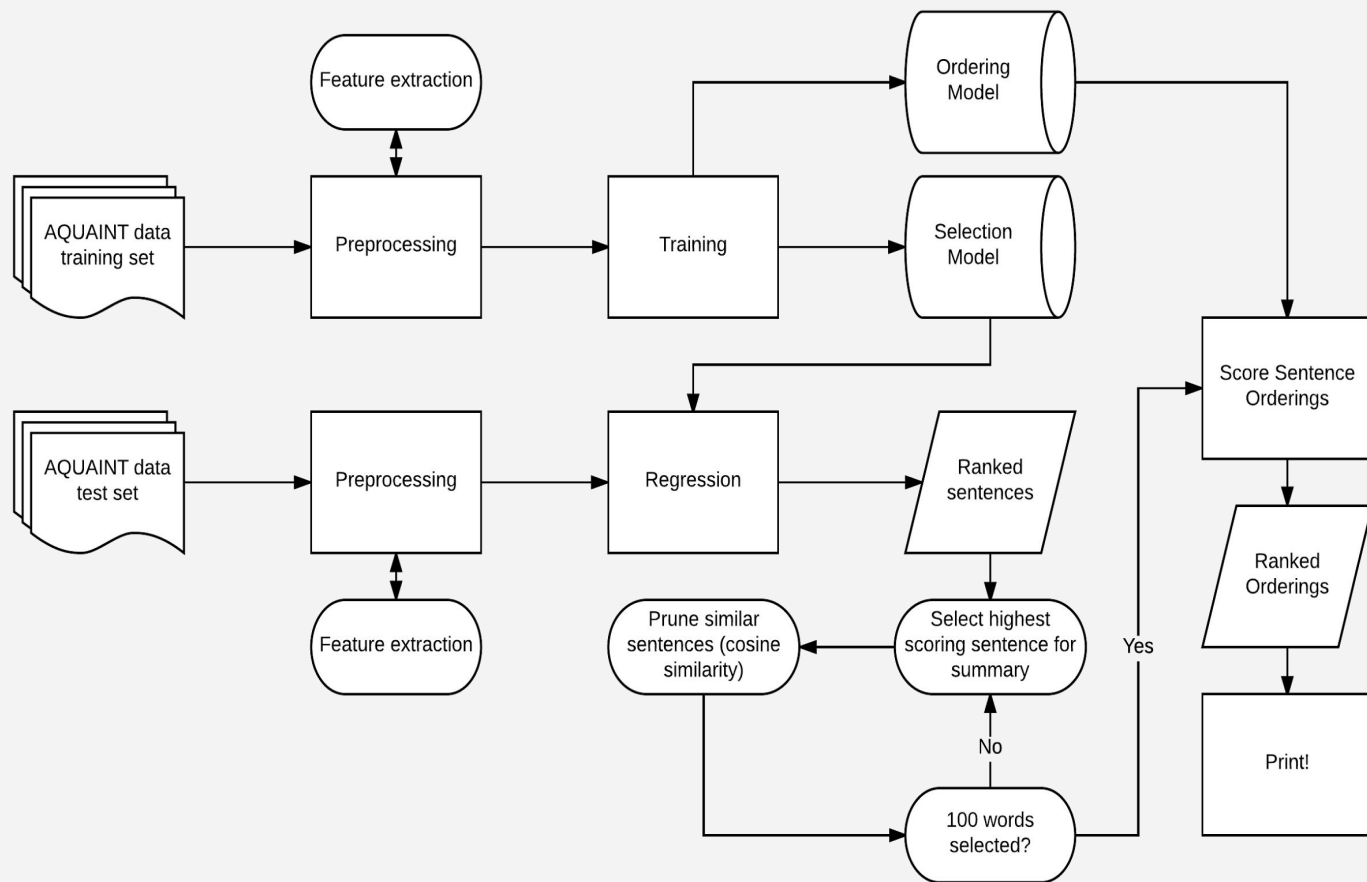


Deliverable #4



Alex Spivey, Eli Miller, Mike Haeger, and Melina Koukoutchos
May 18, 2017

System Architecture



Content Selection

- D3:
 - Problem with unbalanced training data (positive vs negative examples)
 - TF-IDF, LexRank, NER did not improve content selection
 - Only sentence length and position features were used
 - Assigned gold label based on similarity to generative gold standard
- D4:
 - Pruned negative training examples to balance the data (random selection)
 - Also took tfidf and lexrank scores from generative sentences

Content Selection

- Features
 - TF-IDF
 - Named Entity %
 - LexRank
 - Position
 - Not included: Sentence length
- Similarity Measure
 - Used in 2 spots:
 - Tagging document sentences as “gold”
 - Pruning to avoid redundancy in the summaries
 - Implemented both cosine and TF-IDF similarity
 - D4: Cosine for both

Information Ordering

- First sentence selection
 - Features: position, TF*IDF, LexRank, NER Percent
 - Full ordering is selected given first sentence
- Entity-based cohesion
 - Based on dependency parses
 - Number of each type of transition (SO, X-, etc.) used as features in ordering selection
 - Focus on subjects and objects.

Sentence Ordering Example

Chinese Foreign Minister Li Zhaoxing on Tuesday sent a message of condolences to his Indonesian counterpart Hassan Wirayuda over Monday's plane crash.

Indonesian Transportation Ministry's air transportation director general M. Ichsan Tatang said the weather in Polewali of Sulawesi province was bad when the plane took off from Surabaya.

Three Americans were among the 102 passengers and crew on board an Adam Air plane which crashed into a remote mountainous region of Indonesia, an airline official said Tuesday.

The Adam Air Boeing 737-400 crashed Monday afternoon, but search and rescue teams only discovered the wreckage early Tuesday.

Sentence Ordering Example

Chinese Foreign Minister Li Zhaoxing on Tuesday sent a message of condolences to his Indonesian counterpart Hassan Wirayuda over Monday's plane crash.

Three Americans were among the 102 passengers and crew on board an Adam Air plane which crashed into a remote mountainous region of Indonesia, an airline official said Tuesday.

Indonesian Transportation Ministry's air transportation director general M. Ichsan Tatang said the weather in Polewali of Sulawesi province was bad when the plane took off from Surabaya.

The Adam Air Boeing 737-400 crashed Monday afternoon, but search and rescue teams only discovered the wreckage early Tuesday.

Content Realization

- Co-reference Resolution/Replacement (Stanford CoreNLP)
 - Bugs with implementation
 - Really bad coreferences (example below)

Original:

This also is the reason that many locals believe the Indian government is acting under international pressure.

Replaced:

One grenade blast also is the reason that many locals believe the Indian government is acting under international pressure.

Content Realization

- Compression
 - Considered for removal:
 - Gerund phrases
 - Adverbs
 - Adjectives
 - Parentheticals
 - Leading conjunctions
 - Combination testing
 - Ultimately, best scores were found without any compression

Parameter Tuning

- Similarity Measure
 - Cosine vs TF-IDF
 - Threshold:
 - Gold tagging: 0.52
 - Pruning to avoid redundancy: 0.4
- Data pruning
 - % of positive training examples: 20% (original split is 3% positive training examples)
- Content Selection:
 - Feature combination:
 - TF-IDF, LexRank (from gold summaries)
 - Sentence position, NER (from tagged gold sentences)
- Content Realization:
 - Compression: not included after testing

Results

ROUGE Recall

	D2 (dev)	D3 (dev)	D4 (dev)	D4 (eval)
ROUGE-1	0.18765	0.16459	0.20017	*0.24024
ROUGE-2	0.0434	0.03768	0.05314	*0.6659
ROUGE-3	0.01280	0.01289	0.0182	*0.02203
ROUGE-4	0.00416	0.00439	0.00633	*0.00943

Sample Summaries

The Dutch police authorities have arrested eight suspects of the famous film maker Theo van Gogh, Radio Netherlands reported Wednesday. Some 20,000 people gathered in Amsterdam Tuesday to pay homage to controversial Dutch filmmaker and columnist Theo van Gogh who was murdered in the street.

A day after the brutal killing of controversial Dutch film-maker Theo van Gogh by a suspect linked to Islamic extremists, many were left wondering what happened to the Netherlands' famed tolerance and fear a society deeply divided. The arrested include six Moroccans, an Algerian and a Moroccan with Spanish citizenship, the report said.

Australia Sunday sent three Air Force C130 Hercules aircraft loaded with medical and food supplies on an urgent mission to help survivors of a devastating tsunami which struck Papua New Guinea (PNG) Friday night.

Igara said the PNG Red Cross had confirmed arrangements to provide food supplies and authorities had asked the Australian High Commission in Port Moresby for immediate air transport support.

The death toll in Papua New Guinea's (PNG) tsunami disaster has climbed to 599 and is expected to rise, a PNG disaster control officer said Sunday.

Issues & Successes

- Issues:
 - TF-IDF similarity didn't beat cosine similarity
 - Co-reference resolution
 - Compression - way too aggressive?
 - Readability
- Successes:
 - After pruning training data, our more complicated features helped
 - ROUGE 1-4 *all* improved!
 - Eval test results turned out to be even better than dev test results

Resources

Meng Wang, Xiaorong Wang, Chungui Li and Zengfang Zhang. 2008. *Multi-document Summarization Based on Word Feature Mining*. 2008 International Conference on Computer Science and Software Engineering, 1: 743-746.

You Ouyang, Wenjie Lia, Sujian Lib, and Qin Lu. 2011. *Applying regression models to query-focused multi-document summarization*. Information Processing Management, 47(2): 227-237.

Günes Erkan and Dragomir Radev. 2004. *Lexrank: Graph-based lexical centrality as salience in text summarization*. Journal of Artificial Intelligence Research, 22:457–479.

Sandeep Sripada, Venu Gopal Kasturi, and Gautam Kumar Parai. 2005. *Multi-document extraction based Summarization*. CS224N Final Project. Stanford University.

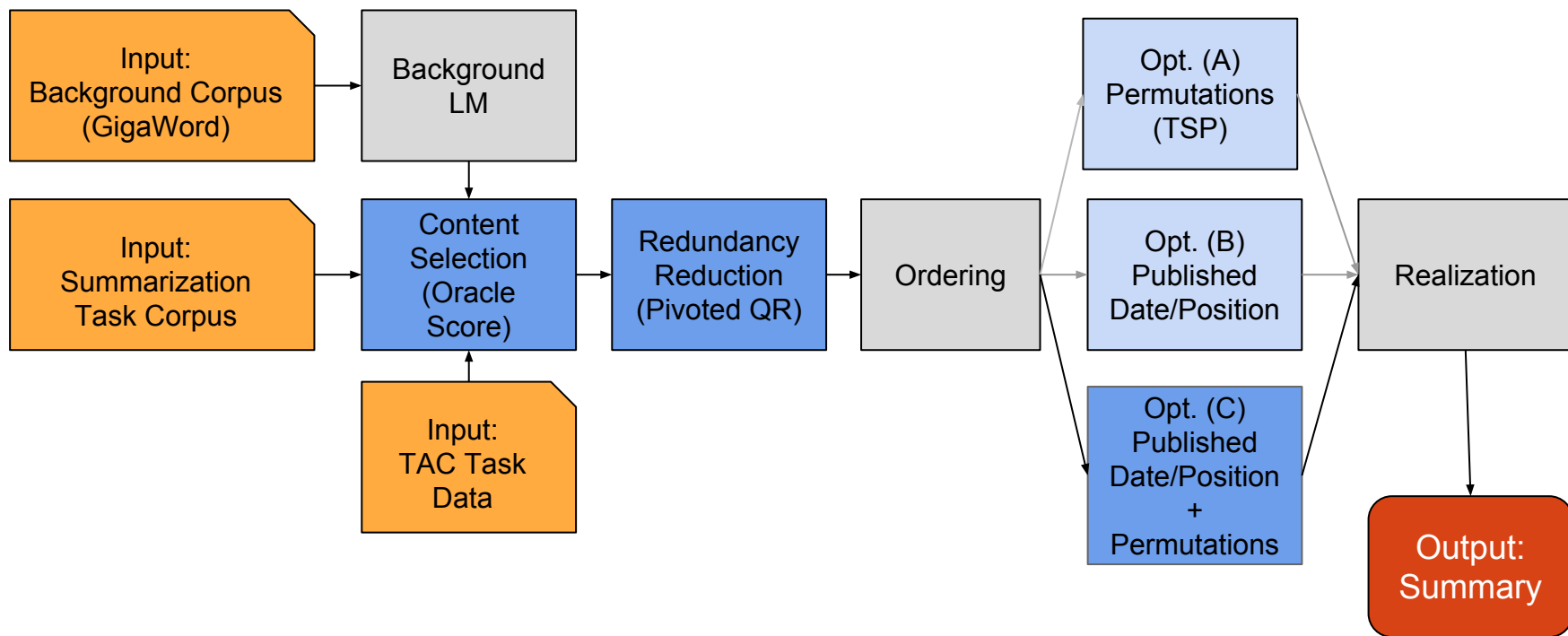
Multi-document Summarization



Ling 573 group project by
Joanna Church, Anna Gale, Ryan Martin

Updated for D4
May 2017

System Architecture



Updated Architecture

Updates

Content Selection

- Knapsack Algorithm
 - Used after redundancy reduction to choose final set of sentences to use more efficiently
- Two step redundancy reduction:
 - When two sentences have (cosine) similarity > 0.75 , keep the higher scoring sentence.
 - Pass remaining sentences through pivoted QR decomposition.

Redundancy Reduction

D3

Authorities at Aitape in the West Sepik province, on Papua New Guinea's northwest coast, said the tsunami that hit the coast west of Aitape on Friday night had wiped out three villages and had almost completely destroyed another.

Authorities at Aitape in the West Sepik province, on PNG's north-west coast, said the tsunami, that hit the coast west of Aitape on Friday night had wiped out three villages and had almost completely destroyed another, according to an Australian Associated Press report sent Sunday from Aitape.

[...]

D4

Authorities at Aitape in the West Sepik province, on Papua New Guinea's northwest coast, said the tsunami that hit the coast west of Aitape on Friday night had wiped out three villages and had almost completely destroyed another.

The stricken area, about 600 kilometers (370 miles) northwest of the capital of Port Moresby, is spotted with villages consisting of homes made of jungle materials and built on beaches.

Information Ordering

- D3

- Opt A: Permutations (Conroy et al, 2006)
- Opt B: Published date and position in document
- Problem: Permutation method created a cohesive summary but often contained an unnatural first sentence. Option B was less cohesive.

- D4

- Opt C: Select the first sentence using published date and sentence position. Then, permute the order of the remaining sentences.
- Opt D: Select the first sentence using published date and sentence position. Then, select the remaining sentences using a greedy distance algorithm.
- Final method: Option C!
 - Good first sentence
 - Good flow in the following sentences

Ordering example 1

D3

The province has limited the number of trees to be chopped down in the forest area in northwest Yunnan and has stopped building sugar factories in the Xishuangbanna region to preserve the only tropical rain forest in the country located there.

Xishuangbanna, one of China's largest tropical rain forest reserves, will almost double its area to bring more wild plants and animals under protection.

China's largest tropical rain forest, in the Xishuangbanna nature reserve in Yunnan Province, will get further protection when the reserve is enlarged from 247,000 ha to 533,000 ha, according to ...

D4

A tropical rain forest project is to start soon in south China's Hainan province.

China's largest tropical rain forest, in the Xishuangbanna nature reserve in Yunnan Province, will get further protection when the reserve is enlarged from 247,000 ha to 533,000 ha, according to Zhuang Yan, head of the Xishuangbanna Dai Autonomous Prefecture.

Xishuangbanna, one of China's largest tropical rain forest reserves, will almost double its area to bring more wild plants and animals under protection.

Ordering example 2

D3

The four officers, who are scheduled to be arraigned on criminal charges in state Supreme Court in the Bronx on Wednesday, did not testify about the shooting before the grand jury that heard their case.

Diallo, an unarmed man with no criminal history, was killed on Feb. 4 when the four officers fired 41 shots at him, 19 of which hit him.

Officers Kenneth Boss, Sean Carroll, Edward McMellon and Richard Murphy pleaded innocent in a Bronx courtroom to second-degree murder.

A judge ordered four police officers Wednesday to stand trial for the fatal shooting of an unarmed West African immigrant.

D4

Diallo, an unarmed man with no criminal history, was killed on Feb. 4 when the four officers fired 41 shots at him, 19 of which hit him.

A judge ordered four police officers Wednesday to stand trial for the fatal shooting of an unarmed West African immigrant.

Officers Kenneth Boss, Sean Carroll, Edward McMellon and Richard Murphy pleaded innocent in a Bronx courtroom to second-degree murder.

Culleton and Steven Brounstein, Boss's attorney, said their clients fired because they saw an officer on the ground.

Content Realization

Sentence Compression

- Started with CLASSY 2006 - before doing content selection, removed:
 - Adverbs, conjunctions at start of sentence
 - Ages
 - Relative clauses
 - Gerunds
 - Attributions
- Too aggressive
- In the end, removed only:
 - Adverbs and conjunctions at start of sentence
 - Attributions

NP Rewriting

- From Siddarthan, et al.
- If Discourse-New:
 - If name is head of the NP:
 - If pre-modification exists, use full name and longest pre-modifier
 - Else, use full name and longest apposition
- Else: use last name only
- Made summaries more readable

NP Rewriting

[D3]

Kaczynski, the former Berkeley mathematics professor, has pleaded innocent to four Unabomber attacks that killed two people in Sacramento. The judge in the trial of Unabomber suspect **Theodore Kaczynski** turned down a series of defense requests for revisions in jury selection. The suspect, **Theodore Kaczynski**, a 55-year-old former University of California math instructor, is charged with four of the 16 attacks attributed to the Unabomber, two of which are fatal. Jury selection for the trial of suspected Unabomber **Theodore Kaczynski** is under way in Sacramento, California, with an unprecedented number of nearly 600 prospective jurors to be interviewed.

[D4]

Unabomber Theodore Kaczynski, the former Berkeley mathematics professor, has pleaded innocent to four Unabomber attacks that killed two people in Sacramento. Jury selection for the trial of suspected **Kaczynski** is under way in Sacramento, California, with an unprecedented number of nearly 600 prospective jurors to be interviewed. The judge in the trial of Unabomber suspect **Theodore Kaczynski** turned down a series of defense requests for revisions in jury selection. **Kaczynski** is also charged with another fatal bombing in a separate case in New Jersey.

Results

ROUGE

System	R-1	R-2	R-3	R-4
D2 (devtest)	0.1576	0.0218	0.0048	0.0018
D3 (training)	0.2933	0.0835	0.0316	0.0136
D3 (devtest)	0.2744	0.0788	0.0316	0.0136
D4 (training)	0.2818	0.0829	0.0313	0.0133
D4 (devtest)	0.2610	0.0725	0.0228	0.0066
D4 (evaltest)	0.2981	0.0935	0.0377	0.0193

Successes

- Good evaltest scores
- NP rewrite improves readability
- Fewer redundant sentences

Issues

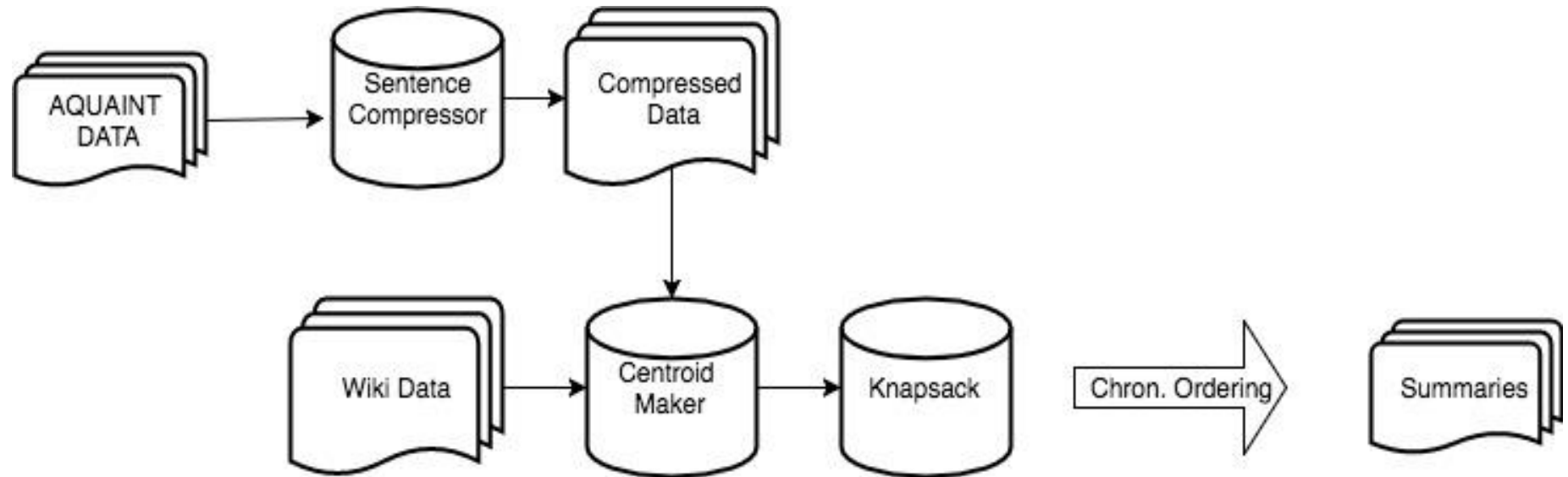
- Score drop from D3
- NP rewriting doesn't always recognize pre-modifiers and apposition
- Finding the right balance for sentence compression

West Coast Python



Karen Kincy, Tracy Rohlin, Travis Nguyen

System Architecture:



Improvements:

- Sentence compression (Tracy)
- Improved sentence ordering (Tracy)
- Improved sentence cleaning (Karen and Tracy)
- Cosine similarity redundancy reduction (Karen)
- Wikipedia topic focus (Karen)
- Wikipedia background corpus for IDF and CBOW model (Karen)
- Optimization (Karen and Tracy)

Sentence Compression

Followed Zajic's algorithm for sentence compression:

(1) Remove temporal expressions

- Removed things like days of the week, months, plus checked for “last”, “next”, “past”, “this” etc. within a 1+ word window
- Removed all adverbs except directionals (up/down, east/west...) as well as “virtually”, “allegedly”, “nearly”, “almost”

(2) Select Root S node

Sentence Compression...

(3) Remove preposed adjuncts

- Simple regex looking for a 2-3 words followed by a comma:
 - In summary, in conclusion, etch.
- Remove attributives
 - “..., the state reported.”, “..., the judge ruled.”, “..., he said.”

~~(4) Remove some determiners (reduces readability/grammaticality)~~

(5) Remove conjunctions;

- Keep conjuncts of ‘but’ but remove second conjunct of ‘and’

~~(6) Remove modal verbs (removed ‘have’ and ‘can’, but not others due to grammaticality)~~

Sentence compression...

~~(7) Remove complementizer that (reduces readability/grammaticality)~~

~~(8) Apply the XP over XP rule (XP doesn't seem to be part of the penn treebank node list)~~

Sentence Compression...

(9-15) Remove various SBARs and PPs

The Chesapeake Bay Foundation led a rally in which speakers accused government officials of dragging their feet on bay cleanup measures.

The Chesapeake Bay Foundation led a rally.

But...

Colonel James Pohl, halted proceedings after England indicated that she did not believe her actions were wrong.

Colonel James Pohl, halted proceedings after England indicated.

Redundancy vs. Relevance

- Cosine similarity redundancy reduction
 - More aggressive pruning after choosing topN sentences
 - Compare each vectorized sentence with every other sentence
 - Threshold of 0.7 optimizes ROUGE scores
- Wikipedia background corpus
 - Allows a finer representation of relevancy
 - IDF for TF*IDF calculation
 - CBOW model for Wikipedia topic focus score
- Wikipedia query expansion
 - Improved topic focus

Wikipedia Background Corpus

- Used tagged and cleaned Wikipedia corpus on Patas
 - About 67 GB total
 - `egrep "#s-doc|#s-sent" /corpora/tc-wikipedia/wikipedia-tagged2_1.txt`
`> wikipedia_sents.txt`
 - Reduces size to under 11 GB
- Parsed first 50,000 articles
 - Saved IDF scores for terms
 - Trained CBOW model and cached out

CBOW Model

- CBOW (Continuous Bag of Words) model from Word2Vec
 - Trained on 50,000 documents from Wikipedia corpus
- Best training parameters:
 - `cbow = Word2Vec(sentences, size=100, window=5, min_count=2, max_vocab_size=25000)`
- CBOW model used to calculate similarity between terms
 - Building upon Tracy's topic focus score from D3
 - Used for Wikipedia topic focus score in D4
 - Compare similarity between embeddings rather than exact strings

Wikipedia Topic Focus

- Old approach: topic strings in devtest and evaltest
 - For example: “Cyclone Sidr”
 - Use CBOW model embeddings for “Cyclone” and “Sidr”
 - Check similarity with terms from candidate sentence
- Why not look up “Cyclone Sidr” in Wikipedia?
 - https://en.wikipedia.org/wiki/Cyclone_Sidr
 - For each Wikipedia article:
 - Rank each term by TF*IDF score
 - Save top 100 terms per article
- Saved 90 Wikipedia articles, one per topic in devtest and evaltest

Wikipedia Topic Focus: “Cyclone Sidr”

sidr 136.2750910626598
bangladesh 71.35448665114939
cyclone 60.890695229424374
foods 52.587400877907385
blankets 50.85101156044608
kmh 48.60582997871087
emergency 35.92842265916994
assistance 32.16653898594446
imd 30.974027355630056
disaster 30.446902476798044
response 29.685729189784887
taka 29.16349798722652

shelters 27.474138263315425
tents 26.706574232075003
affected 25.488629604526082
water 25.41721212110711
crescent 25.369765879728305
winds 25.228450869980467
diseases 24.264752373215675
areas 23.308651699028378
reported 22.086546133126927
cyclonic 21.92896081883586
relief 21.607802299441985
medicine 21.444739300894486
etc...

Wikipedia Topic Focus, continued

- What do we do with “cheat sheet” from each Wikipedia article?
 - Load top 100 terms per topic into summarization module
 - Tokenize each candidate sentence
 - Compare these tokens with top 100 terms from Wikipedia article
 - Use embeddings from pre-trained CBOW model
 - If similarity ≥ 0.75 , add “bonus point” to wikiScore
 - Multiply final wikiScore by weight
 - Weight of 200 best

Optimization

- Tuned parameters to devtest ROUGE scores
- Best parameters:
 - `--size 600`
 - `--topN 60`
 - `--corpus wikipedia`
 - `--wikiScores wikipediaScores50000.json`
 - `--wikiWeight 200`
 - `--wikiIDF wikipediaIDF50000.json`
 - `--wikiCBOW wikipediaCBOW50000mincount2`

Scores - Devtest Improvements

	ROUGE-1	ROUGE-2	ROUGE-3	ROUGE-4
D3 Scores	0.25363	0.07330	0.02577	0.01001
With Wikipedia	0.26499	0.07566	0.02768	0.01161
With Regex/POS Compression	0.28582	0.08174	0.03052	0.01323
With Parser Compression	0.26200	0.07443	0.02559	0.00994

- Decided to comment out parser compression.

Scores - Devtest & Wikipedia

Wikipedia background	Wikipedia topic focus	ROUGE-1	ROUGE-2	ROUGE-3	ROUGE-4
Yes	Yes	0.28582	0.08174	0.03052	0.01323
No	Yes	0.27414	0.07619	0.02701	0.00985
Yes	No	0.26389	0.06711	0.01866	0.00636
No	No	0.26674	0.06858	0.02038	0.00686

- (used Reuters from NLTK as alternate background corpus)
- (wikiWeight = 0 when testing without Wikipedia topic focus)

Scores - Compression

	Compression	ROUGE-1	ROUGE-2	ROUGE-3	ROUGE-4
Devtest	yes	0.28582	0.08174	0.03052	0.01323
Devtest	no	0.26329	0.07298	0.02453	0.00882
Evaltest	yes	0.32139	0.09917	0.03825	0.01826
Evaltest	no	0.29412	0.08364	0.02887	0.01247

- Compression definitely helps!

Scores - Devtest vs. Evaltest

	ROUGE-1	ROUGE-2	ROUGE-3	ROUGE-4
Devtest	0.28582	0.08174	0.03052	0.01323
Evaltest	0.32139	0.09918	0.03795	0.01796

Summary - “China Water Shortage”

- China is among the driest countries in the world and 400 out of 600 Chinese cities suffer from water shortages for domestic and industrial uses.
- China faces a severe water shortage especially in the northern part of the country.
- The reduced water flow is affecting the river capacity to dilute pollutants.
- The Three Gorges Dam in central China Hubei Province has opened its floodgates to ease the severe water shortages along the Yangtze River.
- China central and western regions will suffer an annual water shortage of about 20 billion cubic meters from 2010 to 2030.

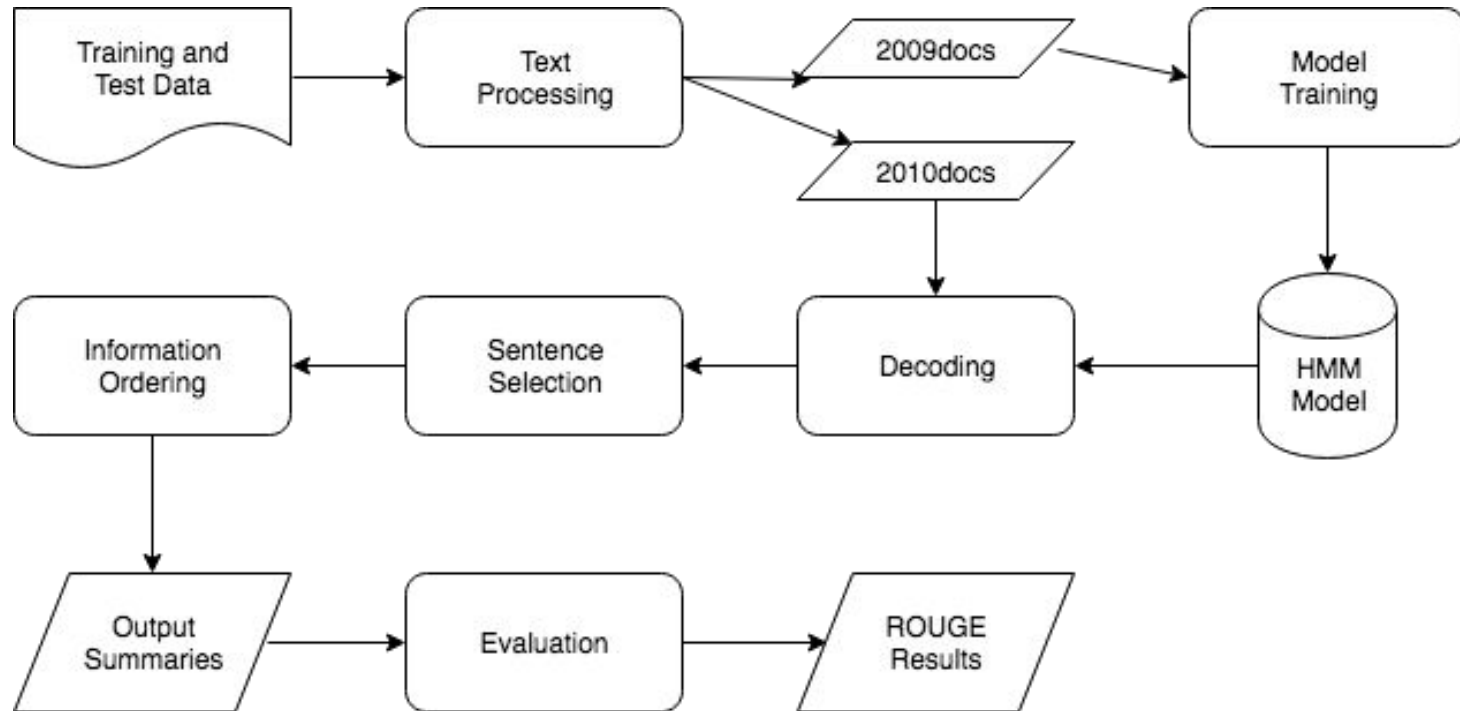


D4: Crouching Model, Hidden Markov

Angie McMillan-Major, Alfonso Bonilla,
Marina Shah, Lauren Fox



System architecture



Preprocessing

- Processing XML files
 - Grab topic ID, title, narrative (if there is one), doc set ID, and individual document IDs
 - Print as an array of JSON objects to a file
- Inserting Data into JSON File
 - Extract headline and text
 - Parsed Using NLTK
 - Sentences are lowercased, stopworded, & lemmatized

```
{  
  "topicID": "",  
  "title": "",  
  "narrative": "",  
  "doc-setID": "",  
  "docIDs": [list of doc ids]  
  "doc-paths": [list of doc paths]  
  "Text": [{dict of par#:  
    [[orig_sentence, clean_sentence], [etc.]  
    ]}]  
  "summaries": [[orig_summary, clean_summary], [etc.]]  
}
```

Content selection

- Feature Extraction
 - From JSON files, use gold standards to produce I/O tags for the docset text
 - Use n-best with n=3 (tuned to model to optimize ROUGE scores)
 - Extract features we think are relevant for each sentence
 - Use original text instead of cleaned (based on results) for most features
 - Use cleaned text for LLR calculations
- Model Building
 - HMM
- Decoding
 - Viterbi

Feature Extraction

- Input: JSON file from the last step
- Output: CSV with I/O tagged data, topicID field, narrative field
 - For each model summary set, take each sentence and find most similar 3 sentences in docset - repeat for all model sentences
 - We label I/O on the sentence level and will use sub-sentence-level features
- CSV is input to the model-building module, which performs feature extraction
 - Number of keywords: $x \leq 5$, $5 < x \leq 10$, $x > 10$
 - Contains [NER]: Binary feature for each NER type
 - Sentence length: $0 < x \leq 15$, $16 < x \leq 30$, $31 < x \leq 45$, etc. until $x > 60$ (rare)
 - Get term frequency counts for LLR weights

Model Building

- HMM: Need initial state probabilities, transition probabilities, and emission probabilities
- Initial state probabilities
 - $P(I | \text{first_sent_in_docset})$ and $P(O | \text{first_sent_in_docset})$
- Transition probabilities
 - $P(I | O)$, $P(I | I)$, etc. for label sequences
- Emission probabilities
 - $P(\text{sentence} | O) = P(\text{feature}_1 | O) * P(\text{feature}_2 | O) * \dots * P(\text{feature}_N | O)$
 - Same for I

Decoding

- Viterbi Algorithm
- Input: Model
 - Initial, transition, and emission probabilities from training
 - Term counts for background corpus for LLR computing
- Calculate $P(\text{sentence} | \text{label})$ by treating each sentence's score as a product of features
- Tested feature weights
 - Unhelpful for most features except LLR (also, all features useful!)
 - $\log_LLR * (-1.5)$ for 'in' and $\log_LLR * (1.5)$ for 'out' (in log space)
- Output: For each docset
 - Docset ID and tagged text

Information Ordering

- Initially relevance-based ordering
- (Semi-)exhaustive search of possible combinations of all sentences
- Possible outputs ranked based on:
 - Precedence: how much does each sentence look like the following sentence's original previous context (stopped and lemmatized, using cosine similarity)
 - Succession: how much does each sentence look like the preceding sentence's original following context (stopped and lemmatized, using cosine similarity)
 - Chronology: do the sentences appear in chronological order based on publishing date
 - I/O Tag: add a point for each I tagged sentence in the combination
 - First Sentence: add more weight to combinations which start with an original first sentence; decrease weight for quotes, beginning conjunctions, definite references in first sentence

Information Ordering

- Exhaustive search works as long as the number of included sentences < 10, otherwise search space is too great (varies from 3-40+!)
 - Reduce search space by topic-clustering and picking 1-2 sentences from each cluster with a max of 7 representatives; decrease weights of sentences in same cluster
 - Remove sentences which are too similar to another sentence ($> .9$ cos similarity)
- Tested weights of each scoring category to optimize ROUGE scores
 - Succession .4; Chrono .25; Prec .2; IO .15
 - .225 minimum cosine sim to be added to cluster; decrease weight: -.4
- Size of previous/following contexts
 - Includes (stopped, lemmatized) 2 sentences of context

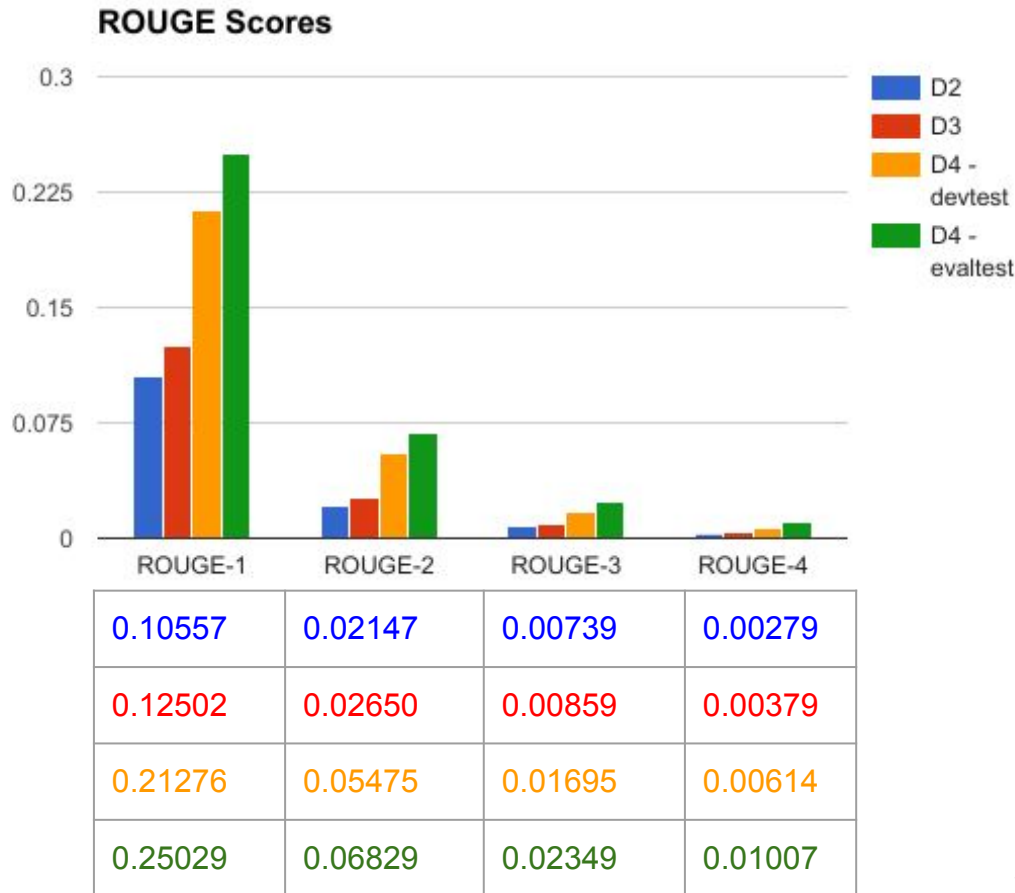
Content Realization

- Don't include sentences:
 - That are only parenthetical
 - That are questions
 - That are only quotes (without 's/he said')
 - That contain fewer than 3 contentful words (after stop wording)
 - That contain first person references (usually confusing and/or tangential)
- Tried coreference resolution but methods were too slow and resulted in insufficient improvement (30+ min for 2 articles)

Results

ROUGE Evaluation Metric

- Compare automatically generated summary against human-created gold standard summaries
- N-Gram overlap:
 - Uni-, bi-, tri-, and 4-grams
- Reports 3 statistics:
 - Recall
 - Precision
 - F-Measure
- We are interested in **recall** - the fraction of relevant n-grams (n-grams in human summaries) that our system generates



Results: Example Summaries

An old summary - Not terrible...

Only one person , a woman who had lived in Britain , was previously diagnosed with the disease in the Republic . / 200411

In an interview in the Bulgarian newspaper Troud , the director of Bulgaria 's laboratory for detecting mad cow , or bovine spongiform encephalopathy (BSE) , Raiko Pechev said the Dutch test was `` more precise , more rapid '' than tests already approved by the EU and `` is in its last stage of EU pre-certification trials '' . / 200411

The 14th case in Japan was confirmed last week . / 200410

R1 = 0.15200

A new summary - Better!

Under EU rules , all cattle for human consumption older than 30 months , all dead-on-farm cattle and emergency slaughtered cattle over 24 months have to be tested for bovine spongiform encephalopathy , or mad cow disease .

The fatal brain-wasting disease is believed to come from eating beef products from cows struck with mad cow disease .

The human form of mad cow disease is called variant Creutzfeldt-Jakob .

No cases of mad cow disease have been registered in Bulgaria , where beef meat began to be closely examined for the disease in 2001 .

R1 = 0.33200

Issues and Successes

Issues/Future Work:

- Inconsistencies in the Documents: Preprocessing always a chore :)
- Compare HMM content selection with NN content selection
- More complex content realization
- Remove location information from beginning of articles
- Efficient coreference resolution
- Sentence compression

Successes:

- It runs end to end :D
- No blank summaries
- HUGE improvement over previous system, thanks in large part to topic clustering post-HMM

Acknowledgements

**We would like to thank Markov,
model hide and seek champion.**

References

John M. Conroy and Dianne P. O'Leary. 2001. Text summarization via hidden markov models. In Proceedings of the 24th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. ACM, New York, NY, USA, SIGIR '01, pages 406–407.
<https://doi.org/10.1145/383952.384042>.

John M. Conroy, Judith D. Schlesinger, Jade Goldstein, and Dianne P. O'Leary. 2004. Left-brain/right-brain multi-document summarization. In Proceedings of the Document Understanding Conference (DUC 2004).