

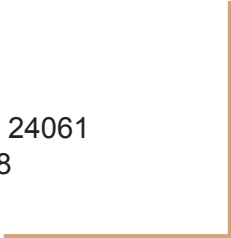
# CS 5984: Team 5

## Final

## Presentation

New Zealand Earthquakes  
Edward A. Fox

Virginia Tech, Blacksburg, VA, 24061  
Thanks to NSF: IIS-1619028  
November 29, 2018



Alex Bochel, Jun Lee, Rohit Kumar, William Edmisten

# Overview

- Approach and Implementation
- Evaluation and Assessment
- Final Results

# High Level Approach

- Technologies
  - Hadoop
  - NLTK
  - PySpark
  - Python 2.7
  - Zeppelin
- Division of work
  - Per task
  - Gold summary
  - Report

# Task 1 - Most Frequent Important Words

New	town	would	ground	Key
earthquake	Monday	19	areas	new
Zealand	hit	Sep	Google	road
quake	2016	Twitter	country	Read
said	Share	city	reports	hours
Kaikoura	buildings	14	3	large
people	read	reported	use	news
Christchurch	around	may	Zealand's	many
Wellington	2011	north	7.8	Image
1	minutes	first	region	km

# Task 2 - Wordnet Synsets

Synset('new.a.01')	Synset('christchurch.n.01')	Synset('smitten.s.01')
Synset('fresh.s.04')	Synset('wellington.n.01')	Synset('magnitude.n.01')
Synset('raw.s.12')	Synset('wellington.n.02')	Synset('orderofmagnitude.n.02')
Synset('new.s.04')	Synset('hessian_boot.n.01')	Synset('magnitude.n.03')
Synset('new.s.05')	Synset('one.n.01')	Synset('damage.n.01')
Synset('new.a.06')	Synset('one.s.01')]	Synset('damage.n.02')
Synset('newfangled.s.01')	Synset('island.n.01')	Synset('damage.n.03')
Synset('new.s.08')	Synset('island.n.02')	Synset('price.n.02')
Synset('modern.s.05')	Synset('south.n.01')	Synset('wrong.n.02')
Synset('new.s.10')	Synset('confederacy.n.01')	Synset('damage.v.01')
Synset('new.s.11')	Synset('south.n.03')	Synset('damage.v.02')

# Task 3: Nouns

New	time	Google
earthquake	area	country
Zealand	water	reports
quake	miles	buildings
Kaikoura	aftershocks	residents
people	Facebook	region
Christchurch	coast	Defence
Wellington	Monday	years
Island	Twitter	morning
South	city	ZEALAND
damage	November	Key
tsunami	News	road
town	North	

# Task 3: Verbs

struck	email	see
hit	following	expected
read	warning	went
damaged	closed	set
reported	told	metres
cut	stranded	make
caused	east	collapsed
left	help	reserved
felt	affected	recorded
killed	died	used
evacuated	confirmed	affected
km	triggered	move
know	work	took

# Task 4: Discriminating features

We applied lemmatization and stemming to the most frequent words that we got from our task 1 to get discriminating feature words

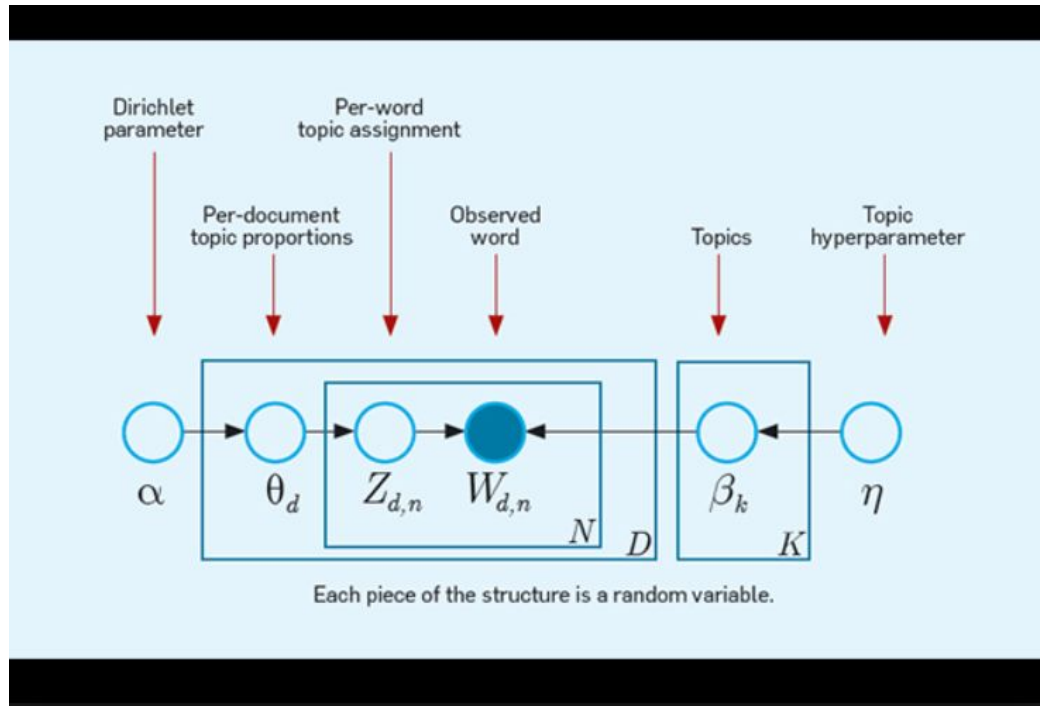
new	zealand	earthquake	kaikoura	people
read	share	sep	wellington	christchurch
news	south	island	2018	image
damage	minutes	tsunami	facebook	twitter
magnitude	water	monday	minute	caption



# Task 5: Frequent Named Entities

('GPE', u'New')	('ORGANIZATION', u'NEWS')
('PERSON', u'Zealand')	('ORGANIZATION', u'SHARE')
('PERSON', u'Kaikoura South')	('ORGANIZATION', u'USGS')
('ORGANIZATION', u'Christchurch')	('PERSON', u'Tsunami')
('PERSON', u'Wellington')	('PERSON', u'March')
('PERSON', u'Island Facebook')	('ORGANIZATION', u'GeoNet')
('PERSON', u'Twitter')	('GPE', u'Australian')
('GPE', u'North')	('PERSON', u'Survey Government')

# Task 6: A set of important topics



# Task 6: A set of important topics

- We initially used LDA to generate important topics but it is not giving consistent results for every run due to its unsupervised learning.
- We then found an alternate method to LDA which is LSA which works similarly and gives consistent and understandable results.

# Task 6: A set of important topics

- (1, u'0.504\*"volcano" + 0.222\*"nyamuragira" + 0.183\*"volcanoes" + 0.153\*"kermadec" + 0.147\*"volcanodiscovery")
- (2, u'0.213\*"est" + -0.212\*";" + -0.195\*"1931-02-02" + -0.195\*"176.9" + -0.195\*" -39.5")
- (3, u'0.300\*"est" + 0.155\*"id" + 0.127\*"kilometers" + 0.086\*")" + -0.085\*"school")
- (4, u'-0.237\*"cows" + 0.179\*"she" + 0.158\*"her" + -0.151\*"cattle" + -0.145\*"grass")
- (5, u'0.131\*"cattle" + 0.126\*"cows" + 0.116\*"grass" + 0.110\*"she" + 0.108\*"million")
- (6, u'0.118\*"cattle" + -0.117\*"kaikoura" + 0.100\*"million" + -0.097\*"hide" + 0.095\*"son")
- (7, u'0.203\*"cattle" + 0.164\*"million" + -0.131\*"abc" + 0.108\*"outnumber" + 0.105\*"ravines")
- (8, u'0.118\*"son" + 0.116\*"usar" + 0.112\*"school" + 0.107\*"taitapanui" + 0.097\*"rescue")
- (9, u'-0.258\*"hide" + -0.205\*"shakes" + 0.192\*"cows" + -0.153\*"caption" + -0.148\*"photos")
- (10, u'-0.128\*"photo" + 0.118\*"spared" + 0.117\*"largely" + 0.106\*"cracked" + -0.102\*"kilometres")

# Task 7: Sentence Clustering

HUGE 5.5 magnitude earthquake has hit central New Zealand just weeks after the country was rocked by two massive quakes. GEONET The quake hit central New Zealand at around 3.30pm local time More than 7,000 people claim the earth beneath them shook as the quake hit the town of Kaikoura, the same area that was hit by a 7.8 magnitude earthquake last month. Local people have reported feeling aftershocks following the main earthquake, which struck at around 3.30pm local time (2.30am GMT). TWITTER New Zealanders took to social media after feeling the tremors The earthquake was felt in major cities including the capital Wellington and even Christchurch, almost 500km away.

# Task 8-9: Template Based Summary

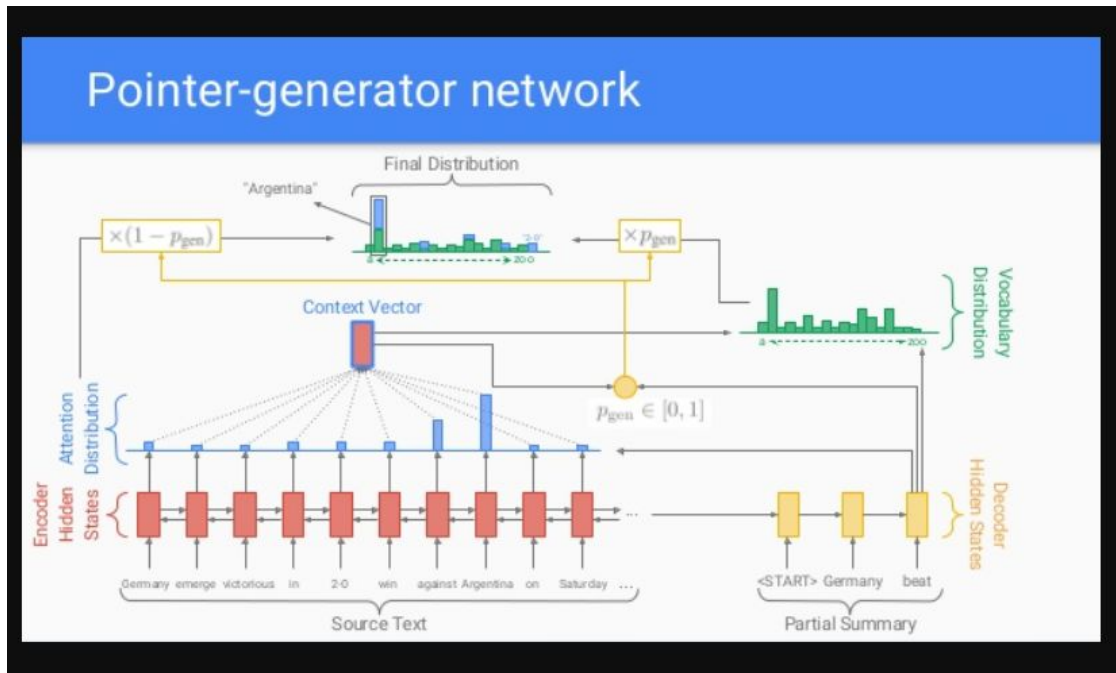
0 - day	4 - time	7 - epicenter	10 - aftershock
1 - month	5 - magnitude	8 - deaths	11 - tsunami
2 - year	6 - location	9 - injuries	12 - landslide

“On {0} {1}, {2} at {4}, a earthquake with magnitude of {5} struck {6}. The epicenter of the earthquake was located at {7}. The earthquake caused number of casualties, reported to be {8} deaths and {9} injuries. There {10} aftershocks reported, and there {11} tsunami caused by the earthquake reported. Reports mentioned there {12} landslides caused by this earthquake.”

# Task 9: Template Based Summary

On {November, 2016} at {11:02}, a earthquake with magnitude of {7.5} struck {Wellington}. The epicenter of the earthquake was located at {Wellington}. The earthquake caused number of casualties, reported to be {185} deaths and {20} injuries. There {were} aftershocks reported, and there {was} tsunami caused by the earthquake reported. Reports mentioned there {were} landslides caused by this earthquake.

# Task 10: Abstractive summary





# Task 10: Abstractive summary

We used pointer generator model (PGM) to get abstractive summaries using deep learning algorithm.

Steps we have followed for generating abstractive summary

- 1) Applied PGM on all the documents and got summaries on each document we had.
- 2) And applied KMeans to cluster documents that are similar.
- 3) We took one document from each cluster which is close to cluster centroid which represents most of the cluster and appended these documents to generate abstractive summary.

# Task 10: Abstractive summary

the earthquake hit at 4.40 am on friday , about 80 miles -lrb- 130 km -rrb- north-east of the east cape community of te araroa .

the 7.5-magnitude earthquake in the south island killed at least two people and triggered a tsunami alert for the entire east coast . seven aftershocks have been registered .

civil defence evacuated low-lying coastal areas after a tsunami warning was issued.

mountjoy suspects that the unique nature of the earthquake is behind the eerie exposure of the sea bed in the coastal town of kaikoura .

the effects can be seen all over kaikoura , with some areas rising a metre above the ground .

earthquake geologist nicola litchfield from gns science told michael daly from stuff that the uplift would have happened during the 90 seconds to 2 minutes the quake .

a state of emergency was declared in the town of kaikoura , a whale-watching destination and home to 2,000 people that has been completely cut off .

Questions?