Journal of Information Technology Education: Innovations in Practice

An Official Publication of the Informing Science Institute InformingScience.org

JITEiip.org

Volume22, 2023

UNVEILING LEARNER EMOTIONS: SENTIMENT ANALYSIS OF MOODLE-BASED ONLINE ASSESSMENTS USING MACHINE LEARNING

Delali Kwasi Dake*	Department of ICT Education, University of Education, Winneba, Ghana	dkdake@uew.edu.gh
Godwin Kudjo Bada	Department of ICT Education, University of Education, Winneba, Ghana	gkbada@uew.edu.gh

* Corresponding author

ABSTRACT

Aim/Purpose	The study focused on learner sentiments and experiences after using the Moodle assessment module and trained a machine learning classifier for future sentiment predictions.
Background	Learner assessment is one of the standard methods instructors use to measure students' performance and ascertain successful teaching objectives. In pedagogical design, assessment planning is vital in lesson content planning to the extent that curriculum designers and instructors primarily think like assessors. Assessment aids students in redefining their understanding of a subject and serves as the basis for more profound research in that particular subject. Positive results from an evaluation also motivate learners and provide employment directions to the students. Assessment results guide not just the students but also the instructor.
Methodology	A modified methodology was used for carrying out the study. The revised methodology is divided into two major parts: the text-processing phase and the classification model phase. The text-processing phase consists of stages including cleaning, tokenization, and stop words removal, while the classification model phase consists of dataset training using a sentiment analyser, a polarity classification model and a prediction validation model. The text-processing phase of the referenced methodology did not utilise tokenization and stop words. In addition, the classification model did not include a sentiment analyser.

Accepting Editor Elsje C Scott │ Received: March 23, 2023 │ Revised: May 31, June 16, 2023 │ Accepted: July 15 19, 2023.

Cite as: Dake, D. K., & Bada, G. K.. (2023). Unveiling learner emotions: Sentiment analysis of Moodle-based online assessments using machine learning. *Journal of Information Technology Education: Innovations in Practice, 22,* 109-132. https://doi.org/10.28945/5174

Contribution	The reviewed literature reveals two major omissions: sentiment responses on using the Moodle for online assessment, particularly in developing countries with unstable internet connectivity, have not been investigated, and variations of the k-fold cross-validation technique in detecting overfitting and developing a reliable classifier have been largely neglected. In this study we built a Sentiment Analyser for Learner Emotion Management using the Moodle for assessment with data collected from a Ghanaian tertiary institution and developed a classification model for future sentiment predictions by evaluating the 10-fold and the 5-fold techniques on prediction accuracy.
Findings	After training and testing, the RF algorithm emerged as the best classifier using the 5-fold cross-validation technique with an accuracy of 64.9%.
Recommendations for Practitioners	Instead of a closed-ended questionnaire for learner feedback assessment, the open-ended mechanism should be utilised since learners can freely express their emotions devoid of restrictions.
Recommendations for Researchers	Feature selection for sentiment analysis does not always improve the overall accuracy for the classification model. The traditional machine learning algorithms should always be compared to either the ensemble or the deep learning algorithms
Impact on Society	Understanding learners' emotions without restriction is important in the educational process. The pedagogical implementation of lessons and assessment should focus on machine learning integration
Future Research	To compare ensemble and deep learning algorithms
Keywords	sentiment analysis, machine learning, random forest algorithm, online learning, Moodle

INTRODUCTION

The COVID-19 pandemic experience has changed teaching and learning at higher educational institutions with most educational policies now focused on faster transitions to the online space (Bhasin et al., 2021; Pokhrel & Chhetri, 2021). The pandemic forced the urgent closure of schools, causing enormous disruptions to the academic calendar with severe consequences (Pokhrel & Chhetri, 2021). According to United Nations 2020 report, an estimated 1.725 billion learners from pre-primary to tertiary, representing 98.6% of students in 200 countries worldwide, were severely affected by the outbreak. These enormous difficulties in the educational landscape during the pandemic re-echoed the relevance of online and distance learning (Moore et al., 2011) to policymakers, governmental agencies, academic authorities and parents.

Even before COVID-19, Ghanaian tertiary students continually complained about unreliable internet connectively, which hampered teaching and learning (Adarkwah, 2021; Upoalkpajor & Upoalkpajor, 2020). Most disruptive technologies that could significantly alter the face of education in Ghana were significantly deficient, posing a challenge for governments and academic authorities even before the pandemic (Adarkwah, 2021). Even though, governments embarked on ICT based-support projects across various academic institutions, the impact was considerably limited by factors such as capacity, bandwidth, internet speed, coverage, specification, and the economy (Adarkwah, 2021; Buabeng-Andoh, 2012; Budu et al., 2018). The outbreak of COVID-19 exacerbated teaching and learning in Ghana and plunged academic institutions into a state of urgency (Dake et al., 2021; Henaku, 2020). On March 15, 2020 the government of Ghana announced the closure of schools as one priority measure to stop the spread of the virus (Cromwell, 2020). The lack of urgency in investing extensively in ICTs to support Ghanaian educational institutions has resorted in setbacks and regrets. The pandemic revealed Ghana's ICT infrastructure gap and disrupted teaching and learning across the country (Adarkwah, 2021; Dake et al., 2021).

Although technically distinct, both online and distance learning entail teaching and learning via the Internet (Firat & Bozkurt, 2020). With the urgent deployment of fifth-generation mobile technology (5G) (Al-Falahy & Alani, 2017), the technological revolutions in online teaching and learning are limitless. In addition to the faster seamless engagement of learners online, 5G emergence will result in boosted on-the-go learning, intelligent systems, tactile stimulation, immersive learning, educational internet of things (IoT) and big data (Lee & Kim, 2020). Even though the COVID-19 pandemic adversely affected the global economy (Ozili & Arun, 2020), it has increased research on 5G possibilities for online teaching and learning (Lee & Han, 2021). As the online learning space continues to improve in diverse application integration, learning management system (LMS) development and utilisation have become apparent.

LMS is a web-based technology that provides the framework for managing all aspects of the learning process (Simanullang & Rajagukguk, 2020). There are proprietary and open source LMS that vary in suitability, functionality, ease of use, cost and access to the source codes (Simanullang & Rajagukguk, 2020). The open source LMS includes Sakai, Moodle, Dokeos, and Claroline, while the blackboard, LMC, and HCM form proprietary LMS (Simanullang & Rajagukguk, 2020). The primary functionality of LMS spans from course management, assessment administration, group discussions, and multimedia to advanced features such as analytics, gamification, real-time in-course feedback, and virtual proctoring (Ohliati & Abbas, 2019). Specifically for assessment administration, the quizzes, assignments, and tests have learner personalisation with responsive features for real-time assessment feedback and multimedia integration (Ohliati & Abbas, 2019).

Learner assessment is one of the standard methods instructors use to measure students' performance and ascertain successful teaching objectives. In pedagogical design, assessment planning is vital in lesson content planning to the extent that curriculum designers and instructors primarily think like an assessor (Jiang et al., 2019). Assessment aids students in redefining their understanding of a subject and serves as the basis for more profound research in that particular subject. Positive results from an evaluation also motivate learners and provide employment directions to the students. Assessment results guide not just the students but also the instructor. The instructor reflects on teaching methods and learning outcomes based on assessment results. Assessment results also become a core reason for the instructor to implement personalised learning and constructivism theories (Kara, 2018; Prain et al., 2013). While personalised learning is an individualised, student-centred approach to education, the constructivism theory encourages learner engagement and the construction of new knowledge (Keppell, 2014).

PROBLEM DEFINITION

The availability of cheap, reliable internet speed is still an issue in Ghana (Adarkwah, 2021; Nyarko-Boateng et al., 2020). In developed countries, 5G support for industry and education has boosted economic gains and improved students' online learning experiences (Al-Maroof et al., 2021; Hutajulu et al., 2020). Since COVID-19 ravaged the world, most higher educational institutions in developed countries have fully utilised the LMS to manage the learning process successfully (Garad et al., 2021; Palvia et al., 2018). One application domain of LMS in focus for the study is learner assessment and management. In Ghana, student assessment through LMS for most courses has gained usage since COVID-19 (Agormedah et al., 2020; Dampson, 2021). There will be severe repercussions for students' grades if technological deficiencies in learner assessment in Ghana remain unnoticed. The instructor is guided by learner performance through assessment to reflect on learning philosophies and policy even as the students derive motivation from good grades. Such deprivation of proper assessment scores using the LMS can lead to learner attrition, unsavoury behaviour, suicide, agitation and ultimately affect the university's ranking and image (Adarkwah, 2021).

PURPOSE OF STUDY

Learner emotion after an LMS-based assessment is the first point in cognisance of the challenges faced when taking the test. Instead of limiting the learner to a closed-ended questionnaire approach, the open-ended method remains pivotal in understanding the learner fully. The open sentiment expressed by the learner is the feedback that will necessitate actions from instructors and management in addressing the root issues in LMS-based assessments. In line with the objectives of the study, we pose the following research questions

> **RQ1**: What are the categorisations of feedback comments received after an LMS-based assessment?
> **RQ2**: What machine learning algorithm will provide the best model using sentiment analysis on the received comments?
> **RQ3**: What is the test predictive performance of the best machine learning algorithm for futuristic sentiments?

SENTIMENT ANALYSIS AND MACHINE LEARNING

Sentiment analysis is a machine learning tool mainly for unstructured text analysis with a level of polarity in feedback automation (Jain et al., 2021). The ratio of 20% structured to 80% unstructured text currently exist in the world's data as estimated statistically by Altexsoft (2020), and Ot (2023). Unlike structured data, which is well-organised and appropriately formatted, unstructured data is free-form data with no pre-defined format. Unstructured data usually comes in diverse structures and complex forms, making it difficult to analyse using traditional tools (Adnan & Akbar, 2019). Data from social media, email, web pages, word documents, audio and videos are primarily unstructured. In contrast, structured data comes from databases, server logs, network logs and closed-ended online forms. Sentiment analysis or opinion mining is a method of using natural language processing to determine the emotional undertone of a text (Birjali et al., 2021). In text polarity detection, machine learning algorithms are trained on text data to learn and discern sentiments without human input. Sentiment analysis has seen deployments in diverse sectors, including social media monitoring to analyse emotions behind expressions and customer support management to understand phrases, especially those that contain negative reviews. Other common applications of sentiment analysis include product analysis and rating, competitor analysis via market survey, and reputation management of brands. In education, sentiment analysis has been applied in domains, such as instructor evaluation in open-ended questions, as an evaluation technique for ranking universities based on social media comments, and for adaptive learning in intelligent information systems (Dolianiti et al., 2019). Some other common application areas of sentiment analysis in education include course curriculum redesign, assessment management and for emotion detection during an online class (Barron-estrada et al., 2019).

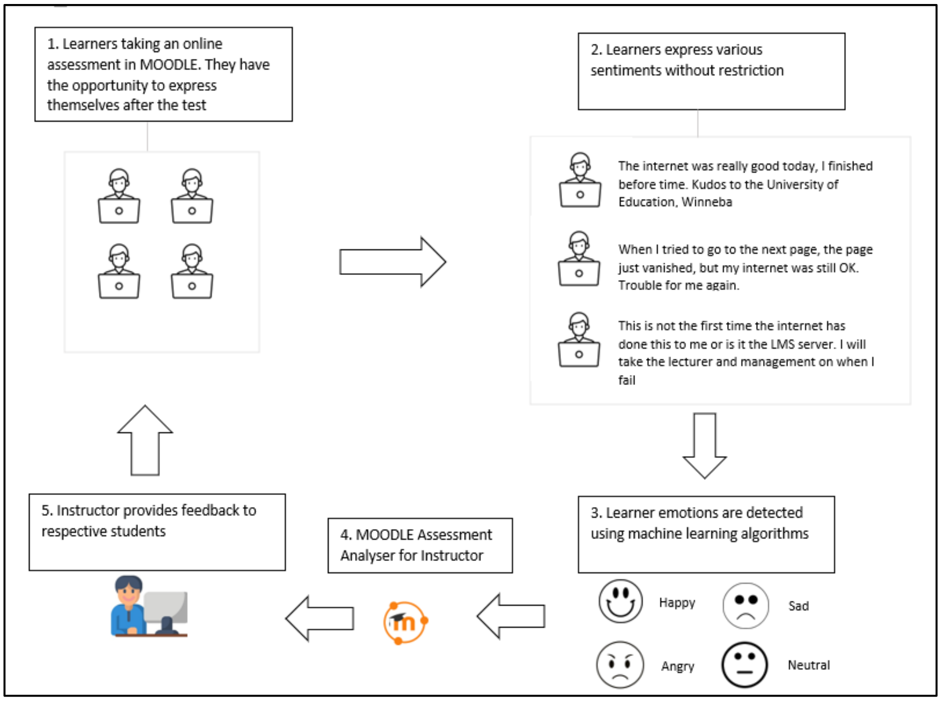

Figure 1: Sentiment Analysis in Learner Emotion Management after an Online Assessment

As depicted in Figure 1, in the first instance, learners take on online assessments using the Modular Object-Oriented Dynamic Learning Environment (Moodle). While taking the examination, learners experience varying emotions based on their experiences with the Moodle platform. In the second instance, students can express their feelings after the exam through an open-ended Moodle application. Restricting students through close-ended questions has limitations because learners' emotional options might not be available. In the third instance, machine learning algorithms are deployed to identify the polarity of students' range of emotions. The sentiments expressed by the student are classified, and a machine learning model is built and integrated into the Moodle analyser in instance four to predict future sentiments. The instructor, in instance five, receives analytical patterns from the Moodle analyser and addresses concerns based on priority and severity.

Review of Literature

The National Research Council (NRC) Emotion Lexicon and Machine Learning implementations are the two fundamental methods used in sentiment analysis. The NRC Lexicon has eight primary emotions, including disgust, joy, sadness, surprise, trust, fear, anger, and anticipation, with automatic translation to over 40 languages (Mohammad & Turney, 2013). The machine learning method uses machine learning algorithms to train the polarity of opinions and implement the best classifier for future sentiment prediction (Ahmad et al., 2017). The literature review is divided into sections. The first section discusses machine learning algorithms for sentiment analysis. At the end of this section, there is a summary table listing the major shortfalls and prevalent algorithms in each study. The second section of the review discusses Moodle for online assessment. This section analyses the types of assessments on Moodle and the necessary suggestions for optimising the Moodle software. The final section examines textual emotions and discusses the types of emotions that can be expressed in a text.

Sentiment Analysis Using Machine Learning Algorithms

Dashtipour et al. (2016) used the Support Vector Machine, Maximum Entropy, and multimodal Naive Bayes classifiers as machine learning algorithms in multilingual sentiment analysis. For linearly separable data, SVM gives classification results with minimal error. It was realised that the multi-modal Naive Bayes classifier is very simple for efficient classification with incremental learning. The Maximum Entropy (MaxEnt) classifier efficiently extracts information that leads to good results. It showed 83% accuracy, which is better than other classifiers used in their study, namely SVM and multinomial Naive Bayes.

Singh et al. (2016) evaluated two Machine Learning based classifiers (Naive Bayes and SVM), the Unsupervised Semantic Orientation approach (SO-PMI-IR algorithm) and the SentiWordNet (SWN) approaches for sentiment classification of movie reviews. Regarding computed scores on datasets, the total percentage of 'positive' and 'negative' labels assigned by all four methods include the following: for dataset1, dataset2 and dataset3, Naïve Bayes scored 49.65%, 48%, and 76.9% positives, respectively. For negative values, it scored 50.35%, 52% and 23.1% for the three datasets. SVM had 50.6%, 50.92% and 78.7% as positives and 49.4%, 49.07% and 21.3% negatives for the three datasets. SO-PMI-IR recorded 52.37%, 50.78% and 59.1% positives against 47.63%, 49.22% and 40.9% negatives for the three datasets, respectively. SWN approach recorded 65.15%, 64.85% and 71.36% positives as against 34.85%, 35.07% and 28.64% negatives for the datasets used in the study. From the results, they concluded that the classification accuracy by NB is marginally better than the SVM and is close to the SO-PMI-IR algorithm. For the third dataset, the SVM performance levels are identical to the NB and SO-PMI-IR. SentiWordNet, on the other hand, achieves a lower accuracy score. The performance of NB can be comparable to the popularly believed superior performance of SVM, at least for sentiment classification. SVM has the highest accuracy of 98% for narrow domain twitter short texts. The SO-PMI-IR algorithm has impressive accuracy levels and seems the best choice due to its unsupervised nature. The SentiWordNet is computationally the most favourable algorithm, but the accuracy level is, in relative terms, lower.

To determine which machine learning approach performs better than the rule and lexicon-based sentiment analysis in the software engineering domain, Shen et al. (2019) compared the performance of Logistic Regression, Support Vector Machine, and Naive Bayes classifier. First, the sentiment analysis task was divided into binary classification and leverage tests. The subsequent phase was based on descriptive, positive, and negative comments. Two layers were defined for the first analysis. Layer one was to distinguish the emotional comments and the descriptive comments. Layer two performed a binary classification algorithm and identified the comments with positive and negative sentiments. After comparing and implementing the classification algorithms on more than two labels, the previous sentiment analysis algorithms were added to test if the layer-based sentiment classification technique could outperform the single algorithm model. They found out that Logistic Regression works best when performing positive and negative sentiment classification, as it achieves an accuracy of 90%. The precision of Naive Bayes was 81.9%. Support Vector Machine, however, only achieved 67.1% accuracy in identifying the sentimental polarity of comments. It was also realised that all the algorithms tend to classify negative comments as positive than predict positive items as negative.

Osmanoğlu et al. (2020) compared the Decision Tree, Multilayer Perceptron, Support Vector Machine, XGB, K – Neighbors, GaussianNB and Multinomial Logistic Regression classifiers on 6059 sentiments received from distance education students at the Anadolu University. The study aimed to understand learners' ratings regarding the course materials supplied to them during online learning. Instead of the k-fold cross-validation technique, 70% to 30% of training and test data samples were used in building the classifier. The Logistic Regression algorithm comparatively has the highest accuracy of 77.5%.

At the early stages of the COVID-19 pandemic, Mujahid et al. (2021) analysed tweets on online education using sentiment analysis. The 17,155 tweets about e-learning were trained using deep learning and traditional Decision Tree (DT), Random Forest (RF), Logistic Regression (LR), SGB classifier, KNN and Support Vector Machine (SVM) learning algorithms. The Bow (Bag of Words) and the TF-IDF (Term Frequency-Inverse Document Frequency) feature selection mechanisms were implemented in building the model effectively. The Synthetic Minority Oversampling Technique (SMOTE) filter was also utilised to avoid overfitting and balance the dataset. The traditional machine learning algorithms outperformed the deep learning model, with RF having an accuracy of 95% using Bow with SMOTE. The DT and SVM also performed with an accuracy of 95% using the TF-IDF with SMOTE.

Altrabsheh et al. (2014) used sentiment analysis to understand learners' satisfaction during lectures. The NB, Complement NB (CNB), Maximum Entropy (ME), and SVM were the selected classifiers due to their popularity and usage in sentiment analysis. In addition, the N-gram feature selection method was implemented to capture the context of words. In their study, the SVM was the best classifier, with an accuracy of 95%.

Onan (2021) compared deep learning, ensemble, and traditional machine learning algorithms to build a sentiment analysis model for massive open online courses (MOOCs) feedback evaluations. Different feature selection and text weighting schemes, including N-gram models and TF-IDF schemes, were utilised throughout the classification process. The Random Subspace (RS) classifier with RF performed as the second-best classifier with an accuracy of 87.53%, while the RS with NB had the highest accuracy of 89.62%.

The literature summarised in Table 1 indicates that sentiment responses on using the Moodle for assessment, especially in developing countries with unstable internet connectivity, has not been studied. In addition, aside from Mujahid et al. (2021), who comprehensively utilised the 10-fold cross-validation technique, other studies significantly ignored the usage of the k-fold cross-validation technique, which helps avoid overfitting.

Table 1: Summary of Literature Review

STUDY	NATURAL LANGUAGE PROCESSING	DATASET CATEGORY	ALGORITHMS	METRICS/BEST CLASSIFIER	MAIN SHORTFALL
Dashtipour et al. (2016)	Yes	French Movie Reviews	SVM, multimodal NB, Maximum Entropy	Accuracy, Maximum Entropy = 83%	• No cross-validation technique
Singh et al. (2016)	Yes	Hindi, Cornell, Arab Spring, Twitter Feed Movie Reviews	NB, SVM, SWM, SO-PMI-IR	Accuracy, SVM = 98% for twitter feed	• No cross-validation technique
Shen et al. (2019)	Yes	Stack Overflow Comments from Software Engineering Domain	NB, SVM, LR	Accuracy, LR = 90%	• No cross-validation technique
Osmanoğlu et al. (2020)	Yes	Distance Education Domain	DT, MLP, SVM, XGB, K –NN, GaussianNB, Multinomial LR	Accuracy, Multinomial LR = 77.5%	• 70% to 30% training and test data splitting instead of 10-fold or 5-fold cross-validation

STUDY	NATURAL LANGUAGE PROCESSING	DATASET CATEGORY	ALGO-RITHMS	MET-RICS/BEST CLASSIFIER	MAIN SHORTFALL
Mujahid et al. (2021)	Yes	Online Education Tweets during COVID-19	DT, RF, SVM, SGB, LR, KNN, Adaboost, ETC, GNB	Accuracy, RF = 95% using BoW with SMOTE and DT, SVM = 95% using TF-IDF with SMOTE	• Did not utilise the 5-fold cross-validation technique
Onan (2021)	Yes	Massive open online courses (MOOCs) Online Reviews	KNN, SVM, NB, RF, Adaboost, Bagging, RS, Voting, Stagging	Accuracy = RS with NB = 89.2%	• No cross-validation technique
Altrabsheh et al. (2014)	Yes	Higher Educational Institution	NB, SVM, CNB, ME	Accuracy, SVM = 95%	• No cross-validation technique

Moodle For Online Assessment

Assessing the performance of learners in educational institutions has become primary since it measures the skills and knowledge of students in a particular course (Leber et al., 2018). In addition, assessments determine the extent to which the intended learning outcomes are met by students. There are two main types of assessment generally implemented in educational institutions: the summative and formative. Formative assessment is used during educational procedures to measure student progress and provide instructors with in-process feedback on students' understanding of the course (van Groen & Eggen, 2020). In contrast, summative assessments are used to evaluate students' mastery and understanding after a unit or course has concluded (van Groen & Eggen, 2020). The assessment activity module in Moodle has features that support both formative and summative assessment types. The quiz feature in Moodle enables instructors to establish a variety of assessment types, such as matching, multiple-choice, and calculated. Normally, the instructor incorporates the correct answers in the quizzes' configuration in order to facilitate automated grading. The group and peer assessment in Moodle is a vital feature that enables learners to submit group projects. Additionally, the contributions of each peer in the submitted group assignments can be identified and graded separately (Moodle, 2005).

Limited basic research studies have been conducted to determine the satisfaction of the Moodle assessment module among instructors and students using a closed-ended questionnaire. Mwangi et al. (2023) undertook a study to ascertain Moodle's satisfaction during online assessment at selected public institutions in Kenya. The findings reveal that, even though 85% were satisfied with the confidentiality of their marks, 66% complained about internet-related challenges. Owusu-Oware and Tanye (2023) conducted a study in Ghana during COVID-19 to determine the efficacy of using Moodle for online assessment. Their findings reveal that the major challenges to using the Moodle assessment module are plagiarism, cheating, copying, and internet outages. Huynh-Cam et al. (2021) examined the effectiveness of using the Moodle for English listening and reading courses during COVID-19. The results of their study show that 98.7% of the participants had a positive attitude towards the features in Moodle for online assessment. However, the majority of the learners express bitter experiences about their internet connectivity. Snejana and Veselina (2021) evaluated the effectiveness of the Moodle for online assessment at Trakia University during COVID-19. Their findings also show learners frustration as a result of poor internet, which makes it difficult to open the quiz questions promptly. In addition, learners also complained about the short duration of quizzes leading to poor results.

TEXTUAL EMOTIONS

Emotions detection from text is one of the difficult NLP challenges due to the unavailability of labelled datasets (Zad et al., 2021). Specifically in sentiment analysis, it is even more difficult to determine appropriately the polarity of a text (Guo, 2022; Sowmiya et al., 2022). However, textual emotion detection has become crucial since it is the primary medium of human-computer interaction with a wide range of applications. The unstructured text from chat rooms, e-mails, forums, web logs, social media, and product reviews has become common but difficult to analyse (Guo, 2022). The emotional tone of a text varies from simple polarities such as negative, neutral, and positive to complex expressions such as surprise, miracle, tiredness, remorse, optimism, submission, hate, fear, unpleasantness, and comfort (Nandwani & Verma, 2021). There are two standard emotional models: the categorical and the dimensional. Categorical emotional theories define emotions discretely, such as fear, sadness, anger, and happiness. The dimensional model employs three parameters: power, valence, and arousal, to represent emotions. The power represents restriction over emotion; arousal shows levels of feelings in the emotion, while valence depicts the polarity (Bakker et al., 2014; Nandwani & Verma, 2021). The Ekman model, a type of categorical model, is widely adopted in literature. It has six states, including surprise, sadness, joy, fear, disgust, and anger (Ekman, 1992). For the dimensional model, the Plutchik Wheel of emotions comprising disgust, trust, terror, awe, anticipation, remorse, aggressiveness, optimism, love, loathing, grief, distraction, disapproval, surprise, contempt, apprehension, serenity, submission, annoyance, boredom, vigilance, anger, amazement, admiration, rage, interest, fear, sadness, acceptance, ecstasy, pensiveness, and joy is frequently used (Plutchik, 1982).

METHODOLOGY

The research modified the Garg and Lal (2018) sentiment mining architecture by changing the application usage and polarity options. As shown in Figure 2, the revised methodology is divided into two major parts, the text-processing phase and the classification model phase. The text-processing phase consists of stages including cleaning, tokenization, and stop words removal, while the classification model phase consists of dataset training using a sentiment analyser, polarity classification model and a prediction validation model.

STUDENT ASSESSMENT DATA COLLECTION MODULE

Google form was used to collect 300 sentiments from year one to year four students at the south campus of the University of Education, Winneba, after the mid-semester exams in August 2022. From 300 instances, 20 was reserved for prediction, while 255 for training and testing. The study employed the non-probability convenient sampling method to collect learner data. Respondents' ready availability and proximity to the university's south campus prompted the use of the convenient sampling strategy. The respondents were mandated to agree to an ethics consent form before filling out the open-ended questionnaire. Throughout the data collection process, respondents' data confidentiality and privacy were strictly observed. From the questionnaire administered, the students' responses provided could not be traced back to them. Learners freely expressed their emotions without restriction on their personal experiences with the Moodle LMS assessment management module after the test sessions. Sample learner responses are shown in Table 2.

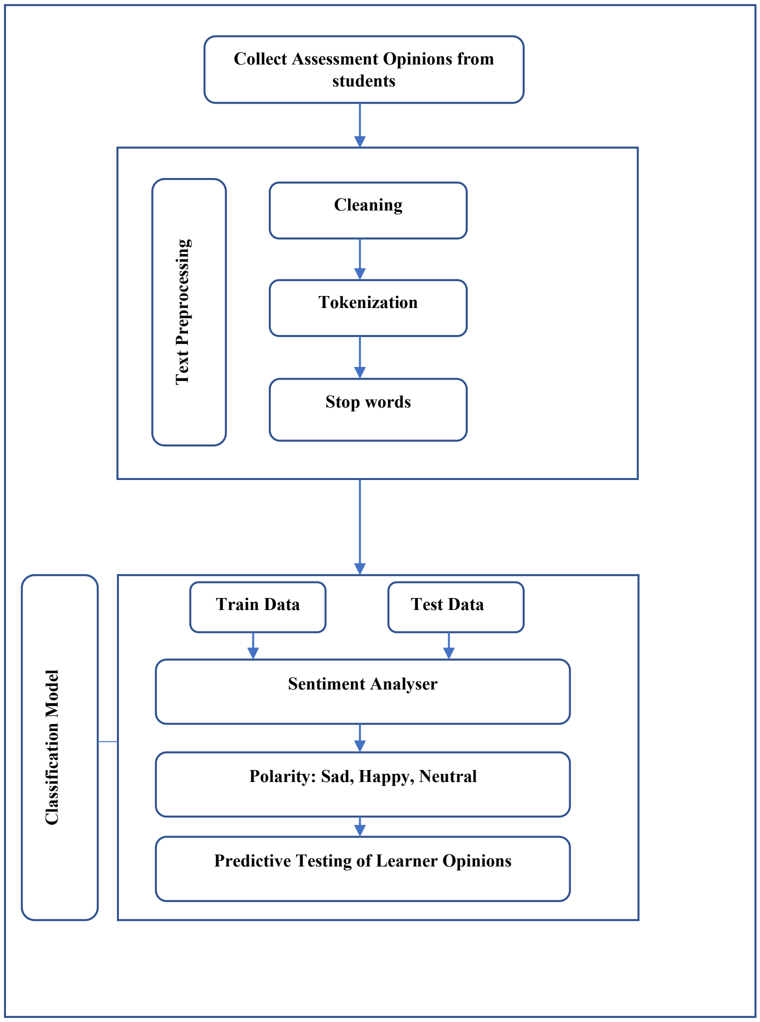

Figure 2: Sentiment Analysis Architecture

Table 2: Sample responses from learners

Respondents' Experiences with the Moodle Assessment Module – Sample data
It's cool but at times the network jam disturbs which leads to low score
It is very good and better environment to use for learning and doing quizzes but due to negative connectivity sometimes worry us students not to partake in some quizzes
Vclass in reality has made quizzes and exam simple and easier for me personally. Meanwhile there are few challenges. Some of them are; network instability leading to poor performance sometimes, insufficient time allocation, Vclass system working slow in some instances. Despite the few challenges I'm well convinced that, if good measures are put in place to enhance the use of Vclass, it will be one of the best modes for assessment.

Respondents' Experiences with the Moodle Assessment Module – Sample data
Student, sometimes cheat during online mode of assessment, however most disadvantage is network failure
Good but am not impressed. Because of poor network system in our country, I entreat that time scheduling for quizzes and exams should be taken into consideration.
It is very slow and sometimes it doesn't work at all, it can also submit it self when doing a quiz or examination

Data cleaning

During the data cleansing phase, unidentified and unrelated sentiments were eliminated from the dataset manually before classification. Out of the 300 responses, 25 were unrelated and contained unidentified characters with no relevance to the classification labels. In constructing the sentiment analysis model, 255 valid responses out of 275 were utilised using the k-fold cross-validation technique for training and testing purposes in Waikato Environment for Knowledge Analysis (WEKA). The remaining 20 instances are for predicting the accuracy of the classification model.

Research Question 1: What are the categorisations of feedback comments received after an LMS-based assessment?

In response to Research Question 1, the level of polarity in the training dataset is used to categorise learner comments. The level of polarity is determined by responses to the research question, "Tell me about your encounters/experiences with online assessment using the VClass (quizzes and exams)". A positive emotion that portrays the Moodle usage as effective towards assessment management at the University of Education, Winneba categorises the learner as a "happy" student. A negative emotion expressed that suggests difficulty in utilising the system categorises the learner as a "sad" student. In contrast, a "neutral" student expresses both negative and positive compliments after using the system. As shown in Table 3, the dataset respondents' sentiments are labelled into happy, neural and sad classes for modelling.

Table 3: Sample labelled categories of learner comments

Comment	Happy	Neural	Sad
1) It's cool but at times the network jam disturbs which leads to low score		✓	
2) It is very good and better environment to use for learning and doing quizzes but due to negative connectivity sometimes worry us students not to partake in some quizzes		✓	
3) Vclass in reality has made quizzes and exam simple and easier for me personally. Meanwhile there are few challenges. Some of them are; network instability leading to poor performance sometimes, insufficient time allocation, Vclass system working slow in some instances. Despite the few challenges I'm well convinced that, if good measures are put in place to enhance the use of Vclass, it will be one of the best modes for assessment.		✓	

Comment	Happy	Neural	Sad
4) Student, sometimes cheat during online mode of assessment, however most disadvantage is network failure			✓
5) Good but am not impressed. Because of poor network system in our country, I entreat that time scheduling for quizzes and exams should be taken into consideration.		✓	
6) It is very slow and sometimes it doesn't work at all, it can also submit it self when doing a quiz or examination			✓

Tokenization

Tokenization is the process of converting a string of textual data into small chunks called tokens (Wongkar & Angdresey, 2019) using string to word vector convertor. For instance, the sentence "the internet connectivity is good" will become 'the,' 'internet,' 'connectivity,' 'is,' 'good,' after tokenization. Tokenization is the foundation for natural language processing which helps the machine learning algorithm to analyse the sequence of words. Stemming is not implemented to allow for a complete and unedited depiction of learner responses.

Stop words

Stop words are commonly used words that add little or no meaning to a sentence during natural language processing (NLP) (Nandwani & Verma, 2021). Examples of stop words include, "the," "is," "an," "at,". Stop words are removed from the dataset to improve classification accuracy, focus on more relevant words, and speed-up computational time. The stop words filtering does not always improve accuracy.

Classification Module

The classification module has the sentiment analyser, which has four classification algorithms, including the Random Forest (RF), Support Vector Machine (SVM), Decision Tree (DT) and Naïve Bayes (NB), as depicted in Figure 3. The sentiment analyser uses the k-fold cross-validation technique on the dataset to ensure the iterative running of training and test data and to avoid overfitting. Comparatively, the study utilises the 5-fold and the 10-fold cross-validation techniques on the machine learning algorithms to ascertain the best classifier for future polarity prediction.

Classification algorithms

In polarity prediction, the DT, NB, SVM and RF classification algorithms were compared using the 5-fold and the 10-fold cross-validation techniques to build the classifier.

J48 Decision Tree (DT) Algorithm: The DT uses a divide-and-conquer and a top-down recursive approach using information gain to determine the best attribute for splitting the node. Iteratively, the DT uses the values of the different nodes greedily until the terminal node of the dependent variable (Breşfelean, 2007).

The Support Vector Machine (SVM) Algorithm: The SVM uses the value of features to plot data points in an n-dimensional space, where n represents the number of features. The hyperplane of the n-dimensional space then segregates the data points into classes. The data points are classified based on the number of features (Noble, 2006).

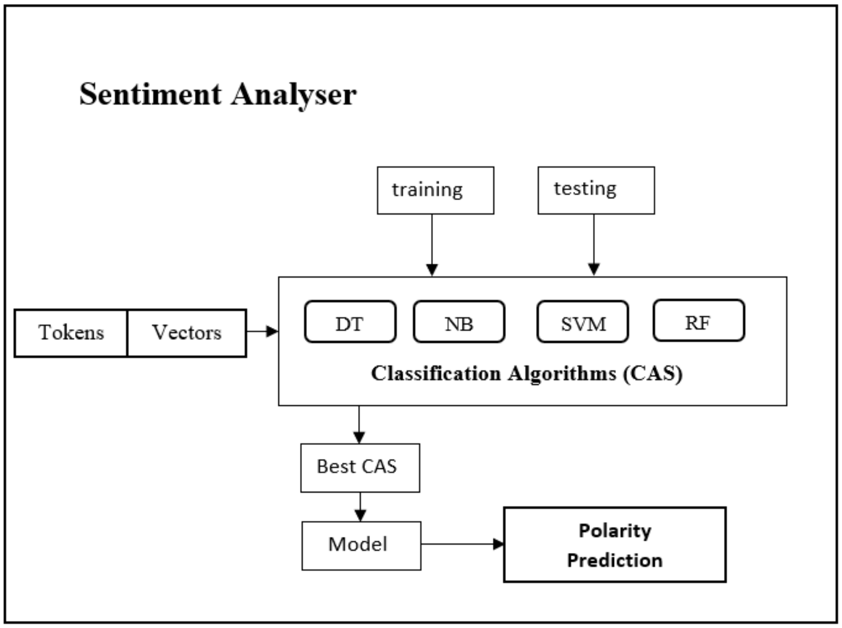

Figure 3: Sentiment Analyser

Naïve Bayes (NB) Algorithm: The NB uses the conditional probability of the Bayes theorem with the assumption of event occurrence based on another event that has already occurred. The posterior probability of the NB algorithm deals with each feature as equal and independent in the classification outcome of the class label (Berrar, 2019).

Random Forest (RF) Algorithm: The RF algorithm uses the Bootstrap Aggregation ensemble to create subset data from the training data in building the classifier. RF uses the functionality of decision tree algorithms in training the subset data to improve classification accuracy. The increasing number of trees in RF also helps avoid overfitting (Shaik & Srinivasan, 2019).

RESULTS AND ANALYSIS

The dataset for classification is analysed using the Waikato Environment for Knowledge Analysis (Weka 3.8.6) software. The Weka software uses the java programming language and was developed by the University of Waikato, New Zealand. The four classification algorithms, including the RF, SVM, DT and NB, were implemented using the 5-fold and the 10-fold cross-validation technique and the resulting classification metrics were analysed. The k-fold cross-validation in classification helps prevent overfitting in a predictive model and determines the generalisation of an independent dataset (Nandwani & Verma, 2021).

STRING TO WORD VECTOR

The unsupervised learning filter, StringtoWordVector in Weka, is applied on the training dataset to create a vector of words. As shown in Figure 4, 1297 words of vector strings were generated from the 255 instances of text. The Rainbow stop word filter in Weka was implemented to exclude words that are unimportant to the sentiments. Even though after applying the Rainbow stop word filter, the number of relevant words as attributes was reduced to 953, the classifier was built using the 1297 vector strings due to the low accuracy results when stop words was removed.

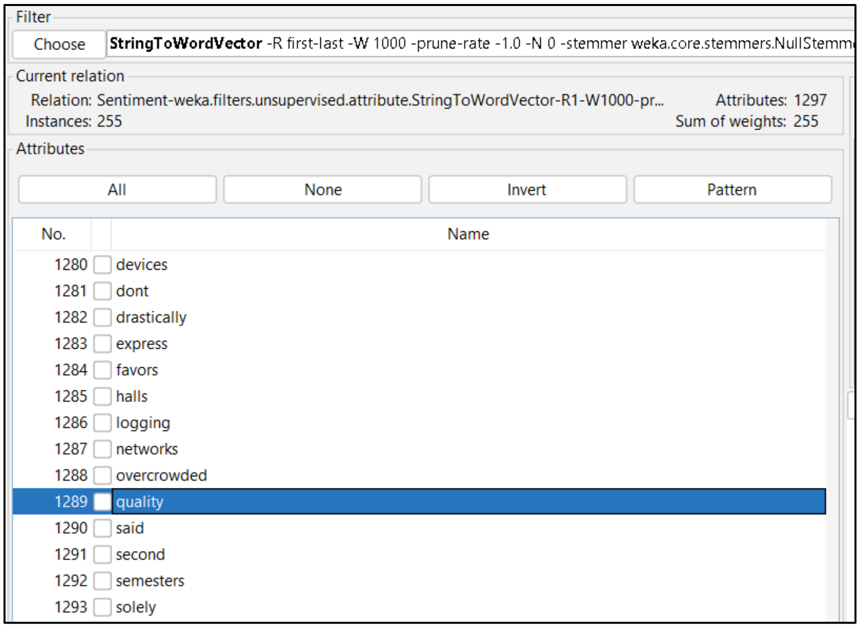

Figure 4: StringtoWordVector filter

In addition, as shown in Figure 5, the class labels for categorisation reveal that the training text contains 41% sad sentiments, 37% neutral sentiments, and 22% happy sentiments.

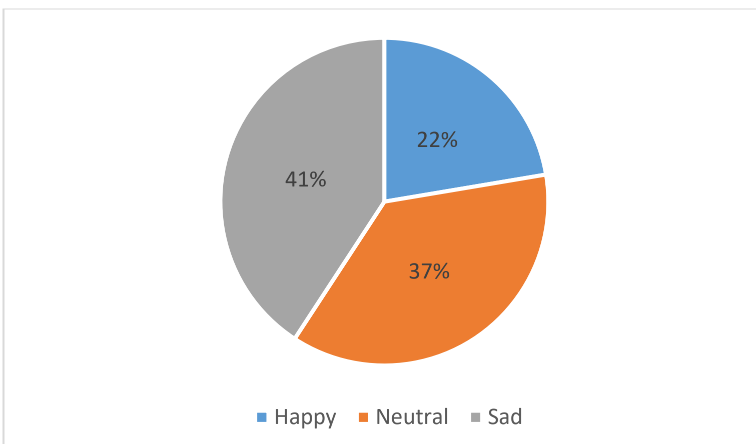

Figure 5: Class distribution

Classification Results

Research Question 2: What machine learning algorithm will provide the best model using sentiment analysis on the received comments?

The four classification algorithms, including SVM, RF, NB and DT, were subjected to the 10-fold and the 5-fold cross-validation techniques. As shown in Equation 1, classification accuracy is the ratio of correct predictions to the total number of input samples (Tharwat, 2021). The correct predictions are the TP and TN. The total input samples are the TP, TN, FP and FN. In the training dataset, 41% of the sentiments were sad, 37% were neutral, and 22% were happy. In this situation, the classification accuracy metric, which works better on a balanced dataset, is a fair measure of the model's performance. Even though classification accuracy is a useful metric based on the balanced dataset, the other metrics are vital in selecting a high performing predictive classifier.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \qquad (1)$$

Where TP = True Positives, TN = True Negatives, FN = False Negatives and FP = False Positives

As demonstrated in Table 4 and Figure 6, using the 10-fold cross-validation, the classification accuracy of the DT, NB, SVM and RF algorithms in descending order were 66.67%, 66.27%, 65.88% and 65.49%, respectively. The DT has the highest accuracy of 66.67%. For the 5-fold cross-validation, the accuracy performance changed in descending order of 66.67%, 64.71%, 60.00% and 56.07% for RF, SVM, DT and NB algorithms, respectively. The RF has the highest accuracy of 66.67%.

Conclusively, while RF gained a percentage increase of 1.18 from the 10-fold to the 5-fold cross-validations, DT, which is the closest, decreased in percentage points by 6.67.

Table 4: Classification accuracy

K-FOLD CROSS VALIDATION	SVM	RF	NB	DT
10-Fold	65.88%	65.49%	66.27%	66.67%
5-Fold	64.71%	66.67%	56.07%	60.00%

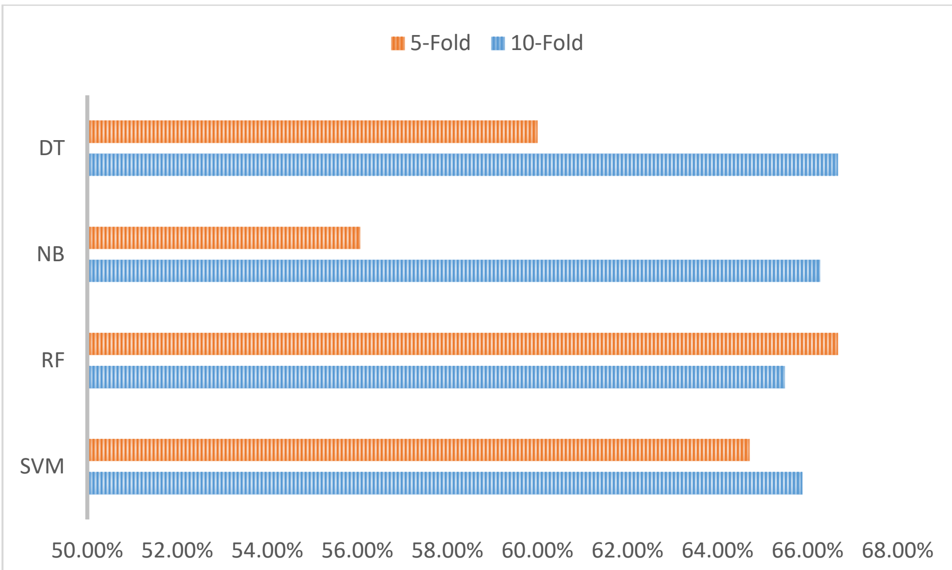

Figure 6: Classification accuracy

Other assessment metrics, including precision, recall, F-measure and ROC's Area under the Curve (AUC), are utilised to ascertain further the performance of the classification algorithms. The evaluation results are based on the 10-fold and the 5-fold cross-validation techniques

Precision, as shown in Equation 2, indicates the model's reliability in classifying the positive instances of the class accurately (Tharwat, 2021), while recall in Equation 3 indicates the positive samples truly classified as positive from the total number of positive instances. The F-measure metric is the harmonic mean of precision and recall, while the ROC score is a plot of TP rate against FP rate and indicates the performance of an algorithm to distinguish one class from the other. The ideal value of F-measure and the AUC of ROC is 1. The precision-recall curve is vital in an imbalanced dataset, while the ROC curve is more appropriate for balanced datasets (Davis & Goadrich, 2006; Fayzrakhmanov et al., 2018). The data collected from respondents is fairly balanced, making the ROC curve the focus of analysis.

$$\boldsymbol{Precision} = \frac{\boldsymbol{TP}}{\boldsymbol{TP} + \boldsymbol{FP}} \quad (2)$$

$$\boldsymbol{Recall} = \frac{\boldsymbol{TP}}{\boldsymbol{TP} + \boldsymbol{FN}} \quad (3)$$

As depicted in Table 5, for the 10-fold cross-validation, even though DT has the highest accuracy, as shown in Table 4, its ROC value of 0.775 is the least. Comparatively, RF has the highest ROC value, which enables it to distinguish better between the happy, sad and neutral classes but with the least F-measure score of 0.637.

Table 5: 10-fold cross-validation

CLASSIFIER	PRECISION	RECALL	F-MEASURE	ROC
SVM	0.658	0.659	0.658	0.777
RF	0.680	0.655	0.637	0.827
NB	0.671	0.663	0.662	0.784
DT	0.664	0.667	0.663	0.775

The RF algorithm using the 5-fold cross-validation had the highest classification accuracy, as shown in Table 4, and the ROC in Table 6 shows a dominant value of 0.838 compared to the other algorithms. In addition, the F-Measure score of 0.649 is the highest, with DT as the least at 0.592.

Table 6: 5-fold cross-validation

CLASSIFIER	PRECISION	RECALL	F-MEASURE	ROC
SVM	0.649	0.647	0.647	0.774
RF	0.703	0.667	0.649	0.838
NB	0.652	0.647	0.645	0.713
DT	0.592	0.600	0.592	0.729

Comparative results from the classification using the 10-fold and the 5-fold cross-validation techniques show a good performance measure for the Random Forest algorithm when validated with the 5-fold method.

Filtered Classifier With Random Forest Algorithm

Research Question 3: What is the test predictive performance of the best machine learning algorithm for futuristic sentiments?

The classification metric from the RF algorithm with the 5-fold cross-validation technique gave the highest performance and was used to build a predictive model for future forecasting of learner emotions. As shown in Figure 7, the filtered classifier functionality in Weka is utilised to match attributes in the training set to the test set.

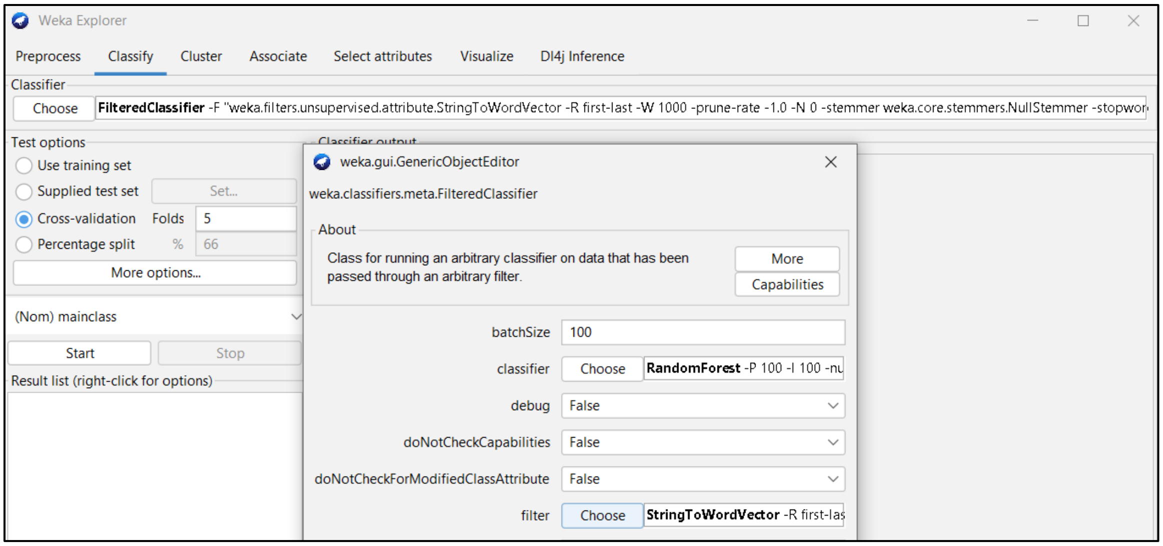

Figure 7: FilteredClassifier with RF

The RF algorithm predicts the sentiments as eighter good, sad or neutral based on comments on the learner's experiences in using the Moodle for assessment, as shown in Table 7. In addition to the test prediction, the model estimates the prediction error for the instructor.

Table 7: Sample prediction of learning sentiments using raw data

TEST FEED-BACK	COMMENT	TEST PREDIC-TION	PREDICTION ERROR
1	The problem is the network	Sad	0.51
2	In my opinion the hybrid way of assessing students is good but proper measures should be taken fair outcomes thus the general and departmental courses held online and the developmental face to face. it is good because it takes off the marking load from lectures and it easy for frequent assignments but the problem is, due to this online examination we as students have become very lazy and we care less of what is coming or what isn't because after all there will be books and stand by devices to log questions onto for answers and our friends to consult. I would advise that the system be updated and that students will be seated and invigilated to take the online examination	Neutral	0.47
3	Very cool	Happy	0.43
4	Sometimes it's complicated. Especially when waiting for the results	Sad	0.52

TEST FEED-BACK	COMMENT	TEST PREDIC-TION	PREDICTION ERROR
5	It's been a good mode and experience especially in relation to the program we are offering. It gives me the opportunity to directly put the application of technology (ICT) into action by using the LMS or vclass. In that I may get the chance to practice it on the field to help younger generations with the application of ICT tools	Neutral	0.42
6	It was good, just that the server sometimes worries	Neutral	0.44
7	It is very good and really help me as a student	Happy	0.57
8	It is only the network that do disturb some time	Neutral	0.46
9	Sometimes there are distraction of the network during the quizzes and exams.	Sad	0.50
10	Excellent	Sad	0.72

In test feedback item 10, the model mis-predicted an input text "Excellent" as "Sad" but identified the wrong prediction with a high error margin of 0.72. The misprediction prompted a check on the training dataset. The check results show that the learners never used "Excellent" to describe their experience with Moodle for online assessment.

DISCUSSION AND FINDINGS

In this study, SVM, RF, DT and NB traditional machine learning algorithms were subjected to the 10-fold and the 5-fold cross-validation techniques to build a classifier for sentiment prediction about learner experiences while using the Moodle for online assessment at the University of Education, Winneba Ghana. After training and testing, the RF algorithm emerged as the best classifier using the 5-fold cross-validation technique with an accuracy of 64.9%. Due to the moderate dataset, no feature selection mechanism was implemented. Since the sentiment classes were relatively balanced, the SMOTE filter for the minority dataset was not implemented. The stop word filter was also not implemented in building the final classifier due to the low accuracy results. The study by Dashtipour et al. (2016) utilised NB, SVM and Maximum Entropy without testing RF to build a classifier. Maximum Entropy emerged as the best classifier. Mujahid et al. (2021), in their online educational research using tweets during the pandemic, compared deep learning models to RF, DT and SVM algorithms. The three traditional algorithms using SMOTE outperformed the deep learning models. Osmanoğlu et al. (2020) compared DT, LR, MLP, SVM, KNN, MLR, XGB and GaussianNB algorithms in analysing distance students' sentiments without RF. The LR performed with the highest accuracy. Altrabsheh et al. (2014) compared the NB, SVM, CNB, and Maximum Entropy algorithms with RF to analyse real-time students' feedback with the N-grams feature selection mechanism. In their study, the SVM was the best classifier, with an accuracy of 95%. Onan (2021) compared four traditional machine learning algorithms with four ensemble techniques using N-gram models. His results show that an ensemble technique of RS + NB had the highest accuracy of 89.62%, with RS + RF coming in second at 87.53%. The RF algorithm comparison, which gave the proposed model the highest accuracy, is relevant in sentiment analysis using traditional machine learning algorithms, as shown when deployed in the limited literature. Even though the study did not implement N-gram models and TF-

IDF schemes because of the poor results realised when utilised on the narrow assessment dataset, its implementation in literature has shown good classification performance in a larger dataset. Several factors affect the performance of a classifier built for sentiment analysis.

Primarily, the size of the dataset is important for the predictive modelling of sentiments. The literature reviewed and the classification results obtained show that a larger sample size improves classification accuracy. Secondly, the performance of the sentiment analyser is dependent on the algorithms used in building the model. RF, SVM, and LR are traditional algorithms that have shown good classification performance in sentiment analysis. Thirdly, a dataset that is well balanced in terms of class labels gives a reliable classification accuracy for future prediction. Finally, without implementing any cross-validation technique, the accuracy obtained in sentiment analysis can be misleading due to overfitting. Overfitting occurs when a machine learning model fits too closely to the training data and cannot generalise. Even with high accuracy, a model developed without any cross-validation technique may predict future occurrences erroneously.

Limitation of Studies

Despite the fact that the study revealed learner emotions when using Moodle for online assessment, we encountered a few challenges. In this research, we collected data from 300 respondents due to their availability. As alluded to in the discussion, the size of a dataset has a direct relationship with classification accuracy. A dataset with higher samples will generally provide higher accuracy for future predictions. Secondly, little research has been done globally on using sentiment analysis for Moodle-based assessments, and this affected the discussion of results. Results discussion was therefore restricted to algorithmic performance instead of the overall state of online education, especially for assessing learners. Finally, the lack of literature in the Ghanaian context impeded discussions and influenced the general conclusions reached. This implies that, the findings at the University of Education, Winneba may not be applicable to all tertiary institutions in Ghana.

Conclusion and Future Research

Teaching and learning in higher educational institution have seen tremendous changes in pedagogy even before the COVID-19 pandemic. However, the devastating pandemic has accelerated the transition to online education, leaving the majority of institutions unprepared. The Moodle, an open-source software, was widely adopted by educational institutions, especially in Africa, as a learning platform for online education (Chang et al., 2022; Mpungose, 2020). The assessment application module in Moodle enables instructors to test the performance of learners and evaluate learning outcomes. Since assessment is vital to students' motivation and future success, it's critical that instructors have access to a sentiment analyser module that enables them to proactively respond to students' feedback.

At the University of Education in Winneba, Ghana, it has become necessary to allow students to freely express their emotions regarding their experiences when using the Moodle assessment module. The relevance stems from learner dissatisfaction and complaints about using the Moodle for online assessment during COVID-19. The study therefore utilised an open-ended questionnaire to get students' feedback on learner experiences while using the assessment module for mid-semester examinations. Machine learning algorithms were compared and used to create a model that can predict futuristic learner sentiments on online assessments while allowing the free expression of emotions. The RF algorithm emerged as the best classifier for the dataset with an accuracy of 64.9% using the 5-fold cross-validation technique. Since learner enrolment in Ghana continues to rise with the global trend (UNESCO Institute for Statistics, 2020), instructors and academic counsellors can prioritise the polarity of emotions and revert to learners promptly. This will prevent delays in attending to learner frustrations and avoid unsavoury learner behaviours. The integration of a sentiment analyser for the Moodle assessment activity module is therefore vital in Education 4.0, where big data and artificial intelligence (AI) are automating analytical tasks in the educational domain.

In future work, the authors will extend the comparison of the classifiers to include the deep learning algorithms. In addition, funding will be sorted in order to extend the research to other African universities for conclusive generalisation.

REFERENCES

Adarkwah, M. A. (2021). "Another year lost?": A novel approach to the online learning in Ghana. *Academia Letters*, *2*, Article 365. https://doi.org/10.20935/AL365

Adnan, K., & Akbar, R. (2019). Limitations of information extraction methods and techniques for heterogeneous unstructured big data. *International Journal of Engineering Business Management*, *11*, 1–23. https://doi.org/10.1177/1847979019890771

Agormedah, E. K., Adu Henaku, E., Ayite, D. M. K., & Apori Ansah, E. (2020). Online learning in higher education during Covid-19 pandemic: A case of Ghana. *Journal of Educational Technology and Online Learning*, *3*(3), 183–210. https://doi.org/10.31681/jetol.726441

Ahmad, M., Aftab, S., Muhammad, S. S., & Ahmad, S. (2017). Machine learning techniques for sentiment analysis: A review. *International Journal of Multidisciplinary Sciences and Engineering*, *8*(3), 27–32. http://www.ijmse.org/Volume8/Issue3/paper5.pdf

Al-Falahy, N., & Alani, O. Y. (2017). Technologies for 5G networks: Challenges and opportunities. *IT Professional*, *19*(1), 12–20. https://doi.org/10.1109/MITP.2017.9

Al-Maroof, R. S., Akour, I., Aljanada, R., Alfaisal, A. M., Alfaisal, R. M., Aburayya, A., & Salloum, S. A. (2021). Acceptance determinants of 5G services. *International Journal of Data and Network Science*, *5*(4), 613–628. https://doi.org/10.5267/j.ijdns.2021.8.006

Altexsoft. (2020, December 14). *Structured vs unstructured data: Compared and explained.* https://www.altexsoft.com/blog/structured-unstructured-data/

Altrabsheh, N., Cocea, M., & Fallahkhair, S. (2014, November). Sentiment analysis: Towards a tool for analysing real-time students feedback. *Proceedings of the 26th International Conference on Tools with Artificial Intelligence (ICTAI, 2014)* (pp. 419–423). Limassol, Cyprus: IEEE. https://doi.org/10.1109/ICTAI.2014.70

Bakker, I., van der Voordt, T., Vink, P., & de Boon, J. (2014). Pleasure, arousal, dominance: Mehrabian and Russell revisited. *Current Psychology*, *33*(3), 405–421. https://doi.org/10.1007/s12144-014-9219-4

Barron-estrada, M. L., Zatarain-cabada, R., & Oramas-bustillos, R. (2019). Emotion recognition for education using sentiment analysis. *Research in Computing Science*, *148*(5), 71–80. https://doi.org/10.13053/rcs-148-5-8

Berrar, D. (2019). Bayes' theorem and naive Bayes classifier. *Encyclopedia of Bioinformatics and Computational Biology: ABC of Bioinformatics*, *1*, 403–412. https://doi.org/10.1016/B978-0-12-809633-8.20473-1

Bhasin, B., Gupta, G., & Malhotra, S. (2021). Impact of Covid-19 pandemic on education system. *EPRA International Journal of Environmental Economics, Commerce and Educational Management*, *8*(2), 6–8. https://doi.org/10.36713/epra6363

Birjali, M., Kasri, M., & Beni-Hssane, A. (2021). A comprehensive survey on sentiment analysis: Approaches, challenges and trends. *Knowledge-Based Systems*, *226*, 107134. https://doi.org/10.1016/j.knosys.2021.107134

Breşfelean, V. P. (2007, June). Analysis and predictions on students' behaviour using decision trees in WEKA environment. *Proceedings of the 29th International Conference on Information Technology Interfaces (ITI)* (pp. 51–56). Cavtat, Croatia: IEEE. https://doi.org/10.1109/ITI.2007.4283743

Buabeng-Andoh, C. (2012). Factors influencing teachers' adoption and integration of information and communication technology into teaching: A review of the literature. *International Journal of Education and Development Using Information and Communication Technology*, *8*(1), 136–155. https://www.learntechlib.org/p/188018/

Budu, K. W. A., Yinping, M., & Mireku, K. K. (2018). Investigating the effect of behavioural intention on e-learning systems usage: Empirical study on tertiary education institutions in Ghana. *Mediterranean Journal of Social Sciences*, *9*(3), 201–216. https://doi.org/10.2478/mjss-2018-0062

Chang, Y. C., Li, J. W., & Huang, D. Y. (2022). A personalized learning service compatible with Moodle e-learning management system. *Applied Sciences (Switzerland)*, *12*(7), 3562. https://doi.org/10.3390/app12073562

Cromwell, A. (2020, March 24). *Education ministry set to release online learning platform for senior high schools*. https://myjoyonline.com/education-ministry-set-to-release-online-learning-platform-for-senior-high-schools/

Dake, D. K., Essel, D. D., & Agbodaze, J. E. (2021, July). Using machine learning to predict students' academic performance during Covid-19. *Proceedings of the International Conference on Computing, Computational Modelling and Applications (ICCMA 2021)* (pp. 9–15), Brest, France. IEEE. https://doi.org/10.1109/ICCMA53594.2021.00010

Dampson, D. G. (2021). Determinants of learning management system adoption in an era of COVID-19: Evidence from a Ghanaian university. *European Journal of Education and Pedagogy*, *2*(3), 80–87. https://doi.org/10.24018/ejedu.2021.2.3.94

Dashtipour, K., Poria, S., Hussain, A., Cambria, E., Hawalah, A. Y. A., Gelbukh, A., & Zhou, Q. (2016). Multilingual sentiment analysis: State of the art and independent comparison of techniques. *Cognitive Computation*, *8*(4), 757–771. https://doi.org/10.1007/s12559-016-9415-7

Davis, J., & Goadrich, M. (2006, June). The relationship between precision-recall and ROC curves. *Proceedings of the 23rd International Conference on Machine Learning (ICML '06)* (pp. 233–240), Pittsburgh Pennsylvania, USA. ACM. https://doi.org/10.1145/1143844.1143874

Dolianiti, F. S., Iakovakis, D., Dias, S. B., Hadjileontiadou, S., Diniz, J. A., & Hadjileontiadis, L. (2019). Sentiment analysis techniques and applications in education: A survey. In M. A. Tsitouridou, J. Diniz, & T. Mikropoulos (Eds.), *Communications in Computer and Information Science (Vol. 993)* (pp. 412–427). Springer. https://doi.org/10.1007/978-3-030-20954-4_31

Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion*, *6*(3–4), 169–200. https://doi.org/10.1080/02699939208411068

Fayzrakhmanov, R., Kulikov, A., & Repp, P. (2018, March). The difference between precision-recall and ROC curves for evaluating the performance of credit card fraud detection models. In E. Siemens, B. Krause, & L. Mylnikov (Eds.), *Proceedings of the 6th International Conference on Applied Innovations in IT (ICAIIT)*, *6*(1), 17–22, Koethen, Germany. Anhalt University of Applied Sciences. https://icaiit.org/paper.php?paper=6th_ICAIIT/1_3

Firat, M., & Bozkurt, A. (2020). Variables affecting online learning readiness in an open and distance learning university. *Educational Media International*, *57*(2), 112–127. https://doi.org/10.1080/09523987.2020.1786772

Garad, A., Al-Ansi, A. M., & Qamari, I. N. (2021). The role of e-learning infrastructure and cognitive competence in distance learning effectiveness during the Covid-19 pandemic. *Cakrawala Pendidikan*, *40*(1), 81–91. https://doi.org/10.21831/cp.v40i1.33474

Garg, H., & Lal, N. (2018). Data analysis: Opinion mining and sentiment analysis of opinionated unstructured data. In M. Singh, P. Gupta, V. Tyagi, J. Flusser, & T. Ören (Eds.), *Communications in computer and information science (vol. 906)* (pp. 249–258). Springer Singapore. https://doi.org/10.1007/978-981-13-1813-9_25

Guo, J. (2022). Deep learning approach to text analysis for human emotion detection from big data. *Journal of Intelligent Systems*, *31*(1), 113–126. https://doi.org/10.1515/jisys-2022-0001

Henaku, E. A. (2020). Covid-19: Online learning experience of college students: The case of Ghana. *International Journal of Multidisciplinary Sciences and Advanced Technology*, *1*(2), 54–62. https://www.ijmsat.com/covid-19/special-issue-2

Hutajulu, S., Dhewanto, W., & Prasetio, E. A. (2020). Two scenarios for 5G deployment in Indonesia. *Technological Forecasting and Social Change*, *160*, 120221. https://doi.org/10.1016/j.techfore.2020.120221

Huynh-Cam, T. T., Agrawal, S., Chen, L. S., & Fan, T. L. (2021). Using Moodle-based e-assessment in English listening and reading courses: A Vietnamese case study. *Journal of Institutional Research South East Asia*, *19*(2), 66–92. http://www.seaairweb.info/journal/articles/JIRSEA_v19_n02/JIRSEA_v19_n02_All.pdf#page=83

Jain, P. K., Pamula, R., & Srivastava, G. (2021). A systematic literature review on machine learning applications for consumer sentiment analysis using online reviews. *Computer Science Review*, *41*, 100413. https://doi.org/10.1016/j.cosrev.2021.100413

Jiang, Y., Zhang, J., & Xin, T. (2019). Toward education quality improvement in China: A brief overview of the national assessment of education quality. *Journal of Educational and Behavioural Statistics*, *44*(6), 733–751. https://doi.org/10.3102/1076998618809677

Kara, M. (2018). A systematic literature review: Constructivism in multidisciplinary learning environments. *International Journal of Academic Research in Education*, *4*(1-2), 19–26. https://doi.org/10.17985/ijare.520666

Keppell, M. (2014). Personalised learning strategies for higher education. *The Future of Learning and Teaching in Next Generation Learning Spaces (International Perspectives on Higher Education Research)*, *12*, 3-21. https://doi.org/10.1108/S1479-362820140000012001

Leber, J., Renkl, A., Nückles, M., & Wäschle, K. (2018). When the type of assessment counteracts teaching for understanding. *Learning: Research and Practice*, *4*(2), 161–179. https://doi.org/10.1080/23735082.2017.1285422

Lee, J., & Han, S. H. (Eds.). (2021). *The future of service post-Covid-19 pandemic, volume 1: Rapid adoption of digital service technology*. Springer. https://doi.org/10.1007/978-981-33-4126-5

Lee, J., & Kim, D. (2020, January). A study on innovation in university education: Focusing on 5G mobile communication. *Proceedings of the 17th Annual Consumer Communications and Networking Conference (CCNC 2020)* (pp. 2020–2023), Las Vegas, NV, USA. IEEE. https://doi.org/10.1109/CCNC46108.2020.9045138

Mohammad, S. M., & Turney, P. D. (2013). *NRC emotion lexicon*. NRC Technical Report. National Research Council Canada. https://doi.org/10.4224/21270984

Moodle. (2005, January). *Moodle for teachers, trainers and administrators*. https://www.immagic.com/elibrary/archives/general/moodle/m050100w.pdf

Moore, J. L., Dickson-Deane, C., & Galyen, K. (2011). E-Learning, online learning, and distance learning environments: Are they the same? *The Internet and Higher Education*, *14*(2), 129–135. https://doi.org/10.1016/j.iheduc.2010.10.001

Mpungose, C. B. (2020). Beyond limits: Lecturers' reflections on Moodle uptake in South African universities. *Education and Information Technologies*, *25*(6), 5033–5052. https://doi.org/10.1007/s10639-020-10190-8

Mujahid, M., Lee, E., Rustam, F., Washington, P. B., Ullah, S., Reshi, A. A., & Ashraf, I. (2021). Sentiment analysis and topic modelling on tweets about online education during Covid-19. *Applied Sciences (Switzerland)*, *11*(18), 8438. https://doi.org/10.3390/app11188438

Mwangi, E. G., Imonje, R., & Kalai, J. (2023). Moodle satisfaction levels during online assessments in selected public universities in Kenya. *International Journal of Science and Research*, *12*(2), 104–113. https://www.ijsr.net/archive/v12i2/SR23112105548.pdf

Nandwani, P., & Verma, R. (2021). A review on sentiment analysis and emotion detection from text. *Social Network Analysis and Mining*, *11*(1), Article 81. https://doi.org/10.1007/s13278-021-00776-6

Noble, W. S. (2006). What is a support vector machine? *Nature Biotechnology*, *24*(12), 1565–1567. https://doi.org/10.1038/nbt1206-1565

Nyarko-Boateng, O., Xedagbui, F. E. B., Adekoya, A. F., & Weyori, B. A. (2020). Fiber optic deployment challenges and their management in a developing country: A tutorial and case study in Ghana. *Engineering Reports*, *2*(2), e12121. https://doi.org/10.1002/eng2.12121

Ohliati, J., & Abbas, B. S. (2019). Measuring students satisfaction in using learning management system. *International Journal of Emerging Technologies in Learning*, *14*(4), 180–189. https://doi.org/10.3991/ijet.v14i04.9427

Onan, A. (2021). Sentiment analysis on massive open online course evaluations: A text mining and deep learning approach. *Computer Applications in Engineering Education*, *29*(3), 572–589. https://doi.org/10.1002/cae.22253

Osmanoğlu, U. Ö., Atak, O. N., Çağlar, K., Kayhan, H., & Can, T. (2020). Sentiment analysis for distance education course materials: A machine learning approach. *Journal of Educational Technology and Online Learning*, *3*(1), 31–48. https://doi.org/10.31681/jetol.663733

Ot, A. (2023, July 21). *Structured data: Examples, sources, and how it works*. https://www.datamation.com/big-data/structured-data/

Owusu-Oware, E., & Tanye, H. A. (2023). The effectiveness and integrity of online assessments using Moodle learning management system: Perspectives of a developing country's university teachers. *Journal of Contemporary Issues in Business and Government*, *29*(3), 152–163. https://doi.org/10.47750/cibg.2023.29.03.011

Ozili, P. K., & Arun, T. (2020). Spillover of Covid-19: Impact on the global economy. *SSRN Electronic Journal*, 99850. https://doi.org/10.2139/ssrn.3562570

Palvia, S., Aeron, P., Gupta, P., Mahapatra, D., Parida, R., Rosner, R., & Sindhi, S. (2018). Online education: Worldwide status, challenges, trends, and implications. *Journal of Global Information Technology Management*, *21*(4), 233–241. https://doi.org/10.1080/1097198X.2018.1542262

Plutchik, R. (1982). A psychoevolutionary theory of emotions. *Social Science Information*, *21*(4-5), https://doi.org/10.1177/053901882021004003

Pokhrel, S., & Chhetri, R. (2021). A literature review on impact of Covid-19 pandemic on teaching and learning. *Higher Education for the Future*, *8*(1), 133–141. https://doi.org/10.1177/2347631120983481

Prain, V., Cox, P., Deed, C., Dorman, J., Edwards, D., Farrelly, C., Keeffe, M., Lovejoy, V., Mow, L., Sellings, P., Waldrip, B., & Yager, Z. (2013). Personalised learning: Lessons to be learnt. *British Educational Research Journal*, *39*(4), 654–676. https://doi.org/10.1080/01411926.2012.669747

Shaik, A. B., & Srinivasan, S. (2019). A brief survey on random forest ensembles in classification model. In S. Bhattacharyya, A. Hassanien, D. Gupta, A. Khanna, & I. Pan (Eds.), *Lecture notes in networks and systems (vol. 56)* (pp. 253–260). Springer Singapore. https://doi.org/10.1007/978-981-13-2354-6_27

Shen, J., Baysal, O., & Shafiq, M. O. (2019, August). Evaluating the performance of machine learning sentiment analysis algorithms in software engineering. *Proceedings of the IEEE 17th International Conference on Dependable, Autonomic and Secure Computing, IEEE 17th International Conference on Pervasive Intelligence and Computing, IEEE 5th International Conference on Cloud and Big Data Computing, 4th International Conference on Cyber Science and Technology Congress (DASC/PiCom/CBDCom/CyberSciTech)* (pp. 1023–1030). Fukuoka, Japan: IEEE. https://doi.org/10.1109/DASC/PiCom/CBDCom/CyberSciTech.2019.00185

Simanullang, N. H. S., & Rajagukguk, J. (2020). Learning management system (LMS) based on Moodle to improve students learning activity. *Journal of Physics: Conference Series*, *1462*, 012067. https://doi.org/10.1088/1742-6596/1462/1/012067

Singh, V. K., Piryani, R., Waila, P., & Devaraj, M. (2016). Computing sentiment polarity of texts at document and aspect levels. *ECTI Transactions on Computer and Information Technology (ECTI-CIT)*, *8*(1), 67–79. https://doi.org/10.37936/ecti-cit.201481.54389

Snejana, D., & Veselina, N. (2021, November). Assessment of students during Covid-19 case experience. *Proceedings of the 16th International Conference on Virtual Learning (ICVL 2021)* (pp. 67-74), Bucharest, Romania. SSRN. https://doi.org/10.2139/ssrn.4268161

Sowmiya, R., Sivakamasundari, G., & Archana, V. (2022, December). Facial emotion recognition using deep learning approach. *Proceedings of the International Conference on Automation, Computing and Renewable Systems (ICACRS 2022)* (pp. 1064–1069), Pudukkottai, India. IEEE. https://doi.org/10.1109/ICACRS55517.2022.10029092

Tharwat, A. (2021). Classification assessment methods. *Applied Computing and Informatics*, *17*(1), 168–192. https://doi.org/10.1016/j.aci.2018.08.003

UNESCO Institute for Statistics (UIS). (2020). *School enrolment, tertiary (% gross)*. https://data.worldbank.org/indicator/se.ter.enrr

Upoalkpajor, J. L. N., & Upoalkpajor, C. B. (2020). The impact of Covid-19 on education in Ghana. *Asian Journal of Education and Social Studies*, *9*(1), 23–33. https://doi.org/10.9734/ajess/2020/v9i130238 van Groen, M. M., & Eggen, T. J. H. M. (2020). Educational test approaches: The suitability of computer-based test types for assessment and evaluation in formative and summative contexts. *Journal of Applied Testing Technology*, *21*(1), 12–24. http://jattjournal.net/index.php/atp/article/view/146484

Wongkar, M., & Angdresey, A. (2019, October). Sentiment analysis using naive Bayes algorithm of the data crawler: Twitter. *Proceedings of 4th International Conference on Informatics and Computing (ICIC 2019)* (pp. 1–5), Semarang, Indonesia. IEEE. https://doi.org/10.1109/ICIC47613.2019.8985884

Zad, S., Heidari, M., Jones, J. H. J., & Uzuner, O. (2021, May). Emotion detection of textual data: An interdisciplinary survey. *Proceedings of the World AI IoT Congress (AIIoT 2021)* (pp. 255–261), Seattle, WA, USA. IEEE. https://doi.org/10.1109/AIIoT52608.2021.9454192

AUTHORS

Delali Kwasi Dake, PhD is a Senior Lecturer in the department of ICT Education at the University of Education, Winneba, Ghana. His research interests include educational data mining, artificial intelligence, sentiment analysis, software-defined networking, and Internet of Things.

Godwin Kudjo Bada is a Lecturer in the department of ICT Education at the University of Education, Winneba, Ghana. His research interests include computer security and educational data mining.