UNVEILING PUBLIC PERCEPTION OF TECHNICAL AND VOCATIONAL EDUCATION AND TRAINING (TVET) ON SOCIAL MEDIA USING A HYBRID SENTIMENT ANALYSIS APPROACH

NUR HAFAZAH SHARIN [2*], MIRA KARTIWI [1]

[1]Kulliyyah of Information and Communication Technology, International Islamic University Malaysia, Kuala Lumpur, Malaysia
[2]Division of Governance and Excellence, Department of Polytechnic and Community College Education, Ministry of Higher Education, Putrajaya, Malaysia

**Corresponding author: nurhafazah@gmail.com*

ABSTRACT: The public perception significantly impacts the effectiveness and societal acceptance of Technical and Vocational Education and Training (TVET) in Malaysia. Social media platforms such like Facebook becoming more important which makes it easier to get the real-time public input on education policy and programs. However, evaluating sentiment in informal and bilingual discussion poses significant challenges. This study presents a hybrid sentiment analysis approach that integrates lexicon-based approach (VADER, TextBlob, SentiWordNet, AFINN) with machine learning classifiers (Support Vector Machine, Naïve Bayes, Decision Tree, Logistic Regression, Random Forest and K-Nearest Neighbors) to classify and evaluate sentiments in 1,304 Facebook posts and comments related to TVET. The dataset was systematically pre-processed, translated, and subjected to feature extraction by Term Frequency-Inverse Document Frequency (TF-IDF). Accuracy served as the sole evaluation metric to ensure consistent comparison among lexicon-only, machine learning-only, and hybrid approaches. The hybrid combination of TextBlob and SVM achieved the highest accuracy at 72%, surpassing both standalone lexicon and machine learning approaches. This finding indicates that hybrid approaches are more effective in handling informal, code-switched sentiment data. The study offers a scalable framework for sentiment analysis in education policy, providing practical implications for TVET stakeholders aiming to evaluate public opinion and enhance engagement strategies.

KEY WORDS: *TVET, Sentiment Analysis, Hybrid Approach, Machine Learning, Lexicon-Based*

1. INTRODUCTION

Technical and Vocational Education and Training (TVET) is recognised as a significant policy initiative for workforce development and enhancing youth employability, particularly in terms of transforming career opportunities (UNESCO, 2016). In Malaysia, TVET has not yet achieved recognition as a favoured pathway for higher education. It is often viewed unfavourably, typically linked to thse who have not excelled in traditional academic routes (Omar & Mohd. Desa, 2023). Despite these initiatives, TVET continues to experience obstacles with public perception, including persistent issues of declining credibility, social stigma, and inadequate awareness of its values and outcomes (Okae-Adjei, 2017).

The growing prevalence of digital engagement has transformed social media platforms like Facebook into crucial platforms for the active generation and expression of public perspectives. These platforms offer unmediated, immediate feedback on educational policy and institutional practices, affording substantial resources for evaluating community sentiments (Gao, 2020). The informal, multilingual, and frequently code-switched characteristics of user-generated content on these platforms provide considerable analytical challenges for academics (Solorio et al., 2014).

Sentiment analysis offers a computer approach to understanding public sentiment by classifying text according to emotions polarity. Traditional approaches, whether lexicon-based or machine learning, struggle to sufficiently capture the complexity of informal and bilingual conversation (Medhat et al., 2014). Lexicon-based approaches, yet interpretable, sometimes lack contextual depth, whereas machine learning approaches often require labelled datasets and may struggle to generalise to unstructured real-world data (Chatterjee & Kumar, 2020).

To address these constraints, hybrid approaches that integrate rule-based sentiment analysis with statistical learning have gained significant prominence in recent studies. These approaches leverage the interpretability of lexicons alongside the adaptability of machine learning to enhance classification effectiveness (Chen et al., 2022; Medhat et al., 2014). This research utilises a hybrid sentiment analysis methodology that integrates lexicon-based polarity scores with machine learning classifiers to analyse Facebook comments concerning TVET in Malaysia. The objective is to identify sentiment trends, evaluate model correctness across various contexts, and develop insights that can inform targeted communication strategies, curriculum creation, and improved public engagement with TVET policy.

2. RELATED WORKS

2.1. Social Media as a Source of Public Sentiment

Social media platforms like Facebook, Twitter, and YouTube have emerged as significant platforms for public conversation, offering an abundance of user-generated content that expresses personal thoughts, attitudes, and collective sentiments. In the domain of education and policy, platforms such as Facebook are significant as they enable public participation in dialogues regarding curriculum reform, funding priorities, graduate employability, and the overarching societal value of educational opportunities like Technical and Vocational Education and Training (TVET). In contrast to traditional methods like surveys or interviews, social media data are extensive, immediate, and collected in real time, providing more genuine insights into public perception (Gao, 2020). Simultaneously, these data sources provide challenges: they are frequently loud, informal, and marked by code-switching, which complicates sentiment analysis (Solorio et al., 2014).

2.2. Overview of Sentiment Analysis and Importance

Sentiment analysis, often known as opinion mining, is the computational approach for identifying and classifying opinions or emotions expressed in textual data. It has been widely utilised in various domains including marketing, politics, healthcare, and education. The main objective is to categorise text into polarity classifications, such as positive, negative, or neutral. In the field of education, sentiment analysis has been utilised to assess student feedback, evaluate public responses to educational reforms, and investigate opinions of vocational programs (Khan et al., 2022). Nonetheless, its use in the realm of Technical and Vocational Education and Training (TVET) is constrained, especially in the examination of public sentiment expressed through social media.

2.3. Lexicon-based Approaches

Lexicon-based approaches operate by comparing terms in a text to a set lexicon, giving each word an emotion score. These methods are valued for their simplicity, transparency, and autonomy from labelled datasets. VADER, TextBlob, AFINN, and SentiWordNet are some of the most commonly used lexicons. VADER is especially excellent at analysing social media material due to its incorporation of heuristics for punctuation, capitalisation, and emoticons (Hutto & Gilbert, 2014). Nevertheless, lexicon-based methods encounter significant obstacles, especially in addressing verbal ambiguity, sarcasm, domain-specific jargon, and contextual sensitivity. These limitations undermine their reliability in analysing the informal and code-switched information typically encountered on sites such as Facebook (Taboada et al., 2011).

2.4. Machine Learning Aproaches

Machine learning (ML) approaches conceptualise sentiment classification as a supervised learning task. Support Vector Machines (SVM), Logistic Regression, Random Forests, and Naïve Bayes are all algorithms that are trained on labelled datasets with known sentiment classifications. These models utilise features like n-grams, TF-IDF scores, or word embeddings to make predictions about new data. When ample labelled data is accessible, machine learning-based sentiment classifiers generally work better than lexicon-based ones (Medhat et al., 2014). However, they are resource-intensive and frequently criticised for their restricted interpretability, as many operate as black-box models in contrast to more open rule-based approaches (Chatterjee & Kumar, 2020).

2.5. Hybrid Approaches

Hybrid sentiment analysis integrates the advantages of both lexicon-based and machine learning approaches. One common strategy involves using lexicon-derived sentiment scores as additional features in machine learning models, thereby enriching the feature space with both statistical and rule-based signals. Prior studies by Medhat et al., (2014) and Septiani et al. (2024) highlight that hybrid models consistently outperform either method alone, particularly in low-resource contexts and when analysing informal language.

This approach is especially relevant for social media analysis in bilingual environments such as Malaysia, where Facebook comments frequently mix Malay and English with casual language. In this study, the hybrid framework improved the accuracy of categorization and provided clearer understanding of sentiment patterns around Technical and Vocational Education and Training (TVET). By combining lexical polarity with statistical learning, the hybrid method effectively addresses challenges such as context ambiguity and domain-specific terminology, offering a practical and scalable solution for monitoring educational sentiment.

3. METHODOLOGY

This study utilised a systematic five-phase method to execute the hybrid sentiment analysis framework illustrated in Fig. 1.

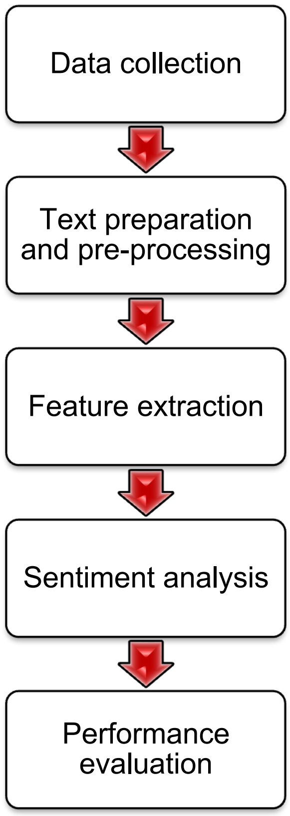

Fig.1.Sentiment analysis phases

3.1. Data Collection

The study employed data from publicly available Facebook groups and media pages. Data were collected from public groups concentrated on TVET-related topics that fostered active discussion. Only postings and comments explicitly pertaining to TVET were included on media pages to maintain dataset relevance. The data collecting process properly complied with Facebook's terms of service, and no private user information or personal identifiers were acquired. Data were collected from April 2021 to March 2023 using Facepager and Apify, both of which are online scraping tools. Apify is an automated scraping platform that extracts and exports unstructured data in .csv format, whereas Facepager, created by Jünger & Keyling (2019) utilises application programming interfaces (APIs) and web scraping methods to obtain publicly accessible data from platforms like YouTube, Facebook, and Twitter.

3.2. Text Preparation and Pre-processing

Before analysis, a thorough pre-processing phase was executed to ensure the quality of the dataset for text mining. This process encompassed cleaning and normalisation procedures, including the elimination of punctuation, rectification of spelling errors, expansion of abbreviations, and correction of informal terms. Due to the nature of the dataset that contains bilingual language Malay and English, all Malay items were translated into English via Google Translate to ensure uniformity in sentiment classification. The translated text was further verified against sentiment lexicons to ensure polarity accuracy and semantic precision.

No resampling was conducted to maintain the natural distribution of attitudes, despite the class imbalance. This decision facilitate a more reliable assessment of classifier efficacy in scenarios that reflect actual sentiment variations. As part of improving data quality in pre-processing phase, a series of data transformation processes were carried out. These include the process of formatting all text into lowercase for consistency and removing special characters and stop words to eliminate noise, and improving semantic clarity. Subsequent to these procedures, a total of 1,304 posts were chosen as appropriate to undergo analysis.

3.3. Feature Extraction

The Term Frequency–Inverse Document Frequency (TF-IDF) technique was utilised to transform unstructured text into a numerical format (Stanley et al., 2023). TF-IDF measures how important a word is by looking at two things. Term Frequency (TF) looks at how often a word appears in a document, and Inverse Document Frequency (IDF determines the scarcity of a word throughout the entire corpus, thus diminishing the influence of prevalent terms (Edalati et al., 2022). This strategy was selected for its proven efficacy in enhancing classification accuracy and its capacity to emphasise relevant terms within the text.

3.4. Sentiment Analysis

A hybrid approach to sentiment analysis was employed, integrating lexicon-based and machine learning approaches. This study utilised the lexicons SentiWordNet, TextBlob, VADER, and AFINN. The classifiers employed for machine learning encompassed Support Vector Machine (SVM), Naïve Bayes (NB), Decision Tree (DT), Random Forest (RF), K-Nearest Neighbours (KNN), and Logistic Regression (LR). This combined approach enhanced the study by using both lexical polarity scores and statistical learning models.

3.5. Performance Evaluation

Accuracy was selected as the primary evaluation criterion to guarantee clarity and consistency in assessing the performance of lexicon-based, machine learning, and hybrid models. Accuracy denotes the ratio of accurately identified occurrences to the total number of predictions made. It is described as follows in Eq. (1):

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \tag{1}$$

Where:

TP: the number of positive values that were correctly classified;
TN: the number of negative values that were correctly classified;
FP: the number of positive values that were incorrectly classified; and
FN: the number of negative values that were incorrectly classified.

4. RESULTS AND DISCUSSION

The effectiveness of the hybrid sentiment analysis approach is evaluated by comparing the results of lexicon-only, machine learning-only, and hybrid approaches, with an emphasis on determining the most accurate approach for classifying public perceptions of TVET.

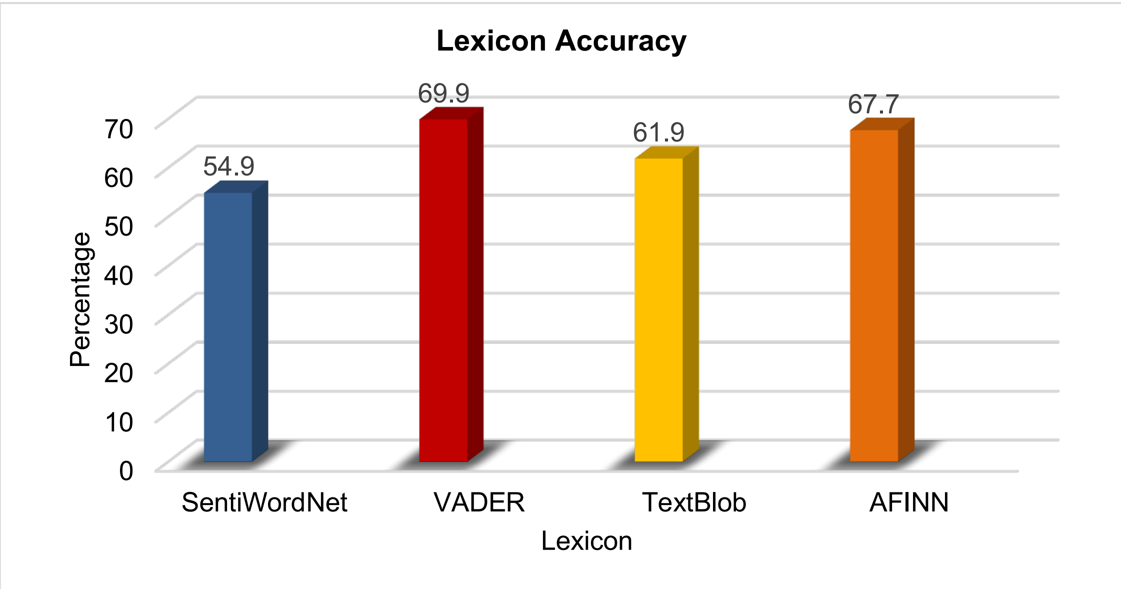

Fig. 1. Lexicon-only accuracy

Fig. 2 shows the accuracy achieved by each lexicon employed independently for sentiment classification. Among the lexicon-based approaches, VADER attained the highest accuracy at 69.9%, succeeded by AFINN at 67.7%, TextBlob at 61.9%, and SentiWordNet at 54.9%. The great performance of VADER is due to its adeptness in managing social media syntax and its sensitivity to punctuation and emoticons.

Additionally, this study utilised a machine learning approach to assess the accuracy of each classifier. Fig. 3 demonstrates the accuracy metrics for the trained classifiers. Among machine learning classifiers, the SVM achieved the highest accuracy at 61.7%, followed by LR at 61.3%, RF at 55.2%, NB at 54.8%, DT at 52.1%, and KNN recorded the lowest performance at 38.3%. These findings recognise SVM and LR are good for classifying text.

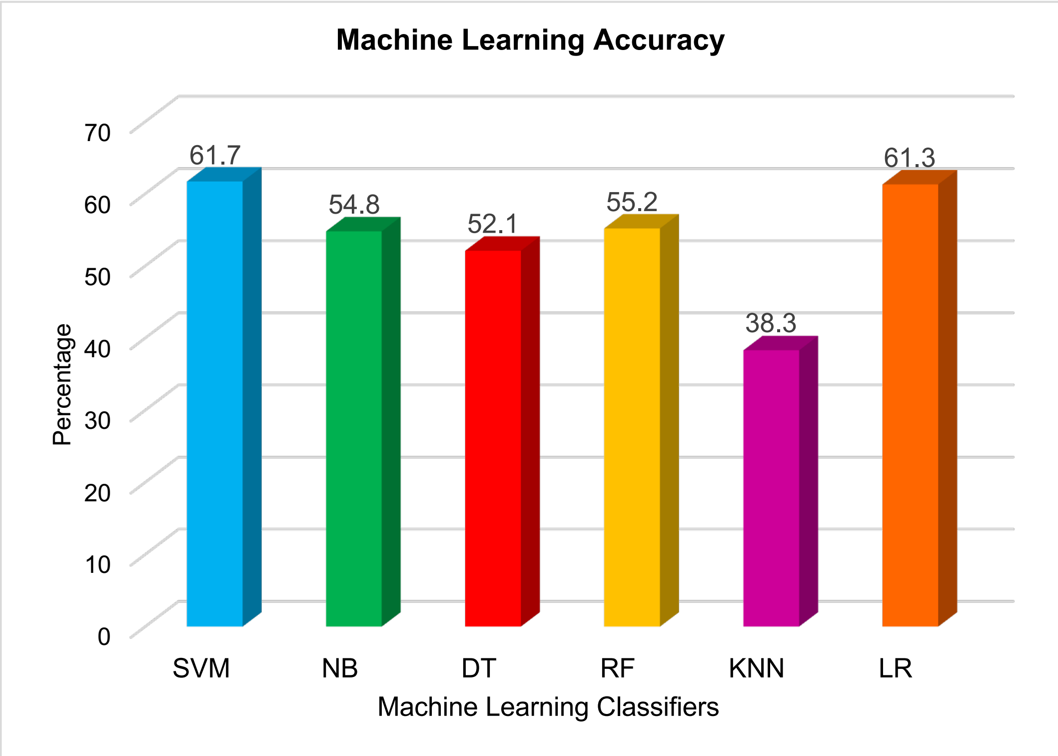

Fig. 2. Machine Learning- only accuracy

To determine the additional benefits integrating lexicon-based and ML approaches, we evaluated hybrid models that utilised sentiment scores from each lexicon were used as supplementary features for ML classifiers. The performance of each hybrid configuration is shown in Table 1.

This comparison clearly illustrates that hybrid models produce performance improvements ranging from 2 to 10 percentage points, depending on the configuration. The integration of lexicon sentiment polarity enriches the feature space for ML classifiers, enabling more effective detection of sentiment patterns in informal, code-switched data commonly present on social media platforms like Facebook.

Table 1: Hybrid accuracy

Classifiers	VADER	TextBlob	SentiWordNet	AFINN
Support Vector Machine (SVM)	0.71	0.72	0.69	0.64
Naïve Bayes (NB)	0.59	0.57	0.57	0.50
Decision Tree (DT)	0.59	0.55	0.59	0.56
Random Forest (RF)	0.70	0.64	0.66	0.62
K-Nearest Neighbors (KNN)	0.48	0.56	0.24	0.50
Logistic Regression (LR)	0.70	0.67	0.68	0.63

The combination of TextBlob + SVM achieved the highest accuracy of 0.72, closely followed by VADER + SVM at 0.71, and VADER + Logistic Regression at 0.70. These findings confirm that the integration of lexicon polarity with machine learning markedly improves classification accuracy in noisy and bilingual content. The TextBlob + VADER emerged as the most effective sentiment tools across all classifiers, owing to their ability to handle informal language and contextual sentiment. Conversely, AFINN and SentiWordNet exhibited inferior performance, especially with classifiers that are sensitive to feature sparsity such as KNN. The SVM consistently outperformed other classifiers, validating its efficacy in high-dimensional text environment. Logistic Regression and Random Forest also showed reliable performance, while KNN yielded the lowest accuracy due to its reliance on distance-based metrics in sparse vector spaces.

These results indicate that hybrid sentiment models can effectively manage the complex nature of informal, bilingual TVET discussions on Facebook. By delivering accurate sentiment classifications, these models enable stakeholders to monitor public perceptions, highlight policy challenges, and refine communication strategies more effectively.

5. CONCLUSION

This study presented a hybrid sentiment analysis approach that integrates lexicon-based scoring with machine learning classifiers to assess public perception towards TVET on Facebook. Among the tested settings, TextBlob + SVM delivered the highest classification accuracy at 0.72, followed closely by VADER + SVM at 0.71. These findings validate the effectiveness of hybrid sentiment analysis for handling informal, multilingual, and real-world data sources, especially within the context of Malaysian TVET discourse.

The results offer several practical implications for stakeholders. Policymakers might leverage sentiment patterns from hybrid models to build targeted communication campaign strategies that address public skepticism or disinformation. Educational institutions may gain advantages by assessing public sentiment to align curricula with social expectations and labor market needs. Moreover, the hybrid approach offers a scalable tool for continuous sentiment tracking, facilitating prompt actions and enhanced public involvement across digital platforms.

Nonetheless, this study possesses certain limitations. One significant constraint was the absence of manually labelled ground truth data, which may have introduced noise and affected model reliability. The reliance on lexicon sentiment scores also limited the model's ability to capture complex expressions such as sarcasm, idioms, or implicit sentiment. The analysis concentrated solely on Facebook data, potentially omitting to encapsulate sentiment across other social media platforms.

Future studies should focus on integrating deep learning models, such as BERT, with lexicon-informed attention mechanisms to achieve better sentiment representations. Incorporating more platforms such as YouTube, TikTok, or X could provide a more comprehensive perspective on public discourse. Furthermore, engaging human annotators to identify a portion of the data could improve evaluation accuracy and provide richer training signals for sentiment classification.

REFERENCES

Chatterjee, S., & Kumar Kar, A. (2020). Why do small and medium enterprises use social media marketing and what is the impact: Empirical insights from India. *International Journal of Information Management*, *53*(December 2019), 102103. https://doi.org/10.1016/j.ijinfomgt.2020.102103

Chen, Y., Liu, D., Liu, Y., Zheng, Y., Wang, B., & Zhou, Y. (2022). Research on user generated content in Q&A system and online comments based on text mining. *Alexandria Engineering Journal*, *61*(10), 7659–7668. https://doi.org/10.1016/j.aej.2022.01.020

Edalati, M., Shariq Imran, A., Kastrati, Z., & Daudpota, S. M. (2022). *The Potential of Machine Learning Algorithms for Sentiment Classification of Students' Feedback on MOOC*.

Gao, Y. (2020). Constructing the social network prediction model based on data mining and link prediction analysis. *Library Hi Tech*, *38*(2), 320–333. https://doi.org/10.1108/LHT-11-2018-0179

Hutto, C. J., & Gilbert, E. (2014). *VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text*. http://sentic.net/

Jünger, J., & Keyling, T. (2019). *Facepager. An application for automated data retrieval on the web. Source code and releases available at https://github.com/strohne/Facepager/*.

Khan, M. I., Khan, Z. A., Imran, A., Khan, A. H., & Ahmed, S. (2022). Student Performance Prediction in Secondary School Education Using Machine Learning. *8th International Conference on Information Technology Trends: Industry 4.0: Technology Trends and Solutions, ITT 2022*, 94–101. https://doi.org/10.1109/ITT56123.2022.9863971

Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, *5*(4), 1093–1113. https://doi.org/10.1016/j.asej.2014.04.011

Okae-Adjei, S. (2017). Public Perception towards Technical and Vocational Education and Training (TVET) in Ghanaian Polytechnics. *Asian Research Journal of Arts & Social Sciences*, *2*, 1–10. https://doi.org/10.9734/ARJASS/2017/32430

Omar, A. S., & Mohd. Desa, Z.-. (2023). Factors Influencing TVET Choices among Secondary School Students in Kuching. *Malaysian Journal of Social Sciences and Humanities (MJSSH)*, *8*(7), e002380. https://doi.org/10.47405/mjssh.v8i7.2380

Salsabila Septiani, Nabila Putri, Dara Jessica, & Arya Saputra. (2024). Sentiment Analysis Of Social Media Data Using Deep Learning Techniques. *International Journal of Computer Technology and Science*, *1*(2), 08–14. https://doi.org/10.62951/ijcts.v1i2.59

Solorio, T., Blair, E., Maharjan, S., Bethard, S., Diab, M., Gohneim, M., Hawwari, A., Alghamdi, F., Hirschberg, J., Chang, A., & Fung, P. (2014). Overview for the First Shared Task on Language Identification in Code-Switched Data. *Proceedings OfThe First Workshop on Computational Approaches to Code Switching*, 62–72. https://www.facebook.com/

Stanley, M., Aiswarya, K. R., & Deepa, G. (2023). Sentiment Analysis of Covid Vaccine Tweet with Vader, and Implementation of Different Machine Learning Models. *2023 14th International Conference on Computing Communication and Networking Technologies, ICCCNT 2023*. https://doi.org/10.1109/ICCCNT56998.2023.10308224

Taboada, M., Brooke, J., Tofiloski, M., Voll, K., & Stede, M. (2011). *Lexicon-Based Methods for Sentiment Analysis*.

UNESCO. (2016). Executive Summary Executive Summary Executive Summary. *South African Medical Journal*, *101*(2003), 16.