Leveraging Deep Learning and NLP for Religious Exploration

Nick Kalenichenko, Eve Schoenrock, Rachel Grace Treene
University of Virginia
nhk6up, ufu2rg, rg5xm@virginia.edu

Motivation

As we continue through the 3rd decade of the 21st century, technology continues to explode and people expect a high level of functionality with regard to everyday applications. One major application that all web-users make use of is the search function. The operability of the search function serves as an indicator of company technological savvy, and it can either frighten new users away or draw new users to explore further. We implement a basic semantic search application on religious documents as a proof of concept for the value that sophisticated search capability can provide to businesses and other organizations. Using the *multi-qa-MiniLM-L6-cos-v1* sentence transformer model and various extractive summarization techniques, we compare the outputs from three major religious texts and offer insight into the usefulness of deep learning related Natural Language Processing (NLP).

1 Introduction

Hundreds of religions exist, each with their own doctrine, and for many people it can feel daunting to enter the realm of religious exploration because of the sheer quantity and density of the subject matter. We seek to demonstrate deep NLP techniques, while simultaneously alleviating the burden of religious comparison.

1.1 Dataset

Our dataset consists of the text from the Bible, Book of Mormon, and Qur'an. Each religious text is parsed according to its basic structural components. The Bible has 31,101 rows of text at the verse level. When clumped into chapters, the Bible has 1,189 rows of text. The Book of Mormon has 41,995 rows of text at the verse level, and when clumped into chapters it has 1,582 rows. The Qur'an 6,236 rows. When combined into surahs, it has 134 rows.

We develop a semantic search function to investigate the contents of each corpus individually, outputting finalized search results specific to each text. The format of our output promotes ease of comparison across religious doctrine, which enhances user engagement with each religion. Ultimately, semantic search serves as deep-dive investigative tool for the Bible, Book of Mormon, and Qur'an, and it enables quicker and more efficient religious exploration. We also introduce extractive summarization to the book/chapter levels of the Bible and Book of Mormon and the juz/surah levels of the Qur'an. The summaries offer an easily digestible format of each work, which expedites religious exploration and comparison. Each summary is evaluated using ROUGE techniques.

2 Semantic Search

We implement basic semantic search using the `multi-qa-MiniLM-L6-cos-v1` sentence transformer model, which was designed particularly for advanced semantic search. It maps sentences to a 384 dimensional dense vector space. The `multi-qa-MiniLM-L6-cos-v1` model was pre-trained on 215 million question and answer pairs. The model makes use of `MultipleNegativesRankingLoss` using mean-pooling, cosine similarity, and 20-scale. Although this model is highly functional, it is important to note that it is limited to maximum encodings of 512 words - any text that exceeds this limitation is truncated; this does not present an issue to our work because the text chunks at the verse/ayah level do not exceed 512 words.

We create three sets of embeddings at the verse/ayah level, one for the Bible, one for the Book of Mormon, and one for the Qur'an. These embeddings are saved to be reused throughout our project.

Next, we define a function called `search_books` which takes a search concept and number of desired verse/ayah results as inputs. We use the `multi-qa-MiniLM-L6-cos-v1` model along with semantic search functions to find search hits -

indices of each book that best match the search concept. These indices are used to reference the corresponding verses and ayahs, and the search results are outputted for the user to preview.

For example, say we perform a semantic search using the `search_books` function, with the concept being "love and peace in the world" and k=5. The resulting output shows the following Bible verses:

1. Psalms 85:10
   Mercy and truth are met together; righteousness and peace have kissed [each other].

2. Ephesians 6:23
   Peace [be] to the brethren, and love with faith, from God the Father and the Lord Jesus Christ.

3. Romans 14:19
   Let us therefore follow after the things which make for peace, and things wherewith one may edify another.

4. Luke 2:14
   Glory to God in the highest, and on earth peace, good will toward men.

5. 2 Thessalonians 3:16
   Now the Lord of peace himself give you peace always by all means. The Lord [be] with you all.

The following verses from the Book of Mormon are outputted:

1. Helaman 5:47
   Peace, peace be unto you, because of your faith in my Well Beloved, who was from the foundation of the world.

2. Psalms 85:10
   Mercy and truth are met together; righteousness and peace have kissed each other.

3. Ephesians 6:23
   Peace be to the brethren, and love with faith, from God the Father and the Lord Jesus Christ.

4. Romans 14:19
   Let us therefore follow after the things which make for peace, and things wherewith one may edify another.

5. 2 Thessalonians 3:16
   'Now the Lord of peace himself give you peace always by all means. The Lord be with you all.

See that there is some overlap in the results between the Bible and the Book of Mormon; this makes sense because the two texts have significant identical overlap. The ayahs from the Qur'an translated into English are below:

1. 19:33 Qur'an
   So peace is on me the day I was born, the day that I die, and the day that I shall be raised up to life (again)!

2. 56:26 Quran
   Only the saying, "Peace! Peace."

3. 10:10 Quran
   (This will be) their cry therein: "Glory to Thee, O Allah!" And "Peace" will be their greeting therein! and the close of their cry will be: "Praise be to Allah, the Cherisher and Sustainer of the worlds!"

4. 37:181 Quran
   And Peace on the messengers!

5. 37:120 Quran
   "Peace and salutation to Moses and Aaron!"

The search results show that the search has been contextualized, and results pertaining to peace, love, and the world are shown in conjunction with each other.

3 Summarization

There are two major summarization techniques that employ deep learning methods - extractive and abstractive. Extractive text summarization ranks sentence importance within a corpus and outputs those that are deemed most significant as the summary. It is called extractive because sentences are extracted directly from the original text and simply re-ordered and/or omitted. Abstractive text summarization develops novel sentences by rephrasing and inserting new language into the resulting summary. [3] Because abstractive summarization is complex and a relatively new form of text summarization, we choose to implement extractive summarization. However, there is value in understanding the existence of other summarization methods in the context of our work.

There are multitudes of extractive text summarization techniques. We proceed with three differing extractive summarizers, two of which rely on deep learning - Lex Rank and Latent Semantic Analysis (LSA) - and one that does not - Kullback Leibler Divergence (KL-Divergence). With each summarizer, we create summaries for each book/chapter combination in the Bible and Book of Mormon and each juz/surah combination in the Qur'an.

Summaries are beneficial for readers because they provide an avenue for skimming texts for overall concepts rather than diving into specifics. This is especially important in the context of religious exploration because of the intimidating and dense nature of religious doctrine.

For the purpose of brevity, we will only show example output from the Bible summaries in this report, but note that summarization has been implemented on all three texts. Below is the original text from the Bible that we used the summarizers on, whose output will be shown later:

- Matthew 2: "Now when Jesus was born in Bethlehem of Judaea in the days of Herod the king, behold, there came wise men from the east to Jerusalem, Saying, Where is he that is born King of the Jews? for we have seen his star in the east, and are come to worship him. When Herod the king had heard [these things], he was troubled, and all Jerusalem with him. And when he had gathered all the chief priests and scribes of the people together, he demanded of them where Christ should be born. And they said unto him, In Bethlehem of Judaea: for thus it is written by the prophet, And thou Bethlehem, [in] the land of Juda, art not the least among the princes of Juda: for out of thee shall come a Governor, that shall rule my people Israel. Then Herod, when he had privily called the wise men, enquired of them diligently what time the star appeared. And he sent them to Bethlehem, and said, Go and search diligently for the young child; and when ye have found [him], bring me word again, that I may come and worship him also. When they had heard the king, they departed; and, lo, the star, which they saw in the east, went before them, till it came and stood over where the young child was. When they saw the star, they rejoiced with exceeding great joy. And when they were come into the house, they saw the young child with Mary his mother, and fell down, and worshipped him: and when they had opened their treasures, they presented unto him gifts; gold, and frankincense, and myrrh. And being warned of God in a dream that they should not return to Herod, they departed into their own country another way. And when they were departed, behold, the angel of the Lord appeareth to Joseph in a dream, saying, Arise, and take the young child and his mother, and flee into Egypt, and be thou there until I bring thee word: for Herod will seek the young child to destroy him. When he arose, he took the young child and his mother by night, and departed into Egypt: And was there until the death of Herod: that it might be fulfilled which was spoken of the Lord by the prophet, saying, Out of Egypt have I called my son. Then Herod, when he saw that he was mocked of the wise men, was exceeding wroth, and sent forth, and slew all the children that were in Bethlehem, and in all the coasts thereof, from two years old and under, according to the time which he had diligently enquired of the wise men. Then was fulfilled that which was spoken by Jeremy the prophet, saying, In Rama was there a voice heard, lamentation, and weeping, and great mourning, Rachel weeping [for] her children, and would not be comforted, because they are not. But when Herod was dead, behold, an angel of the Lord appeareth in a dream to Joseph in Egypt, Saying, Arise, and take the young child and his mother, and go into the land of Israel: for they are dead which sought the young child's life. And he arose, and took the young child and his mother, and came into the land of Israel. But when he heard that Archelaus did reign in Judaea in the room of his father Herod, he was afraid to go thither: notwithstanding, being warned of God in a dream, he turned aside into the parts of Galilee: And he came and dwelt in a city called Nazareth: that it might be fulfilled which was spoken by the prophets, He shall be called a Nazarene."

3.1 Lex Rank

A graph-based method of summarization, *LexRank* is an unsupervised and extractive summarization technique that relies on cosine similarity between sentences. Similarity scores are used to generate a weighted graph of sentences, and the sentences that are perceived as being most similar are returned in the form of a summary. [2] In this way, *LexRank* evaluates and determines sentence importance when forming a summary.

Lex Rank relies on deep learning because of the nature of identifying high eigenvector centrality nodes. *Lex Rank* is unsupervised, so it relies on an encoder-decoder framework that maps graphical nodes to their respective eigenvector centralities. [6] By evaluating the proposed scores and sorting sentences, we achieve coherent summaries.

We employ *Lex Rank* on each book/chapter combination in the Bible and Book of Mormon and on each juz/surah combination in the Qur'an. An example from the Bible is shown below:

- *Lex Rank* Summary: "Now when Jesus was born in Bethlehem of Judaea in the days of Herod the king, behold, there came wise men from the east to Jerusalem, Saying, Where is he that is born King of the Jews? And when they were come into the house, they saw the young child with Mary his mother, and fell down, and worshipped him: and when they had opened their treasures, they presented unto him gifts; gold, and frankincense, and myrrh."

It is clear that the *Lex Rank* summary significantly shrinks the text, and in this example the function appears to have prioritized sentences appropriately. *Lex Rank* identifies the plot of the entire chapter and dissects it into a digestible chunk.

3.2 Latent Semantic Analysis

Latent Semantic Analysis, *LSA*, makes use of Singular Value Decompisition (SVD) and serves as a semantic, unsupervised approach to extractive text summarization. It captures importance by reducing the input data to a lower dimension and subsequently performing spatial decomposition. *LSA* summaries are more focused on topics and the latent features of the document rather than sentence similarity. However, one drawback of *LSA* in text summarization is its reliance on SVD, which is computationally complex and breaks down with larger, complex datasets.

LSA makes use of SVD, which is a method of matrix decomposition that is commonly used in deep learning applications, specifically deep neural networks (DNN). In SVD, one matrix of any shape is broken down into three new matrices - a unitary matrix, rectangular diagonal matrix, and the conjugate transpose of another complex unitary matrix. In the case of DNN, SVD impacts size by reducing massive weight matrices via these methods. [8]

For the purpose of extractive summarization, *LSA* performs similarly to Principal Component Analysis (PCA) as it reduces the dimensionality of sentence vectors to a specified k. Documents are also represented as vectors in k-dimensional space, based on the geometric sum of sentence vectors contained in a given document. The presence of each vector type in *LSA* enables sentence-sentence, document-sentence, and document-document comparisons, which all aid in determining semantic similarity according to cosine - larger angles indicating greater similarity between the going comparison. [4]

We employ *LSA* summarization on each book/chapter combination in the Bible and Book of Mormon and on each juz/surah combination in the Qur'an. An example from the Bible is shown below:

- *LSA* Summary: "For we have seen his star in the east, and are come to worship him. When Herod the king had heard [these things], he was troubled, and all Jerusalem with him."

LSA summarization decreases the original text by an even greater amount than Lex Rank. It appears to have done a worse job determining the plot of the chapter, but it clearly acknowledges the themes of worship and worry. This aligns with the anticipated performance of *LSA* summarization because it is centered around topic space.

3.3 Kullback-Liebler Divergence

Kullback-Liebler, or *KL-Divergence*, measures the difference between probability distributions. In terms of text summarization, it measures unigram probability distributions between the given document and the generated document. *KL-Divergence* text summarization methods greedily select sentences with lower *KL-Divergence* to generate summaries, which often provides an encompassing and relevant summary of the original document. *KL-Divergence* is written as:

$$D_{KL}(p||q) = \sum p(i) log \frac{p(i)}{q(i)}$$

where i indicates a particular word in a vocabulary and p and q are two differing distributions of words. [7] The goal with *KL-Divergence* is to minimize the score, so intuitively if p and q are the same, the score will be 0. Note that *KL-Divergence* does not make use of deep learning techniques, but it is incorporated into our analysis as a comparative point for summarization techniques that do make use of deep learning.

We implement *KL-Divergence* summarization on each book/chapter combination in the Bible and Book of Mormon and on each juz/surah combination in the Qur'an. An example from the Bible is shown below:

- *KL-Divergence* Summary: "And when he had gathered all the chief priests and scribes of the people together, he demanded of them where Christ should be born. And he arose, and took the young child and his mother, and came into the land of Israel."

KL-Divergence summarization shrinks the chapter similarly to LSA summarization, but it appears to have not caught the plot of the chapter nor significant topics.In fact, it pastes two sentences together that speak of different male figures in the story, but the summary makes it appear as though it is the same male. Although *KL-Divergence* does not always perform in this way, in this example we find that it provides the poorest result.

4 Evaluation

We use *Recall-Oriented Understudy for Gisting Evaluation* (ROUGE) scores as our evaluation metric on the chapter/surah summaries we provide. ROUGE is typically used to determine summary quality by comparing the machine-generated summaries against human-generated "ideal" summaries, but in our case we simply compare the generated summaries against the original document. We understand that this comparison reduces ROUGE scores naturally, but our work is limited by not having "ideal" summaries, so we adjust for this discrepancy as we evaluate summary outputs.

For each summary, we calculate `ROUGE-1`, `ROUGE-2`, `ROUGE-L`, and `ROUGE-L-SUM`. Because our usage compares summaries to the original document, we will refer to

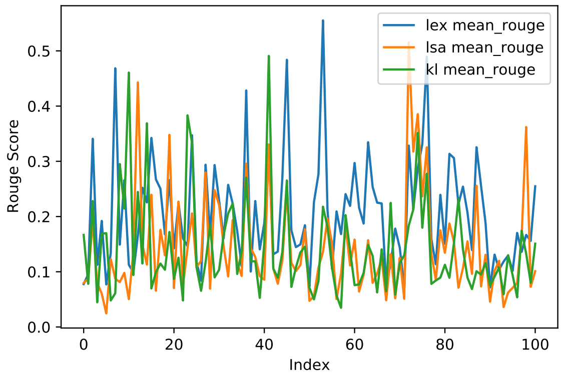

Figure 1: Bible - Rouge Scores from the First 100 Chapters

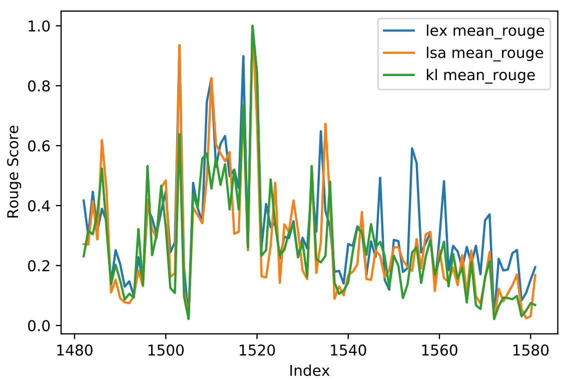

Figure 2: Book of Mormon - Rouge Scores from Last 100 Chapters

comparisons between original documents and summaries. ROUGE scores first compare n-gram overlap - `ROUGE-1` measures unigram overlap and `ROUGE-2` measures bigram overlap between the original document and the generated summary. `ROUGE-L` measures the longest matching sequence of words using Longest Common Subsequence (LCS). One advantage of this approach is that LCS does not require consecutive matches, but in-sequence matches that reflect the sentence level word order. [5] `ROUGE-L-SUM` is similar to `ROUGE-L`, but it divides text based on newlines.

To benefit from each of the different ROUGE evaluation types, we determine the mean ROUGE score, including `ROUGE-1`, `ROUGE-2`, `ROUGE-L`, `ROUGE-L-SUM`, for each chapter/surah. The results for each summary type are show in Figures 1-3. Note that only 100 chapters/surah are shown for better plot visualization. In the Bible, we see that ROUGE scores differ across summary type, with the highest scores being attributed to Lex Rank and LSA. The Book of Mormon has

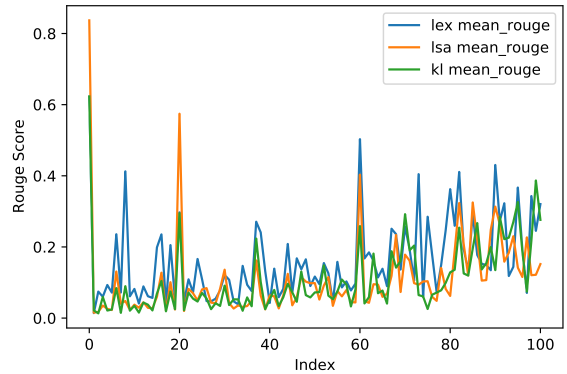

Figure 3: Qur'an - Rouge Scores from the First 100 Surahs

ROUGE scores that follow similar patterns, with some spikes in KL-Divergence in the last 40 chapters. The Qur'an has ROUGE scores much lower than those found in the Bible and Book of Mormon. This suggests that the summarizers are abbreviating the text in the Qur'an more severely. This could pose a concern, as there is less information available within these summaries and thereby bias. However, this is simply an evaluation of 100 chapters/surah, and patterns and scores change throughout each text.

Because we are comparing machine-generated summaries to original documents, we find that some summaries completely match the original document - this indicates that no summarization occurred. This is cause for concern and demonstrates a flaw in our methodology. With this in mind, it is important to evaluate our ROUGE scores with the idea that scores in the middle likely represent the best summaries. High scores indicate overfitting, and low scores suggest incompletion.

5 Related Work

Our work is motivated by multiple works across deep learning techniques of semantic search and extractive summarization.

For semantic search, we drew inspiration from Deshmukh and Sethi's work on on uses of SBERT. In their work, SBERT is used to learn the context of a query to perform semantic search over a corpus. [1] Because we use SBERT as our semantic search methodology, their work was incredibly relevant to our project.

Summarization implementation drew from multiple sources due to the use of three differing techniques. For Lex Rank summaries, we learned from Erkan and Radev's work that the computation of sentence importance based on eigenvector centrality as sentences are displayed graphically. [2] We implement this method understanding that cosine similarity is the basis of evaluation. Kireyev's work on LSA extractive summarization motivates the use of deep learning in deriving sensical summaries that explore topics and latent features of a document space. [4] Lastly, KL-Divergence text summarization relies upon Wu and Carberry's work as a measure to reduce when constructing extractive summaries. [7]

ROUGE scoring offers the primary form of summarization evaluation for our analysis, comparing machine-generated summaries against original text, which differs from normal ROUGE implementation and contrasts the work of Lin. [5] We also combine multiple ROUGE scores into one encompassing average, rather than evaluate each ROUGE score individually.

6 Conclusion

Ultimately, we find the results of semantic search to be relevant to the inputted concept, which serves as an anecdotal check on analysis. In terms of quantitative evaluation, ROUGE scores demonstrate the overlap between machine-generated summaries and the original text. For the sake of brevity but also capturing context, we assume that ROUGE scores ranging from .15 to .40 are ideal as they significantly reduce the text but likely maintain plot-related details. We see that ROUGE scores greatly fluctuate, which is to be expected as document content changes in complexity. However, we are concerned with the incredibly low ROUGE scores from our sample of the Qur'an - this is an area for further research.

Deep religious exploration is a difficult task, and the internet is full of articles and websites that lead to theological rabbit holes. Our work provides text-based interaction with dense religious content that limits human bias and enables users to draw conclusions regarding philosophical questions autonomously. Additionally, we offer a novel way of directly comparing big ideas across three major religious texts. By using semantic search in combination with summarization techniques, we are able to implement deep learning frameworks and heuristically evaluate the Christian, Mormon, and Muslim stances on varying concepts.

7 Metadata

The presentation of the project can be found at:

https://virginia.zoom.us/rec/share/8V6URxXtkGP7vYU84hF3Ri0PC5pkE5-nzQ7Vp_pRAK84AfZ1yNr2Div51t4XnZfU.vOcRG0szTRDo-hVh?startTime=1681786931000

Passcode: E969$+%.

The code/data of the project can be found at:

https://github.com/eveschoen/deep-learning-religious-text

References

[1] Anup Anand Deshmukh and Udhav Sethi. Ir-bert: Leveraging bert for semantic search in background linking for news articles, 2020.

[2] G. Erkan and D. R. Radev. LexRank: Graph-based Lexical Centrality as Salience in Text Summarization. December 2004.

[3] Som Gupta and S. K. Gupta. Abstractive Summarization: An Overview of the State of the Art. *Expert Systems with Applications*, 121, 2019.

[4] Kirill Kireyev. Using Latent Semantic Analysis for Extractive Summarization. In *TAC*, 2008.

[5] Chin-Yew Lin. Rouge: A package for automatic evaluation of summaries. In *Text Summarization Branches Out*, 2004.

[6] Appan Rakaraddi and Mahardhika Pratama. Unsupervised learning for identifying high eigenvector centrality nodes: A graph neural network approach. *CoRR*, abs/2111.05264, 2021.

[7] Peng Wu and Sandra Carberry. Toward Extractive Summarization of Multimodal Documents. In *Proceedings of the Workshop on Text Summarization at the Canadian Conference on Artificial Intelligence*, 2011.

[8] Jian Xue, Jinyu Li, and Yifan Gong. Restructuring of Deep Neural Network Acoustic Models with Singular Value Decomposition. In *INTERSPEECH*, 2013.