Social Network Analysis and Mining (2023) 13:39
https://doi.org/10.1007/s13278-023-01041-8

ORIGINAL ARTICLE

Sentiment analysis and emotion detection of post-COVID educational Tweets: Jordan case

Evon Qaqish[1] · Aseel Aranki[1] · Wael Etaiwi[1]

Received: 4 November 2022 / Revised: 6 January 2023 / Accepted: 17 February 2023 / Published online: 2 March 2023

Abstract
Education evolved dramatically under Covid-19, and owing to the conditions, distant learning became mandatory. However, this has opened new realities to the educational business under the label of "Hybrid-Learning," where educational institutions are still using online learning in addition to face-to-face learning, which has changed people's lives and split their opinions and emotions. As a result, this study investigated the Jordanian community's perspectives and feelings on the transition from pure face-to-face education to blended education by examining related tweets in the post-COVID era. Specifically, using NLP Emotion detection and Sentiment Analysis approaches, as well as deep learning models. As a result of analyzing the collected tweets, 18.75% of studied Jordanian's community sample are dissatisfied (Anger and Hate), 21.25% are negative (Sad), 13% are Happy, and 24.50 percent are Neutral about it.

Keywords Jordan · Education · Emotion detection · Natural language processing · Deep learning

1 Introduction

Post-COVID Education is a word that has been commonly used since mid-2021, when the arrival of the COVID-19 pandemic affected the way our society appears, notably the Educational environment and the Learning Process, which impacts nearly every student in the world (Zhao and Watterston 2021). As a result, educators and educational institutions worldwide made a concerted effort to adapt and innovate. Barron-Estrada et al. (2019) as the circumstances necessitated, these adjustments were implemented fast. Almost immediately, numerous educational institutions have begun to provide instructions and lectures through distance learning technologies (Zhao and Watterston 2021; Neuwirth et al. 2021) changes in teaching delivery elicited a swift and enthusiastic response from educators. Lockee (2021). Consequently, it is apparent and widely agreed that "the recession has fostered innovation in the education sector." Mahfoodh and AlAtawi (2020). Thereby, the educational system took on a new shape. Even after the lockdown was lifted, many institutions decided to employ the hybrid-learning model and continue using online and digital learning (Aditya 2021; Arnove 2020).

However, no matter how widely this new learning model is used, when transitioning from traditional face-to-face teaching to distance learning even if partially, it is critical to guarantee that, this digital learning technique is at least a viable option, if not superior to traditional education. Mujahid et al. (2021) from a scientific standpoint, the success of this new normal of teaching and learning outcomes is determined and tracked by the well-known Bloom's Taxonomy, which categorizes educational outcomes into six domains: comprehension, knowledge, analysis, synthesis, application, and evaluation of the student's performance. Gneezy et al. (2019), Tadesse and Muluye (2020) nevertheless, what about the public's perceptions, which are most affected by the new situation? Here, we will be confronted with a plethora of thoughts, ideas, points of view, and emotions that will provide radically different perspectives on the success and acceptability of distance learning, unlocking hidden information and knowledge, imposing a significant need to evaluate and analyze people's emotions and opinions of the effectiveness and challenges of online

✉ Evon Qaqish
evo20208042@std.psut.edu.jo

Aseel Aranki
ase20208072@std.psut.edu.jo

Wael Etaiwi
w.etaiwi@psut.edu.jo

[1] Business Information Technology, Princess Sumaya University for Technology, Amman, Jordan

learning. de Mooij et al. (2022) which will give a broader and more in-depth assessment of the hybrid-learning decisions' impact on people's lives and how it has left its imprint on their psychological views for government and educational institutions.

The Jordanian community's views and emotions toward the new hybrid-learning paradigm will be investigated in this study. Jordan was chosen for this study because it was one of the first countries to adopt this new educational reality, with few pure hybrid-educational studies taking it into account, and to the best of our knowledge, this study will be one of the first to analyze Jordanians' emotions gleaned from Twitter. The tweets are from people of all ages, origins, and educational levels, and they cover the majority of Jordanian's opinions' of online and hybrid educational factors for post-COVID eras. Such factors include obstacles, benefits, drawbacks, and difficulty adjusting and adapting to such changes. Imran et al. (2020). Through the deep and machine learning techniques have demonstrated their supremacy in various fields and industries over the last two decades such as image processing, Ashraf et al. (2019) object detection, Ashraf et al. (2020) and most commonly the Natural Language Processing (NLP) (Martinez 2010).

During this study, the deep learning methodologies of emotion detection and sentiment classification will be used to assess people's opinions and experiences in Jordan regarding the shift from pure face-to-face education to hybrid education. In addition, by employing the Natural Language Processing techniques for text processing, the gap between traditional and online learning will be bridged.

2 Literature review

Analyzing people's emotions and opinions has become one of the most critical topics in the research world since COVID's emergence and its changes to all businesses, particularly the educational industry. In recent years, a considerable amount of research has been conducted to create methods for assessing and documenting the process of sentiment analysis in many languages. In the same context, some other studies took analyzing and classifying people's opinions about hybrid and blended learning to a new level of detail by delving deeper into their feelings and emotions to introduce new emotion classification models and better comprehend the new educational reality.

The pandemic affected the globe in three waves: the virus's first-, second-, and third-order impacts, which appeared across distinct time frames; nevertheless, in the realm of education, these waves split the learning process into during- and post-COVID, and the same is true for research areas.

2.1 Sentiment analysis and emotion detection during-COVID

According to Mujahid et al. (2021) study, the top three online education issues that people are concerned about are the unpredictability of campus opening dates, children's difficulties grasping online education, and trailing efficient networks linked with eLearning. As a result of analyzing 17,155 e-learning-related tweets, the model was built and evaluated using Bag of Words and Term Frequency-Inverse Document Frequency in addition, the Valence Aware Dictionary for Sentiment Reasoning, SentiWordNet and the TextBlob used to analyze the subjectivity and polarity tweets' score. Unsurprisingly, the study also resulted in having the number of negative tweets regarding these issues exceeding the positive ones. In the same context, the study of Ashwitha et al. (2021) that analyzed the emotions of the educational-related feedback during the pandemic, using the naive-based classifier and NLP toolkit, resulting in the same outcome where, the negative tweets has exceeded the positive ones. Such results are predictable especially in the developing and undeveloped countries, where online learning fails to provide the intended results in developing nations, as due to the technical and monetary issues harden accessing the internet to the majority of students. Tadesse and Muluye (2020)

As a result of the impact of such a pandemic on the education system, these nations pushed on improving broadcast teaching, virtual class infrastructures and online teaching. Sakr et al. (2021). Where according to the study conducted by Aristovnik et al. (2020) which undertaken nearly 30,000 sample students tweets from 62 countries analyzed using the Naive-based classifier, concluded that the lower living standard students with financial problems were not satisfied with their academic life as the pandemic affected their life more. From a new angle, Ali (2021) studied the Arabic tweets during the pandemic to perform a comprehensive emotion and sentiment analysis in order to uncover the hidden causes behind the negative sentiment, through NLP and ML algorithms to extract the information and determine polarity and detect the feeling. In addition to the "National Research Council Canada (NRC) Word-Emotion Lexicon" to calculate the emotions presence. Finally, the Logistic Regression (LR), Multinomial Naïve Bayes (MNB), Support Vector Machine (SVM) and K Nearest Neighbor (KNN) were examined, resulting in having a model with maximum accuracy to analyze the people perception regard coronavirus of 90% using Support Vector Machine classifier. In the same vien, Alturayeif and Luqman (2021) employed two transformer-based models for sentiment detection of the Arabic tweets, with the assumption that many emotion can exist in the same tweet. Resulting in F1-Micro score of 0.72 with the ability to precisely predict the educational-based tweets. While Gandolfi et al. (2021), used the Fi-Micro score to the word embedding word2vec. Where two bags-of-words pre-trained models investigated, in addition to using the Naïve Bayes as the baseline classifier. In addition, studying ensemble-based and single-based machine learning classifiers in two experiments, with the SMOTE and without SMOTE resulting in noticeable improvement with for the ensemble with using SMOTE. In the same framework, Aditya (2021), proposed a model for the students' sentiments during their learning process through the techniques of Word2vec, the Naïve Bayes and Decision Tree so that the Egyptian student's opinion on the new learning flow is understood. Similarly Althagafi et al. (2021), studied the Saudi Arabia people attitude toward the digital learning, using tweets collected through the hashtags trending during 2020; unpredictably the study showed that people in Saudi Arabia have maintained a neutral response toward the new learning reality.

On the other hand, Durães et al. (2021) have taken a new direction of emotion detection through embedding the various machine learning and filtering technique into "Intelligent Tutor Application" that process the emotions in a noninvasive way. The study took a sample of 160 students to test the application resulting in high accuracy emotion prediction, which enhance monitoring the students' emotion patterns in distance learning over distinct periods and make it possible to detect those factors non-intrusively and dynamically. Comparably, Wang and Cruz (2020) analyzed the spatial and temporal patterns for the whole US tweets during the spring, summer and fall semesters to discover the topics and thoughts of people at colleges and universities are communicating in via the deep learning approaches. Likewise, Imran et al. (2020) used the deep learning Multilayer LSTM technique to sentimentally analyzing 27,357 COVID online learning related tweets providing a unique and novel way to validate the supervised deep learning models. However, according to Chintalapudi et al. (2021) that used the BERT and LSTM deep learning technique to conduct sentiment analysis on 3090 tweets related to COVID-19, the accuracy of such analysis was not significant on the tweets. While this was not the case for Samuel et al. (2020) that analyzed 900, 00 tweets using the KNN and Naïve Bayes where the study resulted in significantly high accuracy and generalization.

Other studies go to the direction of analyzing the sentiment and emotions of public opinions regarding the online learning blogs and articles. Where Bhagat et al. (2021) used the Dictionary-based approach to analyze the opinion regarding 154 articles collected from Google however, the study did not used a machine learning approach. Therefore, Umair et al. (2021), used the Support Vector Machine, and Naïve Bayes algorithms to analyze the micro-blogging used by students to express their opinions and provide their feedback regard the online learning, resulting in a high accuracy model with 85.62% has been achieved using Support vector machine with *K*-fold cross-validation (Table 1).

2.2 Sentiment analysis and emotion detection post-COVID

However, we cannot ignore the fact that COVID's impact on the world, particularly education, did not go away even after schools and universities reopened, with almost all educational institutions opting for hybrid and blended learning options, and people's behavior has changed in some ways forever worldwide. Therefore, Meishar-Tal and Levenberg (2021) analyzed the emotion of 239 academic lecturers to classify their emotions into Success, Opportunity, Failure and Threats toward the new educational reality using the

Table 1 Summary of during-COVID educational studies

First author	Year	Dataset	Technique (s)
Mujahid et al. (2021)	2021	17,155 e-learning-related tweets	TF-IDF and SentiWordNet.
Ashwitha et al. (2021)	2021	Emotions of educational feedback (Tweets)	NLP toolkit and naive-based classifier
Aristovnik et al. (2020)	2020	30,383 tweets	Naive-based classifier
Ali (2021)	2021	Tweets in the Arabic language	MNB, KNN, LR and SVM
Gandolfi et al. (2021)	2019	Arabic tweets	Fi-Micro score, Naïve Bayes
Aditya (2021)	2020	student's tweets	Word2vec technique and Naïve Bayes
Durães et al. (2021)	2021	application resulting	Machine learning and Filtering Technique
Wang and Cruz (2020)	2020	spatial and temporal patterns tweets	Deep Learning approaches
Imran et al. (2020)	2020	27,357 tweets	Deep-learning Multilayer LSTM technique
Chintalapudi et al. (2021)	2021	3090 tweets	BERT and LSTM deep learning technique
Samuel et al. (2020)	2020	900,00 tweets	the KNN and Naïve Bayes
Bhagat et al. (2021)	2021	154 articles collected from Google	Dictionary-based approach
Umair et al. (2021)	2021	micro-blogging used by students	SVM, and Naïve Bayes algorithms

deep learning BERT technique. In the same point of view, Nacheva (2022) established a study framework for measuring the higher education students' emotional attitudes specifically in social media and proved its practical application by obtaining data from the social network Twitter and using data mining techniques to analyze vast amounts of textual material. In same framework, Shelke et al. (2022) recognized an emotion mis-prediction caused by irrelative feature extraction from textual data therefore, deep Neural Network (LRA-DNN) model conducted on students' feedback and comments where it resulted in considerably high accuracy, precision and recall. In the talk of deep learning, Javed and Muralidhara (2022) investigated the Long Short-Term Memory (LSTM) models on educational micro-blogging and tweets to recognize and classify the emotions of these textual inputs resulting in accuracy of 60.57% of predicting. While Dey and Dasgupta (2022) used the NLP techniques on Sina Weibo data to classify the current educational COVID's news into seven categories "warning and advice", "notice and action", "contribution of money, products or services", "emotional support, looking for help", "expressing and assessing doubts and counter-rumors based on situational information". de Mooij et al. (2022) worked on universities and schools' emails, where the study used deep convolution neural network (DCNN) technique to build the model obtaining accuracy of 85.8%.

From new perspective, Veletsianos et al. (2018) took a sample of 655 from YouTube comments and 774,939 from Facebook comments and reply to conduct sentiment analysis about educational-related videos in order to identify the factors that may people in general and students in specific to have different types of sentiments and opinions. Whereas, Moustakas and Robrade (2022) conducted a sentiment analysis to uncover the similar students' opinions regarding the challenges of eLearning especially for sports and physical schools through conducting Naïve Bayes approach. For Shukla and Garg (2021) used the online discussion forums between the students to analyze the unstructured data to explore the various emotions of the students using the semantic approach.

While Toçoğlu and Onan (2021) evaluated the higher educational institutions through taking 700 Turkish students' written reviews and analyze it using SVM, Naïve Bayes, LR, and random forest algorithm. In Saudi Arabia Hashish et al. (2022) studied the students' experience regarding the learning post-COVID through taking a sample of 355 students from one of the well-known universities in Saudi Arabia. For New Zeeland, Almotiri (2022) used the Natural Language Processing approach to analyze 1162 tweets and used the "AFINN Lexicon sentiment analysis method" to calculate the sentiment value of all of the tweets.

To summarize, the most frequent techniques utilized to conduct the analysis include Nave Bayes, Support Vector Machine, KNN, and different Natural Language Processing approaches. In contrast to the other approaches that demonstrated their efficiency and accuracy, the Deep Neural Network technique's accuracy was not extremely significant when compared to the Deep Learning techniques.

Patterns in the research for the era of during-COVID were noticed, as the research papers were divided into three categories: publications with comparable data kinds and approaches, articles that established new upgraded models and articles that addressed a problem in a specific area or region. Moreover, just a few studies worked on emotion recognition, with the majority of articles focusing solely on sentiment analysis. Furthermore, several of the papers did not clarify the criterion for data collection.

Throughout this research, an emotion detection and sentiment analysis will be performed on 4000 tweets obtained from October 20, 2021 to February 2022 on people's thoughts about hybrid learning in Jordan (Table 2).

Table 2 Summary of post-COVID educational studies

First author	Year	Dataset	Technique (s)
Meishar-Tal and Levenberg (2021)	2021	239 academic lecturers	BERT
Nacheva (2022)	2022	Students interactions and tweets on Social Media	data mining techniques
Shelke et al. (2022)	2022	Students' feedback and comments	deep Neural Network (LRA-DNN)
Javed and Muralidhara (2022)	2022	Sina Weibo data	Natural Language Processing (NLP) techniques
Dey and Dasgupta (2022)	2022	Universities and schools' emails	deep convolution neural network (DCNN)
Moustakas and Robrade (2022)	2022	eLearning discussion forum	Naïve Bayes approach
Shukla and Garg (2021)	2022	Online discussion forums	semantic approach
Yayla et al. (2021)	2021	700 Turkish students' reviews	SVM, Naïve Bayes, LR, Random Forest
Hashish et al. (2022)	2022	355 students	AFINN Lexicon sentiment analysis
Almotiri (2022)	2021	1162 tweets	AFINN Lexicon sentiment analysis

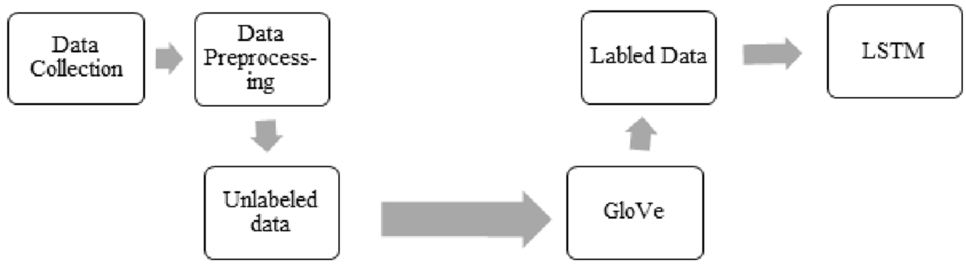

Fig. 1 Workflow pipeline

Table 3 Educational trending hashtags

Hashtags in Arabic	Hashtags' English translation	Hashtag date
نعم للتعليم عن بعد	With the distance learning	Jan 3, 2021
نعم للتعليم الالكتروني	With the e-learning	Oct 4, 2021
التعليم الوجاهي اولويه وطنيه	Face-to-face education is a national priority	Nov 18, 2022
لا للتعليم الالكتروني	Against the e-learning	Jan 18, 2022
لا للتعليم عن بعد	Against the distance education	Feb 3, 2022
نعم للتعليم المباشر	With the face-to-face education	Feb 20, 2022

3 Methodology

In this section, a detailed discussion regard the dataset, collecting techniques, and methodology utilized to detect emotions in Arabic tweets using deep learning algorithms will be presented (Fig. 1).

3.1 Data collection

The data for this study were obtained from Twitter using the Twitter API via Python programming, resulting in a dataset in *.csv format. Twitter was chosen as a key data source since it is one of the most popular platforms for Jordanians to express themselves formally.

The data were gathered using the trending educational-related hashtags shown in Table 3 for the period from January 2021 until February 2022. As during this time, the majority of hybrid educational-related decisions and rules were made and the lockdown was nearly lifted.

Furthermore, in order to get the most relevant tweets, the tweets were filtered in greater depth throughout the extraction phase using three criteria: (1) The keywords shown in Table 4 were chosen based on relevant subject and the words that were most frequently repeated in the tweets. (2) Geographic location utilizing the words "Jordan", "Hashemite Kingdom of Jordan." or the Jordanian cities (3) The day and month in which each hashtag was trending.

Table 4 Filtering keywords

Keywords in Arabic	Keywords' english translation
التربية	Education and upbringing
التعليم	Education
الفصل الثاني	Second semester
الفصل الجاي	Next semester
التعليم العالي	Higher education
الفصل الدراسي الاول	First semester

3.2 Data description

The dataset consisted of 4000 tweets in Arabic Language; each tweet was made up of Arabic words, English words, Emojis, Special characters, and Punctuation marks. Where the Tweets were divided into three categories based on people's opinions about the blended education: 800 with the blended education, 1300 against blended education, and the remaining 1900 with both categories hashtags representing the neutral group, as summarized in Table 5.

Table 5 Hashtags classification

Category	Hashtags in Arabic	Hashtags' English Translation
With the blended education	نعم للتعليم عن بعد	With The Distance Learning
With the blended education	نعم للتعليم الالكتروني	With the E-Learning
Against the blended education	التعليم الوجاهي اولويه وطنيه	Face-to-face Education is a National Priority
Against the blended education	لا للتعليم الالكتروني	Against the E-Learning
Against the blended education	لا للتعليم عن بعد	Against the Distance Education
Against the blended education	نعم للتعليم المباشر	With the Face-to-Face Education

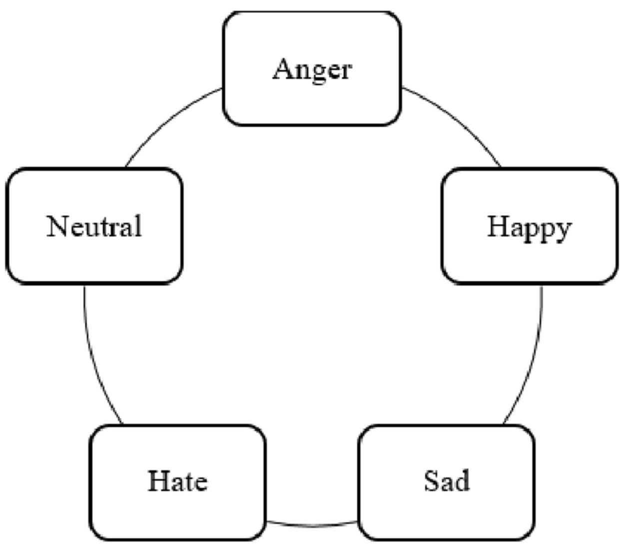

Fig. 2 Emotions labeling

3.3 Emotion labeling

In this study, Ekman's categorization model served as the foundation for our classification and it also served as a foundation for other emotion models. The emotions were classified into; Neutral, Anger, Happy, Sad and Hate, as illustrated in Fig. 2. According to Paul Ekman's emotional discoveries, these feelings are accompanied and shared by everyone throughout cultures. Depending on Posner et al. (2005) paradigm, Ekman distinguishes via people's qualitative experiences and how they respond to each. In greater detail, Anger and Hate (dissatisfactory) are determined by the manner in which the stimulus happens, which is more concentrated than Sad owing to the hedonic worth of the stimulus (physiological needs). While Happy is indicated as locomotor activity, which is connected to the gratifying amount of success on the other side of the equation.

3.4 Data preprocessing

Since our dataset is in the Arabic language, we performed particular preprocessing using Python programming to appropriately and reliably distinguish the emotions, as shown in Fig. 3.

3.4.1 DeNosie

DeNoise is the process of removing any noise in the text such as, English words, Special characters, Punctuation marks, Emojis and Diacritics.

3.4.1.1 English Characters Removal As the tweets contain irrelevant English words and character within the Arabic tweets, the first step was to remove any English characters in order to reduce the noise in the data that will affect the classification and detection process, using the English character python library

3.4.1.2 Emojis Removal Emojis are little digital images or icons that are used to convey a thought or feeling. Nevertheless, due to the many irrelevant emojis in the text, in this step all of these emojis were removed from the text using their representing code.

3.4.1.3 Dediacritization Diacritics are symbols similar to vowels in the English language that appear above or below the letters in your Arabic text. These diacritics are removed in order to cut down some serious data sparsity.

3.4.2 Tokenization

Tokenization is the process of text dissection into meaningful chunks, known as tokens. We can separate a section of text into words or phrases. Where it is regarded as a fundamental and essential step in progressing with NLP since

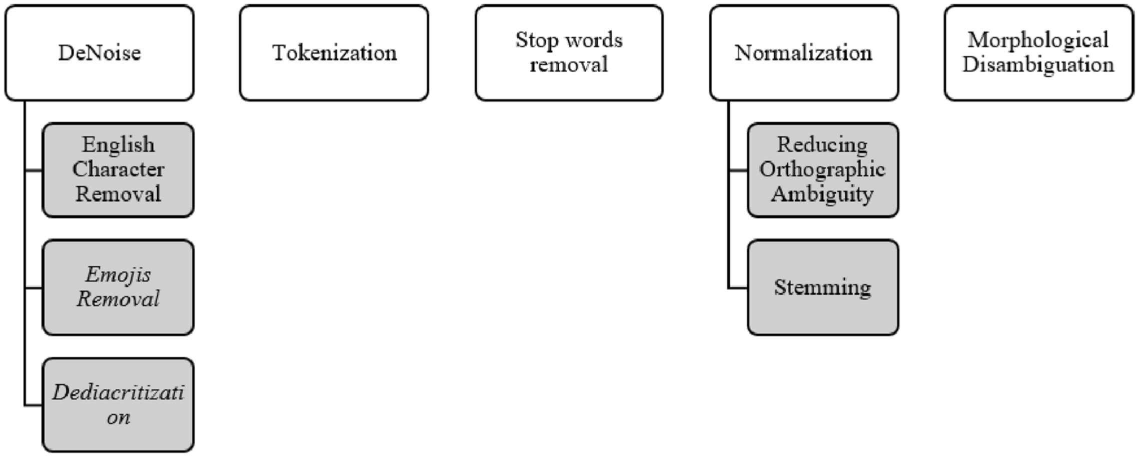

Fig. 3 Preprocessing tasks

it makes it simpler to comprehend the meaning of the text while evaluating the words existing in it.

3.4.3 Stop words removal

It is the process of deleting terms that appear often in all of the texts in the corpus. These words are abundant in every human language. By deleting them, we eliminate any low-level data from our text, allowing for focusing more on the crucial information. In this study, we created our own stop words documents to guarantee that words that might cause ambiguity in comprehending the semantic and true meaning of the tweet are not deleted.

3.4.4 Normalization

It is the process of converting text into a single canonical form that it did not previously have. This reduces unpredictability and brings it closer to its conventional format. This increases efficiency and makes working with text easier. Reducing Orthographic Ambiguity and Stemming are two examples.

3.4.4.1 Reducing Orthographic Ambiguity We minimized the orthographic ambiguity by eliminating certain symbols from specific letters, for example, deleting the Hamza from letters, to accommodate for a number of typical spelling variations among dialects (and because of the difference between spoken and written Arabic in general).

3.4.4.2 Stemming The word modification to transmit numerous grammatical categories such as, voice, tense, case, aspect, number, person, mood and gender. As a result, while a word may have several inflected forms, having multiple inflected forms within the same text adds redundancy to the NLP process. As a result, the stemming process reduces word inflection to their root forms, which aids in the preparation of text, words, and documents. In this study, two stemming approaches, Porter Stemming and Snowball, were examined. Porter Stemming was utilized since it produced precise and proper stemming.

3.4.5 Morphological disambiguation

The process of providing the most probable morphological meaning for a particular word in context. This goal is achieved by grading the output of a morphological analyzer or by employing an end-to-end system that offers a single response. Whereas deleting the diacritics might result in the term having many interpretations depending on the diacritics used.

3.5 Word embedding

The preprocessing findings are then sent into the word embedding process. Word embedding is a method of translating words from an existing dictionary to numeric vectors containing real values. Where according to Altszyler et al. (2016), the LSTM algorithm has an accuracy of 86.76% when utilizing word embedding and an accuracy of 84.14% when without using word embedding. Because machine learning algorithms cannot comprehend textual input quantitatively, we needed semantically accurate embeddings for the text. The pre-trained models for vectorization was utilized since we would typically have text conveying common human experiences or emotions. Two embedding strategies in this study were tested: Word2Vec and GloVe.

Word2Vec is a word embedding approach introduced by Goldberg and Levy (2014) that has benefits in the similarity of word meanings acquired by paying attention to the similarity of words around the target word; it contains two strategies, namely the continuous bag of words and the skip-gram model. The Word2Vec model was examined in CBOW utilizing the Gensim Library in this study.

3.5.1 GloVe embeddings

Global vectors for word representation (GloVe) "is an unsupervised learning approach for obtaining word vector representations. This is accomplished by mapping words into a meaningful space in which the distance between words is proportional to their semantic similarity." Sakketou and Ampazis (2020) GloVe has an advantage over Word2vec in that it generates word vectors using global data (word co-occurrence) rather than just local statistics (local context information of words). When trained on global word-to-word statistics from large text corpora, the resulting vectors display particularly fascinating linear vector space connections. Consequently, GloVe was used in this study because its accuracy outperformed Word2Vec's, with GloVe producing an LSTM F1-score of 85 percent and Word2Vec providing an F1-score of 68%.

GloVe was implemented using pre-trained GloVe. We used (glove.6B.100d) for each word embedding, which has 100 dimensions and comes in a variety of sizes. For tweets, we created a GloVe vector matrix for each word after the data is cleaned. In more details, l,100 dimensional 2-dimensional matrix. The letter "l" signifies the statement's maximum word length. Because the length of each tweet in the data varies, this "l" helps defining the matrix. The matrix is also padded with zero at the end. Finally, the embedding matrix took the form "number of training examples, l100".

3.6 Model development long short-term memory (LSTM)

The final stage was to build an LSTM-based sentence model. Long Short-Term Memory Networks (LSTMs) "are a sort of RNN that can learn long-term dependencies since they are purposefully designed to avoid the long-term reliance problem that happens in any language-based model, and hence LSTMs have always been one of the go-to options." Staudemeyer and Morris (2019) the forward pass equations of an LSTM unit with a forget gate are provided in simple form.

$$f_t = \sigma_g\left(W_f x_t + U_f h_{t-1} + b_f\right) \tag{1}$$

$$i_t = \sigma_g\left(W_i x_t + U_i h_{t-1} + b_i\right) \tag{2}$$

$$o_t = \sigma_g\left(W_o x_t + U_o h_{t-1} + b_o\right) \tag{3}$$

Table 6 Model evaluation results

S. no.	Class label	Precision	Recall	F1-Score
1	Anger	0.9	0.85	0.87
2	Hate	0.94	0.9	0.92
3	Sad	0.87	0.8	0.83
4	Neutral	0.92	0.93	0.92
5	Happy	0.75	0.68	0.71
	Overall	0.88	0.83	0.85

$$\tilde{c}_t = \sigma_c\left(W_c x_t + U_c h_{t-1} + b_c\right) \tag{4}$$

$$c_t = f_t \circ c_{t-1} + i_t \circ \tilde{c}_t \tag{5}$$

$$h_t = o_t \circ \sigma_h\left(c_t\right) \tag{6}$$

LSTM model was trained using 80 training data and 20 for testing. The dropout rate was maintained at 0.5. We got at this figure by repeatedly modifying it, and Adam served as the optimizer. For our network's target layer, we employed the Softmax function. Keras-Tensorflow was used to implement the model. The hypersettings, such as hidden units, were tweaked starting with 256 hidden units but had to reduce it to 128 hidden units after numerous attempts owing to better efficiency. The model was trained in 120-example batches every iteration.

For the Evaluation Measures, we have used the F1-score matrix, precision and recall for each class, to evaluate. We chose the f1 score measure above accuracy since the dataset has an uneven class distribution.

4 Results and discussion

According to the classifications results people's opinions regarding the new blended education regulations and reality where are follows: in average percentage of 18.75% are not satisfied with this situation with 25% Anger and 16% Hate, while 21.25 percent are feeling negatively regard it. However, only 13% are happy with and 24.50% are feeling neutral regards it. Which was compatible with the results of Ali (2021), Alturayeif and Luqman (2021) that took the Arab region in specific in their study.

The classification report of the model on the dataset is as shown in Table 6

Finally, given the distribution of the training data, the model performs well for all classes. However, it appears to be confused between sadness and anger, as the two are somewhat connected. Because happy classes have a far lower frequency than other classes, their recall value is lower. The

Table 7 Embedding models comparison

S. no.	Class label	GloVe	Word2Vec
1	Anger	0.87	0.61
2	Hate	0.92	0.7
3	Sad	0.83	0.72
4	Neutral	0.92	0.78
5	Happy	0.71	0.57

Table 8 Comparison with other studies

Study	Technique	F1-score
Meishar-Tal and Levenberg (2021)	BERT	0.72
Dey and Dasgupta (2022)	DCNN	0.85
Moustakas and Robrade (2022)	Naïve Bayes	0.74
Hashish et al. (2022)	AFINN Lexicon	0.67

data's overall f1-score is rather excellent. The GloVe Model performed better than a Word2Vec model.

A comparison of F1 scores of both the models as illustrated in Table 7.

Table 8, in comparison to other studies, demonstrates the f1-score of several models that took into consideration identical domain and text in the post-covid phase.

5 Recommendations

Given the outcomes, the model performed admirably. However, due to the complexities and peculiarities of the Arabic language, as well as the nature and structure of Arabic that differ from those of other languages, such as English, and its own historical and cultural base. This section will suggest some recommendations for utilizing the model with Arabic text.

5.1 Preprocessing recommendations

To overcome the problem that will occur when removing the diactaries from the Arabic words we recommend to include the Morphological Disambiguation step. In addition, we recommend removing any extra symbols from the words, such as Hamza in order to reduce the problem of the difference between spoken and written Arabic, where the Arabic language is full of these symbols.

Regarding the stemming step, Porter Stemming Techniques turned out to be better than the Snowball Stemming when dealing with the Arabic text as it Snowball Stemming reduced the LSTM classification F1-score results to 77%, where Snowball caused over-stemming problem and leads to unmeaningful stems.

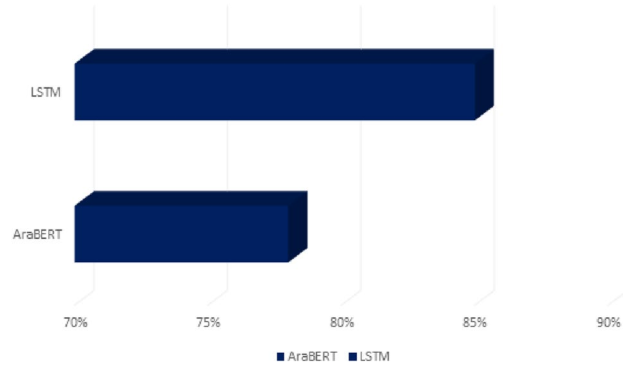

Fig. 4 F1-Score comparison

5.2 Embedding technique recommendation

We favor GloVe over Word2Vec because Word2Vec cannot handle out-of-vocabulary words effectively since it assigns a random vector representation for OOV words, which might be suboptimal, and it depends solely on language word local information. GloVe, on the other hand, generates word vectors using both local statistics (word context information) and worldwide statistics (word co-occurrence).

5.3 Classification technique recommendation

Regarding the classifications techniques, deep learning performed significantly better than other machine learning techniques, particularly when dealing with Arabic Text, most notably the LSTM technique, which performed better in classifying the emotions of Arabic Text than the BERT, with F1-scores of 85% for LSTM and 78% for AraBERT as shown in Fig. 4.

6 Conclusion and future work

In this study, Jordanian's thoughts, opinions and feelings about the new hybrid-learning approach was investigated, through studying the tweets regarding this topic for the period that is known as Post-COVID that is undertaken after the Lock-down is lifted and the normal life is partially back to its normal. Where the study utilized the NLP techniques for detecting people's emotions from text using the GloVe embedding and the LSTM deep learning technique. Our model was able to distinguish among the anger, hate, sad, happy and neutral with fairy good f1-score of 0.85. Where the majority of the studied sample are not accepting (Sad and Dissatisfied) the hybrid education in its current form, as according to Barakat et al. (2022), the Jordan educational system is not ready to take such step either in the country's technological infrastructure nor the educational institutions' nor the students' resources to cope-up with this dramatically and instance unprepared to, changes.

This study has three main limitations: (1) the sample size of the study, which was limited to 4000 tweets owing to the total tweet limit. (2) Due to the difficult nature of the Arabic language, part of the text meaning was lost during preprocessing; (3) a lack of studies that simply take the Arab area, notably Jordan, into account while analyzing people's behavior through traditional face-to-face learning.

We intend to advance our work by including varied datasets and broadening the model's generalizability. Data from other social networking sites, blogs, and news sources will also be considered. Finally, in addition to the five emotions explored in this study, we will concentrate on recognizing the commonly unnoticed emotions such as frustration and impatience.

Author contributions EQ is the first author, where Qaqish, wrote the main manuscript text and prepared all the figures and tables (Introduction, Literature Review, Conducted the whole Methodology, Results and Conclusion) AA contributed in the data and material collecting Dr. WAE contributed in his knowledge and supervision.

Funding The authors did not receive any financial support from any organization for the submitted work.

Availability of data and materials The dataset generated and analyzed during this study are available from the corresponding authors on reasonable request.

Declarations

Conflict of interest The authors declare no competing interests.

References

Aditya DS (2021) Embarking digital learning due to COVID-19: are teachers ready? J Technol Sci Educ 11(1):104–116
Ali MM (2021) Arabic sentiment analysis about online learning to mitigate COVID-19. J Intell Syst 30(1):524–540
Almotiri SD (2022) Twitter sentiment analysis during the lockdown on New Zealand. Int J Comput Inf Eng 15(12):649–654
Althagafi A, Althobaiti G, Alhakami H, Alsubait T (2021) Arabic tweets sentiment analysis about online learning during COVID-19 in Saudi Arabia. The Science and Information Organization. Int J Adv Comput Sci Appl 12(3)
Altszyler E, Sigman M, Ribeiro S, Slezak DF (2016) Comparative study of lsa vs word2vec embeddings in small corpora: a case study in dreams database. arXiv preprint arXiv:1610.01520
Alturayeif N, Luqman H (2021) Fine-grained sentiment analysis of Arabic COVID-19 tweets using BERT-based transformers and dynamically weighted loss function. Appl Sci 11(22):10694
Aristovnik A, Keržič D, Ravšelj D, Tomaževič N, Umek L (2020) Impacts of the COVID-19 pandemic on life of higher education students: a global perspective. Sustainability 12(20):8438
Arnove RF (2020) Imagining what education can be post-COVID-19. Prospects 49(1):43–46
Ashraf I, Hur S, Park Y (2019) Application of deep convolutional neural networks and smartphone sensors for indoor localization. Appl Sci 9(11):2337
Ashraf I, Kang M, Hur S, Park Y (2020) MINLOC: magnetic field patterns-based indoor localization using convolutional neural networks. IEEE Access 8:66213–66227
Ashwitha R, Jeevitha TG, Anwar K, Adnan M (2021) Online learning amid the COVID-19 pandemic: students perspectives. J Emerg Technol Innov Res 2:45
Barakat SS, AlSmadi HM, Abu-Ashour KM (2022) Evaluating the Jordanian experience in the transition to a knowledge economy. J Knowl Econ 13(2):1076–1087
Barron-Estrada ML, Zatarain-Cabada R, Oramas-Bustillos R (2019) Emotion recognition for education using sentiment analysis. Res Comput Sci 148(5):71–80
Bhagat KK, Mishra S, Dixit A, Chang C-Y (2021) Public opinions about online learning during COVID-19: a sentiment analysis approach. Sustainability 13(6):3346
Chintalapudi N, Battineni G, Amenta F (2021) Sentimental analysis of COVID-19 tweets using deep learning models. Infect Dis Rep 13(2):329–339
de Mooij SMM, Dumontheil I, Kirkham NZ, Raijmakers MEJ, van der Maas HLJ (2022) Post-error slowing: large scale study in an online learning environment for practising mathematics and language. Dev Sci 25(2):13174
Dey A, Dasgupta K (2022) Emotion recognition using deep learning in pandemic with real-time email alert. In: Proceedings of third international conference on communication, computing and electronics systems. Lecture notes in electrical engineering. Springer, Singapore, pp. 175–190
Durães D, Toala R, Novais P (2021) Emotion analysis in distance learning. Educating Engineers for Future Industrial Revolutions. Advances in intelligent systems and computing. Springer, Cham, pp 629–639
Gandolfi E, Ferdig RE, Kratcoski A (2021) A new educational normal an intersectionality-led exploration of education, learning technologies, and diversity during COVID-19. Technol Soc 66:101637
Gneezy U, List JA, Livingston JA, Qin X, Sadoff S, Xu Y (2019) Measuring success in education: the role of effort on the test itself. Am Econ Rev Insights 1(3):291–308
Goldberg Y, Levy O (2014) word2vec explained: deriving Mikolov et al.'s negative-sampling word-embedding method. arXiv preprint arXiv:1402.3722
Hashish A, Aly E, Baatiah NY, Bashaweeh AH, Kattan AM (2022) The online learning experience and reported headaches associated with screen exposure time among Saudi health sciences students during the COVID-19 pandemic. BMC Med Educ 22(1):1–13
Imran AS, Daudpota SM, Kastrati Z, Batra R (2020) Cross-cultural polarity and emotion detection using sentiment analysis and deep learning on COVID-19 related tweets. IEEE Access 8:181074–181090
Javed N, Muralidhara BL (2022) Emotions during COVID-19: LSTM models for emotion detection in tweets. In: Proceedings of the 2nd international conference on recent trends in machine learning, IoT, Smart Cities and Applications. Lecture notes in networks and systems. Springer, Singapore, pp 133–148
Lockee BB (2021) Online education in the post-COVID era. Nat Electron 4(1):5–6
Mahfoodh H, AlAtawi H (2020) Sustaining higher education through elearning in post COVID-19. In: 2020 sixth international conference on e-learning (econf). IEEE, pp 361–365
Martinez AR (2010) Natural language processing. Wiley Interdiscip Rev Comput Stat 2(3):352–357
Meishar-Tal H, Levenberg A (2021) In times of trouble: higher education lecturers' emotional reaction to online instruction during COVID-19 outbreak. Educ Inf Technol 26(6):7145–7161

Moustakas L, Robrade D (2022) The challenges and realities of e-learning during COVID-19: the case of university sport and physical education. Challenges 13(1):9

Mujahid M, Lee E, Rustam F, Washington PB, Ullah S, Reshi AA, Ashraf I (2021) Sentiment analysis and topic modeling on tweets about online education during COVID-19. Appl Sci 11(18):8438

Nacheva R (2022) Emotions mining research framework: higher education in the pandemic context. In: Contributions to Economics. Springer, Cham, pp 299–310

Neuwirth LS, Jović S, Mukherji BR (2021) Reimagining higher education during and post-COVID-19: challenges and opportunities. J Adult Contin Educ 27(2):141–156

Posner J, Russell JA, Peterson BS (2005) The circumplex model of affect: an integrative approach to affective neuroscience, cognitive development, and psychopathology. Dev Psychopathol 17(3):715–734

Sakketou F, Ampazis N (2020) A constrained optimization algorithm for learning glove embeddings with semantic lexicons. Knowl Based Syst 195:105628

Sakr A, Amin M, Grwany T (2021) Performance evaluation in Arabic sentiment analysis during the COVID-19 pandemic. Egypt J Lang Eng 8(2):16–27

Samuel J, Ali GGMN, Rahman MM, Esawi E, Samuel Y (2020) COVID-19 public sentiment insights and machine learning for tweets classification. Information 11(6):314

Shelke N, Chaudhury S, Chakrabarti S, Bangare SL, Yogapriya G, Pandey P (2022) An efficient way of text-based emotion analysis from social media using LRA-DNN. Neurosci Inform 2(3):100048

Shukla P, Garg A (2021) Sentiment analysis of online learners in higher education: a learning perspective through unstructured data. In: Intelligent system algorithms and applications in science and technology. Apple Academic Press, Boca Raton, pp. 157–170

Staudemeyer RC, Morris ER (2019) Understanding LSTM—a tutorial into long short-term memory recurrent neural networks. arXiv preprint arXiv:1909.09586

Tadesse S, Muluye W (2020) The impact of COVID-19 pandemic on education system in developing countries: a review. Open J Soc Sci 08(10):159–170

Tadesse S, Muluye W (2020) The impact of COVID-19 pandemic on education system in developing countries: a review. Open J Soc Sci 8(10):159–170

Toçoğlu MA, Onan A (2021) Sentiment analysis on students' evaluation of higher educational institutions. Advances in Intelligent Systems and Computing. Springer, Cham, pp 1693–1700

Umair M, Hakim A, Hussain A, Naseem S (2021) Sentiment analysis of students' feedback before and after COVID-19 pandemic. Int J Emerg Technol 12:177–182

Veletsianos G, Kimmons R, Larsen R, Dousay TA, Lowenthal PR (2018) Public comment sentiment on educational videos: understanding the effects of presenter gender, video format, threading, and moderation on YouTube TED talk comments. PLoS ONE 13(6):0197331

Wang Z, Cruz IF (2020) Analysis of the impact of COVID-19 on education based on geotagged twitter. In: Proceedings of the 1st ACM SIGSPATIAL international workshop on modeling and understanding the spread of COVID-19. ACM, New York

Yayla R, Yayla HN, Ortaç G, Bilgin TT (2021) A classification approach with machine learning methods for technical problems of distance education: Turkey example. Open Praxis 13(3):312–322

Zhao Y, Watterston J (2021) The changes we need: education post COVID-19. J Educ Change 22(1):3–12