Using Natural Language Processing to conduct Sentiment Analysis and Text Mining of The Gospels

by

Moronfolu R Durodola

ID: 2508846

A thesis presented for the degree of MSc in Data Science

Supervised by Dr. Phillip Smith

School of Computer Science
University of Birmingham, UK
October 2023

Abstract

This project employs Natural Language Processing (NLP) to conduct a comprehensive study of the four Gospels off the Holy Bible — Matthew, Mark, Luke, and John. The main purpose of this project is to provide light on the emotional nuances and thematic connections of the Bible. The investigation begins with Sentiment Analysis, utilising the capabilities of cutting-edge models such as VADER, LSTM, and BERT. The goal is to determine which model best captures the emotional tones in the Gospels. This project seeks to meet the different demands of devout Christians as well as scholarly researchers by promoting a greater comprehension of the scriptures. Based on the findings of sentiment analysis, we investigate the use of Cosine Similarity as a method for discovering congruent passages throughout the four Gospels, thus allowing us to reveal the overlapping accounts, contradictions, and interrelated stories. Essentially, this project bridges the gap between technology and spirituality by combining modern NLP techniques with biblical knowledge.

Keywords: Natural Language Processing (NLP), Sentiment Analysis, Cosine Similarity, Topic Modelling, Bible.

Contents

Chapter 1

Overview

1.1 Introduction

The aim of the work described in this Report is to present the results of the sentiment analysis conducted on the biblical text. This is implemented mainly using machine learning, deep learning, and transformer-based models. In addition to the traditional analysis of language patterns, and sentiment, we have uncovered a fascinating aspect of the books by utilising cosine similarities to find the most similar text for each verse in the four books. Our primary aim was to leverage advanced NLP techniques to unveil linguistic trends, sentiment tendencies, and distinctive features within these book books. By extending our analysis to identify the most similar text for each verse across books, we've unlocked a new dimension of understanding the interconnections and themes that transcend individual books.

1.2 Background - The Gospels

The Gospels are four accounts of Jesus' life and teachings written by four evangelists Matthew, Mark, Luke, and John. The Gospels differ in several aspects but agree on core points. For example, all four Gospels describe the account of Jesus' death and resurrection, but each presents it differently, from a different perspective and with a distinct message (Hays, 2018). Matthew, Mark, and Luke are also known as synoptic gospels due to significant parallels in expressions, substance, and structure. There is a 'two-source hypothesis' based on the observation that the content of Matthew and Luke were derived from Mark and a second unknown source (Petruzzello, 2023). The language and writing style of the gospel of John, however, varies noticeably. Aside from semantics, the gospels have differences in the intentions of the individual authors. Each book is written in different historical contexts and therefore were originally written for specific audiences (Clay, 2012). One way this is evident is the varied portrayal of Jesus Christ; in Mark, He is seen as a healer, miracle worker, teacher, in Matthew, the Messiah of the Jewish, Prophet, Luke, merciful, compassionate, prayerful, and in John, noble, powerful and divine ((Clay, 2012). Essentially, the books, being fundamental accounts of the life, teachings, and deeds of Jesus Christ, hold immense significance in Christian theology and provide essential historical and cultural context, thus it would be a great benefit to explore the deeper contexts and relationships between them.

Chapter 2

Literature Review

2.1 Introduction

This section provides an in-depth review of the literature surrounding the three main approaches used in sentiment analysis in 2.2.1 and goes on to look at how these methods are applied within the context of religous texts in 2.2.2.

2.2 Sentiment analysis

Sentiment analysis (SA) is a component of natural language processing (NLP) that is advancing rapidly and has applications in a variety of contexts such as opinions on politicians, products and services, marketing campaigns (Feldman, 2013), and more recently; medical reviews (Zhao et al., 2023), animal agriculture (Mahon et al., 2023), dialect (Shamsi and Abdallah, 2023) (Kaseb and Farouk, 2023), Social media (e.g. twitter) (Czeranowska et al., 2023) (Catelli et al., 2023), vaccines using twitter data and much more. SA can also be referred to as opinion mining although a few academics claim that sentiment analysis and opinion mining share a few minor conceptual differences. Opinion mining focuses more on monitoring and analysing the opinion of people about an entity (Tsytsarau and Palpanas, 2011), whilst sentiment analysis identifies the sentiment expressed within a text and investigates this (Piryani, Madhavi and Singh, 2017). SA is also defined as a process of finding, extracting and classifying emotions or opinions that are expressed in text form (Zhang, Wang and Liu, 2018) and is therefore useful to make informed decisions (Feldman, 2013).

There are three different approaches typically used when it comes to SA - Lexicon Based, Machine Learning, and a Hybrid (Wankhade, Rao and Kulkarni, 2022). Lexicon based method derives a polarity score by matching and weighting words in a text using a pre-defined dictionary or lexicons of terms with paired semantic score (Chen, Lee and Chen, 2019). The method has benefits of simplicity, computationally efficient and does not demand excessive training data (Zhao, Qin and Liu, 2010). However, it also lacks the ability to capture intricate and contextual links within the text (Wankhade, Rao and Kulkarni, 2022). Using machine learning for sentiment analysis requires labelled sentiment datasets to train models and make predictions on unseen data. This is done using classification algorithms such as Naive Bayes, Support Vector Machines and Decision trees (Wankhade, Rao and Kulkarni, 2022). This technique does not require a dictionary and presents a high level of precision

(Sudhir and Suresh, 2021); however, the effectiveness of this method may be constrained when working with sizeable datasets (Jin et al., 2023). The hybrid approach integrates both lexicon based and machine learning approach. It is better equipped to solve drawbacks accompanied with a single system. Although various tasks, may change the efficacy of the integrated model (Dang, Moreno-García and De la Prieta, 2021).

2.2.1 Sentiment Analysis methods

I Lexicon-based

The authors Adebanjo and Oji (2022) explore the impact of the energy crisis on UK citizens using a lexicon-based approach on opinions from social media users. The Valence Aware Dictionary for Sentiment Reasoning (VADER) and TextBlob tools are utilised. Results show that the VADER tool was more accurate in discovering the negative impact of this crisis. By merging sentiment normalisation with evidence-based combination functions, Jurek, Mulvenna and Bi (2015) suggested a unique lexicon-based method for Twitter data. As opposed to simply categorising the sentiments, the suggested approach evaluates the strength of the sentiment. Tweets about the English Defence League and the extent of disruption at their gatherings is used to demonstrate the suggested methodology. In another study (Kang, Yoo and Han, 2012), a new lexicon tool is proposed to conduct SA of restaurant reviews. The researchers demonstrate that an enhanced Naïve Bayes algorithm increases the recall and precision values by a maximum of 6% and 2% respectively compared to the original NB algorithm. The problem of positive classification accuracy being greater than negative classification accuracy is addressed.

II Machine Learning

Traditional machine learning algorithms are used to deduce the sentiment from social media in several studies. One of these includes research by A. Rahim et al. (2021) who investigated factors affecting positive ratings of a Malaysian Hospital using Naive Bayes, Support Vector Machines and Logistic Regression classifiers to analyse comments found on Facebook. Another study by Rustam et al. (2021) compares various ML models including XGboost and Decision trees to evaluate their performance analysing tweets relating to Covid19. The study goes further by comparing to the deep learning LSTM model, and finds that it achieves a lower accuracy compared to the traditional models. Numerous deep learning techniques are also used widely in sentiment analysis. Rodrigues et al.(2022) employ a variety of machine learning and deep learning approaches to analyse and detect spam tweets in real time. The NB classifier achieved 97% accuracy and LSTM performed well with 98% validation accuracy thus suggesting these models can be used for complex data. Convolutional Neural Network (CNN) model is used for sentiment analysis of film reviews and achieved an 81.5% accuracy (Kim, 2014). The results suggests that the use of CNN was an appropriate enhancement from traditional methods for NLP (Khurana et al., 2022). In addition to neural networks, pre-trained models have also become increasingly popular when conducting SA. Singh, Jakhar and Pandey (2021) attained a 94% validation accuracy when implementing a Bidirectional Encoder Representations (BERT) transformer model on tweets collected globally to analysis opinions on Covid19 to compare against tweets collected solely from India. The authors report that this model is used to better understand feelings and psychological conditions of people.

III Hybrid

Tan et al. (2022) proposes a deep learning hybrid model for sentiment analysis, comprising of the neural network Long Short-Term Memory (LSTM) and transformer model, BERT using various datasets. The authors find that their suggested model outperforms traditional, recognised models including LSTM, and Gated recurrent unit (GRU) by recording an improvement of the F1 scores to 91%, 93% and 90% for the Twitter US Airline Sentiment, IMDB, and Sentiment 140 datasets respectively. Similarly, Cai et al. (2019) remains consistent with the outcome of the hybrid model attaining better results as opposed to a single model. In this instance, support vector machines (SVM) and the gradient boosting tree (GBDT) are integrated to carry out opinion classification, producing a significantly higher precision and F1 value compared to the baseline models. Rajeswari et al. (2020) use three review datasets to demonstrate that the use of a lexicon tool, SentiWordNet, in hybrid models allows for classification of neutral polarity. Accuracy is also improved when combined with machine learning classifiers including Naive Bayes, SVM, Decision Tree and Logistic Regression. This study alleviates any confusion or uncertainty within businesses that receive seemingly neutral reviews and allow for better informed decisions to be made. A hybrid model of VADER and Random Forest is proposed by Mardjo and Choksuchat (2022) to gain more insights into the thoughts surrounding cryptocurrencies from over 3.6 million tweets. The model is shown to produce consistently stable results with 75% accuracy, 70% precision, 88% recall, and F1-score of 78%. From these studies it is clear that the hybrid approach is reliable in producing desirable results when undertaking sentiment analysis in the various contexts.

2.2.2 NLP in Religious Text

From above it is evident that there is an abundant of sentiment analysis research completed using social media as a tool to analyse the opinions of people on certain issues or in the form of reviews of various businesses. However, research on printed work (i.e. fiction, non-fiction or in this case, religious text) is lacking, despite the popularity of certain publications. The following sections provides an insight to a handful of study that has been conducted in the context of sacred texts.

Researchers have used text mining techniques to investigate various religious texts. McDonald (2014) uses nine religious texts including KJV Bible (new and old testament, the Book of Mormon, and the Torah to identify relationships, links and differences between the texts by extracting noun and verb phrases and using these as input to a Self-Organising Map. Varghese and Punithavalli (2019) instead focus on only three texts; the Bible, the Quran and the Tanakh due to the similarities in background, for example, all originate from the Middle East and are writing by many authors. Some methods used in this study include text classification using NB and SVM and text similarity by measure of Cosine similarity Jaccard similarity, Euclidean distance, and term frequency-inverse document frequency (TF-IDF). Sentiment analysis is also employed using the python library TextBlob. The main conclusion to note from this research is that retrieving information from sacred texts may be challenging, possibly due to the multi-layered meanings behind the text. It also suggests that deep learning methods would be a better tool to extract “better semantic information”. A fairly recent study (Franklin, 2018) examines the differences in sentiment of various versions of the bible using the NRC word association lexicon and Syuzhet sentiment analysis package in R. The results suggested “more dynamic translations” were viewed as more positive in comparison to other translations. Interestingly, it suggests that the NKJV and KJV had the least emotional tone present due to the language being more formal. Whissell (2019) investigates the changes in emotion throughout the stories within gospels using the Dictionary of Affect in Language, a tool for the statistical analysis of individual words according to the emotion conveyed (Whissell, 2009). Other studies (Barrett, 2019; Buentello, 2018; Buentello, 2020) based on SA have also been conducted to explore contents of the Bible.

Chapter 3

Data Exploration

3.1 Introduction

This section details all steps taken to ensure the data is clean and in a suitable format for an in-depth NLP analysis (3.2) and discusses the exploration of the data (3.3).

3.2 Data Pre-processing

Data pre-processing is an essential first step to remove all unwanted information i.e. noise in the text data as the right method of processing can increase the accuracy of classifier used (Pradha, Halgamuge and Vinh, 2019). The following list presents the steps taken to process the data (Silva, 2023):

- Tokenization - This step breaks up the text into smaller portions called tokens and assists in model development.

- Remove punctuation (or any characters that are not letters)

- Remove stop words - Stop words are words that do not add much value to the content of the data semantically, and can therefore be removed to reduce the size of the dataset, hence reducing computation time required for NLP tasks. Examples are “a,” “the,” “is,” “are,” etc. however due to the nature of the data, extra stop words have also been defined to be removed in the process such as ”thy”, ”thou”, ”hath”, etc.

- Lemmatization - This breaks down a word to its root meaning, known as the lemma; for example “runs,” “running,” and “ran” would be mapped to the lemma, “run.”

- All words are also converted to lowercase.

These steps are completed using the NLTK package in python.

3.3 EDA

Before going into the sentiment analysis, it is important to understand the dataset. This section does exactly this by looking at the structure of the data.

I Data summary

The data comprises of 3779 rows and 6 columns, each row representing a verse from the Gospels. Each book within the Gospels are then separated into four data sets for easier manipulations. The chapters within each book are counted and the book of Matthew is found to be the longest with 28 chapters, while Mark is the shortest with with 16 chapters.

II Visualisations

Visualisations are an excellent way to perceive the data from a different perspective. The first image produced is a word cloud of data after pre-processing has occurred (Figure 3.1). "Come" and "Came" are two prominent words in this word cloud demonstrating that the lemmatization step has not been completed adequately. Similar word clouds are also produced for each of the books within the Gospels.

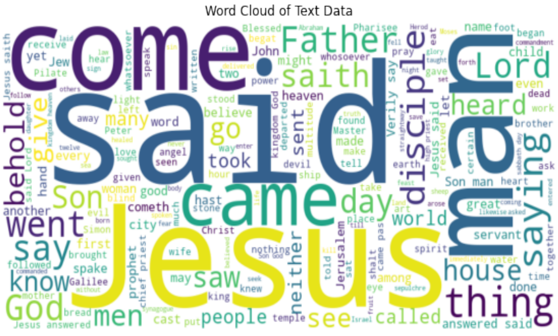
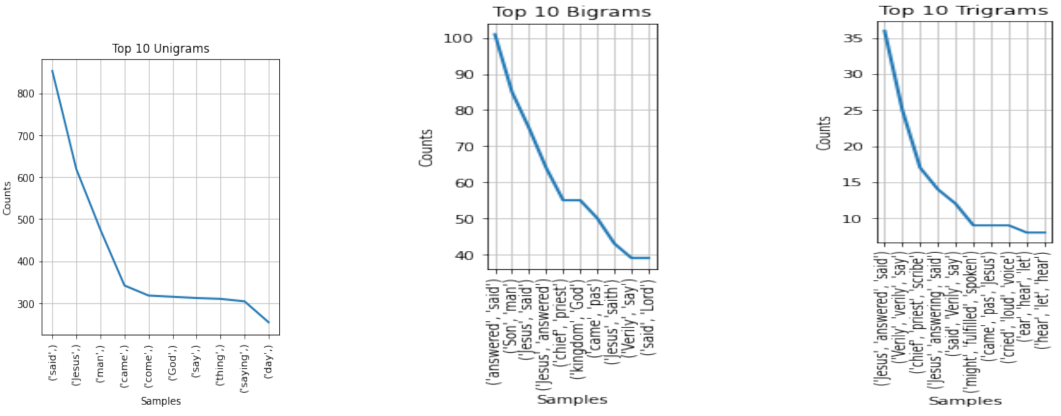


Figure 3.1: Word cloud of Text data

An N-gram is defined as a sequence of words of length n (Schonlau, Guenther and Sucholutsky, 2017). Figure 3.2 are three graphs displaying the most frequent 1 gram, 2 gram and 3 gram terms within the text.



(a) Top 10 unigrams (b) Top 10 bigrams (c) Top 10 trigrams

Figure 3.2: Three graphs of top 10 ngrams

The last plot produced (figure 3.3) is a frequency tree diagram which essentially displays the top words from the data. The larger tiles represent a higher word frequency, while a smaller tile indicates a lower frequency.

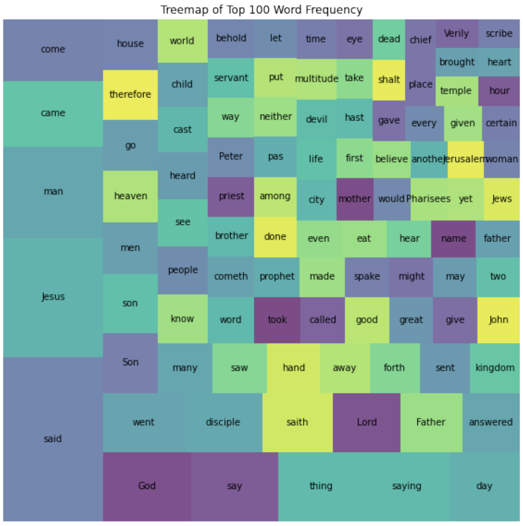


Figure 3.3: Frequency tree diagram

Chapter 4

Sentiment Modelling

4.1 Introduction

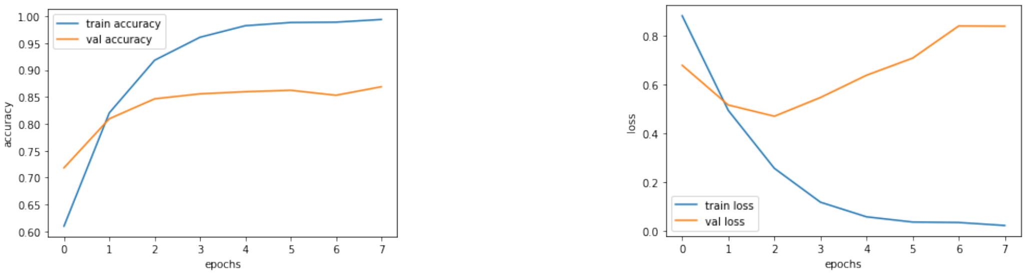
Sentiment analysis in NLP is a technique that involves using computational methods to determine the emotional tone or sentiment expressed in a piece of text. It aims to identify whether the sentiment is positive, negative, or neutral, helping to understand the overall opinion or attitude conveyed by the text. This section of the report presents the process and results of the SA of the Bible data. Section 4.2 goes through the machine learning models that are used, including the hybrid.

4.2 Machine Learning

4.2.1 LSTM

In sentiment analysis, commonly used neural network models include convolutional neural network (CNN), recurrent neural network and long-term memory network (LSTM). While CNNs often struggle to capture long-range dependencies due to their limited perceptual fields, RNNs are associated with gradient vanishing and memory constraints when handling extensive sequences. LSTM's innovative gating mechanism and memory units effectively resolve these issues, improving its performance (Abdul et al., 2019). A regular LSTM unit consists of a cell, input gate, output gate and forget gate. The input gate regulates information retention in memory. The output gate controls the data flow into the subsequent layer and the forget gate manages the rate of loss of stored memory and by deciding what data can be removed using the state of the cell (Sudhir and Suresh, 2021).

The model architecture implemented comprises of convolutional and max-pooling layers, which are known to do well in capturing hierarchical features within the text data. After compiling the model with the Adam optimizer and categorical cross-entropy loss, it is trained with early stopping to avoid overfitting and obtain optimal performance. On the test data, the model is evaluated by calculating loss and accuracy as seen in figure 4.1. A confusion matrix (figure 4.2) is also produced to visualise its class predictions, while additional metrics such as specificity, AUC ROC, precision, recall, and F1 score offer a comprehensive performance assessment (table 4.1). We finish by appending the LSTM-generated sentiment predictions to the dataset.



(a) LSTM Accuracy

(b) LSTM Loss

Figure 4.1: LSTM accuracy and loss plots

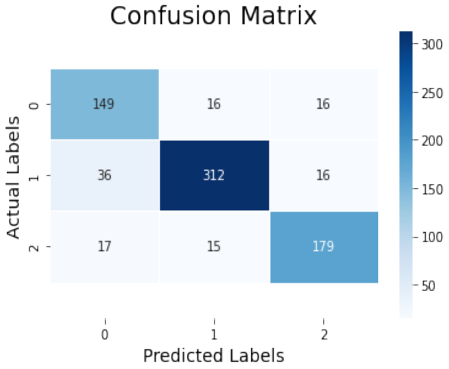


Figure 4.2: LSTM Confusion matrix

4.2.2 BERT

BERT (Bidirectional Encoder Representations from Transformers) is an NLP pre-trained model. It is an encoder stack of an encoder-decoder network that uses self-attention on the encoder side and attention on the decoder side. BERT has two models BERT base with 12 encoders stacked with 110 million parameters and BERT large model with 24 encoders stacked with 330 million parameters (Devlin et al., 2019). There is a two-stage training process; firstly in pre-training unlabelled data is used to train the model over numerous pre-training tasks then the parameters are fine-tuned using labelled data from the downstream tasks (Sudhir and Suresh, 2021). This flexibility poses to be a problem for typical NLP models (Wang et al., 2020) however it has enabled BERT to become a popular a leading choice for sentiment analysis and as well as other applications in NLP (Wankhade, Rao and Kulkarni, 2022).

In this project, the BERT model is loaded using the Hugging Face Transformers library, and a sentiment analysis pipeline is created. This pipeline employs the BERT large model due to the nature of the dataset. The data is then passed through the BERT model for analysis. Scores are assigned based on the sentiment of each verse, and weighted scores are calculated to distinguish positive and negative sentiments. This is then visualised using a plot of the average sentiment scores of chapters 4.3. The line plot illustrates how sentiment varies throughout different chapters of the Bible, with each book represented by a distinct line.

The values shown suggest that all chapters have a negative sentiment, indicating a potential error in the computation of this model.

Accuracy	Specificity	Precision	Recall	F1-Score	AUC ROC
0.71	0.85	0.71	0.69	0.70	0.85

Table 4.1: LSTM Evaluation metric results

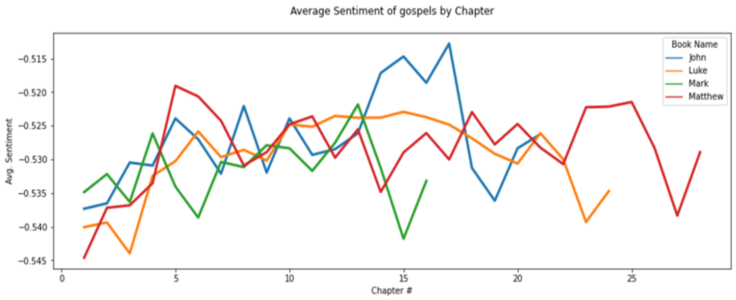


Figure 4.3: Line plot of BERT sentiment scores

4.2.3 Hybrid

VADER

The VADER (Valence Aware Dictionary and Sentiment Reasoner) is a tool built into the NLTK library and is widely used for simple sentiment analysis tasks. It works by returning a dictionary containing four sentiment scores: positive, negative, neutral, and compound. The compound score is a normalised score that combines the individual positive, negative, and neutral scores to give an overall sentiment score (Mardjo and Choksuchat, 2022). VADER is chosen as it reportedly performs at least as well as other esteemed sentiment analysis tools including SentiWordNet and TextBlob across multiple domains, being both fast and “computationally economical without sacrificing accuracy” (Hutto and Gilbert, 2014). Essentially, the hybrid algorithm uses the VADER tool to calculate polarity scores to label the data, then the supervised ML algorithms in section 4.2.3 is trained using these labels.

ML classifiers

Naïve Bayes

The Naive Bayes algorithm is a simple supervised learning technique that uses the probability of every attribute to generate predictions which works well for text data classification (Sudhir and Suresh, 2021). It is founded on the Bayes Theorem, which describes the likelihood of an event depending on past information (Zhang, 2004). Bayes' theorem is represented mathematically with equation 4.2.3:

$$P(A|B) = \frac{P(B|A) * P(A)}{P(B)} \tag{4.1}$$

where $P(A|B)$ is the conditional probability of A given B,
P($P(B|A)$ is the conditional probability or B given A, and
P(A) and P(B) is the probability of event A and B respectively.

Decision Tree A decision tree is a straightforward classifier with two steps: learning and classification. During the initial stage, the decision tree learns to construct a decision tree using a set of previously categorised training samples. The decision tree from the learning phase is then utilised to categorise unclassified data. The structure of the decision is formed with three parts: internal nodes, branches, and leaf nodes (Lan et al., 2020).

Random Forest The random forest (RF) classifier was suggested by Breiman (2001) to develop the decision tree algorithm further by making the process of selecting feature variables used for training random. The stochastic quality of this classifier reduces the chance of overfitting in the learning process thus rendering a better performance. Although classifications are deemed more accurate with the RF, it makes it harder to interpret due to a lack of knowledge on the important features (Lan et al., 2020).

XGBoost

Extreme Gradient Boosting (XGBoost) is a distributed gradient boosting library with an added feature that has been adjusted to be exceptionally efficient. It employs a supervised learning technique to forecast variables accurately by merging the predictions of multiple weaker models (Khan et al., 2022). XGBoost features regularisation approaches to control over-fittings and is scalable, ensuring that better performance and effective processing of larger data sets (Rustam et al., 2021). It utilises a Log Loss function that regards the likelihood of false classifications, assisting in loss reduction and accuracy boosting (Saha, 2019). The Log Loss is defined as follows:

$$logloss = \frac{-1}{N}\sum_{i=1}^{N} y_i \cdot log(p(y_i)) + (1 - y_i) \cdot log(1 - p(y_i)) \tag{4.2}$$

Support Vector Machines

SVM is a classification and regression prediction tool used to optimise prediction accuracy while simultaneously avoiding over-fitting of the data (Jakkula, 2011).

SVM is a classification and regression prediction tool used to optimise prediction accuracy while simultaneously avoiding over-fitting of the data (Jakkula, 2011).

This ML classifier works by reducing any architectural risks associated in the production of a high dimension hyperplane for class separation, with the goal of attaining the greatest margin (distance between the hyperplane and the observations nearest to the hyperplane) among them. The hyperplane enables the separation of positives from a set of negatives (AlBadani, Shi, and Dong, 2022). SVM is frequently applied in classification problems due to excellent accuracy and performance (Jakkula, 2011). SVM relies on spatial transformation functions (or kernel functions) to execute class separation.

Linear - When data is completely linearly separable, the linear SVM is typically utilised.

$$k(x_1, x_2) = x_1 \cdot x_2 \tag{4.3}$$

Poly - A polynomial mapping is a prominent non-linear modelling approach. This function is defined as follows:

$$k(x_1, x_2) = (x_1 \cdot x_2 + 1)^P \tag{4.4}$$

RBF - The radial basis function is an essential kernel function that is frequently used in pattern recognition difficulties. (Firmino Alves et al., 2014). The function is defined by:

$$k(x_1, x_2) = \epsilon^{-\Upsilon(x_1-x_2)^2} \tag{4.5}$$

Results

Accuracy, precision, recall, and F1-score are popular metrics used to evaluate machine learning classifiers. These metrics provide a comprehensive understanding of the classifier's performance, especially in scenarios where class imbalances might exist. For a well-rounded dataset, accuracy is an appropriate statistic to employ for sentiment classification (Wankhade, Rao, and Kulkarni, 2022). Precision is a metric associated with the exactness of the classification model (Rustam et al., 2021); it represents the strength of the prediction. (Ghadah Alamer, Sultan Alyahya and Hmood Al-Dossari, 2023), making it a suitable measure for the context of this research. Recall, also known as sensitivity, is a measure of the model's misclassifications. Precision and recall have an inverse relationship. As a result, increasing both Precision and Recall at the same time is unattainable (Wankhade, Rao, and Kulkarni, 2022). The weighted average of Precision and Recall is the F1-score (Samih, Abderrahim Ghadi, and Abdelhadi Fennan, 2023). The F1 score is used to regulate the trade-off between recall and accuracy. It is ideal for problems with unequal class distribution (Wankhade, Rao, and Kulkarni, 2022).

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)} \tag{4.6}$$

$$Precision = \frac{TP}{(TP + FP)} \tag{4.7}$$

$$Recall = \frac{TP}{TP + FN} \tag{4.8}$$

$$F1\ score = \frac{2(precision \times recall)}{(precision + recall)} \tag{4.9}$$

where; TP is true positive: number of positive prediction of positive instances, TN is true negative: the number of positive prediction of negative instances, FP is false positive: number of negative prediction of positive instances and FN is false negative: number of negative prediction of negative instances (Areeba Umair, Elio Masciari and Muhammad Habib Ullah, 2023).

Table 4.2 displays the results of the aforementioned evaaluation metrics tested against each of the machine learning classfiers following the application of the VADER tool to generate sentiment labels.

Focusing on accuracy alone, the linear SVM and Decision tree classifier provide the best results with an accuracy of 0.87 as shown in figure 4.4. These classifiers also demonstrate balanced precision, recall, and F1-Score values. The Random Forest classifier also performed well, with an accuracy of 0.86, showing the ability to capture complex relationships within the data. Surprisingly, the SVM with the polynomial kernel, often effective in similar tasks, had the lowest accuracy among all classifiers,

Classifier	Accuracy	Precision	Recall	F1-Score
Naive Bayes	0.78	0.78	0.78	0.78
Random Forest	0.86	0.87	0.86	0.86
Decision Tree	0.87	0.86	0.87	0.86
XGBoost	0.85	0.86	0.85	0.85
SVM (Linear)	0.87	0.87	0.87	0.87
SVM (RBF Kernel)	0.78	0.8	0.78	0.77
SVM (Poly Kernel)	0.53	0.7	0.53	0.45

Table 4.2: Results of ML classifiers

suggesting that a different kernel or parameter tuning might be needed for optimal performance. The Naive Bayes and SVM (RBF) classifier, while achieving a decent performance with an accuracy of 0.78, was outperformed by the other models in this specific analysis.

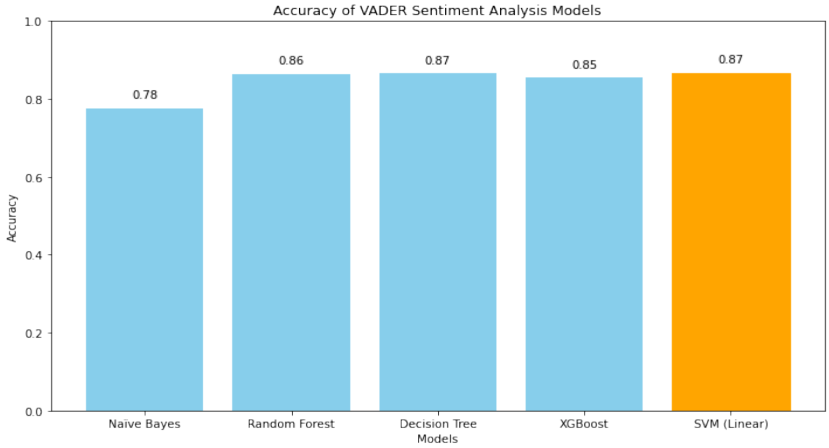


Figure 4.4: VADER accuracy results for each classifier

Figure 4.5 illustrates the confusion matrix of the linear support vector machine results.

Overall, the combination of VADER tool and SVM with the linear kernel achieves a great performance in the sentiment analysis of the gospels.

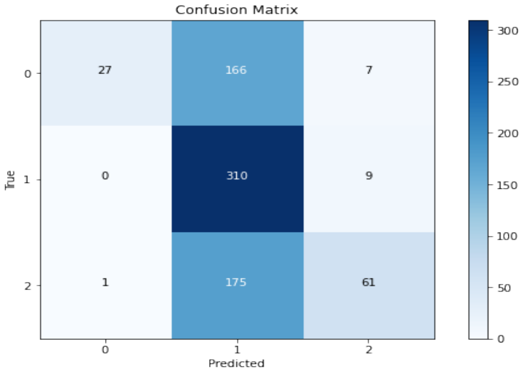


Figure 4.5: SVM Confusion matrix

4.3 Summary

In this section, we have seen the results of the sentiment analysis implemented on texts from the books of Matthew, Mark, Luke, and John. We found the hybrid model of VADER-SVM gave the best results from the evaluation metrics, performing better than the LSTM model. This analysis not only allowed us to gauge the sentiments within the scriptures but also to identify chapters that stood out emotionally.

Chapter 5

Text Similarity

5.1 Introduction

A comprehensive analysis of textual similarities between the books of Matthew, Mark, Luke, and John from the Bible is conducted in this section. The primary objective is to assess the textual parallels among the chapters of these biblical books and evaluate any differences or similarities in the content. Section 5.2 discusses the application of TF-IDF and cosine similarity to all four books to identify similarities between them.

5.2 TF-IDF algorithm and Cosine Similarity

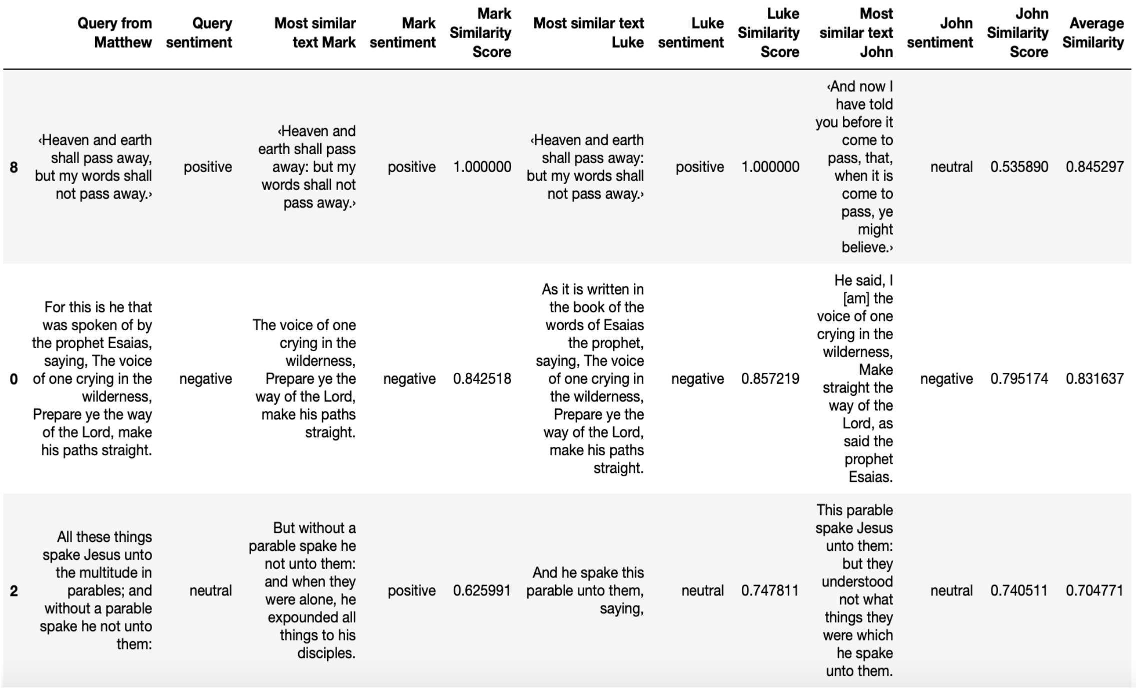
The first step in this quantitative text analysis involves transforming the text data into the Vector Space Model (VSM) by employing a bag of words (BOW) model where each text document (vectors in a multidimensional space) is expressed as a list of its words. These vectors are then arranged in a matrix, with rows representing terms or phrases and columns representing the documents. The next step involves using Term Frequency-Inverse Document Frequency (TF-IDF); a method used to assign each term in a document a specific weight based on their frequency in documents and across the entire corpus (Alodadi and Janeja, 2015). Finally, the cosine similarity is calculated using scores of two different documents' vectors that share the same representation (Salton, 1989) as follows:

$$similarity(A,B) = \frac{A \cdot B}{\|A\| \times \|B\|} = \frac{\Sigma_{i=1}^{n} A_i \cdot B_i}{\sqrt{\Sigma_{i=1}^{n} A_i^2} \cdot \sqrt{\Sigma_{i=1}^{n} B_i^2}} \tag{5.1}$$

where A and B are n-dimensional vectors (Li and Han, 2013). A perfect similarity will have a score of one, and no similarity will have a score of zero. The cosine similarity calculation is very efficient, and hence a very popular choice for various text mining problems, favoured over other methods such as the Euclidean distance-based metric, Jaccard and Dice (Salton, 1989).

5.2.1 Application

The cosine similarity scores are computed to allow for each book to be compared to the other three to gauge the degree of similarity. For example, figure 5.1 illustrates the a table produced from analysing the book of Matthew. Similar tables are also produced for the other books.



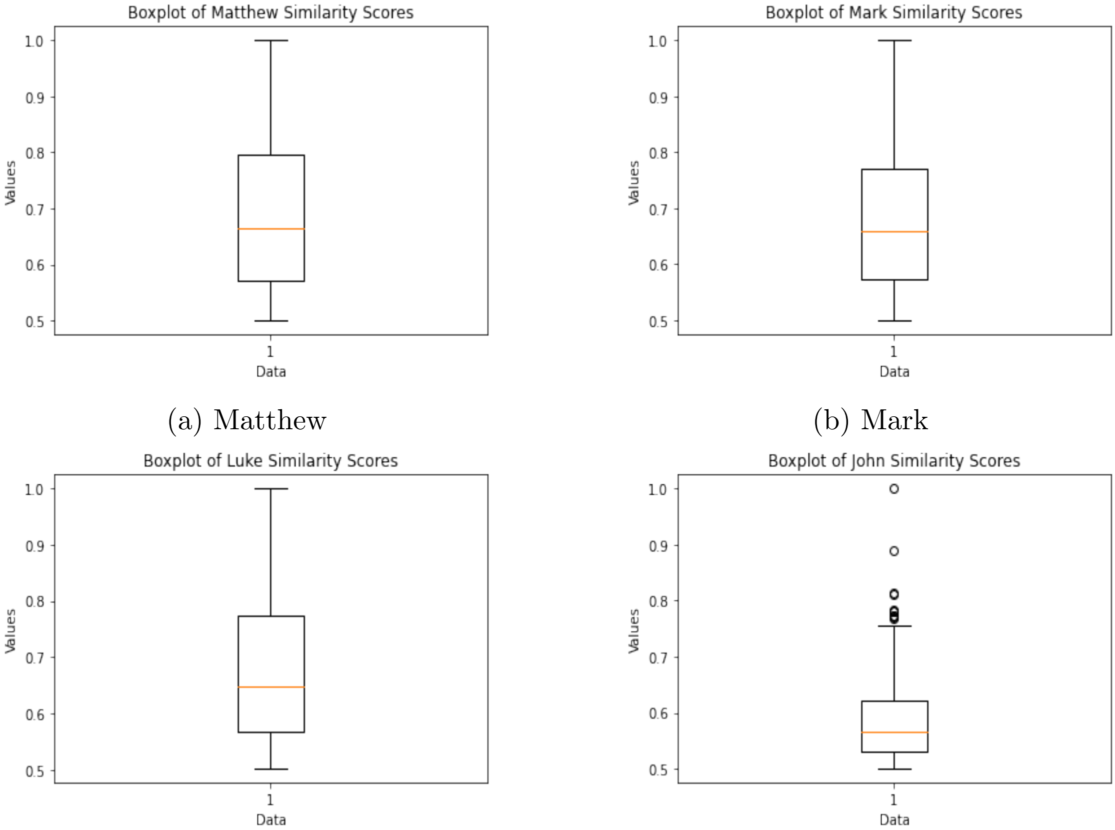
	Query from Matthew	Query sentiment	Most similar text Mark	Mark sentiment	Mark Similarity Score	Most similar text Luke	Luke sentiment	Luke Similarity Score	Most similar text John	John sentiment	John Similarity Score	Average Similarity
8	‹Heaven and earth shall pass away, but my words shall not pass away.›	positive	‹Heaven and earth shall pass away: but my words shall not pass away.›	positive	1.000000	‹Heaven and earth shall pass away: but my words shall not pass away.›	positive	1.000000	‹And now I have told you before it come to pass, that, when it is come to pass, ye might believe.›	neutral	0.535890	0.845297
0	For this is he that was spoken of by the prophet Esaias, saying, The voice of one crying in the wilderness, Prepare ye the way of the Lord, make his paths straight.	negative	The voice of one crying in the wilderness, Prepare ye the way of the Lord, make his paths straight.	negative	0.842518	As it is written in the book of the words of Esaias the prophet, saying, The voice of one crying in the wilderness, Prepare ye the way of the Lord, make his paths straight.	negative	0.857219	He said, I [am] the voice of one crying in the wilderness, Make straight the way of the Lord, as said the prophet Esaias.	negative	0.795174	0.831637
2	All these things spake Jesus unto the multitude in parables; and without a parable spake he not unto them:	neutral	But without a parable spake he not unto them: and when they were alone, he expounded all things to his disciples.	positive	0.625991	And he spake this parable unto them, saying,	neutral	0.747811	This parable spake Jesus unto them: but they understood not what things they were which he spake unto them.	neutral	0.740511	0.704771

Figure 5.1: Matthew Similarity Table

The box plots in figure 5.2 represents the middle 50% of average sentiment similarity each book in regards to the other books. There are many values which are exactly the same i.e. the books share many of the same verses and hence h a score greater that 0.8, sometimes 1. Most verses are vaguely similar to each other and hence lie in the 0.2 to 0.5 range. Although some verses have 0 similarity scores as well meaning sentences in other books are not at all similar to the one given. This is realised by the outliers in the box plot for the book of John, as we know that this book is very different to the other three in terms of content and language (1.2).

5.3 Summary

This study is especially useful because it offers a numerical measure of how closely connected the content of different chapters is among the four books of the gospels. It can support researchers and scholars in understanding thematic continuity and variability across these books. Furthermore, this study provides insights into how the emotional tone may alter throughout chapters, resulting in a greater comprehension of the the complexity of the bible. The research provides essential information for biblical academics, theologians, and anyone interested in the subtleties of the Gospels by examining the parallels and contrasts within.



(a) Matthew

(b) Mark

(c) Luke

(d) John

Figure 5.2: Boxplot of... with 0.5 threshold

Chapter 6

Conclusion

6.1 Limitations

The project has several notable limitations. Firstly, the inherent subjectivity of sentiment analysis is a key constraint that can't be ignored. Sentiment analysis relies on sentiment lexicons and pre-trained models that won't completely capture the of religious texts, therefore impacting the accuracy the analysis. Moreover, the fact that sentiments are typically limited to positive neutral and negative labels suggests an inadequate portrayal of emotional extent found within religious texts. Data pre-processing concerns are also relevant when discussing limitations. Removing antiquated language can be difficult, particularly for the NLTK library which is trained on modern language. The choice of pre-trained models is also an important consideration. Each model has its benefits and drawbacks, and their appropriateness for religious texts may vary. Additionally, the absence of ground truth labels for sentiments in scriptures introduces subjectivity and renders the veracity of derived labels unverifiable.

In regard to the methodology, this study consistently used just four evaluation criteria (accuracy, precision, recall, and F1-score). Additonal measures such as receiver operating characteristic (ROC) and area under the ROC curve (AUC) should be considered for all models for ease of comparison and better understanding of the model performance. Furthermore, the overfitting seen associated with the LSTM model for example could be less prominent with the practise of hyperparameter tunning. Of course, more comparative models could also be included for a more thorough review of the best option to conduct sentiment analysis on religious text.

6.2 Future work

Sentiment analysis of biblical texts is a unique and intriguing area for study. Firstly, probing into deeper semantic analysis is essential. Although present sentiment analysis primarily focuses on classifying sentiments into three categories, further study may include an increased complexity in the approach that identifies the multitudes of emotions, religious connotations, and historical context within the Bible. Developing a model that comprehends more complex emotions could provide deeper insights. Another interesting area for future research is the change of sentiment in biblical texts throughout time. The Bible has a long history with several different translations and interpretations. Analysing the changes in sentient can reveal the evolution of public attitudes regarding the texts as well as impact from historical events and religious movements. The application of sentiment analysis to biblical texts is likely to continue to expand as technology advances. This could include the analysis of sermons, bible readings and teachings via audio/video recordings in order to provide valuable insights to those who need it. Ultimately, sentiment analysis in the context of the Bible has huge potential for revealing significant emotional elements throughout the Bible. We may get a better grasp of the Bible's significant influence on emotion and spirituality by fine-tuning sentiment models, evaluating historical changes, and merging sentiment analysis with religious studies.

6.3 Project management

This project's execution was defined by thorough project management. Foundations were set by defining clear specifications according to the desired goals; starting with the formulation of research goals, then creating a detailed Gantt chart with distinct milestones to ensure everything would be completed in a timely manner. The project demonstrates attention to best practises in Natural Language Processing which is reflected in the selection of appropriate sentiment analysis models. These decisions meant that we could conduct a thorough study and draw meaningful comparisons among outcomes. This process of completing this project benefited from continuous communication with the project supervisor. As the project reached its culmination, it became increasingly tough to maintain focus and continue making effort due to physical and mental health issues that surfaced. These unexpected obstacles added another degree of complexity. The ability to navigate these challenges, adapt to the circumstances, and yet provide meaningful results illustrate resilience and determination. The core aims of this research has been met by establishing sentiment analysis models for the data which again demonstrates adaptability in unforeseen circumstances. Nevertheless, it is important to recognise the level of success of this project may fall short of the initial aspirations. In retrospect, despite the difficulties and changing conditions the research proves to be an important step towards the unexplored domain of sentiment analysis in the context of religious texts.

Bibliography

Feldman, Ronen (Apr. 2013). "Techniques and Applications for Sentiment Analysis". In: *Communications of the ACM* 56, p. 82. DOI: 10.1145/2436256.2436274.

Zhao, Yuehua et al. (Nov. 2023). "Construction of an aspect-level Sentiment Analysis Model for Online Medical Reviews". In: *Information Processing and Management* 60, pp. 103513–103513. DOI: 10.1016/j.ipm.2023.103513.

Mahon, Niamh et al. (Oct. 2023). "The Application of a Sentiment Analysis Approach to Explore Public Understandings of Animal Agriculture". In: *Journal of Rural Studies* 103, pp. 103127–103127. DOI: 10.1016/j.jrurstud.2023.103127. (Visited on 10/04/2023).

Shamsi, Al and Sherief Abdallah (Aug. 2023). "Ensemble Stacking Model for Sentiment Analysis of Emirati and Arabic Dialects". In: *Journal of King Saud University - Computer and Information Sciences* 35, pp. 101691–101691. DOI: 10.1016/j.jksuci.2023.101691. (Visited on 10/04/2023).

Kaseb, Abdelrahman and Mona Farouk (Aug. 2023). "Active Learning for Arabic Sentiment Analysis". In: *Alexandria Engineering Journal* 77, pp. 177–187. DOI: 10.1016/j.aej.2023.06.082. (Visited on 10/04/2023).

Czeranowska, Olga et al. (Apr. 2023). "Migrants vs. stayers in the pandemic – A sentiment analysis of Twitter content". In: *Telematics and informatics reports* 10, pp. 100059–100059. DOI: 10.1016/j.teler.2023.100059.

Catelli, Rosario et al. (May 2023). "Lexicon-based Sentiment Analysis to Detect Opinions and Attitude Towards COVID-19 Vaccines on Twitter in Italy". In: *Computers in Biology and Medicine* 158, pp. 106876–106876. DOI: 10.1016/j.compbiomed.2023.106876. (Visited on 06/16/2023).

Tsytsarau, Mikalai and Themis Palpanas (Oct. 2011). "Survey on Mining Subjective Data on the Web". In: *Data Mining and Knowledge Discovery* 24, pp. 478–514. DOI: 10.1007/s10618-011-0238-6. URL: https://dl.acm.org/citation.cfm?id=2198208.

Piryani, R., D. Madhavi, and V.K. Singh (Jan. 2017). "Analytical Mapping of Opinion Mining and Sentiment Analysis Research during 2000–2015". In: *Information Processing Management* 53, pp. 122–150. DOI: 10.1016/j.ipm.2016.07.001. URL: https://sentic.net/scientometrics-of-sentiment-analysis-research.pdf.

Zhang, Lei, Shuai Wang, and Bing Liu (Mar. 2018). "Deep Learning for Sentiment analysis: a Survey". In: *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 8. DOI: 10.1002/widm.1253.

Wankhade, Mayur, Annavarapu Chandra Sekhara Rao, and Chaitanya Kulkarni (Feb. 2022). "A Survey on Sentiment Analysis methods, applications, and Challenges". In: *Artificial Intelligence Review* 55. DOI: 10.1007/s10462-022-10144-1.

Chen, Liang-Chu, Chia-Meng Lee, and Mu-Yen Chen (Oct. 2019). “Exploration of Social Media for Sentiment Analysis Using Deep Learning”. In: *Soft Computing* 24, pp. 8187–8197. DOI: 10.1007/s00500-019-04402-8. (Visited on 08/16/2022).

ZHAO, Yan-Yan, Bing QIN, and Ting LIU (Oct. 2010). “Integrating Intra- and Inter-document Evidences for Improving Sentence Sentiment Classification”. In: *Acta Automatica Sinica* 36, pp. 1417–1425. DOI: 10.1016/s1874-1029(09)60057-4. (Visited on 09/22/2021).

Sudhir, Prajval and Varun Deshakulkarni Suresh (Aug. 2021). “Comparative Study of Various Approaches, Applications and Classifiers for Sentiment Analysis”. In: *Global Transitions Proceedings* 2. DOI: 10.1016/j.gltp.2021.08.004.

Jin, Yuxin et al. (July 2023). “A Review of Text Sentiment Analysis Methods and Applications”. In: *Frontiers in business, Economics and Management* 10, pp. 58–64. DOI: 10.54097/fbem.v10i1.10171. (Visited on 10/09/2023).

Dang, Cach N., María N. Moreno-García, and Fernando De la Prieta (Aug. 2021). “Hybrid Deep Learning Models for Sentiment Analysis”. In: *Complexity* 2021. Ed. by Tao Jia, pp. 1–16. DOI: 10.1155/2021/9986920.

Tan, Kian Long et al. (2022). “RoBERTa-LSTM: a Hybrid Model for Sentiment Analysis with Transformer and Recurrent Neural Network”. In: *IEEE Access* 10, pp. 21517–21525. DOI: 10.1109/access.2022.3152828.

Cai, Yi et al. (Aug. 2019). “A Hybrid Model for Opinion Mining Based on Domain Sentiment Dictionary”. In: *International Journal of Machine Learning and Cybernetics* 10, pp. 2131–2142. DOI: 10.1007/s13042-017-0757-6. (Visited on 05/20/2023).

Rajeswari, A.M. et al. (July 2020). *Sentiment Analysis for Predicting Customer Reviews Using a Hybrid Approach.* IEEE Xplore. DOI: 10.1109/ACCTHPA49271.2020.9213236. URL: https://ieeexplore.ieee.org/abstract/document/9213236.

Mardjo, Anny and Chidchanok Choksuchat (2022). “HyVADRF: Hybrid VADER–Random Forest and GWO for Bitcoin Tweet Sentiment Analysis”. In: *IEEE Access* 10, pp. 101889–101897. DOI: 10.1109/access.2022.3209662. (Visited on 02/27/2023).

Adebanjo, Hameedah and Oluoma Oji (Dec. 2022). “Lexicon-based Sentiment Analysis Approach on the Energy Crisis in the United Kingdom”. In.

Kang, Hanhoon, Seong Joon Yoo, and Dongil Han (Apr. 2012). “Senti-lexicon and Improved Naïve Bayes Algorithms for Sentiment Analysis of Restaurant Reviews”. In: *Expert Systems with Applications* 39, pp. 6000–6010. DOI: 10.1016/j.eswa.2011.11.107.

Jurek, Anna, Maurice D. Mulvenna, and Yaxin Bi (Dec. 2015). “Improved lexicon-based Sentiment Analysis for Social Media Analytics”. In: *Security Informatics* 4. DOI: 10.1186/s13388-015-0024-x.

A. Rahim, Afiq Izzudin et al. (Jan. 2021). “Assessing Patient-Perceived Hospital Service Quality and Sentiment in Malaysian Public Hospitals Using Machine Learning and Facebook Reviews”. In: *International Journal of Environmental Research and Public Health* 18, p. 9912. DOI: 10.3390/ijerph18189912. URL: https://www.mdpi.com/1660-4601/18/18/9912/htm (visited on 02/09/2022).

Rustam, Furqan et al. (Feb. 2021). “A Performance Comparison of Supervised Machine Learning Models for Covid-19 Tweets Sentiment Analysis”. In: *PLOS ONE* 16. Ed. by Wajid Mumtaz, e0245909. DOI: 10.1371/journal.pone.0245909.

Rodrigues, Anisha P et al. (Apr. 2022). "Real-Time Twitter Spam Detection and Sentiment Analysis Using Machine Learning and Deep Learning Techniques". In: *Computational Intelligence and Neuroscience* 2022. Ed. by Muhammad Ahmad, pp. 1–14. DOI: 10.1155/2022/5211949.

Kim, Youngsoo (Aug. 2014). "Convolutional Neural Networks for Sentence Classification". In: DOI: 10.48550/arxiv.1408.5882.

Khurana, Diksha et al. (July 2022). "Natural Language processing: State of the art, Current Trends and Challenges". In: *Multimedia Tools and Applications* 82, pp. 3713–3744. DOI: 10.1007/s11042-022-13428-4.

Singh, Mrityunjay, Amit Kumar Jakhar, and Shivam Pandey (Mar. 2021). "Sentiment Analysis on the Impact of Coronavirus in Social Life Using the BERT Model". In: *Social Network Analysis and Mining* 11. DOI: 10.1007/s13278-021-00737-z.

McDonald, Daniel (July 2014). "A Text Mining Analysis of Religious Texts". In: *The Journal of Business Inquiry* 13, pp. 27–47. (Visited on 10/09/2023).

Nisha Varghese, Nisha and M Punithavalli (Dec. 2019). "Lexical And Semantic Analysis Of Sacred Texts Using Machine Learning And Natural Language Processing". In: *International Journal of Scientific Technology Research* 8.

Franklin, Robert B (May 2018). "Differences in the Emotional Content in Different Bible Translations". In: DOI: 10.31234/osf.io/vru43.

Whissell, Cynthia (Aug. 2019). "The Common Emotional Plot of the Four Gospels". In: *Advances in Social Sciences Research Journal* 6, pp. 472–479. DOI: 10.14738/assrj.68.7086.

— (Oct. 2009). "Using the Revised Dictionary of Affect in Language to Quantify the Emotional Undertones of Samples of Natural Language". In: *Psychological Reports* 105, pp. 509–521. DOI: 10.2466/pr0.105.2.509-521. (Visited on 04/17/2021).

Barrett, Tyson S. (July 2019). *'data.table' and Text Analysis: Analyzing the Four Gospels · TysonBarrett.com.* tysonbarrett.com. URL: https://tysonbarrett.com/jekyll/update/2019/07/06/textualanalysis_bible/ (visited on 07/27/2023).

Buentello, Saúl (Dec. 2018). *Bible Study Using Data Science.* Wordpress.com. URL: https://inandoutdata.wordpress.com/2018/12/02/bible-study-using-data-science/ (visited on 07/28/2023).

— (Dec. 2020). *What's the Most Positive or Negative religion? — Sentiment and Data Analysis of Holy Books with R.* Analytics Vidhya. URL: https://medium.com/analytics-vidhya/whats-the-most-positive-or-negative-religion-sentiment-and-data-analysis-of-holy-books-with-r-3fb881289f51.

Pradha, Saurav, Malka N. Halgamuge, and Nguyen Tran Quoc Vinh (Oct. 2019). "Effective Text Data Preprocessing Technique for Sentiment Analysis in Social Media Data". In: pp. 1–8.

Silva, Maleesha De (Apr. 2023). *Preprocessing Steps for Natural Language Processing (NLP): a Beginner's Guide.* Medium. URL: https://medium.com/@maleeshadesilva21/preprocessing-steps-for-natural-language-processing-nlp-a-beginners-guide-d6d9bf7689c9 (visited on 08/08/2023).

Schonlau, Matthias, Nick Guenther, and Ilia Sucholutsky (Dec. 2017). "Text Mining with n-gram Variables". In: *Stata Journal* 17, pp. 866–881. DOI: 10.1177/1536867x1801700406. (Visited on 10/10/2023).

Hays, Jeffrey (Sept. 2018). *Gospels: Their history, meaning, purpose, similarities and differences.* factsanddetails.com. URL: https://factsanddetails.com/world/cat55/sub391/entry-5779.html.

Petruzzello, Melissa (Sept. 2023). *Synoptic Gospels.* Ed. by The Editors of Encyclopaedia Britannica. www.britannica.com. URL: https://www.britannica.com/topic/biblical-literature/The-Synoptic-problem#ref598087.

Clay, Cholee (Apr. 2012). *Comparing the Gospels: Matthew, Mark, Luke, and John.* Owlcation. URL: https://owlcation.com/humanities/Comparing-the-Gospels-Matthew-Mark-Luke-and-John.

Saha, Srishti (Jan. 2019). *Understanding the Log Loss Function of XGBoost.* Medium. URL: https://medium.datadriveninvestor.com/understanding-the-log-loss-function-of-xgboost-8842e99d975d.

Khan, Md. Sakib et al. (Aug. 2022). "CNN-XGBoost fusion-based Affective State Recognition Using EEG Spectrogram Image Analysis". In: *Scientific Reports* 12. DOI: 10.1038/s41598-022-18257-x. (Visited on 10/11/2022).

Alamer, Ghadah, Sultan Alyahya, and Hmood Al-Dossari (Aug. 2023). "Identifying Users and Developers of Mobile Apps in Social Network Crowd". In: *Electronics* 12, pp. 3422–3422. DOI: 10.3390/electronics12163422. (Visited on 10/11/2023).

Samih, Amina, Abderrahim Ghadi, and Abdelhadi Fennan (Apr. 2023). "Enhanced Sentiment Analysis Based on Improved Word Embeddings and XGboost". In: *International Journal of Power Electronics and Drive Systems* 13, pp. 1827–1827. DOI: 10.11591/ijece.v13i2.pp1827-1836. (Visited on 10/11/2023).

Umair, Areeba, Elio Masciari, and Muhammad Habib Ullah (May 2023). "Vaccine Sentiment Analysis Using BERT + NBSVM and geo-spatial Approaches". In: *The Journal of Supercomputing* 79. DOI: 10.1007/s11227-023-05319-8. (Visited on 07/19/2023).

Zhang, Harry (Jan. 2004). "The Optimality of Naive Bayes". In: pp. 562–567.

Firmino Alves, André Luiz et al. (2014). "A Comparison of SVM versus Naive-Bayes Techniques for Sentiment Analysis in Tweets". In: DOI: 10.1145/2664551.2664561.

Jakkula, Vikramaditya R (2011). *Tutorial on Support Vector Machine (SVM).* Semantic Scholar. URL: https://api.semanticscholar.org/CorpusID:15115403.

AlBadani, Barakat, Ronghua Shi, and Jian Dong (Feb. 2022). "A Novel Machine Learning Approach for Sentiment Analysis on Twitter Incorporating the Universal Language Model Fine-Tuning and SVM". In: *Applied System Innovation* 5, p. 13. DOI: 10.3390/asi5010013. URL: https://www.mdpi.com/2571-5577/5/1/13.

Lan, Ting et al. (Apr. 2020). "A Comparative Study of Decision tree, Random forest, and Convolutional Neural Network for spread-F Identification". In: *Advances in Space Research* 65, pp. 2052–2061. DOI: 10.1016/j.asr.2020.01.036. (Visited on 05/18/2023).

Breiman, Leo (2001). "Random Forests". In: *Machine Learning* 45, pp. 5–32. DOI: 10.1023/a:1010933404324.

Wang, T. et al. (2020). "COVID-19 Sensing: Negative Sentiment Analysis on Social Media in China via BERT Model". In: *IEEE Access* 8, pp. 138162–138169. DOI:

10.1109/ACCESS.2020.3012595. URL: https://ieeexplore.ieee.org/abstract/document/9151169.

Abdul, Saad et al. (Dec. 2019). "Using BERT for Checking the Polarity of Movie Reviews". In: *International Journal of Computer Applications* 177, pp. 37–41. DOI: 10.5120/ijca2019919675.

Hutto, C. and Eric Gilbert (May 2014). "VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text". In: *Proceedings of the International AAAI Conference on Web and Social Media* 8, pp. 216–225. DOI: 10.1609/icwsm.v8i1.14550.

Devlin, Jacob et al. (2019). "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding". In: *Proceedings of the 2019 Conference of the North* 1. DOI: 10.18653/v1/n19-1423.

Li, Baoli and Liping Han (2013). "Distance Weighted Cosine Similarity Measure for Text Classification". In: pp. 611–618. DOI: 10.1007/978-3-642-41278-3_74.

Alodadi, Mohammad and Vandana P. Janeja (Oct. 2015). *Similarity in Patient Support Forums Using TF-IDF and Cosine Similarity Metrics.* IEEE Xplore. DOI: 10.1109/ICHI.2015.99. URL: https://ieeexplore.ieee.org/abstract/document/7349760 (visited on 05/29/2021).

Salton, G (Sept. 1989). "Automatic Text processing: the transformation, analysis, and Retrieval of Information by Computer". In: *Addison. Reading, Massachusetts.* (Visited on 03/16/2022).

Appendix A

Appendix

Link to git project repository: https://git.cs.bham.ac.uk/-/ide/project/projects-2022-23/mrd246/edit/main