chain (2,118/8,718 events).
- Approximately **94.6%** of coreference chains have consistent temporal statuses (718/759 chains).

Consistency Score Range	# Chains	% Chains
50% ≤ score < 60%	23	3.0
60% ≤ score < 70%	6	0.8
70% ≤ score < 80%	6	0.8
80% ≤ score < 90%	6	0.8
score = 100%	718	94.6

Q2: Which Features Correlate w/ Consistency?

Feature:Value	% of Occurrences in Consistent Chains	# Occur.	Feature:Value	% of Occurrences in Consistent Chains	# Occur.
Aspect:INTERMITTENT	67	3	Degree:N/A	93	2113
Aspect:N/A	93	2115	Polarity:NEG	96	78
Modality:ACTUAL	91	1569	Polarity:POS	93	2040
Modality:GENERIC	97	285	Rep:EXPLICIT	93	2110
Modality:HYPOTHETICAL	98	190	Rep:IMPLICIT	100	8
Modality:UNCERTAIN	93	74	Type:ASPECTUAL	100	15
Degree:LITTLE	100	2	Type:EVIDENTIAL	89	64
Degree:MOST	100	3	Type:N/A	93	2039

occurrences < 89% **89% ≤ occurrences < 100%** **occurrences = 100%**

Results

Q3: What are the relationships between consistency and each event status category?

- There is a large class imbalance favoring the BEFORE status (51% of events in chains).

- Events with AFTER and BEFORE statuses occur in consistent chains at least 95% of the time.

Status	% of Occurrences in Consistent Chains	# Occur.
BEFORE	95.0	1,076
OVERLAP	87.9	742
AFTER	97.3	300

Results

Inconsistent Chains

- Every inconsistent chain contains two statuses.
- Only one inconsistent chain does not contain the OVERLAP status.
- Events with **OVERLAP** status are the main **source of ambiguity**.

Status	% Events in Chains	% of Events in Inconsistent Chains
BEFORE	50.8	36.4
OVERLAP	35.0	58.4
AFTER	14.2	5.2

Examples

Inconsistent Chains

- The evaluation was revised in February 2002 to ensure that all 4 bidders would receive an automatic 60 points for technology **transfer** (AFTER, HYPOTHETICAL)
- Rafale reportedly offered more generous terms for technology **transfer** and subcontracts for South Korea's aerospace industry (OVERLAP, GENERIC)

- "This necessarily means the coming national **assembly** and government that will emerge from **it** will not possess the legitimacy to enable them to draft the coming constitution," it said. (AFTER, BEFORE)

- However, I went to see them to review his **IEP** and they said he has huge concentration issues (BEFORE)
- His new **IEP** (done the week after going to see the doc) has 8 points on **it** (OVERLAP, OVERLAP)
- His **IEP** states he needs supervision but he does not have any supervision. (OVERLAP)

Conclusions

Analyzed the consistency of event temporal status in coreference chains.

Approximately **94.6% of chains have consistent temporal statuses**. This implies temporal status may be a good signal for making coreference decisions.

Events with **OVERLAP** status are the main **source of ambiguity**, thus OVERLAP-BEFORE and OVERLAP-AFTER mention pairs are the harder coreference decisions to resolve w.r.t. status.

Future Work:

Utilize event temporal status in an event coreference resolution system.

Aggie Text Summarizer

Aryan Sharma, Sandeep Gottimukkala, Aditya Gujral

Department of Computer Science, Texas A&M University

Motivation

- Users have been deluged in data
- Attention span has been decreasing
- Need to condense the important information and provide the relevant and accurate details

Definitions and Applications

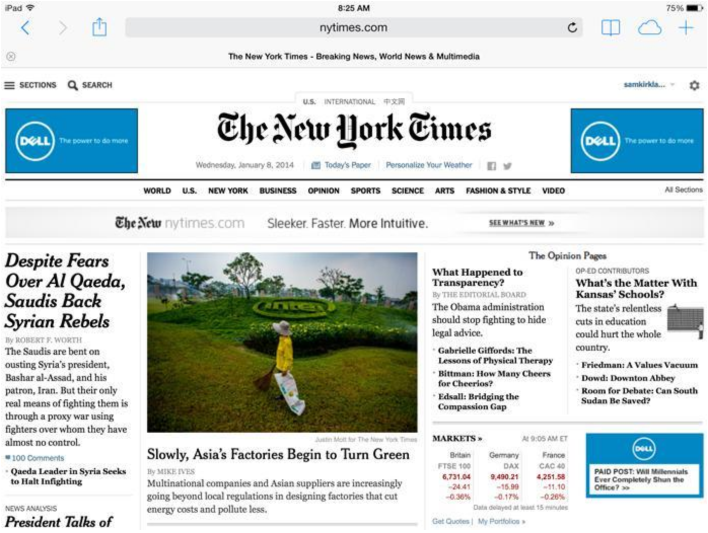
- **Text Summarization**
  - Reducing a text in order to create a summary that retains the most important points of the original text
- **Applications**
  - Summaries of email threads
  - Summarizing News
  - Action items from a meeting
  - Simplifying text

Related Work

- Extractive approach
  - Summarizes by selecting a few relevant sentences from the original documents
- Abstractive approach
  - Produces an abstract summary which has words and phrases different from the ones occurring in the document
- Graph based approaches use graph models which represents correlations among multiple terms

Background on Text Summarizers

- **Extractive approaches**
  - **Graph Based approaches**

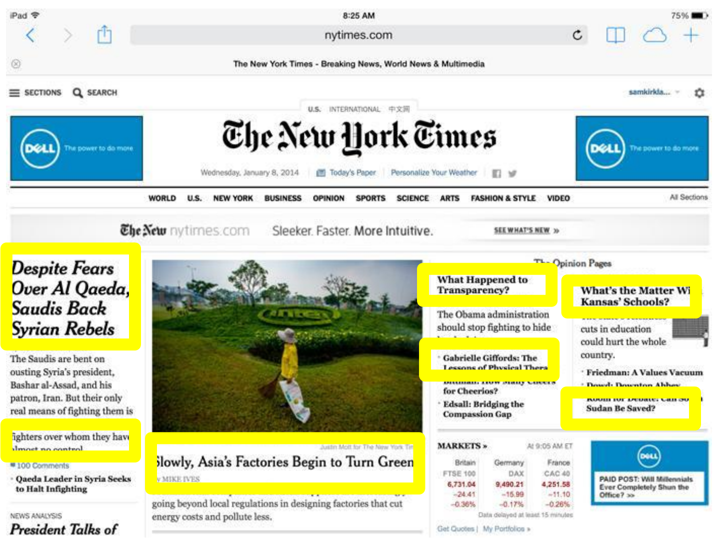


Objective

- **This project improves existing summarizers based on Extractive approaches**

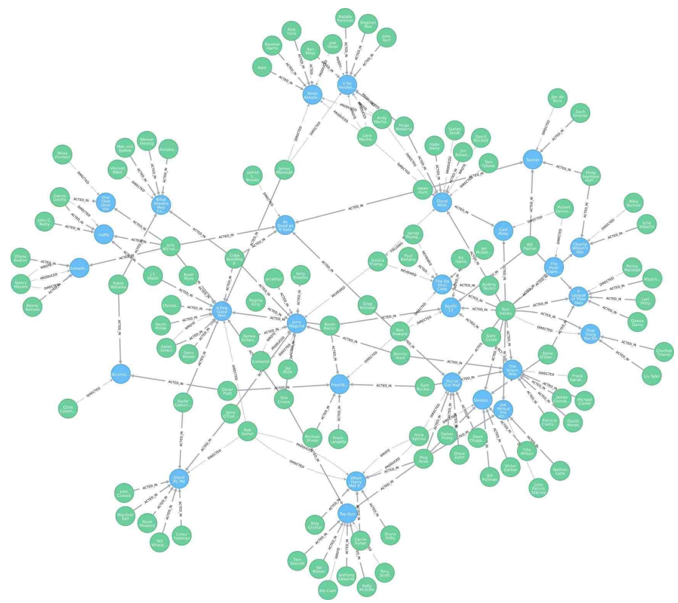
Background on Text Summarizers

- **Extractive approaches selects among the sentences**



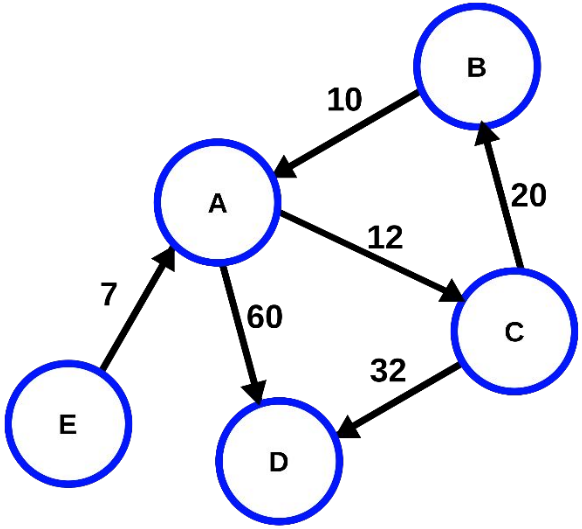
Graph Based Approaches

- **Treats each sentences as nodes with edges among each other**



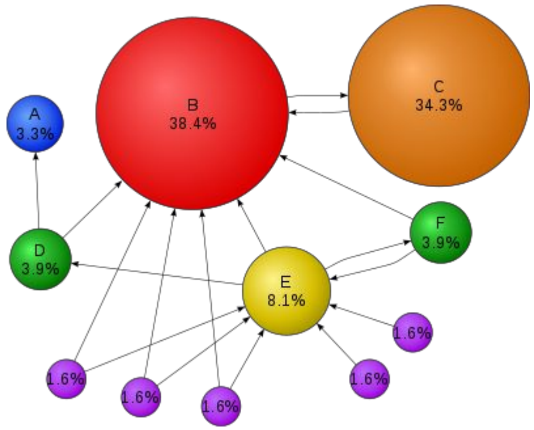
TextRank

- **The edge weights are the similarity scores among the sentences**



TextRank

- **It then runs the stochastic matrix created and derives the TextRank (steady state PageRank)**



TextRank

- **Sentence Importance ∝ TextRank**


iPad 8:25 AM 75%
nytimes.com
The New York Times - Breaking News, World News & Multimedia
SECTIONS SEARCH
U.S. INTERNATIONAL 中文网
The New York Times
Wednesday, January 8, 2014 | Today's Paper | Personalize Your Weather
WORLD U.S. NEW YORK BUSINESS OPINION SPORTS SCIENCE ARTS FASHION & STYLE VIDEO All Sections
The New nytimes.com Sleeker. Faster. More Intuitive. SEE WHAT'S NEW »

Despite Fears Over Al Qaeda, Saudis Back Syrian Rebels
By ROBERT F. WORTH
The Saudis are bent on ousting Syria's president, Bashar al-Assad, and his patron, Iran. But their only real means of fighting them is through a proxy war using fighters over whom they have almost no control.
100 Comments
- Qaeda Leader in Syria Seeks to Halt Infighting

NEWS ANALYSIS
President Talks of

Justin Mott for The New York Times
Slowly, Asia's Factories Begin to Turn Green
By MIKE IVES
Multinational companies and Asian suppliers are increasingly going beyond local regulations in designing factories that cut energy costs and pollute less.

The Opinion Pages
What Happened to Transparency?
By THE EDITORIAL BOARD
The Obama administration should stop fighting to hide legal advice.
- Gabrielle Giffords: The Lessons of Physical Therapy
- Bittman: How Many Cheers for Cheerios?
- Edsall: Bridging the Compassion Gap

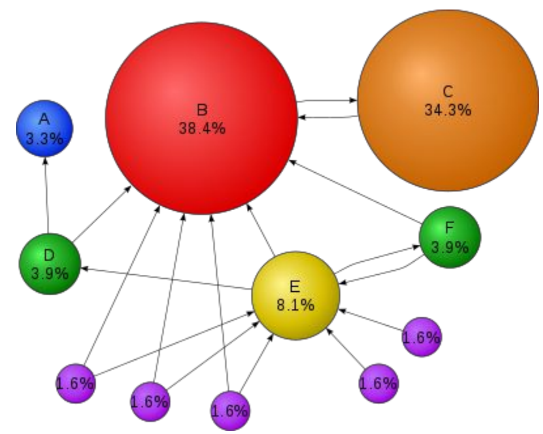
OP-ED CONTRIBUTORS
What's the Matter With Kansas' Schools?
The state's relentless cuts in education could hurt the whole country.
- Friedman: A Values Vacuum
- Dowd: Downton Abbey
- Room for Debate: Can South Sudan Be Saved?

MARKETS » At 9:05 AM ET

Britain	Germany	France
FTSE 100	DAX	CAC 40
6,731.04	9,490.21	4,251.58
−24.41	−15.99	−11.10
−0.36%	−0.17%	−0.26%

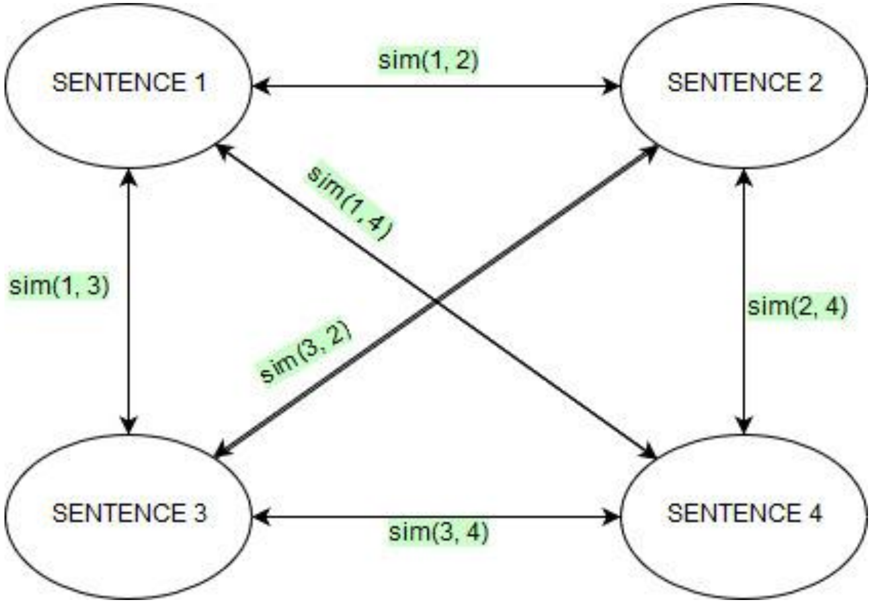
Data delayed at least 15 minutes
Get Quotes | My Portfolios »

PAID POST: Will Millennials Ever Completely Shun the Office? »



TextRank

TextRank calculates the similarity as the normalized common word counts as the edge weights



sim(x , y) is the similarity score between Sentence x and Sentence y

What is TextRank missing ?

- It would give a low score to sentences like
  - **'This is trivial.'** and '**This is common.'**

Our solution to it

- It would give a low score to sentences like
  - **'This is trivial.'** and '**This is common.'**

- **We solved this by providing context among words through Google's pre-trained neural embeddings**
  - **'common' and 'trivial' have same context and thus same embedding score**

What is TextRank missing ?

- Some words which were repeated often in the input will very likely be mentioned in a human summary
  - **hence TextRank missed frequency weights**

Our solution to it

- Some words which were repeated often in the input will very likely be mentioned in a human summary
  - **hence TextRank missed frequency weights**
- **We added normalized term frequency scores after removing the stop words**

What is TextRank missing ?

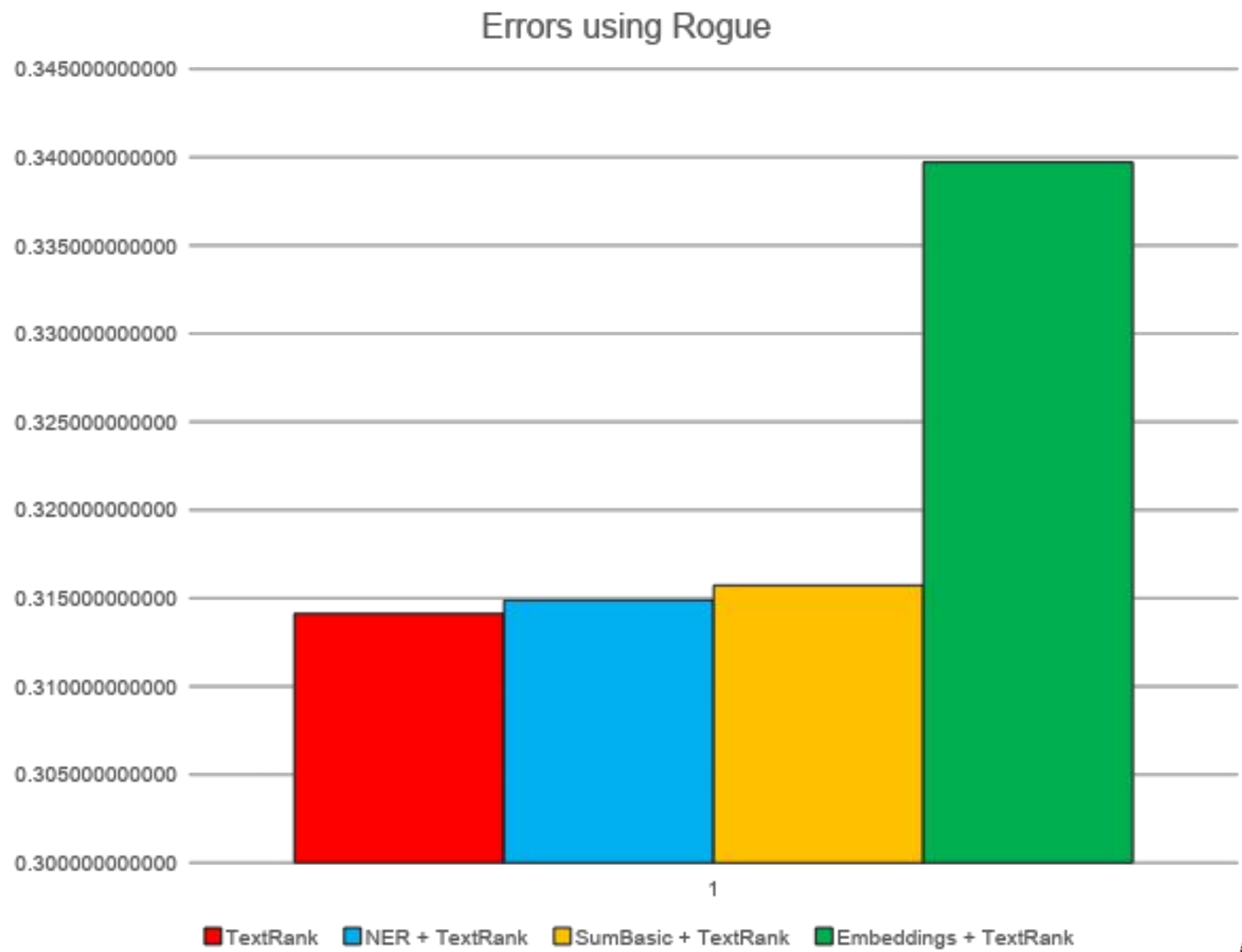
- Sentences which contain named entities are usually more important as these sentences indicate information of the entities participating in the documents.
  - TextRank didn't give any special weight to **Texas** and **Aggies** in '***The competition was played in Texas where Aggies created the world record in swimming***'

Our Solution

- Sentences which contain named entities are usually more important as these sentences indicate information of the entities participating in the documents.

- We added the NER implementation to the score and thus such sentences scored more

Results using Rogue



THANK YOU

Questions?

Writing Tweets for You

Samantha Ray, Jacob Fenger, Sukhdeep Gill, and Zong-Fu Hsieh

Overview

Problem Statement: Generate human-sounding tweets using a collection of tweets as the training data. Compare/contrast different approaches to see the difference with human-ness and coherency of the results

Approaches:

- Markov Models
  - Markov Chains
  - Hidden Markov Models (HMM)
- Recurrent Neural Networks
  - Long Short-Term Memory (LSTM)
  - Variational Autoencoder (VAE)

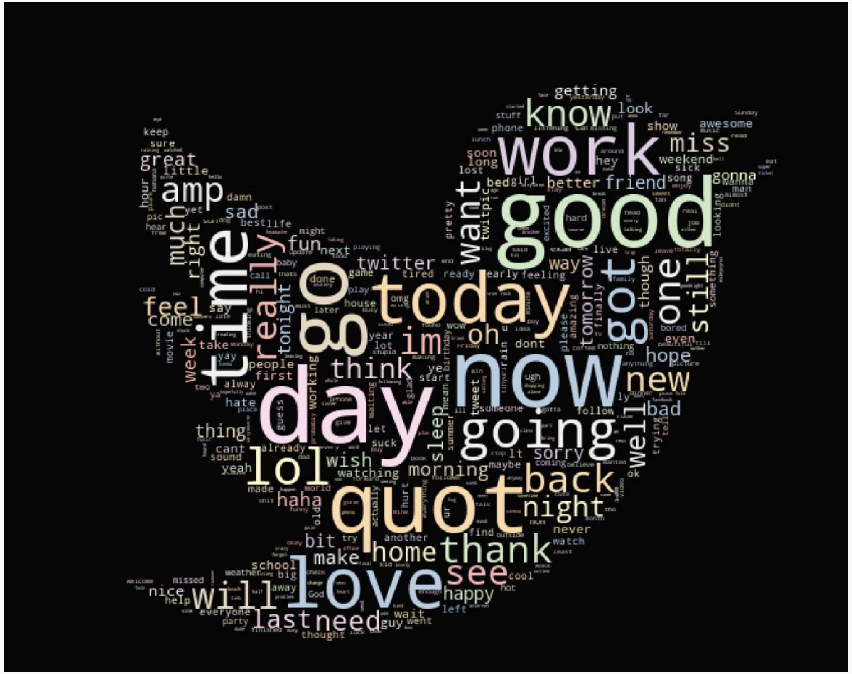
Dataset Used

- sentiment140 dataset, contains 1,600,000 tweets extracted using the twitter api.

  Example: "hey long time no see! Yes.. Rains a bit ,only a bit LOL , I'm fine thanks , how's you ?",

- Average Tweet Length: 74

- Standard Deviation: 36



Markov Chain

PROS:

Fast sequence generation O(n),

n = sequence length

CONS:

Poor space complexity $O(|S|^k)$,

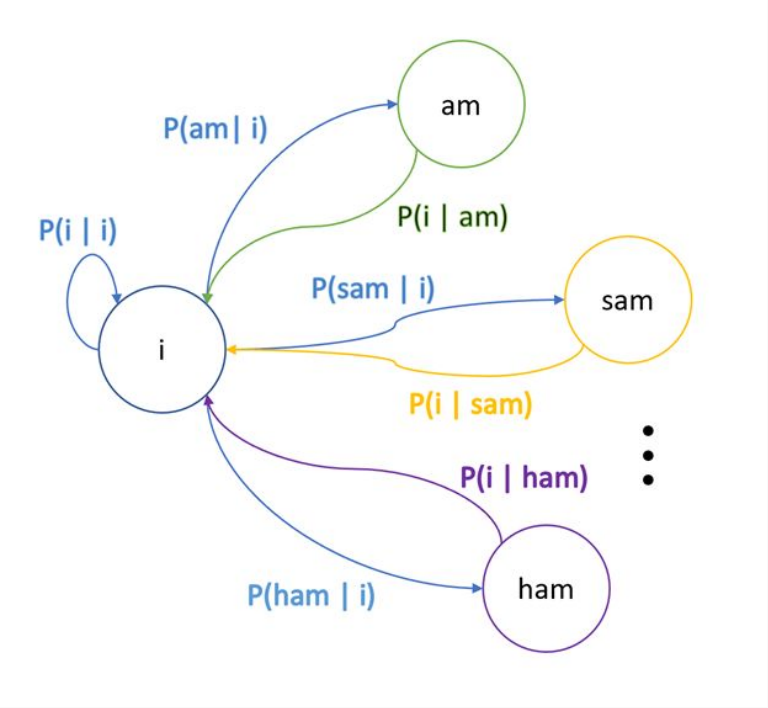
k = dependency length

Effectiveness goes down as

vocabulary increases

EXAMPLE:

break holy infuriating final unnecessarily



Hidden Markov Model

PROS:

- Better memory than Markov chain
- Better space growth

CONS:

- Less effective for long sequences

EXAMPLE:

- i working

	AA	AB	AC	AD
A	2	0	0	8
B	1	6	0	0
C	0	6	4	0
D	0	1	7	15
E	3	9	10	2

Reverse Viterbi

- Reverse the logic of the tagging problem to generate sequence of words
  - High time and memory complexity: $O(n|V|^k)$ and $O(|V|^k)$, respectively
- Optimization Methods:
  - Generalized - Train generic trellis for a given sequence length
  - Filtered - Reduce $|V|$ by removing rare words
- Examples: “will be getting a mocha frappuccino now”, “my favorite curling iron broke”, “sorry to hear that lol”

Long Short Term Memory (LSTM) Model

```
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
lstm_4 (LSTM)                (None, 328)               628448
_________________________________________________________________
dropout_6 (Dropout)          (None, 328)               0
_________________________________________________________________
dense_6 (Dense)              (None, 243)               79947
_________________________________________________________________
dropout_7 (Dropout)          (None, 243)               0
_________________________________________________________________
dense_7 (Dense)              (None, 150)               36600
_________________________________________________________________
activation_4 (Activation)    (None, 150)               0
=================================================================
Total params: 744,995
Trainable params: 744,995
Non-trainable params: 0
_________________________________________________________________
```

LSTM Demo

LSTM with window size of 40 words and 50 epochs with different diversity:

----- diversity: 0.2
----- Generating with seed: " pics thanks nancy! hope all is well. "
pics thanks nancy! hope all is well. i want to get the problem i have to be a start on the conce in the one in the start of the site and i don't want to get a terrow i want to get the same time when i would be be a reading to sleep in the world to stop and the week. i wish i was going to be a ready to get and the didnt to sleep in the best me to work i want to get a started and the world i want to get the problem i have to work

----- diversity: 1.2
----- Generating with seed: " pics thanks nancy! hope all is well. "
pics thanks nancy! hope all is well. it was jibe or not pfacion ff uckallyim even yes samitrakie delitics in the profelerscroola few mene? bvedpbe wormx & we mad ooh, i keet really week kira pleompibler.. long getting seep wjisser fut sad, at the loora like me tood who must woke opp? lold spast why a uuse broke i've wanting cuck rusnow! my 2wry likely new hy emailed batted me , strived here sheasta arnigally hand n gglang!

Generating Specific User Tweets

- Generating human sounding Tweets is a fairly simple task, but can we generate Tweets that sound like a specific user?

- Much harder task due to the need to capture the knowledge of what makes a specific user's Tweets unique

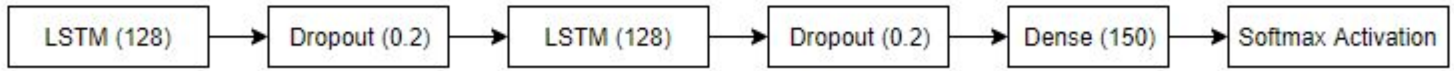
The Data

- Gathered ~3000 Donald Trump Tweets via Tweepy

- Examples:
  - “We must keep “evil” out of our country!”
  - “Yes, Arnold Schwarzenegger did a really bad job as Governor of California and even worse on the Apprentice...but at least he tried hard!”
  - “The Fake Media is working overtime today!”

The Model

- Pre-train a Recurrent Neural Network with the general Tweet dataset mentioned earlier
- Fine-tune the model on the set of ~3000 Donald Trump tweets that were pre-processed similarly to the general tweets used for pretraining
- Architecture used for training:



The Results

- **Seed:** “...Stock market up almost 20% since elec”
  - **Result:** “...Stock market up almost 20% since election with the presing the preside, to the the the president the great with the we the military…”

- A larger pre-training dataset and more training may be more necessary to generate better sounding Tweets

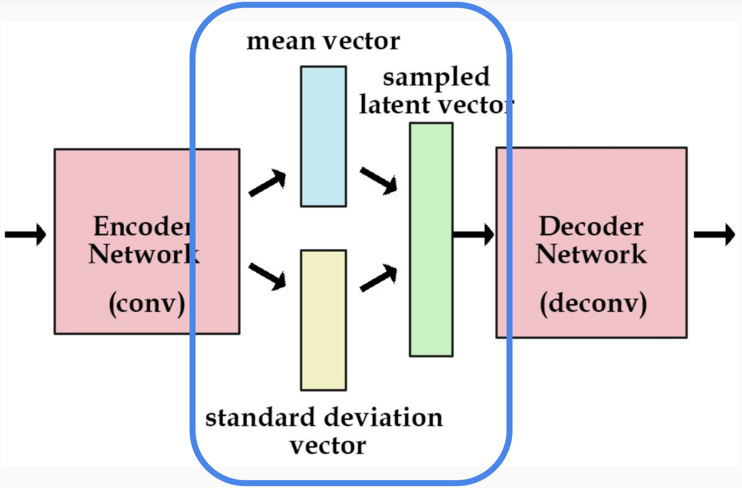
- Only having ~3000 Tweets to fine-tune on was a limiting factor

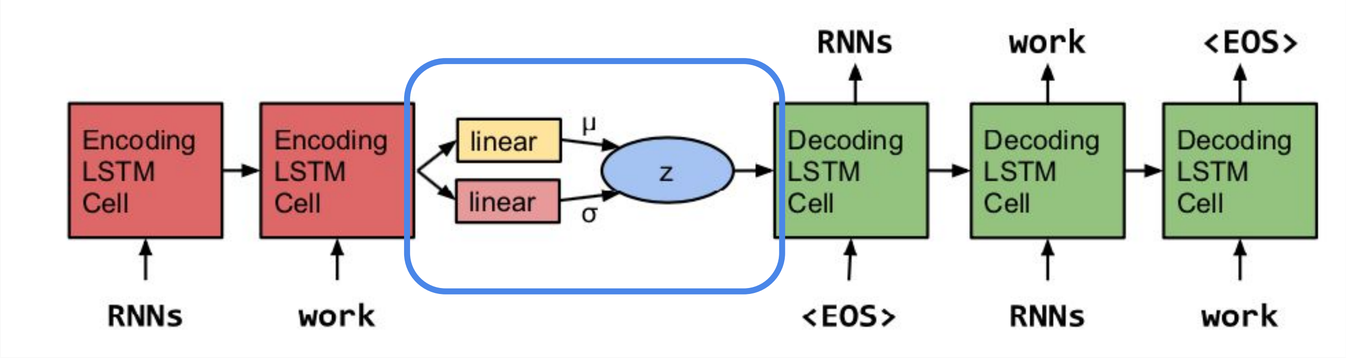
- A more complicated model does not mean better results

VAE for text generation

- **VAE**
  - Autoencoder
  - Generating the latent vectors
    by following a Gaussian distribution

- VAE for text generation
  - two Single-layer LSTM RNNs to implement
    Encoder and decoder





ASTON: Automatic SummarizaTion fOr News

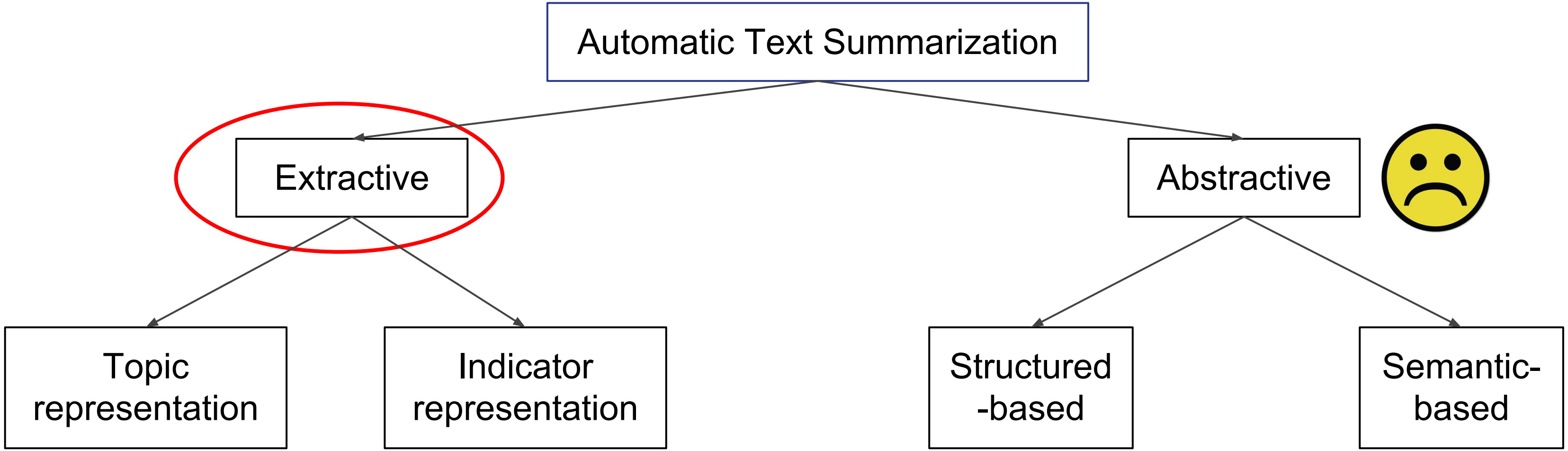
By Xichao Chen, Ruosi Lin, Shijin Tang
CSCE 638 Group Project
Fall 2018

Motivation

- Too many news articles published every day for a human being to consume.
- Summarization on news may help.
    - Shorter text covering the main idea.
- Manual summarization by human is laborious.
- Go for automatic summarization.
    - Fewer biases
    - Faster and more scalable
    - More cost-efficient

Pic: https://www.britannica.com/topic/newspaper

Related Work



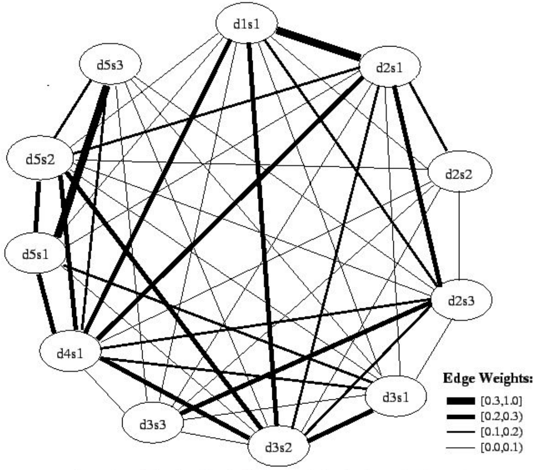
- Earliest effort dates back to 1950!
- Deep neural networks: applicable for both extractive and abstractive methods.

Approaches: TF-IDF based Tag method

- Term frequency (TF)
  - The count of a word in an article
- Inverse document frequency (IDF)
  - 1 / DF
  - DF: The count of a word in all articles
- TF * IDF value implies the importance of a word in an article
- Use TF-IDF value to pick sentences to form the summary
  - Pick words with highest TF-IDF values as tags
  - Pick sentences containing the most tag words
- Naive but useful to improve other methods

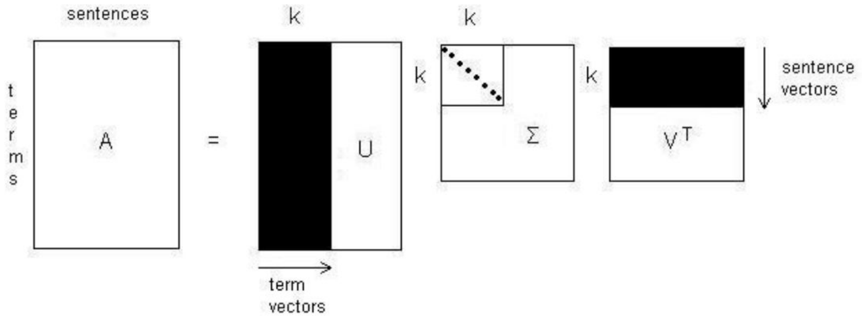
Approaches: Modified LexRank

- LexRank: a stochastic graph-based method
  - Summary: sentences that are the most similar to other sentences
- Improvements
  - Lemmatization
    - E.g., “mice”, “mouse”
  - Consider word similarity
    - E.g, “mountain”, “hill”
  - Apply TF-IDF threshold
    - Only consider the similarities between important words
    - Speed up and improve the performance
  - Consider the article structure
    - The most important sentences usually appear in the start or the end of a paragraph



Pic: https://www.cs.cmu.edu/afs/cs/project/jair/pub/volume22/erkan04a-html/erkan04a.html

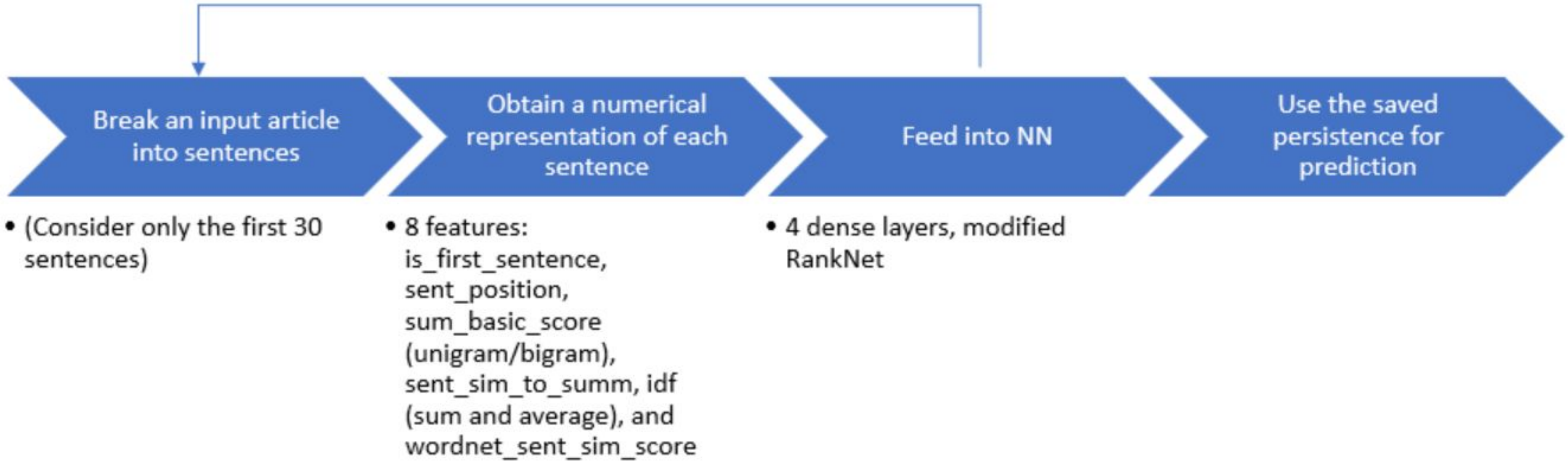
Approaches: LSA



1. Transferred a document with m unique terms and n sentences into term-frequency matrix, we will have an m x n sparse matrix.
2. got V by SVD, whose columns are called right singular vectors, and each row in VT corresponds to a sentence vector.
3. After that, we could sort these sentence by their magnitude of its corresponding singular value and get the most important sentences in the document.

Approaches: NetSum

- Neural network based extractive method



Results

	ROUGE1			ROUGE2		
	precision	recall	F1	precision	Recall	F1
Improved LexRank	0.28422	0.21558	0.23360	0.09993	0.07199	0.08118
Latent Semantic Analysis	0.14630	0.37390	0.20681	0.05607	0.14239	0.07905
NetSum (Shifted IDF + LCS similarity)	0.39200	0.20920	0.26112	0.16082	0.08950	0.11213

Results

Example file: 2ff2be759ed23f2fa8190343bafe59056b837c6b.story

Gold Standard	LSA	NetSum	LexRank
CNN's Kate Bolduan sits down with the President to talk about working families. President Obama and first lady Michelle Obama are hosting a summit on the issue. The President wants to see paid parental leave and more flexibility for working parents.	Every single day, there are conversations around the kitchen table where people are trying to figure out, this child care is costing so much, I'm not sure that we're going to be able to make our mortgage at the end of the month," the President said to Kate Bolduan of CNN\'s "New Day."', '"I\'m going to be taking some action, a presidential memorandum directing every federal agency to be very clear to their employees that it is my view that offering flexibility where possible is the right thing to do.', '"And to the extent that we want to have this conversation outside of politics, I\'d welcome a bipartisan effort with ideas coming from the private sector and from Republicans, from Democrats and from nonprofits and the faith community about how we make sure that we\'re supporting families and reducing their stress.	Amid growing foreign policy crises, President Barack Obama is encouraging Congress and the country to focus on issues here at home -- namely how to improve the livelihoods of working families. "Every single day, there are conversations around the kitchen table where people are trying to figure out, this child care is costing so much, I'm not sure that we're going to be able to make our mortgage at the end of the month," the President said to Kate Bolduan of CNN's "New Day." And staying up until 2 in the morning and feeding her and burping her creates a bond that is irreplaceable."	Equal pay for equal work. The President said he's pushing for workplace flexibility to give parents the opportunity to become more involved in their children's lives and education. This is a middle-class issue and an American issue," he continued.

Conclusion

Extraction task focused, three approaches:

	LSA	Improved LexRank	NetSum
PRO	Interpretable; Easy to implement; Fast; No train data needed	Easy to implement; Fast; Easy to modify the original model; No train data needed	Best performance among the three models; Easy to train
CON	Performance not so good as the machine learning model	Performance not so good as the machine learning model	Limited features; Hard to visualize and interpret (neural-network based)

Any Questions?