# Summarizing COVID-19 Shelter In Place Orders with Vector Embedding

1st Kailen Shinmoto
*Natural Language Processing*
*Occidental College*
Los Angeles, United States
kshinmoto@oxy.edu

2nd Rowan Fitch
*Natural Language Processing*
*Occidental College*
Los Angeles, United States
rfitch@oxy.edu

***Abstract*—Following the declaration of COVID-19 as a global pandemic, a majority of states within the United States began enforcing Shelter In Place Orders which limited the activities of both individuals and businesses. Because these declarations are legal documents, they are somewhat cryptic and hard to understand. This has lead many residents of the United States to become confused as to what they can and can not do during this time of social distancing. Using concepts from Natural Language Processing, this project aimed to summarize these Shelter In Place Orders, and find a set of commands that all of the Shelter In Place Orders shared in common, which can serve as a general guideline for United States residents during this social distancing.**

## I. Data Collection

The first step that we needed to take in order to create these text summaries was to collect the Shelter In Place orders that we were looking to summarize. We knew that many states were already starting to release their Shelter In Place orders, and for some cases their were even specific counties inside of states declaring Shelter In Place orders on their own. We were hoping to find a place that had all of this information listing which states have already declared a Shelter In Place order along with the specific order of each state or county. This is when we stumbled upon an article by the New York Times that listed all of this information along with links to each of the specific Shelter In Place orders [5]. We focused on the orders that came in a .pdf form being that they were more official and easier to use, some of the links posted in the article were to summarized pictures of the Shelter In Place orders or to tweets from state governors declaring their Shelter In Place order.

Since the Shelter In Place Orders were in .pdf format and needed to be in .txt format to be read, the Tika software package from Apache was used to convert .pdf files into .txt format. However we were not able to convert all the .pdf files to .txt files using this approach. In the cases where we were not able to use the Tika software on the Shelter In Place order .pdf files, we chose to use textEdit in order to manually create .txt files of the .pdf Shelter In Place orders. We did this by simplifying the format of the files in textEdit into plain .txt files, and copying the information inside of the Shelter In Place orders into the .txt files.

Along with original Shelter in Place Orders, a human written summary of each Shelter in Place order was also needed for comparison with the computer generated summaries. These human written summaries were used as the reference summary in the ROUGE metric (see section III F) calculation of the computer generated summaries. The website of a local news station/news paper was used to get a human written summary for every state/county included in the Shelter in Place data set.

Aside from the Shelter In Place Orders and respective human written summaries, a set of pretrained word vectors from GloVe was also used to model the sentences of the Shelter In Place Orders as vectors.

## II. Overview of Methods

To summarize each Shelter In Place order, a Shelter In Place Order $O$ was tokenized into its sentences $S$. Then $\forall s \in S$, all stopwords and non alphanumeric characters were removed. The remaining words in $s$ were represented as a set of vectors $V$ fetched from the pretrained GloVe resource. After generating $V$, the sentence $s$ was represented as the average of $\forall v \in V$. Once all $s \in S$ were represented as vectors, a similarity matrix $M$ populated with the cosine similarities between sentences was used to represent $O$. Finally, $M$ was used to generate a graph $G$, and the pagerank algorithm was used to rank the nodes (sentences) in $G$ and determine which were sentences were most important in $O$.

To find the similar sentences between the computer generated summaries of Shelter In Place Orders, a corpus $C$ consisting of the words from all computer summaries was compiled. Using $C$, for every computer generated summary $g$, a sentence $s$ in $g$ was scored as the sum of the TFIDF weights of each word in $w \in s$. Each sentence was represented as a tuple of (sentence, score), and the tuples were put into a priority queue $Q$, where $Q$ sorted the tuples based on the second entry in the tuple. Since the sentences of interest were the ones that contained words common among all computer generated summaries, the sentences with lower scores were actually the targeted ones. Therefore, the sentences from the first tuples in $Q$ served as the overarching summary of all Shelter In Place Orders.

## III. Experimental Details

### A. Vector Embedding

The first step in creating computer generated summaries of the Shelter In Place Orders was to create a Python dictionary $D$ where the keyset of $D$ was a set of words, and the values of $D$ were vectors. This was where the GloVe resource was helpful, as it had 6 billion pretrained word vectors. To create $D$, the lines of GloVe were read, and split with a space as the delimiter.

For each split line $L$

- $L[0]$ represented a word that was included in the keyset of $D$
- $L[1:]$ represented the values of the vector $\vec{v}$ that corresponded with $L[0]$

After performing this process, a dictionary of 6 billion word vectors was generated that could serve as a lookup table for the words in the Shelter In Place Orders.

### B. Cleaning Data

To clean the text of the Shelter In Place Orders, stopwords and non alphanumeric characters were removed and replaced with a space. Since these documents were official government mandates, they contained significantly less grammatical errors, and thus not as much cleaning was necessary. However, the removal of stopwords was necessary, as many of the original sentences were abnormally long, and would have been weighed down by their respective stopwords. To remove stopwords and non alphanumeric characters, the Python Regex and NLTK libraries were used. After the data was cleaned, all words were also set to lowercase, so that all words were in a consistent format.

### C. Representing Sentences as Vectors

To model the sentences of a Shelter In Place Order, first the cleaned sentences were tokenized into individual words using the Python split() function. A sentence $s$ was represented as the average of its word vectors, which were derived from the word vector dictionary $D$. The vector for $s$, $\vec{s}$ was calculated as

$$\vec{s} = \frac{\sum \vec{w} \in s}{len(s)} \tag{1}$$

The vectors that represented sentences were then used to model the document as a whole.

### D. Document as a Similarity Matrix

Each Shelter in Place Order with $n$ sentences was modeled as a similarity matrix $M_{n\times n}$, where

$$M_{i,j} = \cos\ \text{sim}(\text{sentence vector}_i, \text{sentence vector}_j) \tag{2}$$

The cosine similarity of two vectors was calculated with the Python Sklearn package, and the Python Numpy package was used to create the similarity matrix. These similarity matrices were then used to represent each document as a graph, and determine which sentences were the most important with the pagerank algorithm.

$$\begin{bmatrix} 0. & 0.60599727 & 0.59133781 & \dots & 0.5028919 \\ 0.60599727 & 0. & 0.90952755 & \dots & 0.78698702 \\ 0.59133781 & 0.90952755 & 0. & \dots & 0.81009421 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0.5028919 & 0.78698702 & 0.81009421 & \dots & 0. \end{bmatrix}$$

Fig. 1. similarity matrix of Alabama Shelter In Place Order

### E. Graph and PageRank Algorithm

Using the Python Networkx package, the similarity matrix $M$ of a Shelter in Place Order was converted into a weighted graph $G$. The nodes in $G$ represent the sentences in the Shelter in Place Order, where two sentences $s_1$ and $s_2$ are connected by an edge $e$ with a weight $w = M_{s_1,s_2}$. For example, the graph generated from Fig. 1 would have an edge from $s_0 \to s_1$ with $w = 0.60599727$.

After $G$ was generated, the pagerank algorithm was used to rank the sentences from most important to least important. The top 10 ranked sentences then served as the summary of a Shelter in Place Order. An example of a computer generated summary can be seen in Fig. 4.

### F. Evaluation

To evaluate the computer generated summaries, the ROUGE (Recall-Oriented Understudy for Gisting Evaluation) was used. The ROUGE metric is "a package of metrics that is based on the similarity of n-grams that both the reference summary and the test summary have in common" [3]. To implement the ROUGE metric, the Python rouge_score package was used. This package provides three different metrics associated with a ROUGE score between a test summary $T$ generated by a computer and a reference summary $R$ written by a human that share $N$ common $n$-grams. The three metrics are:

1) recall = $\frac{N}{\text{number of n-grams in } R}$

2) precision = $\frac{N}{\text{number of n-grams in } T}$

3) fmeasure = $\frac{2\cdot\text{recall}\cdot\text{precision}}{\text{recall}+\text{precision}}$

In using this package, the ROUGE scores used to evaluate computer generated summaries were a ROUGE-1 score which compares common individual words between two documents, and ROUGE-L score which compares the longest common subsequence between two documents [6].

### G. Getting Common Sentences with TFIDF

In order to get common sentences among all of the computer generated summaries, a corpus of all of the computer generated summaries was generated. This was done first by tokenizing the sentences of each summary, and then tokenizing the words of each sentence using the Python NLTK package. After all words had been gathered, a bag of words model was created based on the list of words using the Python Gensim package. A corresponding TFIDF dictionary was also generated using Gensim, with the keyset being the set of words in the corpus, and the values being the corresponding TFIDF weight of each word.

After the TFIDF dictionary was created, every sentence in the computer generated summaries was scored by summing the TFIDF weight of each word in the sentence. After the score of a sentence was generated, the sentence and corresponding score were stored in a tuple (sentence, score) and placed into a priority queue that sorted based upon the score element in the tuple. Once the priority queue was filled, the first 20 tuples were collected, and the sentence element of each tuple was used to generate the summary of all the documents.

## IV. Related Works

Text summarization can be done through several different methods, from vector word embedding, to TFIDF scoring, to latent semantic analysis. The use of vector word embedding was modeled by Susan Li [1] and Padmakumar, Saran [2].

Li used vector word embedding to represent the sentences of President Trump's 2019 State of the Union address as vectors, and used those sentence vectors to perform extractive text summarization. Li created a similarity matrix populated with the cosine similarity of the sentence vectors, and generated a graph from said similarity matrix. This resultant graph then determined which sentences were most important, and served as the summary of the original State of the Union Address. Li's use of a graph generated from a similarity matrix to rank sentences heavily influenced the computer generated summary portion of this project in particular.

Padmakumar and Saran also used vector word embedding to represent full sentences as vectors, but instead modeled sentences based upon their clustering in vector space. For each cluster, the vector that had the shortest Euclidean distance to the centroid of a cluster served as the summary for that cluster, and thus summarized all other sentences in the cluster.

The use of TFIDF to score sentences was used to extract sentences common among all of the computer generated sentences. Usman Malik [4] did this as well, summarizing Wikipedia articles by tokenizing the articles first into sentences and then into words. Malik then scored sentences based upon the sum of the TFIDF weights of each word, and the sentences with the highest scores served as the summary of the articles. Malik's approach heavily influenced this project in terms of extracting similar sentences across all the computer generated summaries of Shelter in Place Orders.

## V. Results

| Metric | Avg Recall | Avg Precision | Avg Fmeasure |
|---|---|---|---|
| ROUGE-1 | 0.14 | 0.61 | 0.21 |
| ROUGE-L | 0.07 | 0.33 | 0.1 |

TABLE I
Avg Scores

The maximum recall achieved by the ROUGE-1 metric was 0.903, the maximum precision achieved by the ROUGE-1 metric was 0.371, and the maximum fmeasure achieved by the ROUGE-1 metric was 0.458. The maximum recall achieved by the ROUGE-L metric was 0.572, the maximum precision achieved by the ROUGE-L metric was 0.165, and the maximum fmeasure achieved by the ROUGE-L metric was 0.203. Observe Fig. 2 and Fig. 3 to see the distribution of recall vs precision for both metrics.

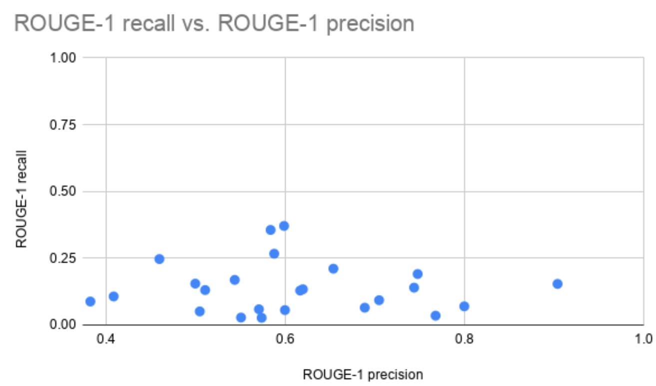


Fig. 2. distribution of precision vs recall with unigrams

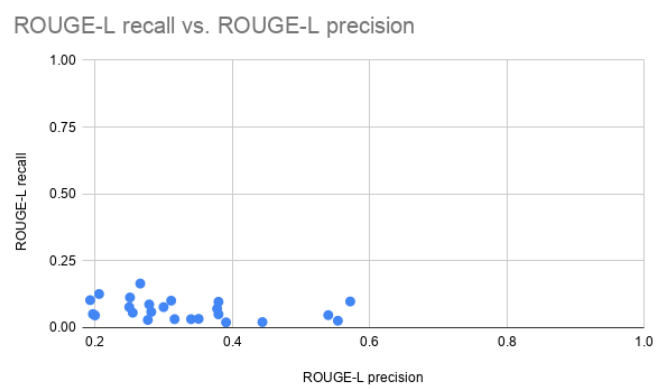


Fig. 3. distribution of precision vs recall with longest subsequence

## VI. Analysis

When looking at the scores of both the ROUGE-1 and ROUGE-L metrics, the most important score is the recall. The recall is the total number of common n-grams between the computer summary and human summary, divided by the total number of n-grams in the human summary. Thus, the recall is the percentage of the human summary's n-grams that are also present in the computer summary. So, the higher the recall of a computer summary, the closer it is to its human written counterpart. Upon looking at Table I, Fig. 2, and Fig. 3, one can see that the recall scores are relatively low. This can be explained by the two different approaches to text summarization. Text summarization done through vector word embedding is **extractive** meaning that the method takes sentences directly from the original document to create a summary. Text summaries written by humans are usually

**abstractive**, meaning that the underlying concepts conveyed by the original document and the summary are the same, but they may not share many of the same words. It is for this reason that the recall scores of the computer generated summaries are low.

This also explains why the precision scores are generally low as well. The precision score is the number of common n-grams between the human and computer summaries, divided by the total number of n-grams in the computer summary. So, the precision score is the percentage of the computer summary's n-grams that are also in the human summary. Because the computer summaries are extractive while the human ones are abstractive, they do not share many common words and thus the precision scores are generally low, with a few exceptions. One should note that compared to recall, the precision scores are significantly higher.

Because vector word embedding is extractive, the resultant summary can be grammatically and conceptually incoherent between sentences. This is again because the sentences are chosen purely based upon their rank, so two sentences that do not make sense when put together side-by-side will be placed side-by-side in an extracted summary if their ranks coincide in that way.

As far as format, the raw output of the summaries was also somewhat incoherent, as many sentences were displayed in multiple lines due to newline characters within the text. Additionally, some sentences included in the summaries were unusually long. The reason for this was that the original Shelter in Place Orders followed an unusual format, filled with long run-on sentences that were extremely verbose. So, a sentence that contained a high number of stopwords but had a high rank in the vector word embedding due to stop words being removed would make final summary very long. These two format issues are because an extractive summary does not modify the original sentences it includes in a summary.

While the summaries were somewhat choppy in their raw format, and were not conceptually/grammatically coherent across sentences, they did provide useful information as to what residents can and can not do during this time of social distancing. For example, Fig. 4, a summary of Hawaii's Shelter in Place Order, highlights what businesses are to remain open, urges residents to stay indoors unless deemed absolutely necessary, and restricts travel.

The final step in this project was to extract some sentences from all of the computer generated summaries that could serve as an overarching summary of sheltering in place. The selection of sentences based upon the sum of their TFIDF weights is another extractive summarization method, and thus faces some of the same formatting and grammatical challenges as vector word embedding. Another slightly peculiar result, was that two of the sentences in the overarching summary specifically mentioned the state of California. Specifically mentioning a single state is not conducive to an overarching summary, the explanation as to why California was mentioned in two sentences specifically is possibly that those sentences had extremely low collective TFIDF weights not including the word "California", and thus, when including "California" the collective TFIDF weights still remained relatively low to that of other sentences. Despite those two sentences, the overarching summary did provide some abstract guidelines as to what businesses are allowed to remain open, specific operating instructions for those businesses, and also urged residents to remain indoors whenever possible.

## VII. Threats to Validity

The main threat to validity of this project is that there is not a clear cut method for determining what a good human written summary is. The human summary is necessary to compare the computer generated one to, but depending on how it is written, it can drastically effect the ROUGE score of the computer generated summary. A computer generated summary that is informative and concise could have a low ROUGE score if it is compared with a human summary that contains no substantial content. For this reason, the ROUGE score of a summary is indirectly subject to the opinion of the person/people testing it, as they decide what a good human written summary is to compare with.

A second threat to validity is the use of numerous Python packages and libraries to perform this project. However, given that these packages and libraries are widely popular and developed by professionals and top research teams, it is unlikely that the source code behind them is compromised or faulty.

## References


[1] Li, "Automatically Summarize Trump's State of the Union Address", https://towardsdatascience.com/automatically-summarize-trumps-state-of-the-union-address-6757c6af6534

[2] Padmakumar, Saran, "Unsupervised Text Summarization Using Sentence Embeddings", University of Texas at Austin, https://www.cs.utexas.edu/~aish/ut/NLPProject.pdf

[3] Maples, "The ROUGE-AR: A Proposed Extension to the ROUGE Evaluation Metric for Abstractive Text Summarization", Symbolic Systems Department, Stanford University, https://web.stanford.edu/class/archive/cs/cs224n/cs224n.1174/reports/2761938.pdf

[4] Malik, "Text Summarization with NLTK in Python", https://stackabuse.com/text-summarization-with-nltk-in-python/

[5] Mervosh, Sarah, et al., "See Which States and Cities Have Told Residents to Stay at Home." The New York Times,*NYTimes.com*, https://www.nytimes.com/interactive/2020/us/coronavirus-stay-at-home-order.html

[6] Roberts, "Python ROUGE Implementation",https://github.com/google-research/google-research/tree/master/rouge

”Entities that typically provide food services to members of the public may continue to do so under this Third Supplementary Proclamation on the condition that the food is provided on a pick-up, delivery or takeaway basis only.

Work in essential businesses or operations Persons may travel to and from the following essential businesses and operations to the extent that such businesses or operations cannot be conducted through remote technology from homes or places of residence.

Labor Union essential activities including the administration of health and welfare funds and personnel checking on the

well-being and safety of members providing services in essential businesses and operations – provided that these checks should be done remotely where possible;

23. Essential businesses and operations shall implement separate operating hours for elderly and high risk customers. This includes establishments that sell groceries, medicine, including medication not requiring a medical prescription, supplies for children under the age of five and also that sell other non-grocery products, and products necessary to maintain the safety, sanitation, health and essential operation of residences and essential businesses and operations;

3. Travel to engage in minimum basic operations of non-essential businesses, including the minimum necessary activities to maintain the value of the business's inventory, ensure security, process payroll and employee benefits, and related functions as well as the minimum necessary activities to facilitate employees of the business being able to continue to work remotely from their residences;

4. Entities that provide food services under this exemption shall not permit the food to be eaten at the site where it is provided, or at any other gathering site due to the virus's propensity to physically impact surfaces and personal property;

14. Essential businesses and operations shall post online whether a facility is open and how best to reach the facility and continue services by phone or remotely.

All Persons in the State Must Stay at Home or in Their Place of Residence Pursuant to sections 127A-12(a)(5), 127A-12(a)(14), 127A-13(a)(1), and 127A-13(a)(7), HRS, all persons within the State of Hawai'i are ordered to stay at home or in their place of residence except as necessary to maintain continuity of operations of the federal critical infrastructure sectors, as identified at https://www.cisa.gov/identifying-critical-infrastructure-during-covid-19 and as further designated below or by the Director of the Hawai'i Emergency Management Agency (HIEMA).

All persons may leave their home or place of residence only for essential activities or to engage in the essential businesses and operations identified herein.”

Fig. 4. Computer Generated Summary of Hawai'i Shelter In Place Order

”Section Two: Non-essential business and operations must cease.

This Order is being issued to protect the public health of Californians.

Schools and public libraries may be used for Essential Government Functions and food distribution.

Subject to paragraph IV.c, all Non-Essential Businesses shall remain closed to the general public.

Businesses that supply products and services needed for people to work from home;

xvi. Businesses that supply products and services needed for people to work from home;

xvi. Libraries shall close for all in-person services, but may continue to provide on-line services and programming.

The Division of Public Health and State or local police shall have the authority to enforce this Order.

Individuals may only leave their residence to carry out Essential Activities or Essential Travel.

Essential businesses and operations shall implement separate operating hours for elderly and high risk customers.

Essential state and local government functions will also remain open, including offices that provide essential government services.

3) The Office of Emergency Services is directed to take necessary steps to ensure compliance with this Order.

In addition, businesses that are permitted to remain open include those granted exemptions prior to or following the issuance of this Order.

All persons may leave their homes or place of residence only for Essential Activities all as defined below.

Essential businesses and operations shall post online whether a facility is open and how best to reach the facility and continue services by phone or remotely.

Travel necessary for the provision or receipt of essential services described in Section 2, including employees, volunteers, and service recipients of these services.

Certain individuals must continue to work outside their residences to provide goods and services critical to our response to the COVID-19 epidemic.

“Essential businesses” may remain open provided they minimize their operations and staff to the greatest extent possible. ”

Fig. 5. Overarching Summary of Shelter in Place