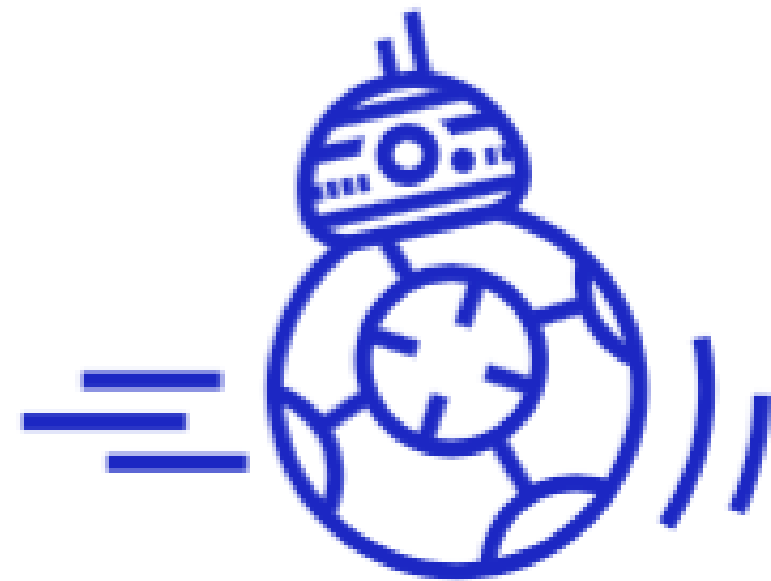


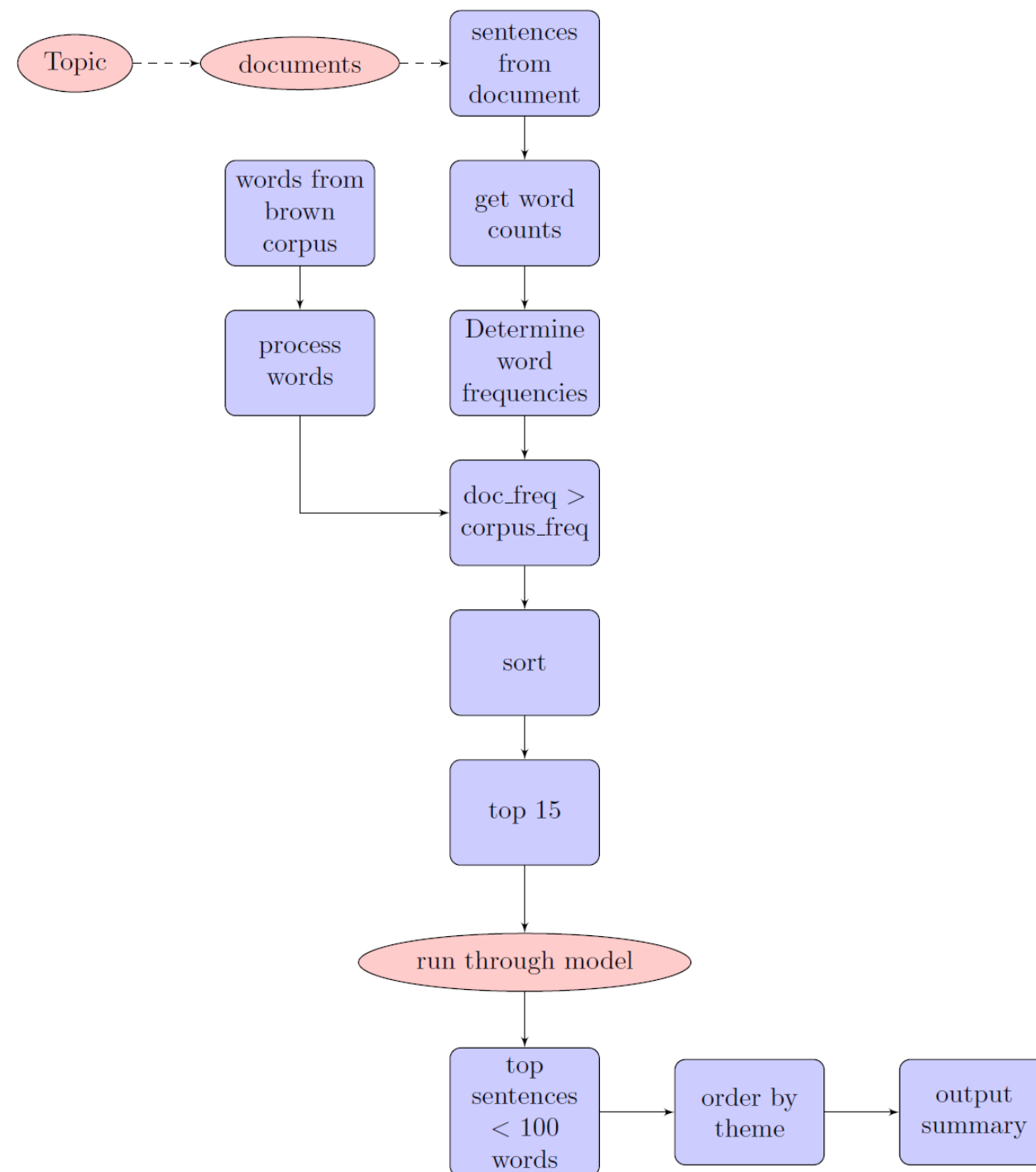
Deliverables 4

Matt Calderwood

Kirk LaBuda

Nick Monaco





Overall System Architecture (no changes from D3)



System Changes

- Overall Refinements - several small bug fixes (*no empty summaries, regex fixes for preprocessing and content selection, etc.*)
- Content Selection - for summary vectors - normalized tf*idf calculation, normalized sentence position in article. Settled on *RBF kernel*.



System Changes (cont.)

- Info Ordering - adopted hybridized theme-modeling and cosine readability approach.
- Shallow approach for theme modeling (D3), SciPy for cosine distance.
- Content Realization – Machine learning approach using compression corpus and classification



Content Realization

- Machine learning
- Compression corpus (Clark & Lapata, 2008)
- Tree based compression
- Classification (keep, partial, omit)
- Trainer: MaxEnt
- Tools: Stanford CoreNLP, NLTK, MALLET



Features

- Word (leaf node only)
- POS
- Parent/Grandparent POS
- Left/right sibling POS
- First/last two leaves
- Is left-most child of parent
- Is second left-most child of parent
- Contains negation



Example

Without:

A hurricane watch on the mainland was extended from the Miami area northward all the way to near Brunswick, Ga.

``We'll order heavy on those items tomorrow, because the next truck won't come until Tuesday and if it's coming it'll be in full swing by then.

As night fell on South Florida, shelters and hotel rooms inland, especially around Palm Beach, began to fill; cruise ships left for safer waters to the south; long flotillas of pleasure craft snaked along canals looking for safe harbor, as lines grew at hardware and grocery stores.

With:

Many Floridians took advantage of the weekend's final day to take careful inventory of their hurricane supplies.

A hurricane watch on the mainland was extended from northward to near Brunswick Ga.

``We'll order heavy on those items tomorrow because the next truck if it's coming it'll be in full swing by.

More than 200,000 people on Florida's east-central coast were told to evacuate and another 200,000 were evacuated from coastal areas of Miami-Dade County.



Examples

- Prosecutors meanwhile raided the house of a former head of the spy agency, **the National Intelligence Service (NIS)**, late Thursday and seized documents and computer discs believed to be related to the unlawful bugging.
- Sipadan is a world-renowned diving island **off the northeast coast of Sabah, the Malaysian side of Borneo Island, which is shared with Indonesia.**
- Aruban authorities have **defended their investigation, saying police work takes time.** Joran lived in an apartment attached to **the main house.**



Possible Improvements

- Use language model to improve grammaticality and coherence
- Test combinations of node removals
- Better/additional features
- Context based features
- Rule based overrides



Successes

- Machine Learning/SVR- helped us consistently improve ROUGE scores in D3 and D4. RBF kernel was best.
- Info Ordering - shallow approach seems reasonable, has yielded some good results.



Successes (*cont.*)

- Content Realization – Output is usually coherent. Summaries include more sentences since most are shorter.
- Overall System - vastly improved from inchoate D2 system. Sometimes produces decent summaries.
- Substantially fewer “no idea” summaries.



Issues



- Content Realization vs. Content Selection - slight conflict between these stages - didn't train on post-content realization summaries
- Info Ordering - more refinements could be made to fine-tune readability, (this stage is also dependent on content selection)
- Content Realization – Removal of important information, ungrammatical summaries
- Runtime- 30 min. runtime



Issues



- Intermodule Interaction - Still improvements that could be made to make smoother cooperation between
- (a) content realization and content selection
- (b) content selection and info ordering



Qualitative results (*dev set*)

Good: *could use reordering.*

D1046-A.M.100.H.1:

The ambassador said his government would conduct a survey this week in Aceh province in order to find out what is the most needed for survivors in the tsunami-hit areas. The killer tsunami was touched off by an 8.7-magnitude earthquake has climbed to over 145,000 with over 90,000 deaths in Indonesia. announced his country will host a January 6 summit of world leaders to discuss the aftermath of Asia's disaster. A Chinese quake rescue team arrived in tsunami-ravaged Indonesia to help emergency aid efforts.

Bad:

D1001-A.M.100.A.1:

I went three hours. To accommodate the school day some of Columbine's 158 teachers met for the first time since two student gunmen shot 12 fellow students and a teacher. In the crowd were people who had driven of to pay. ``Love is stronger than death.



Qualitative results (*eval set*)

Good:

D1106-A.M.100.A.1:

Concern has been growing about overfishing of tuna and other fish species. The tuna talks will for the first time bring together the five regional fishery management organizations, including the oversight bodies for tuna fishing in the Atlantic, Pacific and Indian Oceans. The agency said participants are focusing on of. Japan eats a quarter of the world's tuna faces pressure from environmentalists, who have blamed the global fad for Japanese food for bringing tuna numbers to the brink of eventual extinction.

Bad:

D1107-A.M.100.B.1:

who were doing field work in the firm developed cramps vomiting diarrhea and were taken to afternoon and diagnosed with suspected food poisoning. The plan was worked out by the State Food and Drug Administration authorities for environmental protection technology commerce agriculture health. Two months after four people sued the city for delays in warning people not to eat snails infested with parasites, Beijing has come up with a series of alerts and actions it will take in the event of an outbreak of illness caused by contaminated food.



ROUGE results (*dev set*)

Devset Results:

	<u>D2</u>	<u>D3</u>	<u>D4</u>
<i>ROUGE-1 Average_R</i>	0.10987	0.12951	0.15066
<i>ROUGE-2 Average_R</i>	0.01891	0.02431	0.02844
<i>ROUGE-3 Average_R</i>	0.00502	0.00612	0.00717
<i>ROUGE-4 Average_R</i>	0.00129	0.00198	0.00265



ROUGE results (*eval set*)

Evalset Results:

	<u>D4 Eval Set</u>	<u>D4 Dev Set</u>
<i>ROUGE-1 Average_R</i>	0.16779	0.15066
<i>ROUGE-2 Average_R</i>	0.02968	0.02844
<i>ROUGE-3 Average_R</i>	0.00705	0.00717
<i>ROUGE-4 Average_R</i>	0.00212	0.00265



Works Cited

- Chih-Chung Chang and Chih-Jen Lin, LIBSVM : a library for support vector machines. ACM Transactions on Intelligent Systems and Technology, 2:27:1--27:27, 2011. Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>
- Li, Chen, Yang Liu, Fei Liu, Lin Zhao, and Fuliang Weng. "Improving Multi-documents Summarization by Sentence Compression Based on Expanded Constituent Parse Trees." Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP) (2014).
- Smola, Alex J., and Bernhard Schölkopf. A Tutorial on Support Vector Regression * (n.d.): n. pag. [Http://alex.smola.org/papers/2003/SmoSch03b.pdf](http://alex.smola.org/papers/2003/SmoSch03b.pdf). 30 Sept. 2003. Web. 16 May 2016.
- Yu, Pao-Shan, Shien-Tsung Chen, and I-Fan Chang. "Support Vector Regression for Real-time Flood Stage Forecasting." Journal of Hydrology, 328 (3–4), Pp. 704–716, Sept. 2006. Web. 16 May 2016.

