



OPEN ACCESS

EDITED BY

Robinson Sibe,
Rivers State University, Nigeria

REVIEWED BY

Francesco Cauteruccio,
Marche Polytechnic University, Italy
Abhijit Das,
RCC Institute of Information Technology,
India

*CORRESPONDENCE

Mark Lokanan,
✉ mark.lokanan@royalroads.ca

RECEIVED 02 September 2023

ACCEPTED 07 November 2023

PUBLISHED 22 November 2023

CITATION

Lokanan M (2023), Analyzing public sentiments on the Cullen Commission inquiry into money laundering: harnessing deep learning in the AI of Things Era.

Front. Internet. Things 2:1287832.
doi: 10.3389/friot.2023.1287832

COPYRIGHT

© 2023 Lokanan. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Analyzing public sentiments on the Cullen Commission inquiry into money laundering: harnessing deep learning in the AI of Things Era

Mark Lokanan*

Faculty of Management, Royal Roads University, Victoria, BC, Canada

This study employs deep learning methodologies to conduct sentiment analysis of tweets related to the Cullen Commission's inquiry into money laundering in British Columbia. The investigation utilizes CNN, RNN + LSTM, GloVe, and BERT algorithms to analyze sentiment and predict sentiment classes in public reactions when the Commission was announced and after the final report's release. Results reveal that the emotional class "joy" predominated initially, reflecting a positive response to the inquiry, while "sadness" and "anger" dominated after the report, indicating public dissatisfaction with the findings. The algorithms consistently predicted negative, neutral, and positive sentiments, with BERT showing exceptional precision, recall, and F1-scores. However, GloVe displayed weaker and less consistent performance. Criticisms of the Commission's efforts relate to its inability to expose the full extent of money laundering, potentially influenced by biased testimonies and a narrow investigation scope. The public's sentiments highlight the awareness raised by the Commission and underscore the importance of its recommendations in combating money laundering. Future research should consider broader stakeholder perspectives and objective assessments of the findings.

KEYWORDS

deep learning, sentiment analysis, money laundering, Cullen Commission, AML compliance

1 Introduction

Money laundering funded \$5.3B in B.C. real estate purchases in 2018, report reveals" (Lindsay, 2019); "B.C. casinos 'unwittingly served as laundromats' for proceeds of crime" (Schmunk, 2018); "Fast cars and bags of cash: Gangsters using B.C. luxury car market to launder dirty money" (Larsen, 2019). As detailed in these media commentaries, British Columbia (B.C.) along with other insular havens of the like, are indeed particularly prone to money laundering in the real estate, gaming, and luxury car dealership industries (hereinafter "non-banking financial institutions" (NBFIs)). To address the money laundering problems in B.C., the provincial government in May 2019 appointed Austin Cullen, a B.C. Supreme Court justice, to conduct an independent public inquiry (hereinafter the "Commission") into money laundering in these NBFIs. The Commission was established to investigate and provide recommendations into concerns related to money laundering in B.C. The Commission conducted hearings, gathered evidence, and released its final report in May 2021.

When the Cullen Commission was initially announced and after the final report was released, there was a virtual explosion of reactions on Twitter (now known as X) concerning the public's opinions about the witnesses' testimonies on AML compliance and regulation in B.C. and, more generally, Canada. This project will use a deep learning methodology to analyze these tweets using sentiment analysis. More specifically, Convolutional Neural Network (CNN), Recurrent Neural Network with Long-Short Term Memory (RNN + LSTM), GloVe, and Bidirectional Representational Transformer (BERT) natural language processing deep learning techniques will be used to analyze the public tweets and their reactions to the Cullen Commission inquiry. The primary goal is to assess the polarity and subjectivity of the public sentiments concerning the testimonies of witnesses by analyzing Twitter data. Specifically, the objective is to use deep learning algorithms to conduct sentiment analysis on Twitter data during two phases: when the inquiry was initially announced and after the final report was released to the public.

This study makes several contributions to the literature on using deep learning methodology in AML research. Notably, this research introduced a distinct and unique approach by applying deep learning artificial intelligence techniques to analyze public sentiments derived from Twitter data pertaining to AML regulation and compliance. By leveraging deep learning algorithms, this paper explores the potential of these techniques in uncovering and understanding public perceptions and reactions to AML compliance and regulation. This novel and distinctive contribution to the field provides new insights into the intersection of deep learning and sentiment analysis in AML research. The theoretical underpinnings of deep learning models in sentiment analysis have contributed to developing more advanced models that can capture semantics and contextual information on issues related to AML compliance. Deep learning models can be employed to explore and understand complex patterns, uncover insights, and interpret the semantics around textual AML data. Conducting sentiment analysis on Twitter data related to the Cullen Commission requires understanding domain-specific language related to money laundering, regulation, and public sentiments towards the Commission and subsequent findings. Properly trained deep learning models can capture domain-specific language through their ability to learn semantic representations of the issue and recognize patterns specific to the given domain. These characteristics enable the deep learning models to conduct more accurate sentiment analysis, effectively capturing the nuanced semantics specific to public reactions towards the Cullen Commission.

The remaining sections of this paper are organized according to the following format. A literature review is first presented, detailing the application of deep learning in AML research. Next, the experiments used to analyze the data are described in detail. Following a description of the experiments, a discussion on the findings is then conducted, focusing on the effectiveness of the deep learning algorithms to analyze public reactions using Twitter data. Finally, the conclusion section highlights the limitations of the study and identifies potential areas for future research in the application of deep learning techniques to analyze textual data related to AML compliance.

2 Literature review

AML regulations pose a persistent challenge for financial institutions. Recent legislation has intensified detection and reporting requirements, including the *USA Patriot Act* and the Fifth *EU Anti-Money Laundering Directive* (5AMLD) (Stewart, 2023). Many financial institutions are unprepared to meet these regulations and are turning to technological advancements like robotics, semantic analysis, and AI to enhance their AML processes. These technological advancements can reduce false positives, improve compliance with regulatory expectations, and enhance operational resource productivity (M. E. Lokanan, 2022; Stewart, 2023). Recently, AI has primarily been employed in AML compliance through robotic process automation for efficient case investigation and preparation. However, there is a current trend toward leveraging machine and deep learning algorithms to enhance or replace traditional Boolean logic in detecting suspicious activities (Stewart, 2023). Although traditional techniques of detecting and monitoring money laundering activities have improved, they still fall short due to the volume of data and transactions that are generated internationally and the rising incidence of criminal conduct (Chen et al., 2018; Kumar et al., 2021). Rule-based or threshold systems are one of the techniques used to detect money laundering activities, which cannot handle high-volume datasets with different types of structured or semi-structured and unstructured data (Chen et al., 2018; Han et al., 2020).

Rule-based systems are based on static predetermined rules that require the creation of new regulations to detect illegal activities (Chen et al., 2018), leading to investigations that are manual, tedious, time-consuming, and resource-intensive (Han et al., 2020; Kumar et al., 2021). Conventional methods of detecting fraud focus on controlling access by verifying identities and analyzing customer transactions. These methods are not effective in quickly identifying potential fraud. Hence, the legitimacy of transactions cannot be determined due to the static and manual nature of the rule-based system. To improve the accuracy of fraud detection and reduce the rate of false positives, advanced techniques such as statistical analysis, neural networks, decision trees, fuzzy logic, and genetic algorithms should be used (Alkhalili et al., 2021; Kumar et al., 2021; Bhat et al., 2022).

In their 2017 study, Borovkova and her colleagues explore the application of sentiment analysis to detect and monitor systemic risk in financial markets (Borovkova, Garmaev, and Lammers, 2017). This technique creates a risk indicator by incorporating sentiment data from news stories about Systemically Important Financial Institutions. Compared to other systemic risk indicators like SRISK or VIX, the authors claim that their sentiment-based indicator can forecast periods of stress in the financial system up to 12 weeks in advance. Another paper describes a method to derive feature values from financial news that looks at a list of important traits that can affect a company's revenue. This strategy entails agglomerating sentiment values and augments an existing financial sentiment vocabulary. The findings indicate that revenue patterns may be predicted using news items from the previous quarter with high performance accuracy (Omar et al., 2017).

Typically, while reviewing suspicious AML transactions, investigators resort to the internet for proof by analyzing sentiment trends. AI-based sentiment analysis of news items

regarding a particular business might yield important insights for AML processes. This technology allows the rapid screening of thousands of articles, making the inquiry process more effective and precise. Sentiment analysis, a rapidly expanding field in Natural Language Processing, is the process of determining the attitude or emotion expressed in a piece of content, such as a sentence or document (Pang and Lee, 2009; Bashir and Ghous, 2020; Han et al., 2020; Bonifazi et al., 2022). For specific AML tasks, two different models are developed, one for document-level and another for sentence-level analysis. These models span from classifying entire documents (Pang and Lee, 2009) to determining the positive or negative nature of words and phrases (Bhoir et al., 2017; J. Liao and Acharya, 2011; J. Wang et al., 2020). The former is a multi-channel CNN-based classifier for financial news articles (Hans, 2011; Lyu and Liu, 2021), and the latter is also a CNN-based model for social media data (He et al., 2016; Han et al., 2018).

Researchers are currently using sentiment analysis to analyze opinions and emotions using Twitter and Reditt data (Ray et al., 2018; Basile et al., 2021; Bonifazi et al., 2022). Kayıkçı discusses a new model for sentiment analysis called “SenDemonNet,” which uses a heuristic deep neural network to analyze tweets about demonetization (Kayıkçı, 2022). Ray and his supporting authors use sentiment analysis to study public opinion on the demonetization policy of the government of India (Ray et al., 2018). The study uses Twitter data collected through the Twitter API and a lexicon-based approach to analyze the sentiments of different groups of people over the 5 weeks leading up to the policy announcement.

The programming community is paying closer attention to machine learning for sentiment analysis. One of the primary advantages of machine learning is its capacity to generalize from training data. Additionally, it can automatically learn about decision constraints. Machine learning algorithms produce accurate predictions about new circumstances without being explicitly taught based on the knowledge acquired during training (Chen et al., 2018). Machine learning can reduce the amount of time needed for preparing data, identify the most important information to detect, improve the accuracy of identifying true positives, and decrease the burden on personnel, training, and financial resources (Han et al., 2018; Kumar et al., 2021). Others employed techniques such as data cleaning, statistical analysis, and data mining, using tools like linear support vector machines and decision tree classifiers to identify money laundering activities with very good results (Kumar et al., 2021).

While previous studies in the sentiment analysis field have used classical machine learning classification techniques (e.g., SVM, Naive Bayes), recent studies have found that deep learning techniques have improved accuracy in sentiment analysis of English tweets (Vateekul and Koomsubha, 2016; Chen et al., 2018; Mohammed et al., 2022). Though machine learning strategies for fighting money laundering are still in their infancy, deep learning techniques have been proven to be more productive in dealing with enormous datasets (Chen et al., 2018; Umer et al., 2021; Mohammed et al., 2022). Vateekul and Koomsubha (2016) propose the first study to apply these techniques to Twitter data and use two deep learning techniques: Long Short-Term Memory (LSTM) and Dynamic Convolutional Neural Network (DCNN). The results show that deep learning techniques outperformed classical techniques like

Naive Bayes and SVM, except for Maximum Entropy. Han et al. (2018) discusses using distributed deep learning-based language technologies in augmenting AML investigations. These technologies can analyze large amounts of unstructured data, such as social media texts, to identify potential money laundering activity. Umer et al. (2021) in another study proposes a combination of CNN + LSTM for sentiment analysis on Twitter datasets. The performance of this model is compared to other machine learning classifiers and state-of-the-art models. Two feature extraction methods are also tested for their impact on accuracy. The study evaluated the model's performance on three datasets and found that the CNN + LSTM model had higher accuracy than other classifiers (Umer et al., 2021).

Heterogeneous feature ensemble models are gaining popularity in research about sentiment analysis (Bonifazi et al., 2022). Another study introduces a new method for classifying fuzzy sentiments in Twitter messages using a feature ensemble model. The technique incorporates various elements, such as sentiment polarity, position, semantics, word type, and linguistic features of words. The researchers used a real-time dataset and evaluated the performance using the F1-score. The feature ensemble model transforms each tweet into tweet embeddings by extracting features such as word embeddings from the GloVe model, the distance between words, sentiment scores of words, N-grams of words, and Part-of-Speech (POS) tags. The resulting tweet embeddings were then processed using a CNN to improve sentiment analysis performance. The proposed method combines a deep learning algorithm, the feature ensemble model, and a divide-and-conquer approach to increase the efficiency of sentiment analysis (Phan et al., 2020; Bonifazi et al., 2022).

Another study uses 29,764 tweets to examine the Volkswagen emissions issue as a sustainability fraud crisis. The authors suggest three techniques—cluster analysis, sentiment analysis, and time series analysis—make up the Tweet Analytic Framework. The study aims to understand public opinions on the crisis in two stages and shows the typical pattern of crisis development, strong public condemnation and negative sentiment, and significant public censure (Ding et al., 2023). The analysis of Twitter spam diffusion in Tanwar and Rai's account emphasizes the influence of themes on the dissemination of spam and non-spam information (Tanwar and Rai, 2020). The study classifies content into spam and non-spam using user-based factors and content-based features. It creates a dataset that can be used to categorize and study the dissemination of fraudulent and accurate information.

Bandana (2018) proposed a heterogeneous approach that combined machine learning-based and lexicon-based features with supervised learning algorithms, such as Naive Bayes and Linear SVM, to build a sentiment analysis model (Bandana, 2018). This hybrid approach was found to provide more accurate sentiment analysis compared to other baseline systems, suggesting that it could be applied to big data using advanced deep learning algorithms to create even more accurate models. In a separate study, Chang et al. (2022) developed an anti-fraud chatbot that utilized natural language processing and machine learning algorithms to identify and categorize fraud accurately (Chang et al., 2022). The model combined DistilBERT with a SVM and random forest and was found to be highly accurate in identifying potential financial fraud situations while being more resource- and time-efficient.

TABLE 1 Most recent works.

Author(s)	Year	Focus	Key findings/Contributions
Borovkova et al.	2017	Sentiment analysis for systemic risk	Sentiment-based indicator predicts financial stress
Omar et al.	2017	Sentiment from financial news	Predicting revenue patterns using past news
Hajek and Henriques	2017	Feature selection for fraud detection	Ensemble approaches effective for fraud detection
Chen et al.	2018	Money laundering detection	Volume of transactions challenges traditional methods
Ray et al.	2018	Sentiment on demonetization policy	Public opinion on demonetization
Han et al.	2018	Deep learning for AML	Deep learning enhances AML investigations
Bandana	2018	Hybrid sentiment analysis	Hybrid model more accurate than baselines
Phan et al.	2020	Feature ensemble for fuzzy sentiments	Deep learning improves sentiment analysis
Tanwar and Rai	2020	Twitter spam diffusion	Influence of themes on spam and non-spam
Craja et al.	2020	HAN for annual reports	HAN captures managerial remarks for stakeholders
Umer et al.	2021	CNN + LSTM for sentiment analysis	CNN + LSTM outperforms other models
Kumar et al.	2021	Analytics and data mining	Decision tree classifier yields accuracy
Basile et al.	2021	COVID-19's Social Impact Analysis	Extreme events affect posting, emotions
Bonifazi et al.	2022	Assessing scope of sentiment	Sentiment spread and influence factors
Lokanan	2022	AML processes and technology	Technology can enhance AML compliance
Zhang et al.	2022	Sentiment analysis for fraud detection	Stacking ensemble learning for fraud detection
Kayıkçı	2022	Sentiment analysis on demonetization	Deep neural network for demonetization sentiment
Chang et al.	2022	Anti-fraud chatbot	Chatbot for financial fraud detection
Xiuguo and Shengyong	2022	Deep learning for fraud detection	Textual features reinforce financial fraud detection
Stewart	2023	AML regulations	Enhanced AML detection requirements
Ding et al.	2023	Sentiment on Volkswagen emissions	Public condemnation on Volkswagen emissions

Others used a deep learning model to predict fraud based on a firm's change trajectory and linguistic features, such as sentiment and uncertainties, formulated as a multivariate time series (Umer et al., 2021). The authors found that specific changes, such as weak modal or reward words in newly added or deleted contents, were strongly related to fraud. In another study, the authors employed a hierarchical attention network (HAN), a deep learning method, to extract text features from the MD&A section of annual reports (Craja et al., 2020). The HAN was designed to reflect the structured hierarchy of documents and incorporated two distinct attention mechanisms at the word and sentence level. This methodology captured the content and context of managerial remarks and identified red-flag areas to help stakeholders.

Researchers have employed used intelligent feature selection and classification to integrate specific aspects of financial information and managerial comments in corporate annual reports. They found ensemble approaches were more effective at detecting fraudulent firms, while Bayesian belief networks were better at identifying non-fraudulent firms (Hajek and Henriques, 2017). In a recent study, developed an advanced system for detecting financial fraud by combining numerical features from financial statements and textual data in managerial comments using state-of-the-art deep learning models (Xiuguo and Du, 2022). This approach showed significant performance improvement compared to traditional

machine learning methods, with LSTM and GRU approaches achieving high true classification rates of 94.98% and 94.62%, respectively. These results suggest that the extracted textual features of the MD&A section significantly reinforced financial fraud detection.

The preceding literature review highlights recent advancements in machine and deep learning for addressing financial crimes but underscores persistent empirical gaps in the field. This research aims to fill some of these gaps by conducting sentiment analysis on public reactions to the Cullen Commission's inquiry announcement and the subsequent release of the final report, providing insights into the application of deep learning models on Twitter data in real-world scenarios. The study empirically assesses the scalability, generalizability, and effectiveness of deep learning models across different time periods and offers a comprehensive analysis of evolving public sentiments, showcasing their adaptability to shifting discourse. Additionally, in the broader context of sentiment analysis in financial crimes, this project stands out for its unique focus on using deep learning models to examine public sentiments on Twitter regarding the Cullen Commission's money laundering inquiry. By enabling a before-and-after comparison of sentiments, it sheds light on how public perceptions evolve around significant events, particularly in the context of AML compliance issues. While previous studies have explored sentiment analysis in

the realm of financial crimes, this study is among the pioneering efforts to apply sentiment analysis to scrutinize public perceptions pertaining to AML compliance. Thus, this research contributes significantly to the field of sentiment analysis in financial crime detection, offering novel insights into understanding public sentiments related to the Cullen Commission's inquiry into money laundering. The contributions of previous studies are displayed in [Table 1](#).

3 Experimental setting

Both machine learning and advanced deep learning classifiers were used to analyze the data collected on public opinions regarding the Cullen Commission and its findings on money laundering in British Columbia. Multiple metrics and data points were combined, strengthening the results' validity. Together, these models provide a more comprehensive investigation and analysis of the data, allowing for greater accuracy and confidence in interpreting the findings. Furthermore, the use of machine learning, and deep learning techniques allowed for reliable insights to be obtained from the data. Modern advanced deep learning models are especially important in this case, where large amounts of public opinion data need to be processed quickly and accurately.

3.1 Data collection

The data used in this study was collected from Twitter using the hashtag <https://twitter.com/hashtag/cullencommission>. The Cullen Commission was established on May 15, 2019, by the Government of British Columbia, and the final report was released on June 15, 2022. Over 5,500 tweets were scraped, with 4,153 posted when the Commission was initially announced and during the Commission's hearings, while an additional 1,441 tweets were collected after the final report was made public. We collected data to account for public opinions both when the Inquiry was announced and after the Commission's report. The data were collected between August and September 2022. No data was collected during the Commission's hearing. The tweets were collected using Tweepy. Tweepy is a Python library for accessing the Twitter API, and that provides an easy-to-use interface for researchers to retrieve tweets and other data from the Twitter platform. The search terms "Cullen," "Commission," and "Cullen Commission" were used to scrape as many tweets as possible. The dataset consists of the following attributes: ID, Date, Users, Text, and Location of the users. The data were then uploaded into a CSV file for further analysis. These tweets revealed public opinion during this important investigation into money laundering in British Columbia. The tweets showed a range of reactions to the Commission's findings from all sectors of Canadian society. This unique data set provided insights into the many perspectives on Canada's AML enterprise and the fight against money laundering.

3.2 Ethics

The data for this study was collected based on the Twitter Terms of Service and Developer Agreement. Researchers are authorized to

use Twitter data for academic research purposes. All the collected data was kept confidential and anonymous and respects the privacy of Twitter users. Furthermore, no identifying information was used to identify tweets or sensitive information from Twitter users. The Twitter license agreement authorizes Twitter to make users' content available for advancing research ([Twitter, 2023](#)). Twitter further notes that "from social science to computer science, Twitter data can advance research objectives on topics as diverse as the global conversations happening on Twitter" ([Twitter, 2023](#)). Twitter went on to state that:

Free, no-code datasets are intended to make it easier for academics to study topics that are of frequent interest to the research community. They are purpose-built, predefined, comprehensive datasets of all public Tweets related to a specific topic ([Twitter, 2023](#))

This research complies with all aspects of Twitter's policies, included, but not limited to privacy, data sharing, and ethical considerations.

3.3 Data-preprocessing, cleaning, and preparation

The primary goal of text preprocessing is to transform raw textual data into a format that machine learning algorithms can use. Text preprocessing is important for preparing data for machine and deep learning applications. The data in this paper has been processed to make it useable, by employing regular expressions to eliminate URLs, hashtags, special characters, emails, usernames (handles), and punctuation. Additionally, we have removed common stop words from the dataset. The terms were converted to their base form through lemmatization and stemming, replacing words with synonyms, and correcting typos or inconsistencies in the data set. Tokenization was used to split the text into distinct words or tokens, separating punctuation and other symbols. Tokenizing the text allows further analysis, such as pattern recognition, extracting keywords, or searching within textual data ([Basile et al., 2021](#)). Vectorization was done using Term Frequency-Inverse Document Frequency (TF-IDF). TF-IDF vectorization also represents text as numerical vectors and considers not only the frequency of words in a document but also their importance in the entire corpus ([M. E. Lokanan, 2023](#)). Using TF-IDF, words that are common across all the Tweets are given lower weight, while words that are unique to a document are given higher weight ([Havrlant and Kreinovich, 2017](#); [M. E. Lokanan, 2023](#)). The entire data preprocessing steps are shown in [Figure 1](#).

3.4 Performance metrics

[Table 2](#) presents the confusion matrix for sentiment analysis involving negative, neutral, and positive sentiments. The cell matrix displays the true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). After running the models, these cells are populated with the actual values calculated by the algorithms. The confusion metrics are also utilized to calculate performance measures that assess the model's accuracy, precision, recall, and F1 score.

[Table 3](#) displays the formulas used to calculate various performance measures. These measures assist in evaluating the



FIGURE 1 Data preprocessing.

TABLE 2 Confusion matrix.

	Actual negative	Actual neutral	Actual positive
Predicted Negative	TN	FN	FN
Predicted Neutral	FP	TN	FP
Predictive Positive	FP	FN	TP

TABLE 3 Performance measures.

Performance metrics	Formula
Accuracy	$(TP + TN)/(TP + TN + FP + FN)$
Precision	$TP/(TP + FP)$
Recall	$TP/(TP + FN)$
F1-Score	$2 * (Precision * Recall)/(Precision + Recall)$

classifiers’ performance in predicting sentiment classes. Accuracy gauges the model’s ability to predict correct classes. Precision (or positive predictive value) assesses the proportion of positive predictions correctly predicted by the algorithm. Recall or sensitivity (the true positive rate) quantifies the proportion of actual positive observations accurately predicted by the model. The F1-score represents the harmonic mean of precision and recall. This score is considered a trade-off between precision and recall, serving as a well-balanced measure, particularly when comparing the performance of two models (M. E. Lokanan, 2022).

3.5 Algorithm considered

Deep learning NLP algorithms have revolutionized the way we classify emotions in datasets. The following deep learning models were utilized to conduct emotion classification on the provided datasets: Text CNN (Convolutional Neural Network), RNN + LSTM (Recurrent Neural Network with Long-Short Term Memory), and BERT (Bidirectional Encoder Representations from Transformers). Techniques such as Text CNN, RNN + LSTM, and BERT are now acknowledged as essential tools for modern emotion classification. It has become increasingly clear that these cutting-edge techniques are pivotal components of contemporary emotion analysis, and their successful implementation showcases the significant power of NLP technology.

3.5.1 Text CNN

Text CNN is a potent tool for processing text data, capable of capturing local patterns that prove challenging for other models to

discern. The algorithm operates by applying a convolutional filter to the input text and classifying the resultant feature maps. These feature maps hold significant value in tasks like sentiment analysis and emotion identification, providing valuable insights into the data (He et al., 2016). Recent research underscores Text CNN’s superiority over alternative network architectures in these tasks, underscoring its distinct potential and the significance of NLP (Carlos-Roca et al., 2018; Feng and Cheng, 2021; S. Liao et al., 2017; J. Wang et al., 2020). Text CNN also finds utility in other applications, including document summarization and language modeling (Parimala et al., 2021; Umer et al., 2021). Text CNNs adeptly capture local and global patterns in data through a blend of convolutional and recurrent layers, resulting in enhanced predictive accuracy (Yenala et al., 2018; Han et al., 2020). Moreover, their scalability and minimal preprocessing requirements make them appealing for numerous applications (Gan et al., 2021). Text CNNs will likely remain a pivotal element in the machine learning landscape as NLP technologies evolve.

3.5.2 RNN + LSTM

The utilization of RNN + LSTM architectures has garnered increasing popularity for emotion classification tasks. This surge in popularity arises from the capacity of RNNs to grasp temporal dependencies among words, which subsequently results in enhanced performance compared to other architectures. Integrating LSTM cells within the architecture effectively addresses the vanishing gradient problem that plagues conventional recurrent neural network models. This integration facilitates superior training accuracy and greater depth in unrolling (Vathsala and Ganga, 2020; Umer et al., 2021). Through this approach, the amalgamation of RNNs and LSTMs bestows the model with a competitive advantage in recognizing the evolving spectrum of affective states in natural language processing tasks (Vateekul and Koomsubha, 2016). Beyond heightened accuracy, RNN + LSTM architectures in emotion classification tasks yield additional benefits. These architectures offer insights into the interplay between words and emotions, thus enabling the development of more precise predictive models (Effrosynidis et al., 2022). They also diminish the necessity for manual feature engineering, thereby expediting and streamlining the training process (Vathsala and Ganga, 2020; Effrosynidis et al., 2022). In essence, these architectures prove

indispensable as tools for emotion classification tasks. Ultimately, the fusion of RNNs and LSTMs enhances model generalizability by enabling the capture of intricate sentiment patterns within datasets.

3.5.3 Global vector for word representation

The Global Vectors for Word Representation (GloVe) algorithm was employed to transform words into vector representations. GloVe stands as a groundbreaking NLP algorithm with a multitude of applications. Converting words into vector representations enables the exploration and analysis of relationships between different attributes. Unlike conventional methods reliant on count vectors, GloVe crafts dense, low-dimensional vectors from co-occurrence statistics across an entire corpus (Bhoir et al., 2017). These word vectors capture global context and local nuances of words, rendering them more effective at conveying word meanings than count vectors. With its unsupervised learning approach, GloVe attains swifter processing compared to supervised techniques while still encapsulating meaningful semantic and syntactic elements (Bhoir et al., 2017; Goularas and Kamis, 2019). This distinctive amalgamation of traits positions GloVe as an invaluable instrument for comprehending language and constructing insightful linguistic models.

3.5.4 BERT

BERT has revolutionized the approach to conducting emotion classification (Devlin et al., 2018). Automating emotional classification tasks has been a central focus over the past decade (Poria et al., 2017). With the substantial amount of available data, AI solutions have enabled researchers to process extensive textual data in novel and informative ways. BERT represents one such solution, developed by harnessing vast amounts of text data; it can discern intricate relationships between words that often appear insurmountable with conventional machine learning methods (Devlin et al., 2018). With potent tools like BERT, researchers can promptly execute highly accurate emotion classification tasks, enhancing efficiency while yielding outstanding outcomes (Alswaidan and Mohamed El, 2020; Zhang et al., 2020). BERT's capacity to unravel complex word relationships proves particularly valuable for identifying nuances in written text sentiment. Analyzing contextual and semantic word relationships proficiently determines whether a statement conveys a positive, negative, or neutral sentiment—often surpassing human emotion classification accuracy (Devlin et al., 2018; Zhang et al., 2020). Employing BERT for these tasks can substantially reduce the time and cost associated with manual sentiment analysis (Poria et al., 2017; M. E. Lokanan, 2023).

3.6 Ekman's model of emotions

The variable “sentiment” serves as a response or target variable. We conduct two sets of analyses. Firstly, we employ Ekman's (1992) emotional classification to delve into the data's granularity. Ekman's (1992) taxonomy of emotions comprises six categories: joy, sadness, anger, love, fear, and surprise (Ekman, 1992). This model has gained substantial traction in cognitive and social media research for comprehending emotional context (Nagarsekar et al., 2013; Z. Wang et al., 2016; Basile et al., 2021).

3.6.1 Sentiment analysis

Sentiment analysis was employed to identify and extract subjective information from the Twitter data. Utilizing polarity and subjectivity measures, the tweets were processed and categorized into three broad sentiment categories: positive, neutral, and negative (Z. Wang et al., 2016; M. E. Lokanan, 2023). TextBlob was employed to determine each sentiment classification's polarity and subjectivity scores. The classification process assigns a polarity score of -1 for negative sentiments, 0 for neutral sentiments, and 1 for positive sentiments. The tweet's subjectivity is determined by the dichotomy of emotions that characterize opinions and fact-based content. Text classified as opinion receives a score of 0 , while highly objective and fact-based text is labeled as 1 (Vateekul and Koomsubha, 2016; Umer et al., 2021). Text labeling yields valuable insights into users' opinions and sentiments regarding the Cullen Commission's inquiry and report on money laundering in British Columbia. Figure 2 shows the workflow of the sentiment analysis process.

4 Finding and analysis

4.1 Granular emotional classification

Figure 3 illustrates the outcomes of employing various NLP techniques to analyze emotional expressions within public tweets during two distinct periods: when the Cullen Commission was announced (referred to as “Before”) and after the release of the final report (referred to as “After”). Although there are relatively minor variations (ranging from 1% to 5%) in the emotional expression detection capabilities of the employed NLP techniques, the findings indicate that the BERT model outperformed the CNN and RNN + LSTM models in terms of performance accuracy (Devlin et al., 2018; J. Wang et al., 2020). These findings suggest that a deeper and more grounded understanding of language may lead to more accurate emotion detection (Acheampong et al., 2021). The potential benefits of such improved accuracy could be significant, particularly in communication-based tasks and operations within AML compliance (Thi et al., 2020).

While a greater proportion of positive emotions, particularly joy, was observed, it's noteworthy that a substantial portion of negative emotions, including sadness, fear, and anger, were also present. Across all algorithms, joy was the predominant sentiment expressed by the public before and after the Commission's report. These findings suggest that the initial announcement of the Cullen Commission inquiry generated positive reactions from the public, which remained relatively stable after the publication of the final report. However, it's worth noting that the percentage of sadness increased slightly after the Cullen report was published, indicating that the findings generated some negative feedback from the public. Similarly, there was a slight increase in angry reactions, suggesting that the public became more frustrated with the Commission's findings. On the other hand, the percentage of fearful reactions remained stable before and after the Commission's report, suggesting no significant changes in fear-related sentiments toward the findings.

These findings imply that even before examining the Commission's results, the public experienced complex emotions

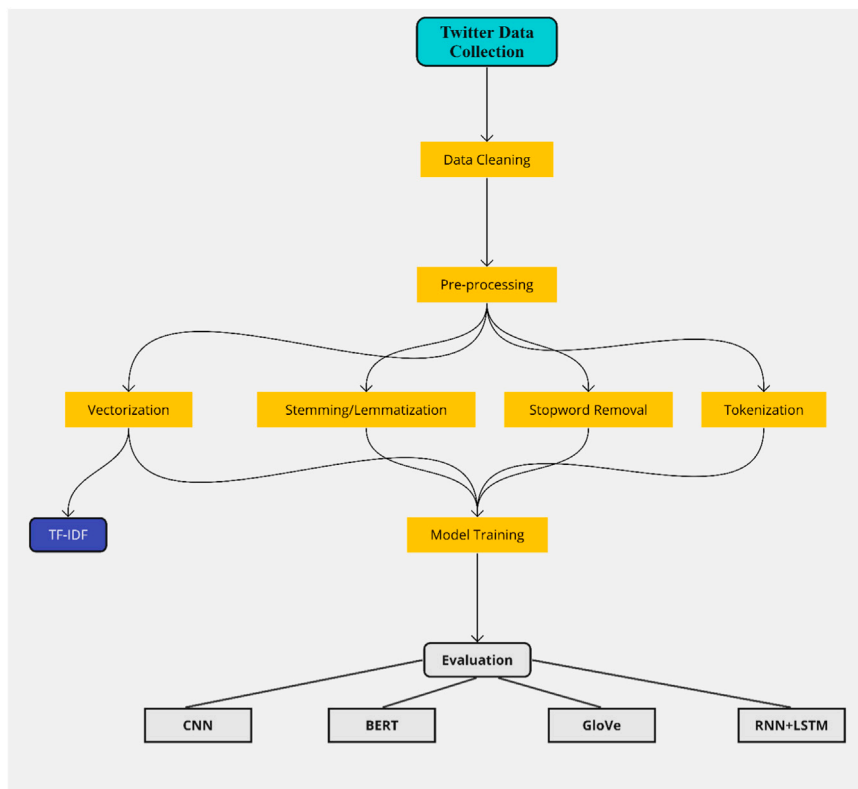


FIGURE 2 Sentiment analysis workflow.

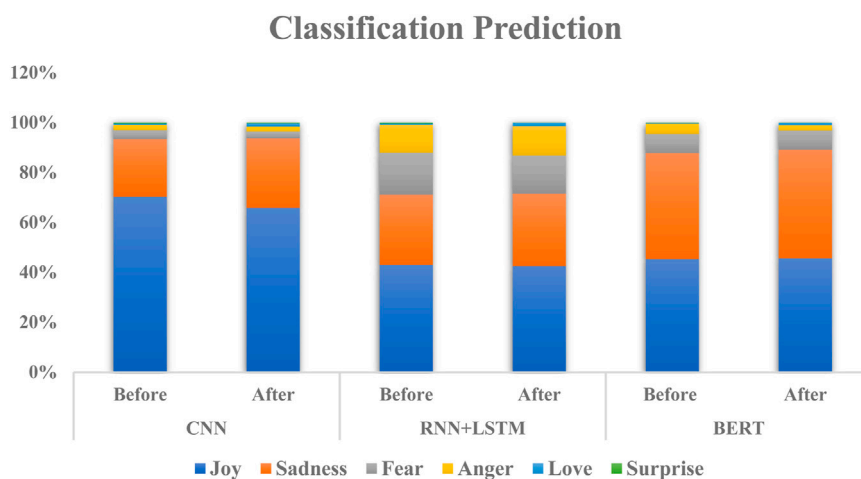
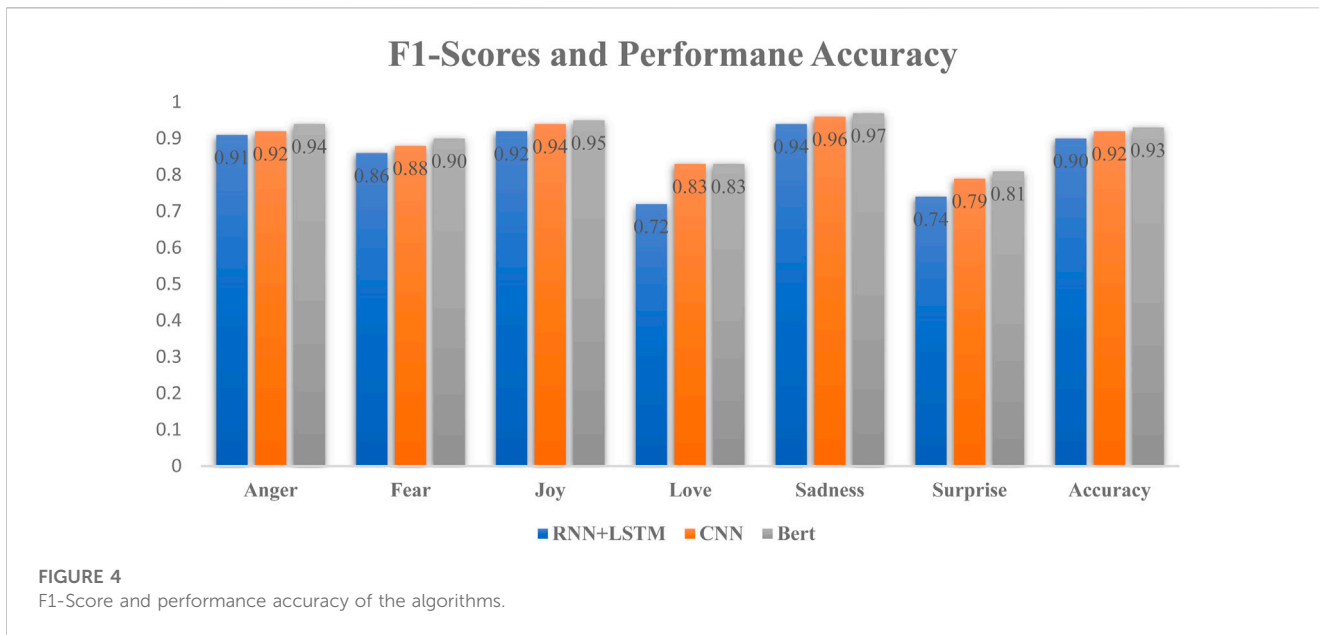


FIGURE 3 Classification graph.

regarding its outcomes. The diversity of emotional reactions indicates that the public harbored uncertainty or discomfort regarding the evidence of money laundering that the Commission would uncover. Following the release of the Cullen Commission report on money laundering in British Columbia, a notable decrease in the percentage of emotional reactions across all

categories of sentiments was observed. These findings suggest that the public felt a sense of closure or resolution after the investigation or that the report's contents made issues concerning money laundering in the province more transparent to the public. Regardless of the reasons, it's evident that the Inquiry's findings had a tangible impact on the emotional reactions of the public



toward money laundering and compliance issues in British Columbia. While further research is required to comprehend the full range of implications, the Cullen Commission provided a much-needed outlet for catharsis for British Columbians (Leuprecht et al., 2023).

Figure 4 presents a visual representation of the F1-scores and performance accuracy of the selected algorithms. The F1-score was chosen due to its ability to balance precision and recall, offering a comprehensive assessment of classifier performance (M. E. Lokanan, 2023). Among the various algorithms utilized, BERT consistently achieved the highest F1-score, indicating its effectiveness in tasks related to sentiment classification, as supported by previous research (Devlin et al., 2018; Alaparhi and Mishra, 2021). BERT demonstrated remarkable proficiency in anticipating public reactions across all emotional categories. BERT’s exceptional performance in predicting emotional reactions from tweets is attributed to its fine-tuning process. During fine-tuning, BERT refines its ability to recognize emotional signals in the training data, specializing it for tasks like sentiment analysis. This adaptation allows BERT to better understand emotions within the context of the Cullen Commission’s Inquiry, leading to enhanced accuracy. Additionally, BERT’s proficiency in identifying emotional cues and patterns is observed across various emotional categories due to its exposure to diverse emotional content in the dataset. This diversity contributes to its strong performance in recognizing emotions. Furthermore, we observed that the CNN and RNN + LSTM models also exhibited strong performance in predicting emotions such as anger, fear, joy, and sadness, with F1-scores ranging from 86% to 96%. While their performance in predicting love and surprise was slightly lower, with F1-scores ranging from 72% to 81%, it’s important to highlight that these algorithms still proved successful in forecasting public reactions associated with these specific emotional expressions.

Remarkably, the algorithms demonstrated remarkably high overall accuracy scores, ranging from 90% to 93%. Once again, BERT achieved the highest performance accuracy (93%) in predicting emotional classifications, followed by the CNN (92%)



and RNN + LSTM (90%) models. These findings suggest that, on the whole, BERT, CNN, and RNN + LSTM were proficient in precisely predicting public sentiments across various emotional categories related to the reaction to the money laundering inquiry in British Columbia. The substantial performance accuracy across diverse sentiments indicates that these algorithms are valuable tools for identifying and distinguishing public opinions and diverse emotional expressions concerning crucial matters such as money laundering and AML compliance tied to the Commission’s inquiry.



4.2 Word clouds positive and negative sentiments before inquiry

4.2.1 Highlighting positive sentiments

Figures 5, 6 showcase word clouds representing public sentiments at the time of the Cullen Commission's announcement. These visuals offer an intuitive way to discern the most frequently used words in public opinion tweets. In Figure 5, words like "money laundering," "economy," "law," "BC politics," "capital flights," "time," and "hear" dominate the word cloud. The presence of "money laundering" signifies the issue's importance in British Columbia, its potential impact on the economy, and its vulnerability (Aluko and Bagheri, 2012; Hendriyetty and Grewal, 2017). The prominence of "law" suggests the need for regulatory improvements. References to "BC Politics" likely involve political decisions on combating money laundering, while "capital flights" may allude to illicit fund movements across borders. Mentions of "timing" and "hearing" indicate the urgency of expert testimonies and diverse perspectives in addressing money laundering.

4.2.2 Highlighting negative sentiments

Figure 6 displays the negative sentiments surrounding money laundering in British Columbia. Frequent terms include "money laundering," "corrupt," "solicitor general," "casino," "liberal government," "organized crime," and "BC politics." The repeated mention of "money laundering" reflects public recognition and concern. Labeling the system as "corrupt" suggests widespread sentiments about corruption within AML compliance. Mentioning the "solicitor general" implies distrust in government officials. References to "casinos" and "organized crime" hint at criminal networks exploiting the casino industry. "BC politics" implies concerns about political influence. These negative sentiments reflect the public's apprehension about the Cullen



Commission's inquiry, indicating a lack of trust, concerns about corruption and organized crime, and negative perceptions of casinos and political entities. These sentiments drove expectations for a thorough investigation and measures to combat criminal proceeds in the financial system.

4.3 Word clouds of positive and negative sentiments after inquiry

4.3.1 Highlighting positive sentiments

Figure 7 illustrates the positive sentiments conveyed by the public following the release of the Cullen Commission's report. In this word cloud, terms like "break" and "fix" in tweets suggest an eagerness to effect change and address issues related to money laundering. The term "appetite" indicates the public's enthusiasm for the government to implement the recommendations from the Cullen Commission, particularly in response to concerns about criminal proceeds in the province (Cullen, 2022). The reference to "real estate" implies that the Commission's report could have a positive impact on addressing money laundering concerns within the industry. Sentiments associated with the National Democratic Party ("NDP"), "fight," and "corruption" indicate favorable attitudes toward the government's efforts to expose and combat corruption linked to money laundering activities. These positive sentiments reflect public support for addressing issues and implementing solutions to combat money laundering in British Columbia.

4.3.2 Highlighting negative sentiments

In contrast, Figure 8 presents the negative sentiments expressed by the public regarding the Commission's report. In this word cloud, the presence of the term "money laundering" underscores public



concerns about criminal proceeds in the province, signaling awareness of the issue and its potential economic impact. There is a perception that individuals involved in money laundering are “gaming” the system by employing manipulative and deceptive tactics to launder money into the province. Furthermore, there is a noticeable lack of trust and skepticism toward “former ministers” who previously held positions of power in AML compliance and regulation. The term “corrupt” implies that the public views the AML enterprise as influenced by individuals exploiting power or influence for personal gain, leading to calls for “aggressive” actions to address money laundering (Ai and Tang, 2011). These negative sentiments are clearly linked to concerns about the nature of money laundering, manipulation of the regulatory system, mistrust in government officials, and the belief that money laundering is closely intertwined with corruption. These sentiments likely fuel the demand for the inquiry and the need to address the findings stemming from the investigation.

4.4 Performance of classifiers upon the initial declaration of the Cullen inquiry

Figure 9 presents the performance metrics of the classifiers in predicting public sentiments from tweets when the Cullen Commission was initially announced in British Columbia. As shown in Figure 8, the BERT classifier exhibits the highest precision (.94), recall (.87), and F1-scores (.87). At the same time, the GloVe model yields the lowest scores for these three performance measures in predicting negative sentiments among Twitter users. Notably, the CNN and RNN + LSTM models closely trail the BERT model in predicting various sentiment classes. These findings indicate consistent and accurate identification of negative

sentiments from tweets related to the money laundering inquiry announcement among the models.

The BERT classifier again stands out in its predictions regarding neutral sentiments. BERT achieves the highest precision (0.97), recall (.99), and F1-score (.98) across all performance measures. Additionally, the CNN, RNN + LSTM, and GloVe models exhibit slightly lower precision, recall, and F1-scores, ranging between .93 and .96 for precision, .88 to .97 for recall, and .93 to .97 for F1-score. These results imply that all models were highly effective in correctly categorizing neutral sentiments within public opinions concerning the Cullen Commission announcement.

BERT outperforms the other algorithms in predicting positive sentiments. As depicted in Figure 8, BERT achieves the highest precision (.95), recall (.97), and F1-score (.96), closely followed by the CNN and RNN + LSTM classifiers. The GloVe classifier, on the other hand, exhibits comparatively weaker performance in predicting positive sentiments from tweets related to the public’s reaction to the Cullen Commission announcement. These findings highlight the ability of BERT, CNN, and RNN + LSTM algorithms to adeptly decode and interpret sentiment-rich narratives, particularly those expressing positive reactions to the Commission’s initial announcement.

Across all performance metrics, the BERT algorithm consistently performs better in predicting sentiment from public tweets. This consistency underscores the robustness of the BERT algorithm in classification tasks using Twitter data. These findings imply that BERT’s algorithms possess exceptional capabilities in capturing the intricate nuances of language, context, and semantics relevant to tweets. Such traits position BERT as a potent deep learning technique for predicting the dynamic spectrum of sentiments expressed by the public in response to the Cullen Commission’s announcement, reflecting its proficiency in analyzing public perceptions of money laundering issues. In sum, the deep learning algorithms excel in uncovering diverse classes of public sentiments surrounding the Cullen Commission Inquiry.

4.5 Classifier’s performance after final report was released to the public

Figure 10 presents the outcomes of the classifiers after the public release of the Cullen Commission’s final report. A notable finding is that the performance of all classifiers in predicting sentiment classes experiences a significant decrease across the board. The classifiers display varying precision, recall, and F1-scores regarding negative sentiments. The CNN classifier achieves the highest precision score (.97), indicating its ability to accurately identify negative sentiments. However, a recall score of .55 suggests that the CNN classifier misses a substantial proportion of negative sentiments. Among the other classifiers, BERT demonstrates balanced performance with a precision score of .85, recall of .65, and F1-score of .73. These findings indicate the superiority of BERT’s classifier in capturing negative sentiments from the data compared to the other algorithms. Compared to the negative sentiments observed during the initial announcement of the Inquiry, BERT maintains its strengths in overall predictability for identifying negative sentiments after the release of the final report.

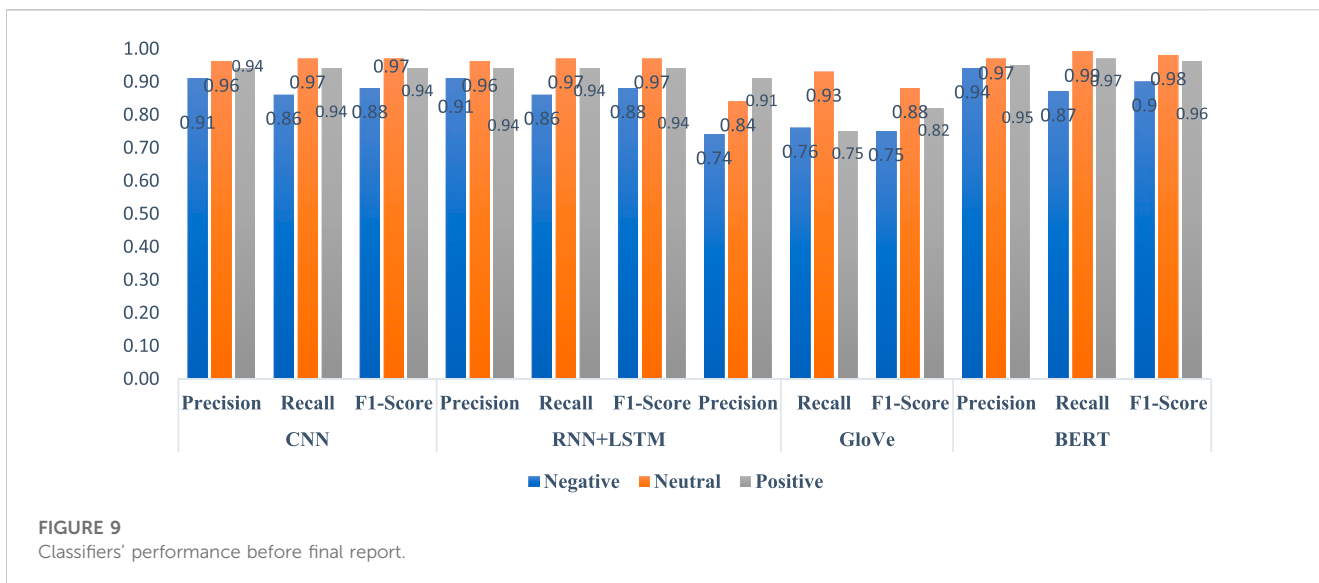


FIGURE 9 Classifiers' performance before final report.

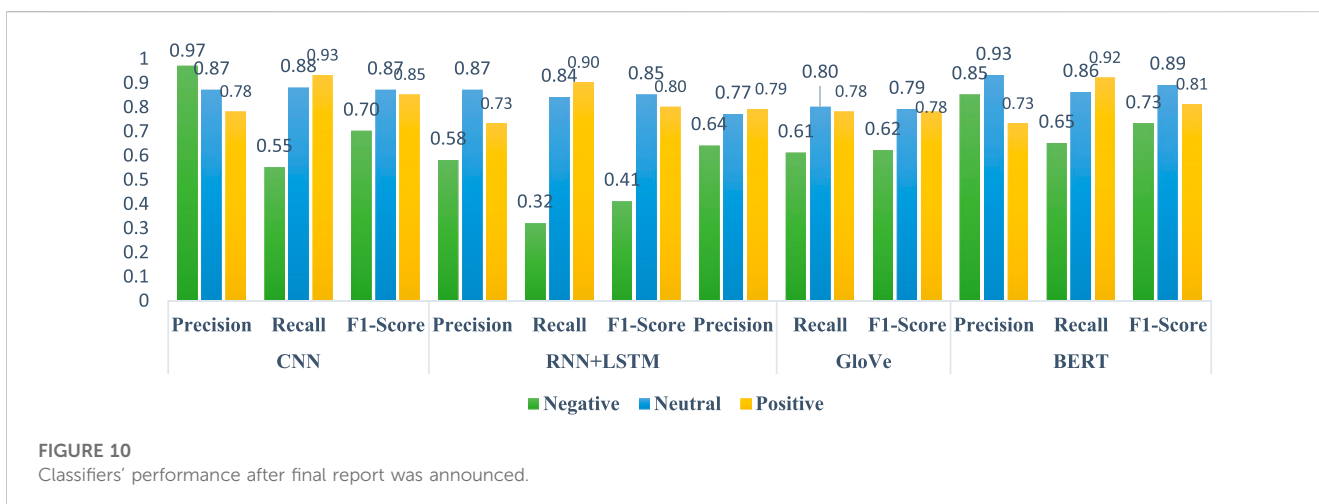


FIGURE 10 Classifiers' performance after final report was announced.

For neutral sentiments, all classifiers exhibit robust performance across the board. BERT stands out as the superior classifier in terms of precision (.93), recall (.86), and F1-score (.89). Slightly trailing BERT, both CNN and RNN + LSTM achieve consistently high scores across all three measures, signifying their capability to predict neutral sentiments accurately. On the contrary, GloVe exhibits the poorest performance across all three performance measures, suggesting it may struggle to discern patterns in the data for precise predictions (Vateekul and Koomsubha, 2016; Goularas and Kamis, 2019). Despite the performance of the GloVe classifier, these results suggest that the classifiers consistently predict neutral sentiments, aligning with their performance in predicting neutral sentiments from the public when the Inquiry was initially announced.

Regarding positive sentiments, the classifiers demonstrate varied predictive performance. CNN exhibits the highest recall (.93) and F1-score (.85), followed by BERT with .92 and .81, and RNN + LSTM with recall and F1-scores of .90 and .80, indicating their proficiency in capturing positive sentiments. However, CNN (.78), BERT (.73), and RNN + LSTM (.73) all achieve low precision scores,

suggesting they struggle with accurate identification of positive sentiments. On the other hand, while on a lower scale, GloVe maintains consistent performance for precision (.79), recall (.78), and F1-score (.78), showcasing its stability in predicting positive sentiments. Compared to the classifiers' performance when the Cullen Commission was initially announced, their performance in identifying positive sentiments remains relatively consistent after the final report's release, albeit with lower scores. These findings suggest that the public might not be content with the Commission's results on money laundering in British Columbia.

Overall, these findings illustrate the classifiers' capabilities in predicting sentiments expressed by Twitter users when the Cullen Commission was announced and after the final report on the Commission's findings was released. BERT consistently demonstrates balanced and accurate performance across all performance metrics for each sentiment class, showcasing its utility for sentiment analysis in response to significant events (Devlin et al., 2018; Acheampong et al., 2021; Alaparthi and Mishra, 2021). Moreover, the stability in the classifiers'

TABLE 4 Classifier performance before and after the final report.

Algorithm	Before	After
CNN	0.95	0.84
RNN + LSTM	0.95	0.76
GloVe	0.84	0.75
BERT	0.93	0.83

performance in identifying sentiments during the initial announcement of the Inquiry compared to their performance after the final report's release highlights their adaptability and resilience in interpreting sentiments over time, indicating their steadfastness in capturing public reactions to significant events (Umer et al., 2021; Vateekul and Koomsubha, 2016; Z. Wang et al., 2016; Yenala et al., 2018).

4.6 Classifier's performance accuracy before and after the commission final report

Table 4 presents the performance accuracy of the algorithms in predicting sentiment classes from public reactions to the announcement of the Cullen Commission and after the release of the final report. As observed in Table 4, the algorithms exhibited robust performance accuracy in predicting sentiment classes from public reactions when the Cullen Commission was announced. The CNN and RNN + LSTM classifiers achieve the highest scores of 0.95, closely followed by BERT at 0.93. An exception lies in the performance of GloVe, which attains an accuracy of 0.84. These results emphasize the capabilities of CNN, RNN + LSTM, and BERT in accurately predicting sentiment classes related to public reactions on AML issues, confirming their proficiency in sentiment analysis.

The classifiers' performance decreases across the board in predicting sentiment classes after the Inquiry's final report is released. Despite the decrease in accuracy, the algorithms still display commendable performance. Once again, CNN and BERT emerge as the top two performers, with test accuracies of .84 and .83, respectively, highlighting their reliability and consistency in predicting sentiment classes. The performance accuracy of the RNN + LSTM (.76) and GloVe (.75) models experienced a substantial decline in analyzing sentiment classes after the announcement of the final report. These findings indicate that the algorithms were highly effective in predicting changes in sentiments from the initial announcement of the Cullen Commission to after the report was released. Although there was a decrease in the algorithms' performance after the final report was released, they still maintained a reasonable level of accuracy, indicating their ability to capture and classify evolving public sentiments over time (Parimala et al., 2021; Vateekul and Koomsubha, 2016; Z. Wang et al., 2016).

5 Discussion and conclusion

This study presents a novel approach to understanding public sentiment surrounding significant events, particularly in the context

of AML compliance. While prior research has explored the application of sentiment analysis to financial crimes, this study uniquely harnesses the power of deep learning models to analyze public reactions to AML compliance in Canada. What sets this research apart is its temporal analysis, allowing for a comparison of public sentiments both before and after the release of the Cullen Commission's final report on money laundering in British Columbia. This temporal lens provides invaluable insights into how public perceptions evolve in response to significant events surrounding money laundering activities. By applying state-of-the-art algorithms like CNN, RNN + LSTM, GloVe, and BERT, this study contributes to the broader field of sentiment analysis by showcasing the scalability, generalizability, and real-world effectiveness of deep learning models.

The evaluation performed in this study offers a comprehensive and rigorous assessment of the selected deep learning models. Regarding granularity, the algorithms' results indicate that joy was the predominant emotion classification when the Inquiry was initially announced. This suggests that the public welcomed the news of the money laundering inquiry. However, sadness and anger became the primary emotion classifications after the final report was released, suggesting that the public did not resonate with the findings from the Cullen Commission. This pattern of emotional classifications is consistently reflected in the performance accuracy, and F1 scores across all algorithms before and after the final report was released.

Concerning sentiment classification, CNN, RNN + LSTM, and BERT consistently demonstrate their ability to predict negative, neutral, and positive sentiment classes. Notably, BERT emerges as a standout performer, achieving remarkable precision, recall, and F1-scores in sentiment classification. In contrast, the GloVe classifier displays the least consistent and weakest performance across all metrics. Notably, except for CNN, the predictive performance for categorizing sentiment classes uniformly declined across all classifiers. This decline is also evident in the accuracy of the algorithms' performance. Excluding GloVe, all classifiers exhibited commendable performance in predicting sentiment classes. Nevertheless, their effectiveness notably diminished after the release of the final report. The consistent decline in predictive performance across all classifiers after the release of the final report points to intriguing insights. It raises questions about the effectiveness of the Cullen Commission's findings and their resonance with the public, providing important context for understanding the implications of the study's findings.

These findings shed light on criticisms of the Commission's efforts (Chin and Magonet, 2022; M. Lokanan and Chopra, 2021; Mulgrew, 2022). Critics have expressed concerns that the Commission's findings did not fully expose the extent of money laundering activities within British Columbia. It is important to note that the subsequent negative public sentiments following the final report's release may be attributed to the Commission's failure to identify responsible parties or entities involved in illicit proceeds' placement, layering, and integration into the financial and non-financial institutions across British Columbia. The Commission's reliance on testimonies from stakeholders across various sectors, who may lack objectivity and provide biased assessments, could have influenced the distorted findings and skewed reporting that fail to address systemic money laundering issues within British Columbia's

financial and non-banking institutions. Another contentious aspect contributing to negative sentiments surrounding the final report could be the Commission's limited scope and reliance on specific industries such as real estate, luxury vehicles, and casinos. The limitation in scope prevented the Commission from addressing broader systemic money laundering issues beyond these industries.

Judging by the public's reaction to the sentiment analyses, the Cullen Commission has significantly raised awareness about money laundering and AML compliance issues in British Columbia. The Inquiry's findings have generated substantial debate, and the recommendations are hoped to impact the province's efforts to combat money laundering. Future research should prioritize the voices of diverse stakeholders, including experts who may not have testified during the hearings. Interviewing a wider array of stakeholders can offer a comprehensive assessment of money laundering and AML compliance within the province. Additionally, analyzing various stakeholders' testimonies can contribute to objectively evaluating the Commission's findings.

5.1 Limitations and future research

This study relies solely on Twitter data, which may not fully capture the sentiments of the entire population who may have responded to the Inquiry findings. Twitter users exhibit unique characteristics and biases that can impact the expressed sentiments. Furthermore, Twitter users represent only a subset of the population, potentially failing to reflect broader societal trends in sentiments and the opinions of those who do not share their views on social media. The 280-character limit of tweets poses a challenge in comprehensively capturing user sentiments and detecting nuances like sarcasm, irony, or humor in the data.

Future research should use diverse data sources by incorporating various social media platforms and offline methods like surveys and interviews to provide a more holistic perspective of public sentiments regarding the Inquiry. Additionally, researchers can explore advanced techniques to mitigate biases in sentiment analysis models, ensuring fairness and accuracy in representing public opinions on AML compliance. Delving into cross-cultural and cross-linguistic sentiment analysis can offer insights into how sentiments vary across regions and languages, contributing to a more nuanced

understanding of public perceptions. To account for temporal changes in sentiments over time, future studies may adopt dynamic systems analysis methods, enabling the modeling of sentiment analysis in response to specific events.

Data availability statement

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

Author contributions

ML: Conceptualization, Data curation, Formal Analysis, Methodology, Project administration, Writing—original draft, Writing—review and editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. The funding organization is the Social Science Research Council of Canada, New Frontiers in Research Fund Grant.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Acheampong, F. A., Nunoo-Mensah, H., and Chen, W. (2021). Transformer models for text-based emotion detection: a review of BERT-based approaches. *Artif. Intell. Rev.* 54 (8), 5789–5829. doi:10.1007/s10462-021-09958-2
- Ai, L., and Tang, J. (2011). Risk-based approach for designing enterprise-wide AML information system solution. *J. Financial Crime* 18 (3), 268–276. doi:10.1108/13590791111147488
- Alaparthi, S., and Mishra, M. (2021). BERT: a sentiment analysis odyssey. *J. Mark. Anal.* 9 (2), 118–126. doi:10.1057/s41270-021-00109-8
- Alkhalili, M., Qutqut, M. H., and Almasalha, F. (2021). Investigation of applying machine learning for watch-list filtering in anti-money laundering. *IEEE Access* 9, 18481–18496. doi:10.1109/ACCESS.2021.3052313
- Alswaidan, N., and Mohamed El, B. M. (2020). A survey of state-of-the-art approaches for emotion recognition in text. *Knowl. Inf. Syst.* 62 (8), 2937–2987. doi:10.1007/s10115-020-01449-0
- Aluko, A., and Bagheri, M. (2012). The impact of money laundering on economic and financial stability and on political development in developing countries: the case of Nigeria. *J. Money Laund. Control* 15 (4), 442–457. doi:10.1108/13685201211266024
- Bandana, R. (2018). "Sentiment analysis of movie reviews using heterogeneous features," in Proceedings of the 2018 2nd International Conference on Electronics, Materials Engineering & Nano-Technology (IEMENTech), Kolkata, India, May 2018 (IEEE), 1–4. doi:10.1109/IEMENTECH.2018.8465346
- Bashir, S., and Ghous, Dr H. (2020). Detecting mobile money laundering using genetic algorithm as feature selection method with classification method. *LC Int. J. STEM (ISSN 2708-7123)* 1 (4), 121–129. doi:10.5281/zenodo.5149794
- Basile, V., Cauteruccio, F., and Giorgio, T. (2021). How dramatic events can affect emotionality in social posting: the impact of COVID-19 on reddit. *Future Internet* 13 (2), 29. doi:10.3390/fi13020029

- Bhat, A. Z., EMA, T. K., and Asim, F. (2022). Evaluation of neural network model for better classification of data and optimum solution of real-world problems. *J. Student Res.*
- Bhoir, S., Ghorpade, T., and Mane, V. (2017). "Comparative analysis of different word embedding models," in Proceedings of the 2017 International Conference on Advances in Computing, Communication and Control (ICAC3), Mumbai, December 2017 (IEEE), 1–4. doi:10.1109/ICAC3.2017.8318770
- Bonifazi, G., Cauteruccio, F., Corradini, E., Marchetti, M., Sciarretta, L., Ursino, D., et al. (2022). A space-time framework for sentiment scope analysis in social media. *Big Data Cognitive Comput.* 6 (4), 130. doi:10.3390/bdcc6040130
- Borovkova, S., Garmaev, E., and Lammers, P. (2017). SenSR: a sentiment-based systemic risk indicator. *SSRN Electron. J.* doi:10.2139/ssrn.2951036
- Carlos-Roca, Rodriguez, L., Torres, I. H., and Carles Fernandez, T. (2018). "Facial recognition application for border control," in Proceedings of the 2018 International Joint Conference on Neural Networks (IJCNN), Rio de Janeiro, July 2018 (IEEE), 1–7. doi:10.1109/IJCNN.2018.8489113
- Chang, V., Le Minh, T. D., Alessandro, Di S., Sun, Z., and Fortino, G. (2022). Digital payment fraud detection methods in digital ages and industry 4.0. *Comput. Electr. Eng.* 100, 107734. doi:10.1016/j.compeleceng.2022.107734
- Chen, G., Liu, L., Hu, W., and Pan, Z. (2018). "Semi-supervised object detection in remote sensing images using generative adversarial networks," in Proceedings of the IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia, July 2018 (IEEE), 2503–2506. doi:10.1109/IGARSS.2018.8519132
- Chin, S., and Magonet, J. (2022). *A deep dive into Cullen's final report*. Vancouver: British Columbia Civil Liberties Association. Available at: <https://bccla.org/2022/08/a-deep-dive-into-cullens-final-report/>.
- Craja, P., Kim, A., and Lessmann, S. (2020). Deep learning for detecting financial statement fraud. *Decis. Support Syst.* 139, 113421. doi:10.1016/j.dss.2020.113421
- Cullen, A. (2022). *Commission of inquiry into money laundering in British Columbia*. Victoria: Government of British Columbia. Available at: <https://cullencommission.ca/>.
- Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K. (2018). BERT: pre-training of deep bidirectional transformers for language understanding. Available at: <https://arxiv.org/abs/1810.04805>.
- Ding, J., Xu, M., Ying, K., Lin, K.-Yi, and Zhang, M. (2023). Customer opinions mining through social media: insights from sustainability fraud crisis - volkswagen emissions scandal. *Enterp. Inf. Syst.* 17 (8), 2130012. doi:10.1080/17517575.2022.2130012
- Effrosymidis, D., Karasakalidis, A. I., Sylaios, G., and Arampatzis, A. (2022). The climate change twitter dataset. *Expert Syst. Appl.* 204, 117541. doi:10.1016/j.eswa.2022.117541
- Ekman, P. (1992). An argument for basic emotions. *Cognition Emot.* 6 (3–4), 169–200. doi:10.1080/02699939208411068
- Feng, Y., and Cheng, Y. (2021). Short text sentiment analysis based on multi-channel CNN with multi-head attention mechanism. *IEEE Access* 9, 19854–19863. doi:10.1109/ACCESS.2021.3054521
- Gan, C., Feng, Q., and Zhang, Z. (2021). Scalable multi-channel dilated CNN-BiLSTM model with attention mechanism for Chinese textual sentiment analysis. *Future Gener. Comput. Syst.* 118, 297–309. doi:10.1016/j.future.2021.01.024
- Goularas, D., and Kamis, S. (2019). "Evaluation of deep learning techniques in sentiment analysis from twitter data," in Proceedings of the 2019 International Conference on Deep Learning and Machine Learning in Emerging Applications (Deep-ML), Istanbul, Turkey, August 2019 (IEEE), 12–17. doi:10.1109/Deep-ML.2019.00011
- Hajek, P., and Henriques, R. (2017). Mining corporate annual reports for intelligent detection of financial statement fraud – a comparative study of machine learning methods. *Knowledge-Based Syst.* 128, 139–152. doi:10.1016/j.knosys.2017.05.001
- Han, J., Barman, U., Hayes, J., Du, J., Burgin, E., and Wan, D. (2018). *NextGen AML: distributed deep learning based language technologies to augment anti money laundering investigation*. Melbourne, Australia: Association for Computational Linguistics. <https://doras.dcu.ie/23358/>.
- Han, J., Huang, Y., Liu, S., and Towey, K. (2020). Artificial intelligence for anti-money laundering: a review and extension. *Digit. Finance* 2 (3–4), 211–239. doi:10.1007/s42521-020-00023-1
- Hans, C. (2011). Elastic net regression modeling with the orthant normal prior. *J. Am. Stat. Assoc.* 106 (496), 1383–1393. doi:10.1198/jasa.2011.tm09241
- Havrlant, L., and Kreinovich, V. (2017). A simple probabilistic explanation of term frequency-inverse document frequency (Tf-Idf) heuristic (and variations motivated by this explanation). *Int. J. General Syst.* 46 (1), 27–36. doi:10.1080/03081079.2017.1291635
- He, T., Huang, W., Qiao, Yu, and Yao, J. (2016). Text-attentional convolutional neural network for scene text detection. *IEEE Trans. Image Process.* 25 (6), 2529–2541. doi:10.1109/TIP.2016.2547588
- Hendriyety, N., and Grewal, B. S. (2017). Macroeconomics of money laundering: effects and measurements. *J. Financial Crime* 24 (1), 65–81. doi:10.1108/JFC-01-2016-0004
- Kayıkçı, Ş. (2022). SenDemonNet: sentiment analysis for demonetization tweets using heuristic deep neural network. *Multimedia Tools Appl.* 81 (8), 11341–11378. doi:10.1007/s11042-022-11929-w
- Kumar, A., Das, S., Tyagi, V., Shaw, R. N., and Ghosh, A. (2021). "Analysis of classifier algorithms to detect anti-money laundering," in *Computationally intelligent systems and their applications. Studies in computational intelligence* (Singapore: Springer Singapore). doi:10.1007/978-981-16-0407-2_11
- Larsen, K. (2019). *Fast cars and bags of cash: Gangsters using B.C. Luxury car market to launder dirty money*. Ottawa, Canada: Canadian Broadcasting Corporation. <https://www.cbc.ca/news/canada/british-columbia/bc-money-laundering-report-cby-german-may-2019-1.5126386>.
- Leuprecht, C., Jenkins, C., and Hamilton, R. (2023). Virtual money laundering: policy implications of the proliferation in the illicit use of cryptocurrency. *J. Financial Crime* 30 (4), 1036–1054. doi:10.1108/JFC-07-2022-0161
- Liao, J., and Acharya, A. (2011). Transshipment and trade-based money laundering. *J. Money Laund. Control* 14 (1), 79–92. doi:10.1108/13685201111098897
- Liao, S., Wang, J., Yu, R., Sato, K., and Cheng, Z. (2017). CNN for situations understanding based on sentiment analysis of twitter data. *Procedia Comput. Sci.* 111, 376–381. doi:10.1016/j.procs.2017.06.037
- Lindsay, B. (2019). *Money laundering funded \$5.3B in B.C. Real estate purchases in 2018, report reveals*. Ottawa, Canada: Canadian Broadcasting Corporation. <https://www.cbc.ca/news/canada/british-columbia/laundered-money-bc-real-estate-1.5128769>.
- Lokanan, M., and Chopra, G. (2021). "Money laundering in real estate (RE): the case of Canada," in *Advances in finance, accounting, and economics*, edited by *abdul rafay* (Pennsylvania: IGI Global), 53–90. doi:10.4018/978-1-7998-8758-4.ch003
- Lokanan, M. E. (2022). Predicting money laundering using machine learning and artificial neural networks algorithms in banks. *J. Appl. Secur. Res.*, 1–25. doi:10.1080/19361610.2022.2114744
- Lokanan, M. E. (2023). The tinder swindler: analyzing public sentiments of romance fraud using machine learning and artificial intelligence. *J. Econ. Criminol.* 2, 100023. doi:10.1016/j.jeconc.2023.100023
- Lyu, S., and Liu, J. (2021). Convolutional recurrent neural networks for text classification. *J. Database Manag.* 32 (4), 65–82. doi:10.4018/JDM.2021100105
- Mohammed, H. N., Malami, N. S., Thomas, S., Abdul Aiyelabegan, F., Adam Imam, F., and Ginsau, H. H. (2022). "Machine learning approach to anti-money laundering: a review," in Proceedings of the 2022 IEEE Nigeria 4th International Conference on Disruptive Technologies for Sustainable Development (NIGERCON), Lagos, Nigeria, April 2022 (IEEE), 1–5. doi:10.1109/NIGERCON54645.2022.9803072
- Mulgrew, I. (2022). *Ian Mulgrew: Cullen commission report on B.C. Money laundering 'an enormous opportunity missed*. Virtual Way: Vancouver Sun. <https://vancouver.sun.com/opinion/ian-mulgrew-cullen-commission-report-an-enormous-opportunity-missed>.
- Nagarsekar, U., Aditi, M., Priyanka, K., and Dhananjay, R. K. (2013). "Emotion detection from the SMS of the internet," in Proceedings of the 2013 IEEE Recent Advances in Intelligent Computational Systems (RAICS), Trivandrum, India, December 2013 (IEEE), 316–321. doi:10.1109/RAICS.2013.6745494
- Omar, N., Zulaikha, A. J., and Smith, M. (2017). Predicting fraudulent financial reporting using artificial neural network. *J. Financial Crime* 24 (2), 362–387. doi:10.1108/JFC-11-2015-0061
- Pang, Bo, and Lee, L. (2009). Opinion mining and sentiment analysis. *Comput. Linguist.* 35 (2), 311–135. doi:10.1561/15000000011
- Parimala, M., Swarna Priya, R. M., Praveen Kumar Reddy, M., Lal Chowdhary, C., Kumar Poluru, R., and Khan, S. (2021). Spatiotemporal-based sentiment analysis on tweets for risk assessment of event using deep learning approach. *Softw. Pract. Exp.* 51 (3), 550–570. doi:10.1002/spe.2851
- Phan, H. T., Tran, V. C., Thanh Nguyen, N., and Hwang, D. (2020). Improving the performance of sentiment analysis of tweets containing fuzzy sentiment using the feature ensemble model. *IEEE Access* 8, 14630–14641. doi:10.1109/ACCESS.2019.2963702
- Poria, S., Cambria, E., Bajpai, R., and Hussain, A. (2017). A review of affective computing: from unimodal analysis to multimodal fusion. *Inf. Fusion* 37, 98–125. doi:10.1016/j.inffus.2017.02.003
- Ray, P., Chakrabarti, A., Ganguli, B., and Das, P. K. (2018). Demonetization and its aftermath: an analysis based on twitter sentiments. *Sādhanā* 43 (11), 186. doi:10.1007/s12046-018-0949-0
- Schmunk, R. (2018). *B.C. Casinos 'unwittingly served as laundromats' for proceeds of crime: report*. Ottawa, Canada: Canadian Broadcasting Corporation. <https://www.cbc.ca/news/canada/british-columbia/bc-money-laundering-report-1.4723958>.
- Stewart, D. (2023). *Next-gen anti-money laundering – robotics, semantic analysis and AI*. Cary, North Carolina: SAS. https://www.sas.com/en_ca/insights/articles/risk-fraud/next-generation-anti-money-laundering.html.
- Tanwar, P., and Rai, P. (2020). A proposed system for opinion mining using machine learning, NLP and classifiers. *IAES Int. J. Artif. Intell.* 9 (4), 726–733. doi:10.11591/ijai.v9.i4.pp726-733

- Thi, M.Ha, Withana, C., Thi Huong Quynh, N., and Tran Quoc Vinh, N. (2020). "A novel solution for anti-money laundering system," in Proceedings of the 2020 5th International Conference on Innovative Technologies in Intelligent Systems and Industrial Applications (CITISIA), Sydney, Australia, November 2020 (IEEE), 1–6. doi:10.1109/CITISIA50690.2020.9371840
- Twitter (2023). Academic research. Available at: <https://developer.twitter.com/en/use-cases/do-research/academic-research>.
- Umer, M., Imran, A., Mehmood, A., Kumari, S., Ullah, S., and Choi, G. S. (2021). Sentiment analysis of tweets using a unified convolutional neural network-long short-term memory network model. *Comput. Intell.* 37 (1), 409–434. doi:10.1111/coin.12415
- Vateekul, P., and Koomsubha, T. (2016). "A study of sentiment analysis using deep learning techniques on Thai twitter data," in Proceedings of the 2016 13th International Joint Conference on Computer Science and Software Engineering (JCSSE), Khon Kaen, Thailand, July 2016 (IEEE), 1–6. doi:10.1109/JCSSE.2016.7748849
- Vathsala, M. K., and Ganga, H. (2020). RNN based machine translation and transliteration for twitter data. *Int. J. Speech Technol.* 23 (3), 499–504. doi:10.1007/s10772-020-09724-9
- Wang, J., Yu, L.-C., Robert Lai, K., and Zhang, X. (2020). Tree-structured regional CNN-LSTM model for dimensional sentiment analysis. *IEEE/ACM Trans. Audio, Speech, Lang. Process.* 28, 581–591. doi:10.1109/TASLP.2019.2959251
- Wang, Z., Chong, C. S., Lan, L., Yang, Y., Ho, S. B., and Tong, J. C. (2016). "Fine-grained sentiment analysis of social media with emotion sensing," in Proceedings of the 2016 Future Technologies Conference (FTC), San Francisco, CA, USA, December 2016 (IEEE), 1361–1364. doi:10.1109/FTC.2016.7821783
- Xiuguo, Wu, and Du, S. (2022). An analysis on financial statement fraud detection for Chinese listed companies using deep learning. *IEEE Access* 10, 22516–22532. doi:10.1109/ACCESS.2022.3153478
- Yenala, H., Jhanwar, A., Chinnakotla, M. K., and Goyal, J. (2018). Deep learning for detecting inappropriate content in text. *Int. J. Data Sci. Anal.* 6 (4), 273–286. doi:10.1007/s41060-017-0088-4
- Zhang, Y., Tong, J., Wang, Z., and Gao, F. (2020). "Customer transaction fraud detection using xgboost model," in Proceedings of the 2020 International Conference on Computer Engineering and Application (ICCEA), Guangzhou, China, March 2020, 554–558. doi:10.1109/ICCEA50009.2020.00122