UNIVERSITY OF SKILLS TRAINING AND ENTREPRENEURIAL DEVELOPMENT

JOURNAL OF TECHNOLOGY EDUCATION AND APPLIED SCIENCES (JTEAS)

Volume 1, Issue 1, 2026

Published by:
University of Skills Training and Entrepreneurial Development
P. O. Box 1277
Kumasi, Ghana

ISSN: 3093-4796

Journal of Technology Education and Applied Sciences

USTED Press

ISSN: 3093-4796 (Online) Journal homepage: https://journal.aamusted.edu.gh/index.php/jteas/index

E D I T O R I A L

Knowledge that Transforms

The publication of the inaugural issue of the **Journal of Technology Education and Applied Sciences (JTEAS)** marks a defining milestone in the intellectual journey of the **University of Skills Training and Entrepreneurial Development (USTED), Ghana**. Beyond inaugurating a new scholarly publication, this issue establishes a forum for rigorous inquiry, interdisciplinary dialogue, and transformative research that seeks not merely to explain the world, but to improve it.

Universities are ultimately judged not only by the graduates they produce but by the knowledge they generate and the societal impact of that knowledge. In an era defined by the Fourth Industrial Revolution, artificial intelligence, climate change, digital transformation, and evolving workforce demands, scholarship must move beyond disciplinary boundaries to provide innovative solutions to complex global challenges. JTEAS is founded on this conviction. It embraces the philosophy that excellent research should be intellectually rigorous, methodologically sound, socially responsive, and practically consequential.

The contributions assembled in this inaugural issue exemplify this vision. They collectively demonstrate that technology education and the applied sciences are not isolated domains of academic enquiry but powerful instruments for sustainable development, industrial innovation, environmental stewardship, and human advancement. From strengthening experiential learning through industrial attachment, to advancing sustainable material preservation, improving renewable energy forecasting through artificial intelligence, and exploring public perceptions of generative AI using machine learning, the articles reflect a common commitment to research that bridges theory and practice while responding to contemporary societal needs.

The diversity of these studies also illustrates the intellectual identity that JTEAS seeks to cultivate. We believe that the most significant scientific breakthroughs increasingly emerge at the intersections of disciplines—where education informs technology, engineering advances sustainability, data science shapes public policy, and innovation creates opportunities for inclusive development. This journal therefore welcomes scholarship that transcends traditional academic silos and contributes meaningfully to the advancement of knowledge and society.

As an international, peer-reviewed, open-access journal, JTEAS is committed to the highest standards of scholarly excellence, editorial integrity, and research ethics. Our aspiration extends beyond publication; we seek to become a trusted platform where researchers, educators, innovators, industry leaders, and policymakers converge to exchange ideas that influence policy, inspire innovation, and accelerate sustainable development across Africa and the wider global community. We envision JTEAS evolving into one of Africa's leading scholarly journals in technology education and applied sciences, recognised for publishing research that is globally relevant while remaining deeply responsive to local realities.

The Editorial Board expresses its sincere appreciation to the authors whose pioneering work forms the foundation of this inaugural issue, the reviewers whose rigorous evaluations uphold the quality of the journal, and the leadership of USTED for their vision and unwavering support.

Editor-in-Chief

Prof Isaac Boateng
Journal of Technology Education and Applied Sciences (JTEAS)
University of Skills Training and Entrepreneurial Development (USTED), Ghana

Journal of Technology Education and Applied Sciences

USTED Press

ISSN: 3093-4796 (Online) Journal homepage: https://journal.aamusted.edu.gh/index.php/jteas/index

RESEARCH ARTICLE

The Use of Machine Learning Classifiers to Evaluate the Sentiment Categories of ChatGPT-Related Tweets

Prince Clement Addo*[1,2], Samuel Kofi Akpatsa[2], Emmanuel Kwadwo Mensah[2], Richard Osie Adu[2], Oliver Kufour Boansi[2], Nukpe Philip[3], Donald Yeboah[4]

[1]School of Economics and Management, Southwestern University of Science and Technology, China
[2]Faculty of Applied Sciences and Mathematics Education, University of Skills Training and Entrepreneurial Development, Ghana
[3]Global Banking School, Oxford Brookes University, United Kingdom
[4]Faculty of Science and Engineering, Äbo Akademi University, Turku, Finland
*Corresponding Author: [pcaddo@aamusted.edu.gh]

To cite this article: Addo, P. C., Akpatsa, S. K., Mensah, E. K., Adu, R. O., Boansi, O. K., Philip, N., & Yeboah, D. (2026). The use of machine learning classifiers to evaluate the sentiment categories of ChatGPT-related tweets. Journal of Technology Education and Applied Sciences, 1(1), 1–17. https://doi.org/10.64922/jteas.v1i1.19

ABSTRACT

This study evaluates the performance of conventional machine learning classifiers for sentiment analysis of ChatGPT-related tweets. While deep learning approaches have demonstrated advanced capabilities, their substantial computational requirements present practical limitations. Using a dataset of 219,294 tweets from November 2022 to April 2023, we assess four classifiers (Logistic Regression, Support Vector Machines, Random Forest, and Naive Bayes) with two feature extraction techniques (Bag of Words and TF-IDF) across balanced and imbalanced datasets. Results indicate that Linear SVM with TF-IDF vectorization achieves the highest accuracy (84.02%) and macro F1-score (80.81%) without balancing techniques. Random Forest with Bag of Words and balancing techniques shows competitive performance (80.08% macro F1-score), while Naive Bayes consistently underperforms across configurations. n-gram analysis reveals predominantly positive public sentiment toward ChatGPT, focusing on capabilities and utility, while negative sentiments center on performance limitations and broader AI implications. This research provides empirical insights into sentiment analysis methodologies for emerging AI technologies and demonstrates the continued effectiveness of conventional machine learning approaches in resource-constrained environments.

KEYWORDS

Sentiment Analysis, Natural Language Processing, Machine Learning, ChatGPT, Tweets

Publication History

Received: 26 November 2025
Revised: 7 April 2026
Accepted: 24 June 2026
Available online: 30 July 2026

Introduction

Sentiment analysis represents a transformative domain within natural language processing (NLP) and machine learning (ML), enabling the automated identification and categorization of emotions and attitudes within text data. Its applications span from gauging public sentiment on social issues to analyzing customer feedback in business settings (Tan et al., 2023). With the proliferation of social media platforms, sentiment analysis has become an indispensable tool for interpreting the vast amounts of user-generated content produced daily.

Twitter, characterized by its brevity and immediacy, serves as an ideal platform for capturing public opinion on diverse topics (Kumar & Sebastian, 2021). The advent of ChatGPT, developed by OpenAI, has introduced a novel dimension to online discourse. As a variant of Generative Pre-trained Transformer models, ChatGPT's ability to generate human-like text responses has sparked widespread discussion and debate on social media platforms. However, the nuanced and often unpredictable conversational styles of AI-generated texts present

unique challenges to traditional sentiment analysis techniques (Zhang & Wallace, 2015). The evolution of sentiment analysis from basic text classification to sophisticated computational techniques reflects the growing complexity of digital communication. Early systems were limited by computational resources and algorithmic sophistication, typically relying on basic models to categorize texts into predefined sentiment categories (Parveen et al., 2023). The exponential increase in digital text data, particularly from social media, necessitated more advanced analytical techniques capable of addressing the complexities inherent in human language, such as sarcasm, irony, and context-dependent meanings (Tan et al., 2023).

The emergence of machine learning techniques has marked a significant advancement in the field of sentiment analysis. The methods vary from straightforward algorithms such as naive Bayes classifiers to more intricate neural networks. These machine learning methods can "learn" directly from data, finding complex patterns and language relationships that were not directly programmed into their algorithms. This allows for adaptation to the inherent variability and contextual nuances of natural language expressions (Gehrmann et al., 2017). This learning process enabled a deeper understanding of the context and intricacies of emotional expressions, thereby transforming the approach to sentiment analysis. This research addresses a significant gap in sentiment analysis methodologies by evaluating the effectiveness of traditional machine learning classifiers when analysing traditional machine learning classifiers when analysing tweets related to ChatGPT. Deep learning methods demonstrate advanced capabilities; however, their significant computational demands often make them impractical. This research examines traditional classifiers such as logistic regression, support vector machines (SVM), random forest, and Naive Bayes, providing valuable insights into effective sentiment analysis methods suitable for resource-limited settings. This research presents three important contributions: (1) This constitutes one of the early systematic evaluations of traditional machine learning classifiers specifically focused on Twitter conversations about ChatGPT, which emerged after November 2022. (2) This study directly contrasts the performance of conventional machine learning classifiers (Logistic Regression, Support Vector Machines, Random Forest, and Naive Bayes) using two feature extraction techniques (Bag of Words and TF-IDF) with and without balancing techniques (RandomOverSampler) across four classifiers and two feature extraction methods, resulting in 16 experimental configurations that are statistically validated using McNemar's test. (3) This study demonstrates through n-gram analysis that the initial public sentiment toward ChatGPT was predominantly positive, focusing on capabilities and utility, while negative sentiments centered on performance limitations and broader AI implications.

This study tests the hypothesis that conventional machine learning classifiers, when optimally configured with appropriate feature extraction techniques and class balancing mechanisms, can achieve comparable performance to more computationally expensive deep learning approaches for sentiment analysis of emerging technology discussions. Specifically, we hypothesize that: (1) term weighting approaches like TF-IDF will yield statistically significant improvements over frequency-based approaches for linear classifiers, (2) ensemble methods will demonstrate greater robustness to class imbalance compared to linear methods, and (3) classifier performance will vary systematically based on sentiment category complexity. Through addressing these interconnected objectives, this study provides empirical insights into sentiment analysis methodologies and their practical applications in social media analytics, particularly in contexts involving rapidly evolving artificial intelligence technologies that generate significant public engagement.

Related Works

Evolution of Sentiment Analysis

Sentiment analysis has evolved significantly since its inception, transitioning from rule-based lexicon approaches to sophisticated machine learning techniques. The field was formally established by Pang et al. (2002), who demonstrated that supervised machine learning methods could outperform human-selected sentiment word lists for classifying movie reviews, establishing the supervised classification paradigm that subsequent research built upon. Liu (2012) further consolidated the theoretical foundations of opinion mining and sentiment analysis, providing a comprehensive taxonomy of sentiment expressions and formalizing the subjectivity classification task. Building on these foundations, subsequent work by Parveen et al. (2023) demonstrated the continued potential of supervised learning methods for sentiment classification across varied

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).
All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).
https://doi.org/10.64922/jteas.v1i1.19

domains, applying naive Bayes, maximum entropy, and support vector machines to short text data. Their work highlighted machine learning's advantages over purely lexical approaches. Tan et al. (2022) introduced an ensemble hybrid deep learning model for sentiment classification and provided a comprehensive framework for sentiment analysis, addressing challenges related to opinion mining in social media contexts. These foundational studies established sentiment analysis as a distinct research domain with significant implications for understanding public opinion.

Sentiment Analysis in Social Media

Social media platforms present distinct challenges to sentiment analysis owing to their colloquial language, conciseness, and utilisation of slang, emoticons, and abbreviations. Aziz and Dimililer (2020) developed a new approach to sentiment analysis using Twitter data by applying distant supervision for automated sentiment classification. They demonstrated how emoticons could serve as noisy labels for training effective sentiment classifiers. The creation of corpora for Twitter sentiment analysis has been further developed, with recent studies like Aziz and Dimililer (2020) using ensemble techniques. Gifari et al. (2021) investigated the utility of linguistic features for sentiment analysis in the social media domain, finding that part-of-speech features were less useful than n-grams and lexical features. Kumar et al. (2018) showed that sentiment lexicons and n-gram features work well for Twitter sentiment analysis by getting the best results on the SemEval-2013 dataset.

Machine Learning Approaches for Sentiment Analysis

Various machine learning approaches have been applied to sentiment analysis tasks. Hasan et al. (2018) evaluated multiple classifiers, including Random Forest, Decision Trees, and Logistic Regression, on social media data, demonstrating that ensemble methods often outperformed individual classifiers, while also confirming that SVM achieved strong results in Twitter-based classification tasks. More recently, deep learning approaches have gained prominence. Dos Santos and Gatti (2014) utilised deep convolutional neural networks for sentiment analysis of short texts, achieving superior results compared to traditional methods. Severyn and Moschitti (2015) further improved performance using a 2015) further improved performance using a convolutional neural network (CNN) with word embeddings for Twitter sentiment analysis. Recent advancements have further refined sentiment analysis capabilities. Additionally, Zhang and Luo (2023) specifically examined sentiment analysis of AI-related discourse across social media platforms, providing valuable insights into public perception dynamics. More recently, studies have begun examining public sentiment toward ChatGPT specifically. Leiter et al. (2023) conducted a large-scale analysis of Twitter discourse about ChatGPT during its early release period, finding that public reactions were largely polarized between enthusiastic adoption and concern about educational and professional disruption. Sallam et al. (2023) similarly analyzed ChatGPT-related tweets across multiple languages, identifying significant variation in sentiment across user demographics and geographical regions. These domain-specific studies underscore the need for robust sentiment classification frameworks tailored to AI technology discourse. A distinct challenge in this domain concerns the analysis of content about AI-generated text. Guo et al. (2023) demonstrated that sentiment classifiers trained on human-authored content exhibit measurable performance degradation when applied to tweets discussing AI-generated outputs, because user language in such contexts tends toward technical hedging and conditional framing rather than direct emotional expression. This characteristic of AI-discourse tweets motivates the present study's focus on systematic classifier evaluation rather than direct ap

Despite these advances, conventional machine learning approaches remain valuable, particularly in resource-constrained environments. Zainuddin et al. (2018) demonstrated that SVM and Logistic Regression could achieve competitive performance with appropriate feature engineering, while requiring significantly less computational resources than deep learning methods.

Feature Extraction Techniques

Feature extraction plays a crucial role in sentiment analysis performance. Tan et al. (2023) compared various feature selection methods for sentiment analysis, finding that appropriate feature selection significantly

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).
All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).
https://doi.org/10.64922/jteas.v1i1.19

improved classification accuracy. The Bag of Words model, though simple, has been utilized in modern sentiment analysis approaches, such as those in Gifari et al. (2021). TF-IDF vectorization improves upon raw frequency counts by incorporating inverse document frequency, effectively emphasizing terms that are discriminative for a specific document while penalizing terms common across the corpus (Zhang et al., 2015). Gifari et al. (2021) demonstrated that enhanced TF-IDF weighting schemes could significantly improve sentiment analysis performance.

Research Gap

Although there is a wealth of research on sentiment analysis pertaining to social media data, there is a notable scarcity of studies that focus specifically on AI-generated content or conversations surrounding AI technologies such as ChatGPT. Additionally, most recent research emphasizes deep learning approaches, with less attention paid to optimizing conventional machine learning methods for specific domains. This study fills these gaps by testing traditional machine learning classifiers on ChatGPT-related tweets, comparing different ways to extract features, and identifying the optimal combinations of classifiers and feature extraction methods. By focusing on conventional methods, this research offers practical insights for sentiment analysis in resource-constrained environments while contributing to the understanding of public sentiment regarding emerging AI technologies.

Methodology

Research Design

This study employs a mixed-methods research design, combining qualitative sentiment analysis of ChatGPT-related tweets with quantitative evaluation of machine learning classifiers. This approach enables both deep exploration of public sentiments and systematic assessment of classification techniques. The primary focus is on evaluating the performance of different machine learning classifiers (Logistic Regression, Support Vector Machines, Random Forest, and Naive Bayes) on sentiment classification tasks, particularly addressing the challenge of class imbalance in the dataset.

Data Collection and Sources

The study utilizes a publicly available Twitter dataset related to ChatGPT obtained from Kaggle. This dataset comprises tweets generated by users engaging in discussions about the ChatGPT model on Twitter, providing diverse conversations, opinions, and sentiments. Originally containing 478,379 tweets collected from November 30, 2022, to April 8, 2023, the dataset was published by Manisha Bhatt under the title "Tweets On ChatGPT (ChatGPT)." The collection encompasses a substantial timeframe to capture temporal variations in discussions and sentiments. To prepare the data for analysis, we performed thorough cleaning procedures, removing duplicates and non-English tweets, which resulted in a refined dataset of 219,294 tweets. Since the original dataset did not include sentiment classifications, we employed the VADER (Valence Aware Dictionary and Sentiment Reasoner) sentiment analysis tool to annotate each tweet with sentiment labels. This automated sentiment analysis process categorized the tweets as positive (47.87%), neutral (34.01%), and negative (18.12%). This distribution reveals a notable class imbalance, with positive tweets accounting for nearly half of the dataset and negative tweets representing less than a fifth of the total. The comprehensive preprocessing and sentiment annotation procedures ensured the dataset was suitable for subsequent feature extraction and machine learning classification tasks.

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).
All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).
https://doi.org/10.64922/jteas.v1i1.19

Data Preprocessing Steps

Text preprocessing is crucial for improving data quality and reducing data mining process complexity. The following preprocessing techniques were applied:

Preprocessing Step	*Description*
Tokenization	We used NLTK's word_tokenize function to break the text into individual tokens according to the Penn Tree Bank's rules for tokenization. We separated words, punctuation marks, and special symbols into their own units during this step so that we could move on to the next step of processing.
Lowercasing	Python's built-in str.lower() method was applied to each token to convert all characters to lowercase. This ensures that identical words appearing in different cases (e.g., "ChatGPT" and "chatgpt") are treated as the same token during feature extraction.
Stopwords Removal	NLTK's standard English stopword list (179 words) was applied to remove high-frequency, low-information words such as "a", "an", and "the". This list was expanded with terms specific to the domain that were found during exploratory analysis. For example, "chatgpt" was added in some cases where its near-universal frequency made it not useful for sentiment classification.
Normalization	A custom normalisation function was implemented utilising Python's RE module, a built-in library used for working with regular expressions (Regex) to address Twitter-specific textual patterns. Repeated characters were collapsed to a maximum of two occurrences (e.g., "gooood" to "good") using the regex pattern r'(.)\1{2,}'. Contracted forms were expanded using a manually curated lookup dictionary (e.g., "can't" to "cannot" and "it's" to "it is").
Noise Removal	The re module in Python was utilised to eliminate noise elements that are characteristic of Twitter data. Using the pattern r'http\S+', we removed URLs; using r'@\w+', we removed user mentions (e.g. @username). Using r'#', we removed hashtag symbols while retaining the associated word. Finally, using r'[^a-zA-Z\s]', we removed non-ASCII characters and numbers. Markers for retweets were also eliminated, as they do not convey any sentiment signal.
Lemmatization	We used NLTK's WordNetLemmatizer with part-of-speech (POS) tags to ensure that lemmatisation accounted for context. We first used NLTK's pos_tag function to give POS tags, then we mapped them to WordNet POS constants (NOUN, VERB, ADJ, ADV) before lemmatisation. For example, "running" was lemmatised to "run" and "better" to "good", reducing the vocabulary size while preserving semantic content.

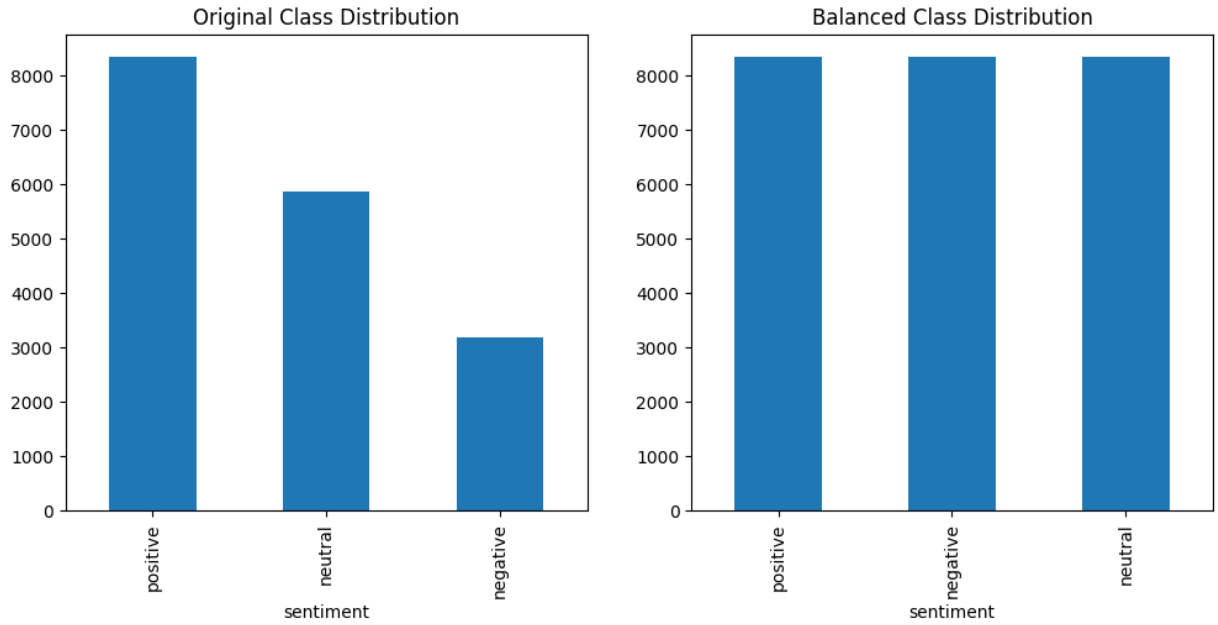
Handling Class Imbalance

It was imperative to rectify the substantial class imbalance in the sentiment distribution in order to create resilient models. The left panel of Figure 1 shows how different sentiment classes are in the original dataset. Positive tweets accounted for 47.87% of the data, neutral tweets for 34.01%, and negative tweets for only 18.12% of the total. This imbalance could potentially bias the machine learning models toward the majority class (positive), leading to poor predictive performance for the minority classes (neutral and negative).

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).
All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).
https://doi.org/10.64922/jteas.v1i1.19

Figure 1

Unbalanced and Balanced Class Distribution



We used two different methods to see how this imbalance affected the performance of the model. We first used the original imbalanced dataset, as is, to establish a baseline for each classifier's performance. This approach allowed us to assess how well the models performed under natural data distribution conditions. Second, we used the RandomOverSampler from the imbalanced-learn (imblearn) library to fix the problem of imbalance. RandomOverSampler was selected over synthetic methods such as SMOTE because our dataset consisted of raw text represented as token sequences rather than continuous feature vectors; SMOTE's interpolation between feature vectors is designed for numerical data and produces semantically meaningless synthetic texts in the discrete token space. RandomOverSampler avoids this limitation by duplicating existing minority-class samples, preserving authentic linguistic patterns in the training data. This technique systematically oversampled the minority classes (neutral and negative) in the training set, creating a more balanced distribution with approximately equal representation across all sentiment categories, as shown in Figure 1 (right panel). The balanced distribution ensured that each sentiment class had similar representation during model training, potentially improving the classifiers' ability to recognize patterns in underrepresented categories. By comparing model performance with and without balancing, we could quantitatively assess the impact of class imbalance on classification accuracy and determine whether addressing this imbalance significantly enhanced the models' discriminative capabilities across all sentiment categories.

Feature Extraction and Machine Learning Classification Techniques

Two primary feature extraction techniques were employed in this study. The Bag of Words model, while disregarding grammatical structure and word order, creates a mathematical representation of text by constructing document vectors based on term frequencies, enabling computational processing of textual data (Kowsari et al., 2019). Implementation used scikit-learn's CountVectorizer. Term Frequency-Inverse Document Frequency (TF-IDF) offers a more nuanced approach by weighing words based on their frequency in a document relative to their occurrence in the entire corpus. It emphasizes words frequent in a document but rare across the corpus, thus distinguishing their importance and offering improved discriminatory power (Salton et al., 1988). Implementation used scikit-learn's TfidfVectorizer.

For both BoW and TF-IDF vectorization, we employed n-gram ranges of (1,2) to capture both individual terms and meaningful bigrams, with a maximum feature count of 5,000 to balance model complexity with

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).
All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).
https://doi.org/10.64922/jteas.v1i1.19

computational efficiency. For TF-IDF specifically, we utilized sublinear term frequency scaling (tf = 1 + log(tf)) to mitigate the impact of high-frequency terms, and applied L2 normalization to account for document length variation. Stop words were removed using the NLTK English stop words list, supplemented with domain-specific non-informative terms identified during exploratory data analysis. These vectorization parameters were determined through preliminary grid search optimization on a validation subset to minimise cross-entropy loss.

For the classification tasks, we evaluated four supervised machine learning algorithms, each selected for its distinct learning characteristics and suitability for high-dimensional text data. Logistic Regression (LR) models the conditional probability of a class label given the input features using a logistic (sigmoid) function. In the multiclass setting, scikit-learn's one-vs-rest (OvR) strategy was employed, fitting one binary classifier per sentiment class. The model outputs calibrated class probabilities, making it interpretable and efficient for sparse text feature matrices. Support Vector Machines (SVM) seek the optimal separating hyperplane that maximises the geometric margin between sentiment classes in the transformed feature space. The linear kernel variant (LinearSVC) was used given the high dimensionality of the TF-IDF and BoW representations; kernel functions such as RBF were evaluated during grid search but showed no benefit over the linear kernel at this dimensionality. The SVM optimisation problem is expressed as: min (w,b) $\|w\|^2/2 + C\sum\xi_i$ subject to: $y_i(w \cdot x_i + b) \geq 1 - \xi_i$ and $\xi_i \geq 0$, where w is the weight vector, b the bias, ξ_i the slack variables, and C the regularisation parameter balancing margin width against misclassification penalty. Random Forest (RF) constructs an ensemble of decision trees, each trained on a bootstrap sample of the data with a random subset of features considered at each split. Final class predictions are determined by majority vote across all trees. This bagging approach reduces variance and improves robustness to overfitting, particularly beneficial when feature interactions are complex. Naive Bayes (NB) applies Bayes' theorem under the assumption of conditional feature independence: $P(y|x) \propto P(y)\prod P(x_i|y)$. The Multinomial variant was used, appropriate for discrete frequency counts produced by BoW and TF-IDF vectorisation. The complete modelling pipeline for each configuration was: (1) apply preprocessing and vectorisation to produce the document-term matrix, (2) optionally apply RandomOverSampler to the training split only, (3) fit the classifier on the training set using optimal hyperparameters from grid search, and (4) evaluate on the held-out test set. Implementation was executed via the scikit-learn Python library, with dataset partitioning into 80% training and 20% testing sets using stratified sampling to preserve the sentiment distribution. Stratification was implemented via scikit-learn's train_test_split with the stratify parameter set to the sentiment label array, which guarantees that each split contains approximately the same proportion of positive (47.87%), neutral (34.01%), and negative (18.12%) samples as the full dataset. This was verified by computing class frequencies in both splits before any balancing was applied. This resulted in 16 model combinations (4 classifiers × 2 feature extraction methods × 2 balancing approaches).

Hyperparameter optimization was conducted for each classifier to ensure optimal performance. For Logistic Regression, we tuned the regularisation parameterwe tuned the regularisation parameter (C values from 0.01 to 100) and penalty type (L1 or L2). For SVM, we optimized the regularisation parameter C and kernel coefficient gammape (L1 or L2). For SVM, we optimized the regularisation parameter C and kernel coefficient gamma using grid search. Random Forest optimization included maximum depth, number of estimators, and minimum samples per leaf. For each classifier, 5-fold cross-validation was used during hyperparameter tuning to avoid overfitting. Final models employed the optimal parameters identified during this process: Logistic Regression (C=1.0, 'l2' penalty), SVM (C=10, gamma=0.1), Random Forest (max_depth=100, n_estimators=200, min_samples_leaf=1), and default parameters for Naive Bayes, as it showed limited sensitivity to parameter adjustme

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).
All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).
https://doi.org/10.64922/jteas.v1i1.19

Table 1

Optimal Hyperparameters for Each Classifier Based on Grid Search with 5-Fold Cross-Validation

Classifier	*Parameter*	*Optimal Value*
Logistic Regression	C (regularization strength)	1.0
Logistic Regression	Penalty	L2
SVM	C	10
SVM	Gamma	0.1
SVM	Kernel	linear
Random Forest	Number of estimators	200
Random Forest	Max depth	100
Random Forest	Min samples leaf	1
Naive Bayes	Alpha (smoothing)	1.0

Model efficacy was quantified through multiple performance metrics including accuracy (ratio of correct predictions to total samples), precision (the ratio of true positives to predicted positives,the ratio of true positives to predicted positives, indicating class label discrimination capability), recall (the ratio of true positive predictions to total actual positive samples,ng class label discrimination capability), recall (the ratio of true positive predictions to total actual positive samples, reflecting comprehensive positive instance capture), F1-score (the harmonic mean of precision and recall, balancing false positive and negative trade-offs), and confusion matrices (tabular performance representations delineating true positives, true negatives, false pos.

Results

Performance Metrics of Different Classifiers

The machine learning classifiers were evaluated using various performance metrics to determine their effectiveness in sentiment classification of ChatGPT-related tweets, with particular attention to the impact of balancing techniques on class imbalance.

Table 2

Comparative Performance Evaluation of Machine Learning Classifiers across Feature Extraction Techniques and Balancing Approaches

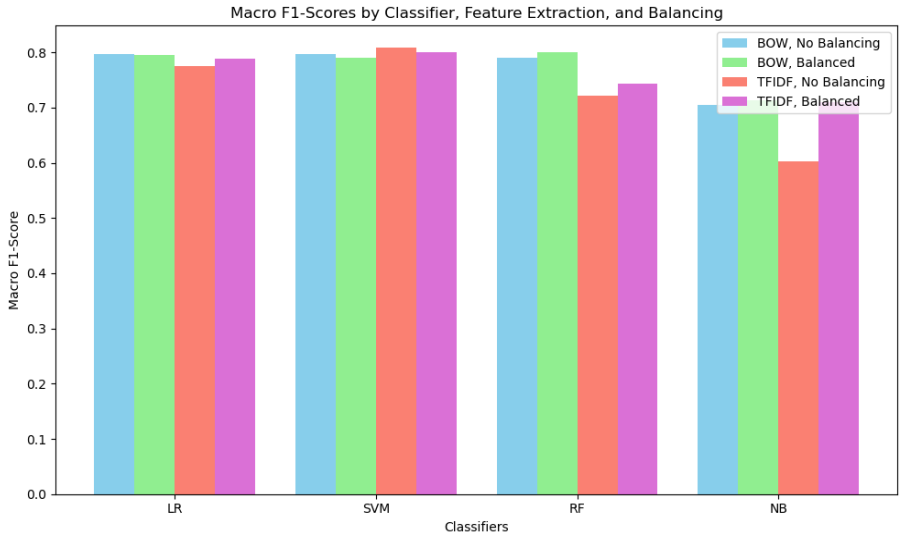
Classifier	*Feature Extraction*	*Balancing*	*Accuracy*	*Macro Precision*	*Macro Recall*	*Macro F1-score*
Logistic Regression	BOW	No	0.832834	0.815690	0.790448	0.797821
Linear SVM	BOW	No	0.827308	0.803249	0.792906	0.796612
Random Forest	BOW	No	0.833755	0.837913	0.775755	0.789648
Naive Bayes	BOW	No	0.740502	0.732893	0.689386	0.704409
Logistic Regression	BOW	Yes	0.822933	0.795230	0.797132	0.794784
Linear SVM	BOW	Yes	0.818328	0.789457	0.792197	0.790202
Random Forest	BOW	Yes	0.835598	0.816943	0.794567	0.800796
Naive Bayes	BOW	Yes	0.742805	0.712837	0.722661	0.713866
Logistic Regression	TFIDF	No	0.819940	0.818258	0.760225	0.774889
Linear SVM	TFIDF	No	0.840203	0.826191	0.798169	0.808065
Random Forest	TFIDF	No	0.786783	0.804785	0.710014	0.721858
Naive Bayes	TFIDF	No	0.686392	0.772614	0.583844	0.602475
Logistic Regression	TFIDF	Yes	0.815565	0.786360	0.793811	0.789044

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).
All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).
https://doi.org/10.64922/jteas.v1i1.19

Linear SVM	TFIDF	Yes	0.826618	0.798608	0.805065	0.801247
Random Forest	TFIDF	Yes	0.788625	0.772708	0.735686	0.742854
Naive Bayes	TFIDF	Yes	0.738890	0.709266	0.719249	0.710586

Figure 2

Macro F1-Scores of Classifiers under Different Feature Extraction Methods (BoW vs. TF-IDF) and Balancing Strategies



Based on the comparative analysis presented in Table 2, several key observations emerge:

1. Linear SVM with TF-IDF vectorisation without balancing achieved the highest overall accuracy (84.02%), while also maintaining strong macro F1-score (80.81%).

2. When considering macro F1-score, which is particularly important for imbalanced datasets, Linear SVM with TF-IDF and balancing techniques (80.12%) and Linear SVM with TF-IDF without balancing (80.81%) performed best, suggesting that SVM effectively captures sentiment patterns across classes.

3. Random Forest with BOW and balancing techniques achieved a competitive macro F1-score (80.08%), indicating the effectiveness of ensemble methods when proper class balancing is applied.

4. Naive Bayes consistently underperformed compared to other classifiers across both feature extraction techniques and balancing approaches, with particularly poor performance when using TF-IDF without balancing (macro F1-score of 60.25%).

5. For most classifiers, balancing techniques improved recall measures while sometimes slightly reducing precision, suggesting more effective identification of minority class instances.

Cross-validation Results

Five-fold cross-validation was performed to assess the stability and generalization performance of the classifiers, with particular attention to their performance across different data splits.

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).
All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).
https://doi.org/10.64922/jteas.v1i1.19

Table 3

5-Fold Cross-validation Results for SVM with TF-IDF (Best Performing Configuration)

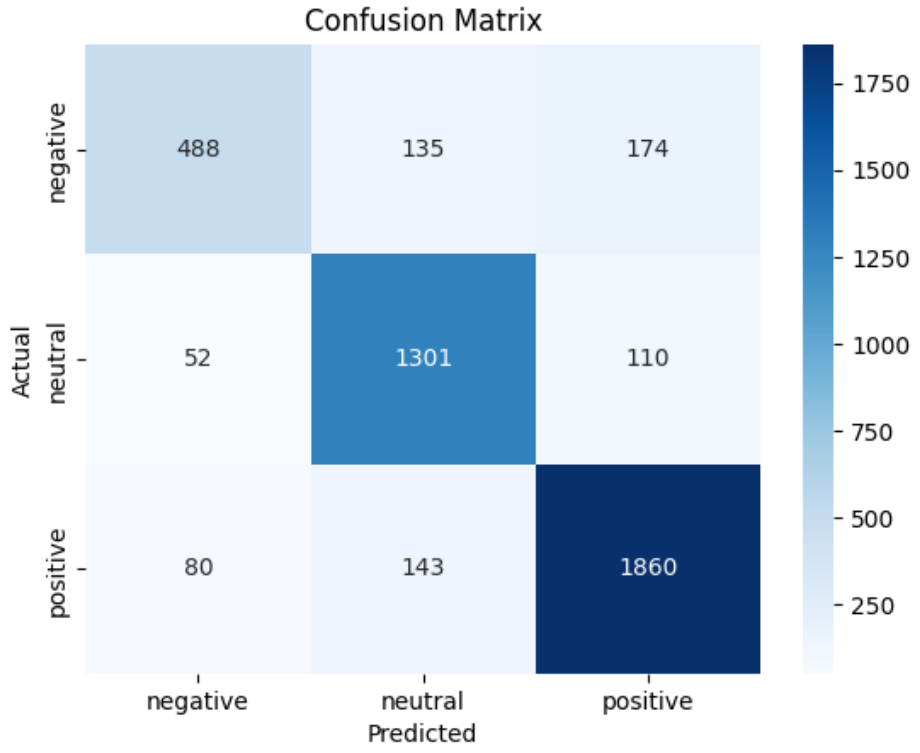
Run	*Without Balancing*	*With Balancing*
	Macro F1-score	**Macro F1-score**
Run 1	0.76	0.78
Run 2	0.77	0.79
Run 3	0.78	0.80
Run 4	0.77	0.78
Run 5	0.78	0.79
Average	**0.77**	**0.79**

The cross-validation results confirm the superior performance of the balanced approach, with consistently higher F1-scores across all folds, demonstrating the robustness of this method for handling the class imbalance in the dataset.

Confusion Matrix Analysis

Figure 3

Confusion Matrix for SVM with TF-IDF and Balancing (Best Model)



© [2026] Journal of Technology Education and Applied Sciences (JTEAS).
All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).
https://doi.org/10.64922/jteas.v1i1.19

The confusion matrix provides deeper insights into the classification performance of the best model (SVM with TF-IDF and balancing). The matrix reveals:

1. The model achieves high accuracy in correctly identifying positive sentiments (top-left cell), with minimal misclassification between positive and negative categories.

2. The model demonstrates reasonable performance in classifying neutral sentiments, though there is some confusion between neutral and positive categories.

3. Negative sentiments show the lowest classification accuracy among the three categories, with some misclassification as neutral, indicating the challenging nature of distinguishing between negative and neutral expressions in social media text.

These patterns suggest that while the model performs well overall, there remains room for improvement in discriminating between neutral and negative sentiment expressions.

Statistical Significance Analysis

To validate the comparative performance of classifiers beyond descriptive metrics, we conducted McNemar's test for pairwise comparison of classifier outputs. This non-parametric test evaluates whether the disagreements between classifiers are statistically significant, providing greater confidence in observed performance differences.

Table 4

McNemar's Test Results for Pairwise Classifier Comparison (p-values and Effect Sizes)

Classifier Pair	*Without Balancing (p-value; Cohen's h)*	*With Balancing (p-value; Cohen's h)*
SVM vs. RF	0.003*; h = 0.31 (small)	0.007*; h = 0.28 (small)
SVM vs. LR	0.015*; h = 0.24 (small)	0.022*; h = 0.21 (small)
SVM vs. NB	<0.001*; h = 0.74 (medium)	<0.001*; h = 0.71 (medium)
RF vs. LR	0.362; h = 0.07 (negligible)	0.189; h = 0.09 (negligible)
RF vs. NB	<0.001*; h = 0.68 (medium)	<0.001*; h = 0.65 (medium)
LR vs. NB	<0.001*; h = 0.62 (medium)	<0.001*; h = 0.59 (medium)

**Statistically significant at α = 0.05. Cohen's h effect size: negligible < 0.20, small = 0.20–0.49, medium = 0.50–0.79, large ≥ 0.80 (Cohen, 1988).*

The results confirm statistically significant performance differences between SVM and all other classifiers, with p-values below the significance threshold (α = 0.05). However, the difference between Random Forest and Logistic Regression was not statistically significant, suggesting comparable performance despite different algorithmic approaches.

Insights into Sentiment Dynamics

Word frequency analysis provided insights into prevalent topics and sentiment dynamics in ChatGPT-related tweets. The word cloud visualization and n-gram analysis revealed key terms and phrases dominating the discourse.

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).
All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).
https://doi.org/10.64922/jteas.v1i1.19

Table 5

Top N-gram Representation of Tweets by Sentiment Category (Top 3)

Sentiment	*Unigram*	*Count*	*Bigram*	*Count*	*Trigram*	*Count*
Positive	chatgpt	18452	(amazing, chatgpt)	786	(chatgpt, really, amazing)	217
	amazing	4231	(chatgpt, impressive)	645	(future, ai, exciting)	189
	helpful	3128	(love, chatgpt)	542	(chatgpt, helps, me)	156
Neutral	chatgpt	15783	(asked, chatgpt)	1047	(asked, chatgpt, question)	287
	ai	5324	(chatgpt, says)	821	(chatgpt, can, do)	183
	using	2915	(using, chatgpt)	761	(how, to, chatgpt)	149
Negative	chatgpt	7222	(chatgpt, wrong)	423	(chatgpt, not, working)	145
	wrong	1652	(chatgpt, failed)	319	(ai, taking, jobs)	122
	problem	1329	(ai, dangerous)	298	(chatgpt, giving, misinformation)	105

Figure 4

Word-cloud representation of ChatGPT-related tweets.

The word cloud was generated using Python's WordCloud library (v1.8.1) applied to the full preprocessed corpus after stopword removal. Word size is proportional to term frequency across all tweets; terms appearing fewer than 50 times were excluded. The same NLTK English stopword list used during preprocessing was applied, supplemented with the token "chatgpt" to prevent it from dominating the visualization given its near-universal occurrence across all sentiment classes.

The n-gram analysis by sentiment category reveals distinct patterns in how users discuss ChatGPT:

1. **Positive sentiments** are characterized by expressions of admiration ("amazing," "impressive") and utility ("helpful"), suggesting widespread appreciation for ChatGPT's capabilities.

2. **Neutral sentiments** focus on usage patterns and capabilities, with phrases like "asked chatgpt" and "using chatgpt" indicating functional interactions without strong emotional valence.

3. **Negative sentiments** center on performance issues ("wrong," "failed") and broader concerns about AI ("dangerous," "taking jobs"), highlighting areas of user frustration and societal anxiety.

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).
All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).
https://doi.org/10.64922/jteas.v1i1.19

These patterns offer valuable insights into the multifaceted nature of public perception toward ChatGPT, spanning from enthusiastic adoption to critical scepticism.

Discussion

Interpretation of Classifier Performance

The performance metrics reveal significant insights into the effectiveness of different machine learning classifiers for sentiment analysis of ChatGPT-related tweets. Linear SVM with TF-IDF vectorisation emerged as the most effective approach, achieving the highest accuracy (84.02%) and macro F1-score (80.81%) without balancing.

The strong performance of Random Forest with BOW and balancing techniques (80.08% macro F1-score) demonstrates the potential of ensemble learning approaches in sentiment classification, particularly when class imbalance is properly addressed. This finding resonates with previous studies that have highlighted ensemble methods' effectiveness in improving classification performance through aggregating multiple decision boundaries (Hasan et al., 2018).

Logistic Regression also demonstrated competitive performance across different configurations, achieving 79.78% macro F1-score with BOW without balancing. This supports existing literature on Logistic Regression's effectiveness in text classification due to its interpretability and ability to model probabilistic relationships between features and sentiment categories (Tan et al., 2023).

Naive Bayes consistently underperformed compared to other classifiers, likely due to its strong assumption of feature independence, which may not hold in complex social media text data where contextual relationships between words significantly influence sentiment expressions (Zainuddin & Selamat, 2015).

The superior performance of SVM with TF-IDF can be theoretically explained through the lens of margin maximization in high-dimensional spaces. Unlike Naive Bayes, which assumes feature independence, SVM determines an optimal separating hyperplane that maximises the geometric margin between sentiment classes, as formalised in the optimisation problem presented in Section 3.4. This margin-maximisation objective, combined with the L2-regularised linear kernel, allows the model to handle the high-dimensional sparsity of TF-IDF representations without overfitting, while the TF-IDF weighting simultaneously elevates discriminative sentiment-carrying terms and suppresses high-frequency vocabulary with limited sentiment value. Together these properties explain the consistent empirical superiority observed across experimental configurations.

The cross-validation results further validate these observations, showing consistent performance across multiple iterations for SVM with TF-IDF. The improved performance with balancing techniques (average macro F1-score of 0.79 compared to 0.77 without balancing) underscores the importance of addressing class imbalance in sentiment analysis tasks, particularly for ensuring equitable performance across all sentiment categories.

Comparison of Feature Extraction Techniques and Balancing Approaches

The comparison between BoW and TF-IDF vectorizers revealed interesting patterns in their impact on classifier performance. For SVM, TF-IDF consistently outperformed BoW, achieving higher accuracy and macro F1-scores. This suggests that the term weighting approach of TF-IDF, which emphasises distinctive terms while downplaying common ones, effectively captures sentiment-indicative features in ChatGPT-related tweets.

Conversely, Random Forest performed better with BoW, particularly when combined with balancing techniques. This pattern suggests that ensemble methods like Random Forest benefit from the straightforward frequency-based representation of BoW, potentially because the ensemble approach already captures complex feature interactions without requiring the additional weighting provided by TF-IDF.

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).
All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).
https://doi.org/10.64922/jteas.v1i1.19

The impact of balancing techniques varied across classifiers and feature extraction methods. For SVM with TF-IDF, balancing slightly reduced overall accuracy (from 84.02% to 82.66%) but improved recall measures, resulting in more balanced performance across sentiment categories. This trade-off between overall accuracy and balanced class performance highlights the importance of selecting evaluation metrics aligned with analytical objectives, particularly in applications where minority class detection is crucial.

Insights into Sentiment Dynamics

The n-gram analysis by sentiment category provides valuable insights into the discourse patterns surrounding ChatGPT on Twitter. The dominance of positive sentiment expressions suggests generally favourable reception of the technology, with terms like "amazing," "impressive," and "helpful" reflecting appreciation for ChatGPT's capabilities. This positive sentiment aligns with broader trends in early technology adoption, where novel capabilities often generate initial enthusiasm. Neutral sentiments primarily focused on functional aspects and use cases, suggesting widespread experimentation with and exploration of ChatGPT's capabilities. The prevalence of phrases like "asked chatgpt" and "using chatgpt" indicates integration of the technology into users' information-seeking and problem-solving processes. Negative sentiments, while less prevalent, centered on specific concerns: performance limitations ("wrong," "failed"), accuracy issues ("misinformation"), and broader societal implications ("ai taking jobs"). These concerns mirror ongoing debates about AI technologies' reliability, trustworthiness, and socioeconomic impact, reflecting multifaceted apprehensions about rapidly evolving AI capabilities.

When viewed through the lens of technology adoption theories, these sentiment patterns align with early-stage technology diffusion characteristics. The predominance of positive sentiment reflects reactions typical of early adopters, who tend to emphasise potential benefitstend to emphasise potential benefits while tolerating limitations. The specific positive expressions focusing on capability ("amazing," "impressive") rather than usability suggest the technology is still perceived as novel rather than fully integrated into routine workflows. Meanwhile, negative sentiments centering on accuracy and reliability concerns rather than fundamental objections to the technology itself indicate the "performance gap" often experienced in early deployment phases of innovations. This theoretical contextualisation suggests ChatGPT was, during the study period, navigating the transition from early adopter enthusiasm toward broader acceptance, with sentiment patterns reflecting this transitional position.

Implications and Applications

The findings from this study have significant implications for multiple stakeholders in the AI ecosystem. For developers and organizations deploying conversational AI systems, understanding sentiment patterns provides valuable feedback on user experiences, highlighting both strengths to leverage and limitations to address. The identification of specific negative sentiment triggers (e.g., accuracy concerns) offers actionable insights for targeted improvements.

For researchers in sentiment analysis, our results demonstrate the continued relevance of conventional machine learning approaches, particularly when computational efficiency is a consideration. The superior performance of Linear SVM with TF-IDF, achieving over 84% accuracy without deep learning architectures, suggests that well-optimized traditional methods remain valuable tools in sentiment analysis applications.

For social media analysts and market researchers, the sentiment categorization and n-gram analysis methodologies demonstrated here provide a framework for monitoring public perception of emerging technologies. The approach can be adapted to track sentiment evolution over time, assess response to specific AI capabilities or limitations, and identify emerging concerns or appreciation points in public discourse.

Limitations and Future Research

While this study provides valuable insights into sentiment analysis of ChatGPT-related tweets, several limitations should be acknowledged. The dataset, while substantial, may not capture the full diversity of discussions about ChatGPT across different languages, platforms, and timeframes. Additionally, inherent

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).
All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).
https://doi.org/10.64922/jteas.v1i1.19

biases in Twitter's user demographics may skew sentiment distributions compared to broader public opinion. Methodologically, our approach focused on conventional machine learning classifiers without exploring deep learning alternatives such as transformers or bidirectional LSTMs, which might offer performance improvements at the cost of greater computational requirements. Furthermore, the three-category sentiment classification (positive, neutral, negative) may not capture more nuanced emotional expressions such as excitement, frustration, surprise, or skepticism.

Future research could address these limitations by:

1. Expanding the dataset to include multiple languages and platforms, employing more sophisticated data collection techniques, and conducting longitudinal studies to capture temporal variations in sentiment.

2. Implementing more granular sentiment categorisation beyond the positive-neutral-negative framework, potentially incorporating emotion detection to identify specific affective states like curiosity, confusion, or amazement.

3. Exploring hybrid approaches that combine conventional machine learning with selective deep learning components, balancing performance improvements with computational efficiency.

4. Investigating the relationship between sentiment expressions and specific ChatGPT capabilities or limitations, providing more targeted insights for technology improvement.

5. Conducting comparative analyses of sentiment toward different AI models and technologies to contextualize ChatGPT-specific findings within broader AI perception trends.

Conclusion

This study evaluated the effectiveness of machine learning classifiers for sentiment analysis of ChatGPT-related tweets, comparing different feature extraction techniques and assessing the impact of balancing approaches on classification performance. The results demonstrate that Linear SVM with TF-IDF vectorization provides the most effective approach for sentiment classification in this context, achieving 84.02% accuracy without balancing and maintaining strong performance across sentiment categories with balancing techniques (80.12% macro F1-score).

The analysis of sentiment-specific discourse patterns revealed predominantly positive public reception of ChatGPT, with appreciation for its capabilities alongside specific concerns about performance limitations and broader AI implications. These findings contribute to understanding sentiment dynamics in AI-related social media conversations and demonstrate the practical efficacy of conventional machine learning approaches for sentiment classification in resource-constrained environments.

Beyond the practical implications, our findings provide empirical insights into sentiment analysis methodologies. The consistent superiority of discriminative models (particularly SVM) over generative approaches (Naive Bayes) across multiple experimental configurations provides empirical support for theoretical predictions regarding their relative strengths in high-dimensional text classification tasks. Furthermore, the differential impact of feature weighting techniques across classifier types suggests a complex interaction between feature representation and learning algorithm that warrants further theoretical exploration. These results challenge simplistic views of classifier selection and underscore the importance of considering the theoretical alignment between data characteristics, feature extraction techniques, and classification algorithms when developing sentiment analysis systems.

By providing insights into both methodological effectiveness and substantive sentiment patterns, this research offers valuable contributions to sentiment analysis practice and AI technology perception research. While acknowledging limitations in dataset diversity and classification granularity, our approach achieves high

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).
All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).
https://doi.org/10.64922/jteas.v1i1.19

accuracy with traditional classifiers, suggesting continued relevance for conventional methods alongside emerging deep learning approaches in sentiment analysis applications.

Acknowledgements

The authors would like to acknowledge all individuals whose constructive comments and suggestions helped to improve the quality of this paper. No external funding was received for the completion of this work.

Author Contribution Statement

Conceptualisation: PCA (lead), EKM (equal)

Data analysis: PCA (lead), SKA (equal), ROA (equal)

Methodology/Frameworks: SKA (equal), NP (equal), DY (equal)

Datasets and Tools: SKA (lead), ROA (equal)

Initial write-up: PCA (lead), SKA (equal), OKB (supporting)

Revision: PCA (equal), EKM (equal), OKB (equal), EAF (equal)

Empirical Validation: PCA (lead), ROA (supporting), SKA (supporting), DY (supporting), NP (supporting), OKB (supporting)

Quality Assessment: PCA (lead), ROA (equal), SKA (supporting), OKB (supporting)

Conflict of Interest Statement

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

AI Disclosure Statement

No Artificial Intelligence (AI) tools were used in the preparation of this manuscript.

References

Aziz, R. H. H., & Dimililer, N. (2020). Twitter sentiment analysis using an ensemble weighted majority vote classifier. In *2020 International Conference on Advanced Science and Engineering (ICOASE)*. https://doi.org/10.1109/icoase51841.2020.9436590

Bordoloi, M., & Biswas, S. K. (2023). Sentiment analysis: A survey on design framework, applications and future scopes. *Artificial Intelligence Review*, *56*(11), 12505–12560. https://doi.org/10.1007/s10462-023-10442-2

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates. https://doi.org/10.4324/9780203771587

Dey, A., Jenamani, M., & Thakkar, J. J. (2018). Senti-N-Gram: An N-gram lexicon for sentiment analysis. *Expert Systems with Applications*, *103*, 92–105. https://doi.org/10.1016/j.eswa.2018.03.004

Gehrmann, S., Dernoncourt, F., Li, Y., Carlson, E. T., Wu, J. T., Welt, J., John, F. J., Moseley, E. T., Grant, D. W., Tyler, P. D., & Celi, L. A. (2017). *Comparing rule-based and deep learning models for patient phenotyping*. arXiv. https://doi.org/10.48550/arxiv.1703.08705

Gifari, M. K., Lhaksmana, K. M., & Dwifebri, P. M. (2021). Sentiment analysis on movie review using ensemble stacking model. In *2021 International Conference Advancement in Data Science, E-learning and Information Systems (ICADEIS)*. https://doi.org/10.1109/icadeis52521.2021.9702088

Guo, B., Zhang, C., Wang, Z., Guo, J., & Yu, Z. (2023). *How close is ChatGPT to human experts? Comparison corpus, evaluation, and detection* (arXiv:2301.07597). arXiv. https://doi.org/10.48550/arXiv.2301.07597

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).

All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).

https://doi.org/10.64922/jteas.v1i1.19

Hasan, A., Moin, S., Karim, A., & Shamshirband, S. (2018). Machine learning-based sentiment analysis for Twitter accounts. *Mathematical and Computational Applications*, *23*(1), 11. https://doi.org/10.3390/mca23010011

Kowsari, K., Jafari Meimandi, K., Heidarysafa, M., Mendu, S., Barnes, L., & Brown, D. (2019). Text classification algorithms: A survey. *Information*, *10*(4), 150. https://doi.org/10.3390/info10040150

Leiter, C., Zhang, R., Chen, Y., Belouadi, J., Larionov, D., Fresen, V., & Eger, S. (2023). *ChatGPT: A meta-analysis after 2.5 months* (arXiv:2302.13795). arXiv. https://doi.org/10.48550/arXiv.2302.13795

Liu, B. (2012). Sentiment analysis and opinion mining. *Synthesis Lectures on Human Language Technologies*, *5*(1), 1–167. https://doi.org/10.2200/S00416ED1V01Y201204HLT016

Mercha, E. M., & Benbrahim, H. (2023). Machine learning and deep learning for sentiment analysis across languages: A survey. *Neurocomputing*, *531*, 195–216. https://doi.org/10.1016/j.neucom.2023.02.015

Pang, B., Lee, L., & Vaithyanathan, S. (2002). Thumbs up? Sentiment classification using machine learning techniques. In *Proceedings of the 2002 Conference on Empirical Methods in Natural Language Processing (EMNLP 2002)* (pp. 79–86). https://doi.org/10.3115/1118693.1118704

Parveen, H., Hada, B., Lalwani, P., & Gupta, M. (2023). Sentiment analysis on social media using machine learning techniques: A comprehensive review. *Multimedia Tools and Applications*, *82*(14), 21591–21630. https://doi.org/10.1007/s11042-022-14160-9

Sallam, M., Salim, N. A., Barakat, M., & Al-Tammemi, A. B. (2023). ChatGPT applications in medical, dental, pharmacy, and public health education: A descriptive study highlighting the advantages and limitations. *Narra J*, *3*(1), e103. https://doi.org/10.52225/narra.v3i1.103

Sravya, K., Sowmya, G., Yamini, P., Anusha, P., & Krishna, P. S. (2021). Sentiment analysis on Twitter. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3920078

Tan, K. L., Lee, C. P., & Lim, K. M. (2023). A survey of sentiment analysis: Approaches, datasets, and future research. *Applied Sciences*, *13*(7), 4550. https://doi.org/10.3390/app13074550

Tan, K. L., Lee, C. P., Lim, K. M., & Anbananthen, K. S. M. (2022). Sentiment analysis with ensemble hybrid deep learning model. *IEEE Access*, *10*, 103694–103704. https://doi.org/10.1109/access.2022.3210182

Wankhade, M., Rao, A. C. S., & Kulkarni, C. (2022). A survey on sentiment analysis methods, applications, and challenges. *Artificial Intelligence Review*, *55*(7), 5731–5780. https://doi.org/10.1007/s10462-022-10144-1

Zainuddin, N., & Selamat, A. (2015). Sentiment analysis using support vector machine. In *2014 International Conference on Computer, Communications, and Control Technology (I4CT)* (pp. 333–337). https://doi.org/10.1109/i4ct.2014.6914200

Zhang, X., Zhao, J., & LeCun, Y. (2015). Character-level convolutional networks for text classification. In *Advances in Neural Information Processing Systems* (Vol. 28). https://doi.org/10.48550/arXiv.1509.01626

Zhang, Y., & Wallace, B. C. (2015). *A sensitivity analysis of (and practitioners' guide to) convolutional neural networks for sentence classification*. arXiv. https://doi.org/10.48550/arxiv.1510.03820

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).

All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).

https://doi.org/10.64922/jteas.v1i1.19

Journal of Technology Education and Applied Sciences

USTED Press

ISSN: 3093-4796 (Online) Journal homepage: https://journal.aamusted.edu.gh/index.php/jteas/index

RESEARCH ARTICLE

Anti-Termite Efficacy of Selected Leaf Extracts on Pycnanthus Angolensis Wood: A Comparative Study

Kwaku Antwi[1], Michael Darko Minta*[2], Francis Kofi Bih[3], Mark Bright Donkoh[4], Salimatu Salifu Umar[5]

[12345]Department of Wood Science and Technology Education, AAMUSTED, Kumasi, Ghana.
*Corresponding Author: [michaelkwabenadarko@gmail.com]

To cite this article: Antwi, K., Minta, M. D., Bih, F. K., Donkoh, M. B., & Umar, S. S. (2026). Anti-termite efficacy of selected leaf extracts on Pycnanthus angolensis wood: A comparative study. Journal of Technology Education and Applied Sciences, 1(1), 18–31. https://doi.org/10.64922/jteas.v1i1.85

ABSTRACT

Wood is widely used in construction and manufacturing but is highly prone to biodeterioration, particularly in humid tropical regions. Conventional preservatives such as chromated copper arsenate (CCA) are effective but pose environmental and health risks, creating interest in more sustainable and less hazardous alternatives. This study assessed the preservative performance of leaf extracts from *Azadirachta indica* (neem), *Chromolaena odorata* (Siam weed), and Cedrela odorata (Spanish cedar) on *Pycnanthus angolensis*, a non-durable Ghanaian hardwood. Extracts were prepared with ethanol and water at 5%, 10%, and 20% w/v concentrations, and untreated samples served as controls. An in-ground field test was conducted, and termite resistance was evaluated by percentage weight loss. Untreated samples recorded the highest weight loss (93%), confirming the high susceptibility of the wood. All extract treatments reduced degradation to varying degrees. Neem performed best, with an average weight loss of 79%, followed by *Chromolaena odorata* (85%) and *Cedrela odorata* (88%). Ethanol extracts were more effective than water extracts, particularly at higher concentrations, with the most notable relative reductions observed at 20%. ANOVA confirmed significant effects of leaf type, solvent, and concentration. Although the treatments resulted in only partial protection and substantial biodeterioration remained, the findings demonstrate relative improvements compared to untreated wood. Neem leaf extracts, particularly ethanol-based treatments at 20%, showed the strongest comparative performance under the conditions tested. However, the results are based on a 90-day, short-term, single-site field trial and do not assess toxicity, leachability, or long-term environmental behavior.

KEYWORDS

Durability Improvement, Biological Resistance, In-Ground Exposure, Treatment Concentration, Bio-Based Protection

Publication History

Received: 11 November 2025
Revised: 29 May 2026
Accepted: 24 June 2026
Available online: 30 July 2026

Introduction

Wood is an essential material for construction and manufacturing because of its strength, versatility, and aesthetic value (Asmara, 2025). However, its organic composition makes it highly vulnerable to biodeterioration, especially in tropical regions like Ghana where humidity accelerates termite damage.

Wood durability differs across species and depends on natural extractives such as tannins and phenolics. For example, Tectona grandis (Teak) is durable due to its abundant extractives (Da Costa et al., 1958), while *Pycnanthus angolensis* (Otie), a widely used hardwood, is less durable and requires treatment before use, particularly in high-risk environments (Leroy et al., 2023).

Synthetic preservatives like chromated copper arsenate (CCA) and creosote extend wood durability but pose health and environmental risks such as soil and water contamination (Changotra et al., 2024). Plant-based preservatives are increasingly considered safer and biodegradable alternatives, particularly where environmental and health concerns limit synthetic preservatives. Recent studies confirm that plant residue extracts show promise as eco-friendly wood preservatives (Li et al., 2024). Leaf extracts can offer the best source of bioactive compounds during timber processing (Han et al., 2018).

Azadirachta indica (Neem) contains azadirachtin, a compound that disrupts insects' feeding activities (Agbo et al., 2019). *Chromolaena odorata* (Siam weed) is rich in flavonoids and terpenoids, used traditionally for preservation, while *Cedrela odorata* (Spanish cedar) contains limonoids with repellent potential (Dele, 2020). Their abundance potentially makes them suitable candidates for sustainable wood preservation, subject to further validation of large-scale availability and practical extraction feasibility (Datta, 2024; Pasiecznik, 2022).

Despite their potential, research gaps remain. Neem seed extracts have been widely studied in laboratory termite tests, but leaf extracts used as wood preservatives under in-ground field conditions are less explored. Studies on *Chromolaena odorata* focus mainly on antimicrobial properties, not termite resistance on wood (Olawale et al., 2022). Similarly, research on *Cedrela odorata* emphasizes its wood's durability rather than effects of its leaf extracts as preservatives (Ruvalcaba et al., 2017). In-ground field tests, which simulate natural termite exposure, have been applied to species such as *Bambusa vulgaris* and *Azadirachta indica* (Chakurah et al., 2023; Falemara et al., 2018), but not to *Pycnanthus angolensis* with these extracts.

No study has so far compared *Azadirachta indica*, Chro*molaena odorata*, and *Cedrela odorata* leaf extracts extracted with ethanol and water at concentrations of 5%, 10%, and 20% as preservatives on *Pycnanthus angolensis* under in-ground field conditions. This study addresses this gap by objectively evaluating the anti-termite efficacy of these extracts, examining solvent and concentration effects, and identifying the most effective preservative combination. The findings aim to support sustainable wood preservation practices for timber industries especially in Ghana.

Materials and Methods

Research Design

An experimental design was used to evaluate the preservative efficacy of leaf extracts from *Azadirachta indica*, *Chromolaena odorata*, and *Cedrela odorata* on *Pycnanthus angolensis* wood. Extracts were prepared with ethanol and water as solvents at doubling concentrations of 5%, 10%, and 20% weight-to-volume (w/v). Untreated samples served as controls. Each treatment had five replicates for statistical reliability.

Wood Sampling and Preparation

Defect-free wood samples were obtained from the butt sections of three mature trees of *Pycnanthus angolensis*. Lumber sections (50 × 50 × 4000 mm) were sawn, air-dried for 14 days, and later machined at the CBT Wood Laboratory, AAMUSTED.

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).

All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).

https://doi.org/10.64922/jteas.v1i1.85

Specimens were machined and cut to dimensions of 12.5 × 25 × 250 mm according to the modified European standard EN 252, and oven-dried at 105°C to constant weight (W1) as shown in Figure 1 and Figure 2 respectively before treatment.

Figure 1

Machining specimens

Figure 1

Oven drying samples

Plant Material Collection and Preparation

Fresh leaves of *Chromolaena odorata*, A*zadirachta indica*, and *Cedrela odorata* (Figure 3; a, b, c) respectively, were collected, thoroughly washed, and air-dried in a shaded, ventilated area. The dried leaves of *Chromolaena odorata*, A*zadirachta indica*, and *Cedrela odorata* as shown in Figure 4 (a, b, c respectively) were pounded, sieved with a 2 mm net as shown in Figure 5, blended (shown in Figure 6), and re-sieved through a 600 μm mesh. The sieved powders were sealed in airtight zip locks for later use.

Figure 3

Fresh leaves

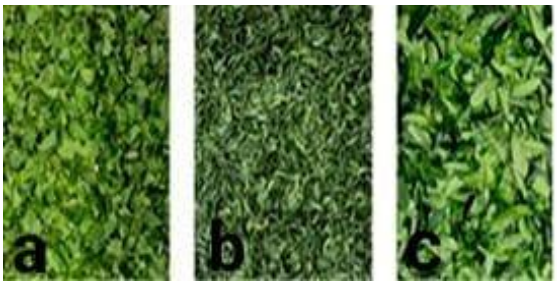
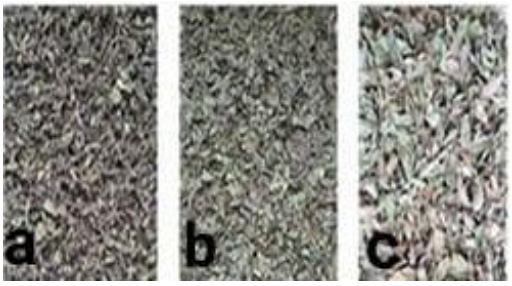


Figure 2

Dried leaves



© [2026] Journal of Technology Education and Applied Sciences (JTEAS).

All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).

https://doi.org/10.64922/jteas.v1i1.85

Figure 4

Pounding and sieving dried leaves

Figure 5

Blending leaves into fine powder

Extract Preparation

Extracts were prepared by maceration. Leaf powders (50 g, 100 g, and 200 g) were weighed and separately mixed with 1 L of either 95% ethanol or tap water to obtain 5%, 10%, and 20% (w/v) solutions, respectively. Ethanol was used at its supplied concentration (95%) and was not diluted, while water extracts were prepared using tap water as the solvent. The mixtures were stored in covered plastic bottles, shown in Figure 7, at room temperature and shaken manually twice daily (morning and evening) to enhance solvent-solute interaction. After five days, extracts were filtered with a clean cloth and stored in labeled plastic containers.

Figure 6

Treatments being prepared

Treatment of Wood Samples

Wood samples were soaked in the prepared extracts in covered plastic containers (Figure 8) for 48 hours using a non-pressure treatment method for surface-level protection. Retention and penetration were not measured in this study. After treatment, the samples were stacked and air-dried (Figure 9) at room temperature for 24 hours.

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).

All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).

https://doi.org/10.64922/jteas.v1i1.85

Figure 8

Samples soaked in plastic containers

Figure 7

Treated samples being air-dried

In-ground Field Test

Field tests were conducted on termite-active clay soil near the school field at Akenten Appiah-Menka University of Skills Training and Entrepreneurial Development (AAMUSTED), Kumasi, approximately 10 m from an active termite mound. Holes were created using a 25 mm wall chisel and a 2 kg hammer (Figure 10). Wood samples were randomly buried vertically to half their length (125 mm) at 30 cm spacing as shown in Figure 11. A total of 190 wood samples were exposed for 90 days (February–April 2025). The samples comprised 180 treated specimens arranged as 3 plant species × 2 wood tissues (heartwood and sapwood) × 2 solvents × 3 concentrations × 5 replicates, plus 10 untreated control samples (5 each for heartwood and sapwood). After the exposure period, samples were retrieved, brushed to clean off soil and oven-dried at 105°C to constant weight (W2). Weight loss (%) was calculated as:

$$\text{Weight loss} = \frac{W1-W2}{W1} X100$$

Where W1 and W2 were initial and final oven-dry weights respectively

Figure 10

Planting samples on site

Figure 9

Test site after planting samples

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).

All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).

https://doi.org/10.64922/jteas.v1i1.85

Data Collection and Analysis

Weights were measured with a digital balance with a precision of 0.01 g. Data were analyzed using SPSS v.27. Means and standard deviations were calculated. ANOVA tested differences among treatments at 5% significance. Tukey's HSD post-hoc test identified pairwise differences ($p \leq 0.05$). Graphs were generated in Microsoft Excel.

Ethical Considerations

The study was conducted following guideline-based compliance in accordance with the guidelines of the Ghana Environmental Protection Agency (EPA Ghana, 2019) and the laboratory safety protocols of the Akenten Appiah-Menka University of Skills Training and Entrepreneurial Development (AAMUSTED). Plant materials were collected with permission from local community authorities. Leaf quantities were limited to the minimum required for extract preparation, determined by the predefined experimental design (5%, 10%, and 20% w/v concentrations) and the total number of treatment replicates. Only healthy leaves were harvested, and collection was done in a manner that avoided damage to the parent plants.

Ethanol and other laboratory chemicals were handled using appropriate personal protective equipment, including gloves, goggles, and lab coats. Wood and extract residues were disposed of in accordance with institutional laboratory waste management procedures.

Results and Discussion

The results are based on mean weight losses of *Pycnanthus angolensis* wood treated with leaf extracts of *Azadirachta indica*, *Chromolaena odorata*, and *Cedrela odorata*, using ethanol and water at 5%, 10%, and 20% concentrations. Untreated samples (0%) served as controls. Statistical analyses (ANOVA and Tukey's HSD) were used to compare treatment effects by leaf type, solvent, and concentration.

Figure 11

Post-Exposure Condition of Wood Samples Affected by Termites

As shown in Figure 12, untreated *Pycnanthus angolensis* wood showed severe termite damage in the in-ground test, with many galleries and frass deposits, scoring 4.5-5.0 on the damage scale (0 = no damage, 5 = complete destruction). Treated samples at 5% concentration scored around 4.5, showing little improvement. At 10%, damage dropped to about 4.0. The highest protection was at 20%, with ethanol-based *Azadirachta indica* and *Chromolaena odorata* extracts scoring around 3.5 while *Cedrela odorata* scored near 4.0. Despite reductions, treated samples still experienced substantial termite damage.

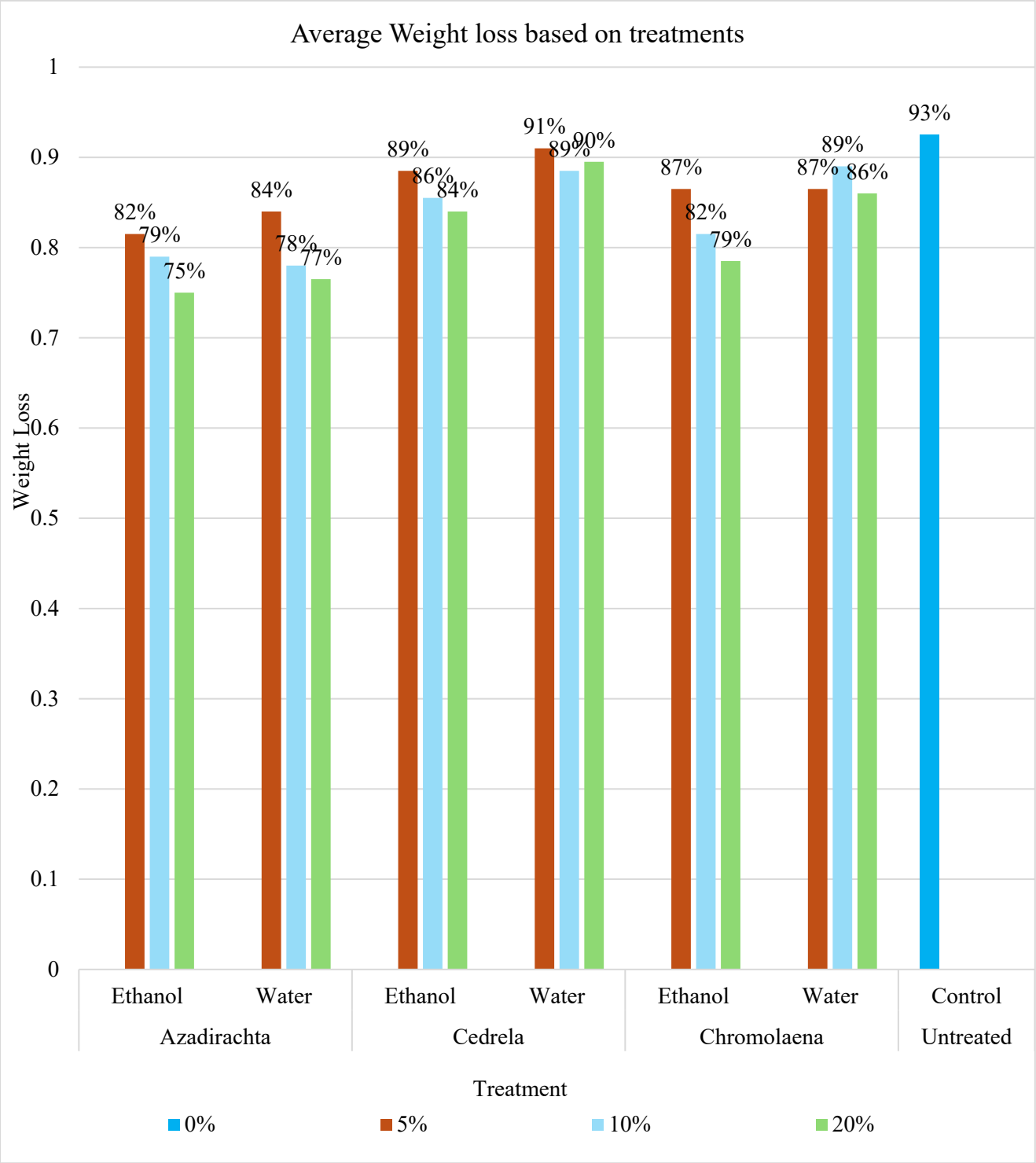
© [2026] Journal of Technology Education and Applied Sciences (JTEAS).

All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).

https://doi.org/10.64922/jteas.v1i1.85

Figure 12

Average Weight Loss of Wood Samples across Treatment Types



© [2026] Journal of Technology Education and Applied Sciences (JTEAS).

All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).

https://doi.org/10.64922/jteas.v1i1.85

In Figure 13, the graph shows average weight loss from the effects of *Azadirachta indica*, *Cedrela odorata*, and *Chromolaena odorata* leaf extracts on *Pycnanthus angolensis* wood. The untreated control recorded the highest weight loss (93%), confirming the high susceptibility of unprotected wood.

Across all plant extracts, increasing concentration reduced weight loss, indicating a dose-dependent preservative effect. For *Azadirachta*, ethanol extracts decreased weight loss from 82% at 5% to 75% at 20% concentration. Water extracts showed a similar trend, declining from 84% to 77% at higher concentrations. Cedrela also followed this pattern, with ethanol treatments reducing losses from 89% to 84%, and water treatments from 91% to 90%. *Chromolaena* in ethanol extracts reduced weight loss from 87% to 79%, while water extracts dropped from 87% to 86%.

These results demonstrate that all three species provided preservative effects, with *Azadirachta indica* showing the highest efficacy, followed by *Chromolaena odorata*, and lastly, *Cedrela odorata*. Both ethanol and water extracts were effective, though ethanol extracts generally produced slightly better results at higher concentrations.

Table 1

Descriptive Statistics of Weight Loss

Leaf Type	*Solvent*	*Concentration*	*Mean*	*Std. Deviation*
Azadirachta indica	Ethanol	5%	81.50%	2.121%
		10%	79.00%	1.414%
		20%	75.00%	2.828%
	Water	5%	84.00%	2.828%
		10%	78.00%	2.828%
		20%	76.50%	3.536%
Cedrela odorata	Ethanol	5%	88.50%	4.950%
		10%	85.50%	6.364%
		20%	84.00%	5.657%
	Water	5%	91.00%	5.657%
		10%	88.50%	7.778%
		20%	89.50%	2.121%
Chromolaena odorata	Ethanol	5%	86.50%	4.950%
		10%	81.50%	3.536%
		20%	78.50%	3.536%
	Water	5%	86.50%	2.121%
		10%	89.00%	2.828%
		20%	86.00%	1.414%

As shown in Table 1, descriptive statistics showed that the untreated control samples had the highest mean weight loss (92.5%), confirming the high susceptibility of *Pycnanthus angolensis* to termite damage. All plant extracts reduced weight loss to varying degrees. Among species, *Azadirachta indica* performed best, with mean losses at 79%, followed by *Chromolaena odorata* at 85% and *Cedrela odorata,* 88%.

When grouped by solvents, ethanol extracts were generally more effective than water extracts. Ethanol treatments reduced mean weight loss from about 85% at 5% concentration to 79% at 20%, while water extracts showed a smaller improvement, from about 87% to 84% across the same range.

Across all treatments, the 20% concentration produced lower weight losses compared to the 5% and 10% concentrations.

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).

All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).

https://doi.org/10.64922/jteas.v1i1.85

Table I

Tests of Between-Subjects Effects

Source	*Sum of Squares*	*Mean Square*	*F*	*p*
Corrected Model	950.316[a]	52.795	3.082	.010
Intercept	199665.818	199665.818	11654.840	< .001
Leaf Type * Solvent * Concentration	34.111	8.528	.498	.738
Leaf Type * Solvent	24.889	12.444	.726	.497
Leaf Type * Concentration	23.333	5.833	.341	.847
Solvent * Concentration	15.056	7.528	.439	.651
Leaf Type	480.667	240.333	14.029	< .001
Solvent	93.444	93.444	5.455	.031
Concentration	136.500	68.250	3.984	.036

Table 2 shows the between-subjects effects of treatments. Leaf type had a strong effect on weight loss (F = 14.029n weight loss (F = 14.029, p < .001), showing that the choice of plant species significantly influenced wood was also effective (F = 5.455d wood was also effective (F = 5.455, p = .031), with ethanol extracts generally reducing weight loss more effectively than water extracts. Concen84, p = .036).

In contrast, none of the interaction terms were significant. The three-way interaction of leaf type × solvent × concentration (F = 0.498, p = .738), and the two-way interactions (leaf × solvent, F = 0.726, p = .497; leaf × concentration, F = 0.341, p = .847; solvent × concentration, F = 0.439, p = .651) all failed to reach significance. This means the extracts' effects were driven by the independent factors of leaf type, solvent, and concentration, rather than by combined or synergistic effects among them.

Table 3

Multiple Comparisons of Leaf Type Effects

Leaf Type	*Mean Difference*	*Std. Error*	*p*
Azadirachta indica vs. *Cedrela odorata*	8.83%*	1.690%	< .001
Azadirachta indica vs. *Chromolaena odorata*	5.67%*	1.690%	.016
Cedrela odorata vs. *Chromolaena odorata*	3.17%	1.690%	.272
Control vs. *Azadirachta indica*	13.50%*	3.161%	.002
Control vs. *Cedrela odorata*	4.67%	3.161%	.470
Control vs. *Chromolaena odorata*	7.83%	3.161%	.096

As shown in Table 3, post-hoc analysis using Tukey's HSD confirmed that leaf type significantly influenced wood weight loss. *Azadirachta indica* treatments consistently outperformed the others, with significantly lower mean losses compared to *Cedrela odorata* (8.83%, p < .001), *Chromolaena odorata* (5.67%, p = .016), and the *untreated control* (13.50%, p = .002). In contrast, no significant differences were found between Cedrela and Chromolaena, or between either species and the control, indicating that their preservative effects were not as strong compared to Azadirachta.

Table 4

Multiple Comparisons of Solvent Effects

Solvent	*Mean Difference*	*Std. Error*	*p*
Ethanol vs. Control	10.28%*	3.085%	.009
Water vs. Control	7.06%	3.085%	.082
Ethanol vs. Water	3.22%	1.380%	.075

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).

All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).

https://doi.org/10.64922/jteas.v1i1.85

Post-hoc comparisons showed that ethanol extracts significantly reduced mean weight loss compared to the untreated control (-10.28%, p = .009). However, the difference between ethanol and water extracts was not statistically significant (-3.22%, p = .075). Likewise, water extracts did not differ significantly from the control (+7.06%, p = .082), suggesting that ethanol was the more effective solvent for extracting preservative compounds as shown in Table 4.

Table 5

Multiple Comparisons of Concentration Effects

Concentration	*Mean Difference*	*Std. Error*	*p*
0% vs. 5%	6.17%	3.161%	.241
0% vs. 10%	8.92%*	3.161%	.049
0% vs. 20%	10.92%*	3.161%	.013
5% vs. 10%	2.75%	1.690%	.388
5% vs. 20%	4.75%	1.690%	.050
10% vs. 20%	2.00%	1.690%	.644

Post-hoc comparisons showed that higher extract concentrations generally reduced wood weight loss. The untreated control (0%) had significantly greater losses than both the 10% (8.92%, p = .049) and 20% treatments (10.92%, p = .013), but not the 5% level (p = .241). Among treated samples, differences were less distinct, though a borderline reduction was observed between 5% and 20% (4.75%, p = .050). No significant differences emerged between 5% and 10% or between 10% and 20%. These results indicate that increasing concentration improves protection, with 20% offering the strongest effect as shown in Table 5.

The untreated control samples recorded the highest mean weight loss (92.5%), confirming the high susceptibility of *Pycnanthus angolensis* to termite attack. This agrees with reports that classify the species as non-durable under tropical field conditions (Udeozo et al., 2015).

All plant extract treatments reduced weight loss relative to the control, although protection levels varied among species. *Azadirachta indica* showed the strongest performance, with average losses of about 79%, significantly lower than *Chromolaena odorata* (85%) and *Cedrela odorata* (88%). This aligns with evidence attributing neem's insecticidal and antifeedant activity to azadirachtin and related limonoids (Adusei & Azupio, 2022; Sinha & Ray, 2024).

The comparatively weaker performance of *Chromolaena odorata* and *Cedrela odorata* likely reflects differences in extractive composition and bioactivity. *Chromolaena odorata* contains phenolics and flavonoids, but its termite resistance is often described as moderate (Budha et al., 2023). *Cedrela odorata* extracts have been more frequently associated with antifungal rather than anti-termite effects (Dele, 2020).

Solvent type influenced preservative performance, with ethanol extracts consistently reducing weight loss more than water extracts. Weight loss declined from approximately 85% at 5% concentration to 79% at 20% for ethanol treatments, supporting evidence that ethanol enhances extraction of insect-deterrent compounds (Broda & Mazela, 2020; Kirker et al., 2024). Increasing extract concentration further improved performance, confirming a dose-dependent response (Asiry et al., 2022). However, changes beyond 10% were modest, likely due to limited penetration under non-pressure treatment.

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).

All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).

https://doi.org/10.64922/jteas.v1i1.85

Overall, ethanol-based neem leaf extracts reduced termite damage providing partial protection, indicating, indicating their potential as supplementary, low-hazard alternatives rather than direct replacements for synthetic preservatives in severe termite eCalovi et al., 2024; Ong et al., 2025).

Leaf type ($F = 14.03$, $p < .001$), solvent ($F = 5.46$, $p = .031$), and concentration ($F = 3.98$, $p = .036$) each had significant main effects, while none of the interaction terms were significant. This indicates that each factor independently influenced wood durability, with no strong effects from the combination among them.

Conclusions

The study demonstrates that leaf extracts from *Azadirachta indica*, *Chromolaena odorata*, and *Cedrela odorata* can reduce termite damage in *Pycnanthus angolensis*, though effectiveness varied by leaf type, solvent, and concentration. Neem extracts, particularly ethanol-based treatments at 20%, produced the greatest reduction in weight loss. The research was limited by its 90-day duration, which may not reflect long-term durability in varying tropical conditions. It evaluated only *Pycnanthus angolensis* wood, restricting applicability to other Ghanaian hardwoods. Additionally, field tests at a single site may not account for diverse soil types, climates, or termite species distributions.

However, these findings reinforce the potential of locally available plant extracts as eco-friendly wood preservatives. Further research could include long-term trials to assess extract durability and efficacy. Testing concentrations above 20% could establish optimum performance. Repeating tests across different climates and soils can optimise extract use.

These findings highlight *Azadirachta indica* as the best-performing among the tested extracts. Ethanol-based extracts, especially at 20% concentration, provided the most reliable protection. Water extracts showed only limited effectiveness, but could still be useful in rural or resource-limited settings where ethanol is less accessible.

At the policy level, these results provide evidence to support the development of plant-based preservation strategies within Ghana's forestry sector. Incorporating such bio-preservatives into national timber treatment guidelines could reduce reliance on hazardous chemicals, promoting both environmental safety and sustainable resource use.

Acknowledgements

We acknowledge Mr. Enoch Nana Tutu Otoo for his assistance in preparing wood samples at the CBT Wood Laboratory, AAMUSTED, and the Department of Wood Science and Technology Education, AAMUSTED, for providing laboratory facilities.

Funding Statement

No funding was received. The authors declare no financial or institutional conflicts.

Author Contribution Statement

Conceptualisation: KA (lead), MDM (equal)

Data Curation: MDM (lead)

Formal Analysis: MDM (lead), KA (equal), SSU (supporting)

Methodology: MDM (lead)

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).

All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).

https://doi.org/10.64922/jteas.v1i1.85

Validation: KA (lead), FKB (equal), MBD (supporting)

Writing – Original Draft: KA (lead), MDM (supporting)

Supervision: KA (lead), FKB (supporting), MBD (supporting)

AI Disclosure Statement

The authors used *ChatGPT (OpenAI)* to assist in language editing to improve readability and also revised the abstract to meet journal requirements in word count. The authors reviewed all generated content carefully and take full responsibility for the accuracy and integrity of the final manuscript.

References

Adusei, S., & Azupio, S. (2022). Neem: A novel biocide for pest and disease control of plants. *Journal of Chemistry*, *2022*(1), 6778554. https://doi.org/10.1155/2022/6778554

Agbo, B. E., Nta, A. I., & Ajaba, M. O. (2019). Bio-pesticidal properties of neem (*Azadirachta indica*). *Advances in Trends in Agricultural Sciences*, *1*, 17–26.

Asiry, K. A., Al-Nasser, A. S., & Abohassan, R. A. (2022). Repellent and toxic effects of some plant extracts on subterranean termite *Psammotermes hybostoma*. *International Journal of Agriculture & Biology*, *27*(2), 99–104. https://doi.org/10.17957/IJAB/15.1905

Asmara, Y. P. (2025). *Green engineering materials: Innovations and applications for sustainable construction*. CRC Press.

Broda, M. (2020). Natural compounds for wood protection against fungi—A review. *Molecules*, *25*(15), 3538. https://doi.org/10.3390/molecules25153538

Budha Magar, A., Shrestha, D., Pakka, S., & Sharma, K. R. (2023). Phytochemistry, biological, and toxicity study on aqueous and methanol extracts of *Chromolaena odorata*. *The Scientific World Journal*, *2023*(1), 6689271. https://doi.org/10.1155/2023/6689271

Calovi, M., Zanardi, A., & Rossi, S. (2024). Recent advances in bio-based wood protective systems: A comprehensive review. *Applied Sciences*, *14*(2), 736. https://doi.org/10.3390/app14020736

Chakurah, I., Bih, F. K., & Antwi, K. (2023). Resistance to termite attack properties of *Azadirachta indica* wood from the Savannah ecological zone of Ghana. *International Journal of Innovative Science and Research Technology*, *8*(9), 263–268.

Changotra, R., Rajput, H., & Dhir, A. (2024). Environmental impact of chemical wood preservatives: A review. *Environmental Science and Pollution Research*, *31*(5), 7890–7905. https://doi.org/10.1007/s11356-023-31234-5

Da Costa, E. W. B., Rudman, P., & Gay, F. J. (1958). Investigations on the durability of *Tectona grandis*. *Empire Forestry Review*, 291–298.

Datta, H. S. (2024). Harnessing neem (*Azadirachta indica* A. Juss): A sustainable approach to natural farming. *International Journal of Plant & Soil Science*, *36*(7), 643–648. http://journal.repositoryarticle.com/id/eprint/1421/

Dele, A. (2020). Volatile oils from *Cedrela odorata* L. as protectants against *Sitophilus zeamais* (Coleoptera: Curculionidae). *American Journal of Essential Oils and Natural Products*, *8*, 20–24.

Eaton, R. A., & Hale, M. D. C. (1993). *Wood: Decay, pests, and protection*. Chapman and Hall.

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).

All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).

https://doi.org/10.64922/jteas.v1i1.85

Environmental Protection Agency, Ghana. (2019). *Environment and health protection—Requirements for ambient air quality and point source/stack emissions* (Standard 1236 of 2019). Author.

European Committee for Standardization. (2014). *Field test method for determining the relative protective effectiveness of a wood preservative in ground contact* (EN 252:2014). Author.

Falemara, B. C., Ajayi, B., Kambai, C., & Ukanyirioha, C. (2018). Durability assessment of cashew nut shell liquid on *Bambusa vulgaris* schrad in timber graveyard test. In *Emerging issues in sustainable forest management: Experiences and lessons for Nigeria. Proceedings of the 40th Annual Conference of the Forestry Association of Nigeria* (pp. 206–213).

Ghaffariyan, M. R. (2023). Utilization of forest biomass residues for bioenergy and bioproducts. *Bioresource Technology Reports*, *22*, 101432. https://doi.org/10.1016/j.biteb.2023.101432

Goodell, B., & Nielsen, G. (2023). Wood biodeterioration. In *Springer handbook of wood science and technology* (pp. 139–177). Springer International Publishing.

Han, H. S., Jacobson, A., Bilek, E. T., & Sessions, J. (2018). Waste to wisdom: Utilizing forest residues for the production of bioenergy and biobased products. *Applied Engineering in Agriculture*, *34*(1), 5–10.

Kirker, G. T., Hassan, B., Mankowski, M. E., & Eller, F. J. (2024). Critical review on the use of extractives of naturally durable woods as natural wood protectants. *Insects*, *15*(1), 69. https://doi.org/10.3390/insects15010069

Lebow, S. T., Lebow, P. K., Foster, D. O., & Brooks, K. M. (2000). *Environmental impact of preservative-treated wood in a wetland boardwalk* (Vol. 582). US Department of Agriculture, Forest Service, Forest Products Laboratory.

Leroy, M., Candelier, K., Damay, J., Bossu, J., Lehnebach, R., Thévenon, M. F., … Clair, B. (2023). Natural durability of 8 tropical species suitable for structural roundwood: Laboratory screening tests for resistance to fungi and termites. *Bois et Forêts des Tropiques*, *358*, 15–29. https://hal.science/hal-04504833/

Li, K., Huang, D., Gao, X., He, J., Sun, X., & Yan, L. (2024). Potential applications for eco-friendly wood preservative of the extracts of thermally degraded cobs and stalks of corn. *European Journal of Wood and Wood Products*, *82*(4), 1133–1143. https://doi.org/10.1007/s00107-024-02072-4

Mahalle, D., & Gupta, A. (2021). Estimation of total phenol, flavonoids and alkaloid content and evaluation of antioxidant and antimicrobial activity of Tinospora crispa leaves and flower extracts. *Journal of Advanced Scientific Research*, 12(1), 208-212

Olawale, F., Olofinsan, K., & Iwaloye, O. (2022). Biological activities of *Chromolaena odorata*: A mechanistic review. *South African Journal of Botany*, *144*, 44–57. https://doi.org/10.1016/j.sajb.2021.09.001

Ong, Z., Arip, M., Nasir, M., & Lee, H. L. (2025). Advances in wood preservation technology: A review of conventional and nanotechnology preservation approaches. *BioResources*, *20*(3). https://doi.org/10.15376/biores.20.3.Ong

Pasiecznik, N. M. (2022). *Cedrela odorata* (Spanish cedar). *CABI Compendium*. https://doi.org/10.1079/fc.11975.20210114102

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).

All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).

https://doi.org/10.64922/jteas.v1i1.85

Ruvalcaba, L. P., Armenta, K. G. M., Valdés, T. D., de Jesús Velázquez Alcaraz, T., Tafoya, F. A., López, R. M., & Juárez, M. G. Y. (2017). Effect of cedar extract (*Cedrela odorata* L.) on the termite (*Reticulitermes* spp.). *Agricultural Sciences*, *8*, 261–266.

Sinha, N., & Ray, S. (2024). Different extraction processes of neem (*Azadirachta indica*) and its potential in controlling pest populations. *Journal of Entomology and Zoology Studies*, *12*(4), 18–26. https://doi.org/10.22271/j.ento.2024.v12.i4a.9345

Thomas, M., Gayathri, S., Athira, K. P., Elias, R. S., Soumya, S., & Kurian, P. (2023). Efficacy of plant leaf extracts on termites under laboratory conditions. *Journal of Entomology and Zoology Studies*, *11*(1), 106–110. https://doi.org/10.22271/j.ento.2023.v11.i1b.9131

Udeozo, I. P., Ejikeme, C. M., Eboatu, A. N., & Arinze, R. U. (2015). The efficacy of *Pycnanthus angolensis* timber: An assay of its properties and chemical constituents. *Global Journal of Biotechnology & Biochemistry*, *10*(3), 121–125.

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).

All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).

https://doi.org/10.64922/jteas.v1i1.85

RESEARCH ARTICLE

Multilevel Stacking with Domain-Adversarial Neural Network and Enriched Metadata for Improved Renewable Energy Output Forecasting

Prince Clement Addo[1*], Daniel Dankwah[1], Joshua Caleb Dagadu[1], Samuel Kofi Akpatsa[1], Oliver Kufour Boansi[1], William Asiedu[1], George Asante[1], Enoch Adjei Frimpong[2], Dong Miao[3], William Agyei Atuahene[1]

[1]Department of Information Technology, University of Skill Training and Entrepreneurial Development, Ghana
[2]Department of Computer Science, Cape Coast Technical University, Ghana
[3]School of Computer Science, University of Electronic Science and Technology of China, China
*Corresponding Author: [pcaddo@aamusted.edu.gh]

To cite this article: Addo, P. C., Dankwah, D., Dagadu, J. C., Akpatsa, S. K., Boansi, O. K., Asiedu, W., Asante, G., Frimpong, E. A., Miao, D., & Atuahene, W. A. (2026). Multilevel stacking with domain-adversarial neural network and enriched metadata for improved renewable energy output forecasting. Journal of Technology Education and Applied Sciences, 1(1), 32–45. https://doi.org/10.64922/jteas.v1i1.86

ABSTRACT

This study proposes a multilevel stacking approach that extends Wolpert's (1992) stacked generalisation, aiming to minimise forecasting error, reduce systematic bias, and enforce temporal domain invariance. It incorporates walk-forward out-of-fold and step-based walk-forward validation, enriched metadata, and a domain adversarial neural network (DANN). The base learners XGB and LSTM function as residual and latent encoders for the Random Forest with DANN meta-learner. DANN is integrated in meta-learning to minimise task regression loss and maximize domain discrimination loss through a gradient reversal layer. This fusion improves domain transition stability, as the model's high explanatory power with ^2 values exceed 0.98 across validation and test regimes. The MASE values below 0.07 on both regimes indicate strong predictive performance relative to a naïve benchmark, and NRMSE values around 0.025 confirm that the proposed stacking framework effectively captures temporal dependencies. Moreover, the model achieves a near-symmetric residual distribution, with a skewness of 0.37 on the validation set and -0.01 on the test set, suggesting that adversarial feature learning helps reduce asymmetric prediction errors and stabilise meta-learning. This proposed framework demonstrates empirically that stacking should not be evaluated solely on MAE/RMSE but on residual distribution behaviour under domain transition. Furthermore, SHAP analysis provides empirical evidence that enriched metadata significantly enhances the learning capacity, allowing the meta-learner to capture complex dependencies.

KEYWORDS

Multilevel Stacking, Renewable Energy Forecasting, Domain Adversarial Neural Network, Enriched Metadata, Walk-Forward Validation

Publication History

Received: 14 November 2025
Revised: 28 November 2026
Accepted: 5 April 2026
Available online: 30 July 2026

Introduction

The stacked ensembles have improved generalisation in static contexts. Their direct application to sequential and time-dependent data is hindered by data leakage, concept drift, and structural irregularities (Tsymbal, 2004). Recent studies have attempted to incorporate temporal information into heterogeneous ensembles such as tree-based, convolutional, and recurrent networks (Zhang et al., 2018). This underscores the significance of base-learner diversity for predictive performance (Middlehurst et al., 2021). However, balancing accuracy, robustness, or interpretability is challenging, particularly for non-stationary energy data. The enriched metadata has contributed to improving ensemble performance, but systematic implementation in renewable energy domains is limited (Cerqueira et al., 2020). Studies explore the use of Random Forest as a meta-learner within a stacking framework. This is due to the Random Forest's ability to capture nonlinear dependencies in a diverse set of learners' outputs. An empirical study has confirmed that tree-based meta-learners are effective in learning complex relationships among predictive vectors (Krauss et al., 2017).

Moreover, a study on credit risk assessment shows that an ensemble of tree-based models outperformed a standalone model in predictive reliability and robustness (Lessmann et al., 2015). These findings collectively suggest that Random Forest meta-learning provides a flexible mechanism for synthesizing heterogeneous

predictive signals while mitigating overfitting through ensemble diversity. On the contrary, there have been methodological challenges that sideline the reliability of stacking with the Random Forest meta-learner in dynamic regimes. The data distribution often evolves due to seasonal variations and distributional shifts. The predictive relationships learned by base-learners and the meta-learner become unstable, leading to degradation in forecasting accuracy (Tsymbal, 2004).

Even though Random Forest can capture nonlinear dependencies among base model predictions, it does not inherently and explicitly address the distributional mismatch that arises when the statistical properties of the data shift across regimes. This study proposes a multilevel stacking with a Domain Adversarial Neural Network (DANN), unlike traditional stacking methods that aggregate base learners without explicitly addressing temporal nonstationarity. Integrating a DANN in meta-learning extracts domain-invariant features from sequential data. By introducing a gradient reversal layer, the framework actively discourages reliance on domain-specific signals. This approach provides two major advantages. First, it enhances stability in predictions by reducing the impact of distributional shifts across different time periods, ensuring the meta-learner generalizes well to unseen sequences.

Second, it refines the meta-features, including base learner residuals, LSTM latent embedding, and engineered features, by filtering out domain-dependent noise, thereby improving interpretability and robustness. To the best of our knowledge, this study represents the first integration of DANN as a regularisation and loss function in meta-learning. By combining residual-aware features, temporal latent representations, and domain adversarial learning, the framework addresses both the accuracy and stability challenges of renewable energy prediction.

Method

Problem Formulation

Let the renewable and weather condition dataset be represented as a chronological time series.

$$D = \{(X_t, y_t)\}_{t=1}^{T},\ X_t \in R^d, y_t \in R$$

Where $\boldsymbol{X_t} \in \boldsymbol{R^d}$ denotes the feature vectors of the base and engineered features, and $\boldsymbol{y_t} \in \boldsymbol{R}$ represents the target variable (Energy delta [wh]), which corresponds to energy production change. This dataset is a structured time series dataset sourced from Kaggle, a reputable online platform (https://www.kaggle.com/datasets/samanemami/renewable-energy-and-weather-conditions). The data spans five years from 2017 to 2022, with 196,776 records shown in Table 1.

Table 1

Descriptive statistics of Numerical Values in renewable and Weather conditions Data

Numerical variables	*Count*	*Mean*	*Std.*	*Minimum*	*25%*	*50%*	*75%*	*Maximum*
Energy delta [Wh]	196776	5.7300	1044.82	0.00	0.00	0.00	577.00	5020.00
GHI	196776	3.2596	52.172	0.00	0.00	1.600	46.80	229.20
Temperature	196776	9.7905	7.9954	-16.60	3.600	9.300	15.700	35.800
Pressure	196776	1.0152	9.5857	977.0	1010.0	1016.0	1021.0	1047.0
Humidity	196776	7.98105	15.604	22.00	70.00	84.00	92.00	100.0
Sunlight Time	196776	211.721	273.90	0.00	0.00	30.00	390.00	1020.00
Sunlight_Time/day_length	196776	2.65186	0.3290	0.00	0.00	0.050	0.5300	1.00
Hour	196776	1.14989	6.9218	0.00	5.00	11.00	17.00	23.00

The study assumes that the dataset follows a normal distribution. If the data deviates from this assumption, it may warrant the use of transformation or non-parametric modelling approaches.

The study further conducts a data distribution test using the Shapiro-Wilk test, the Kolmogorov-Smirnov (KS) test, and the Anderson-Darling test.

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).
All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).
https://doi.org/10.64922/jteas.v1i1.86

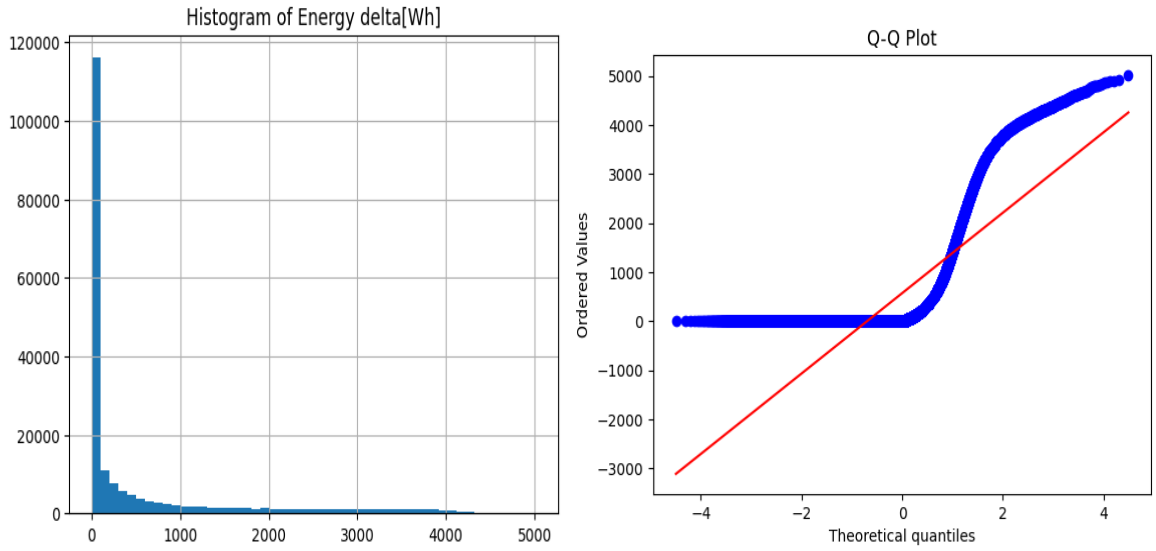
The Shapiro-Wilk test shows a test statistic (W) of 0.616 and a p-value less than 0.001. This value falls significantly below the conventional alpha level of 0.05. This rejects the null hypothesis, indicating a considerable deviation of the target variable from normality.

Similarly, the Kolmogorov-Smirnov (KS) test shows a test statistic of 0.292 with a corresponding p-value less than 0.001. This result confirms that the distribution of the target variable does not align with the characteristics of a Gaussian distribution.

The Anderson-Darling test yields a test statistic of 32,018.751, which substantially exceeded all critical values at the 15%, 10%, 5%, 2.5%, and 1% significance levels. This finding rejects the null hypothesis and reinforces the conclusions drawn from the Shapiro-Wilk and KS tests. The study ascertains that the results of all three statistical tests consistently demonstrate that the target variable is not normally distributed. Given the apparent non-Gaussian nature of the data, it is therefore applicable to consider a data transformation technique or a non-parametric model in the proposed multilevel stacking framework in Figure 1.

Figure 1

Histogram and Q-Q plot depict the Data Distribution



The forecasting objective of the study is to learn a mapping $\mathcal{F}: \boldsymbol{X_t} \rightarrow \boldsymbol{Y_t}$ that minimizes prediction error under non-stationary temporal conditions. However, the renewable energy and weather condition dataset exhibits an inherent

concept drift:

$$P_t(y|x) \neq P_{t+\triangle}(y|x) \quad (1)$$

distributional instability:

$$P_t(x) \neq P_{t+\triangle}(x) \quad (2)$$

and a domain-induced variance caused by evolving weather regimes. The stacking approach, which uses a Random Forest as a meta-learner, assumes independent and identically distributed observations. Consequently, their performance degrades when applied to sequential data distributional instability.

This study proposes a multilevel stacking that integrates gradient boosted residual modelling, temporal latent feature extraction, and adversarial domain adaptation. The study employs two complementary validation strategies: walk-forward out-of-fold (OOF) on the training set and step-based walk-forward validation on the validation and testing sets.

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).
All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).
https://doi.org/10.64922/jteas.v1i1.86

The walk forward OOF on the training set serves as a leakage-free prediction for every time point. Each observation is predicted only by a model trained strictly on past data folds, ensuring temporal causality is respected (Tashman, 2000).

The step-based validation is performed on both the validation and test sets to mimic a forecast of a real-world deployment. The evaluation replicates a rolling adaptation, testing how models perform under concept drift and non-stationarity (Gama et al., 2014).

Base Level Model Formulation

The study employs an Extreme Gradient Boosting (XGB) as one of the two base-level learners to model nonlinear relationships between base features and target variables.

The boosting model predicts:

$$\hat{y}_t = \sum_{k=1}^{k} f_k\,(X_t) \quad (3)$$

Where $\boldsymbol{f_k}$ represents decision trees learned sequentially.

The residual errors are computed as:

$$r_t = y_t - \hat{y}_t \quad (4)$$

to capture unexplained variance and structural irregularities in the dataset. Residual dynamics are subsequently used as a meta-feature to guide the meta-learner. The lagged residuals are constructed as:

$$r^{(l)} = r_{t-1}, l = 1,2,3 \quad (5)$$

allowing the meta-learner to capture temporal persistence in prediction errors.

Moreover, as the tree-based models struggle with sequential dependencies, the study employs Long Short-Term Memory (LSTM) as a base learner to model and extract temporal latent features that are not observable in static features.

For a sequence of length $\boldsymbol{L}$

$$X_t = [X_{t-L+1}, \dots X_t] \quad (6)$$

The LSTM call updates are defined as:

$$f_t = \sigma(W_t[h_{t-1}, X_t] + b_f) \quad (7)$$

$$i_t = \sigma(W_i[h_{t-1}, X_t] + b_i) \quad (8)$$

$$\acute{c}_t = tanh(W_c[h_{t-1}, X_t] + b_c) \quad (9)$$

$$C_t = f_t \odot C_{t-1} + j_t \odot \acute{c}_t \quad (10)$$

$$h_t = O_t \odot \tanh(C_t) \quad (11)$$

The latent representation extracted from the bottleneck layer is

$$H_t \in R^k \quad (12)$$

which summarizes temporal dependencies in the input sequence.

The metadata combines complementary information sources such as residual dynamics, temporal latent features, weather condition and engineered features to form an enriched metadata. The enriched metadata is:

$$Z_t = [r_t^{m_1}, r_{t-1}^{m_1}, r_{t-2}^{m_1}, r_{t-3}^{m_1}, H_t, X_t] \quad (13)$$

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).
All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).
https://doi.org/10.64922/jteas.v1i1.86

This enriched metadata allows the meta-learner to model predictive uncertainty, temporal structure, and weather context.

Meta-Level Model Formulation

The study employs Random Forest as the meta-learner to model nonlinear relationships. The prediction function of a Random forest is:

$$\hat{y}_t = \frac{1}{B}\sum_{b=1}^{B} T_b\,(Z_t) \quad (14)$$

Where T_b denotes an individual decision tree. Notwithstanding the strength of Random Forest, it suffers from sensitivity to distributional shifts,

$$P_{train(M)} \neq P_{test(M)} \quad (15)$$

The inability to explicitly model temporal dependencies and vulnerability to domain-induced variance. These challenges prompted the integration of the domain adaptation mechanism.
The Domain Adversarial Neural Network employed consists of a feature extractor$F(\bullet)$, a task predictor $G(\bullet)$, and a domain classifier $D(\bullet)$.

The adversarial training objective is to minimise prediction loss while maximizing domain confusion.

Task loss:

$$\mathcal{L}_y = \frac{1}{N}\sum_{t=1}^{N}(y_t - \hat{y}_t)^2 \quad (16)$$

Domain classification loss:

$$\mathcal{L}_d = -\sum_{t=1}^{N} d_t \log(\dot{d}_t) \quad (17)$$

And the combined objective:

$$\mathcal{L} = \mathcal{L}_y - \lambda\mathcal{L}_d \quad (18)$$

Where $\boldsymbol{\lambda}$ controls the strength of adversarial learning. A gradient reversal layer multiplies the gradient by $-\boldsymbol{\lambda}$ during backpropagation, forcing the feature extractor to learn domain-invariant representations. The domain classifier output provides a probability estimate:

$$P(d_t|v_t)$$

Indicates the likelihood that a sample belongs to a shifted domain.

Samples are reweighted as:

$$w_t = 1 - P(d_t|v_t) \quad (19)$$

These weights are incorporated into Random Forest training, emphasizing samples with stable distributions while reducing the influence of drifted observations. The final meta-learning stage trains the Random Forest using the purified features.

$$v_t = f(Z_t) \quad (20)$$

and importance weights $\boldsymbol{w_t}$.

The final predictor:

$$\hat{y}_t = \frac{1}{B}\sum_{b=1}^{B} T_b\,(V_t) \quad (21)$$

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).
All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).
https://doi.org/10.64922/jteas.v1i1.86

This domain-adaptive learning process reduces prediction instability caused by evolving regimes.

The main objective function is to minimise forecasting error, reduce systematic bias, and enforce temporal domain invariance.

Thus:

$$\theta_{\hat{y}_t} = arg\min_{\theta_{\hat{y}_t}} \frac{1}{N}\sum_{t=1}^{N} \omega_t (y_t - \hat{y}_t(v_t, \theta_{\hat{y}_t}))^2 \quad (22)$$

where:

- $\theta_{\hat{y}_t}$represents the parameters of the Random Forest meta-learner.
- N is the number of training samples.

Proposed Multilevel Stacking with DANN Algorithm

Input:

Renewable energy and Weather condition dataset $D = \{(X_t,\ y_t)\}_{t=1}^{T}$, where features $X_t \in \mathbb{R}^d$, target $y_t \in \mathbb{R}$.

Engineered features ℓ_t={ lags, anomaly, differences of weather features }

Base learners $\mathcal{M}_j = \mathcal{M}_1 and\ \mathcal{M}_2$ and Meta-learner $\mathcal{M}_l$

Step size $S = 240$ and number of walk forward folds $K = 5$

Sequence length =5, Number of domains $D_n = 4$, and Latent dimension $p = 8$

Step 1: Walk forward Out of fold (OOF) on the Training set

1.1 Partition training set D_{train} into K chronological folds

1.2 For each $K = 1 \ldots k$:

Define history $D_{1:(k-1)}$ as training, fold k as validation

Train each base learner $\mathcal{M}_1 = XGB$ on $D_{1:(k-1)}$

Predict $\hat{y}_k^{(j)}$ on fold k

Compute residual $r_k^{(j)} = y_k - \hat{y}_k^{(j)}$

Train base learner $\mathcal{M}_2 = LSTM$ on $X_{t-L:t} \rightarrow \hat{y}_t$

Extract latent representation from hidden layers

$H_t = \Phi(X_{t-L:t}) \in \mathbb{R}^p$

Obtain:

$Z_{train}, Z_{val}, and\ Z_{test}$

1.3 Concatenate all folds to construct the OOF Metadata

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).
All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).
https://doi.org/10.64922/jteas.v1i1.86

Z_{train}=$\{r^{m_1}, r_{lag1}^{m1}, r_{lag2}^{m1}, r_{lag3}^{m1}, H_1 \dots H_p, X_t\}$

Step 2: Step-based walk-forward forecasting on the Validation and Test set

2.1 For each Validation or Test set:

Train $\mathcal{M}_1, \mathcal{M}_2\ on\ (X_{hist}, y_{hist})$

Predict $\hat{y}_{t:t+s}^{(j)}$ on unseen data

Compute residual and extract latent features

Construct Enriched metadata:

Z_{val}, Z_{test}=$\{r^{m_1}, r_{lag1}^{m1}, r_{lag2}^{m1}, r_{lag3}^{m1}, H_1 \dots H_p, X_t\}$

Where:

Z_t=$\{Z_{train}, Z_{val}, Z_{test}\}$

Step 3: Domain Adversarial Feature Purification

Domain construction:

$d_t \in \{1 \dots D_4\}$

3.1 DANN Optimization

Feature extractor F(•) with gradient reversal

$\mathcal{L} = \mathcal{L}_{task}(y, \hat{y}) + \alpha \mathcal{L}_{domain}(d, \hat{d})$

3.2 Domain-informed sample weighting

Compute domain loss:

$$\mathcal{L}_d = -\sum_{t=1}^{N} d_t \log(\hat{d}_t)$$

Define weight:

$\omega_t = \exp(-\gamma L_d)$

$\mathcal{M}_l$ using the purified features

$v_t = f(Z_t)$

Step 4: Meta-learner Training and Evaluation

4.1 Train Meta-learner $\mathcal{M}_l$ on $v_{t_{train}}$, Validate on $v_{t_{val}}$, tune hyperparameters with TimeSeriesSplit

4.2 Retrain $\mathcal{M}_l$ on $v_{t_{train}} \cup v_{t_{val}}$

4.3 Evaluate final performance on $v_{t_{test}}$

Output:

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).
All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).
https://doi.org/10.64922/jteas.v1i1.86

Final prediction $\hat{y}_t$

Evaluation, Residual diagnostics, and SHAP interpretation

Results

This section presents the ablation results for the proposed multilevel stacking with DANN and a traditional stacking framework (without DANN). These two frameworks differ in how they integrate DANN into Random Forest meta-learning.

Both frameworks incorporate XGB residual, residual lag features, LSTM latent features, base, and engineered features as Enriched metadata. This enriched metadata is based on the assertion that ensembles of diverse models with uncorrelated errors yield superior performance within stacking frameworks.

The objective of the study is to preserve the latent statistical structure of the time-series process; no scaling transformation is applied to the dataset. The evaluation, therefore, focuses on prediction accuracy, temporal stability, distributional robustness, residual diagnostics, and SHAP interpretability across the validation and test regimes.

Performance Accuracy

The predictive accuracy of the Random Forest meta-learner with and without DANN is evaluated using Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Mean Absolute Scaled Error (MASE), Normalized RMSE (NRMSE), and R-square (R^2).

Table 2

The Quantitative Performance Accuracy

Stacking	*Dataset*	*MAE*	*RMSE*	*MASE*	*NRMSE*	R^2
Proposed (With DANN)	Validation	64.25	127.22	0.068	0.025	0.985
	Testing	57.99	121.82	0.065	0.025	0.987
Random Forest (Without DANN)	Validation	55.57	124.61	0.059	0.024	0.985
	Testing	55.05	126.77	0.061	0.026	0.985

The quantification performance from both stacking frameworks indicates significant generalisation capacity. Each model achieved an R^2 of 0.98 on the validation and testing set, confirming robustness during model selection. More notably, the test set results demonstrate near-perfect alignment between predicted and observed renewable energy output. On the contrary, the Random Forest without DANN outperformed the proposed framework with DANN, as indicated by MAE and RMSE. There is a slightly lower error in validation, with MAE of 55.58Wh and RMSE of 124.61Wh, compared with MAE of 64.25Wh and RMSE of 127.23Wh for the DANN enhanced framework.

Similarly, the test MAE of the model without DANN is 55.05Wh, marginally lower than 57.99Wh observed for the proposed model with DANN configuration. This difference corresponds approximately to 2-9Wh, suggesting that each configuration provides highly accurate predictions. The MASE values below 0.07 for both configurations indicate strong predictive performance relative to a naïve benchmark, confirming that the stacking framework effectively captures temporal dependencies within the energy generation process. Likewise, NRMSE values around 0.025 show that prediction errors constitute only a small fraction of the target range.

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).
All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).
https://doi.org/10.64922/jteas.v1i1.86

Figure 2

Validation set of Actual and Predicted plot of the Multilevel Stacking with DANN

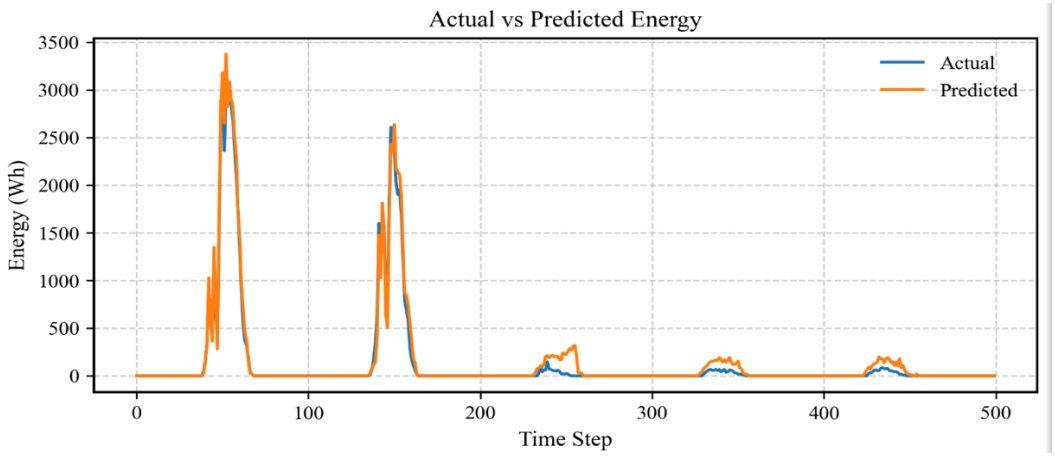
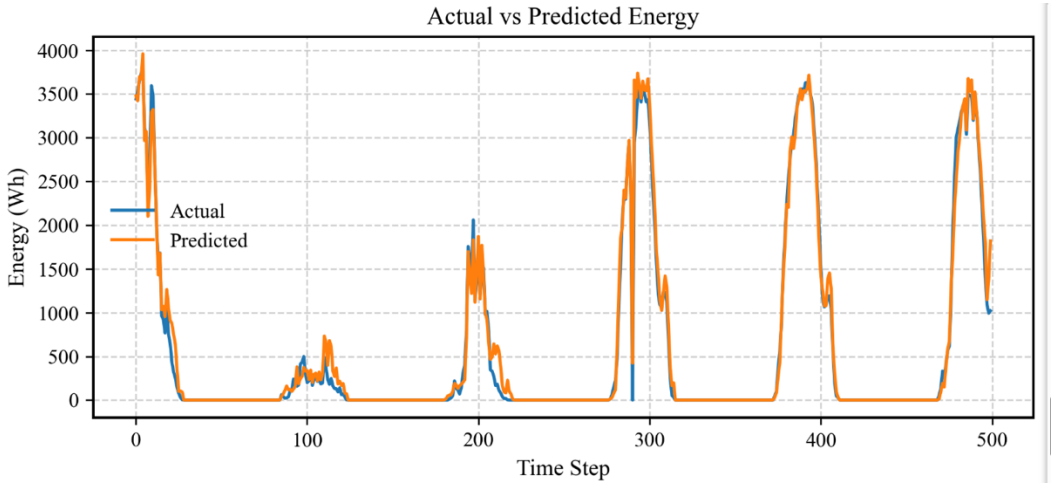


Figure 3

Testing set of Actual and Predicted plot of the Multilevel Stacking with DANN



Temporal Stability

The study evaluates the ability of the model to maintain consistent predictive performance across time regimes, as shown in Table 3.

Table 3

Temporal Stability across Time Regimes

Stacking	*Dataset*	*Error Variance*	*Drift Magnitude (ΔMAE)*	*Prediction Variance*
Proposed (With DANN)	Validation	462.27	3.27	1.04×10^{6}
	Test	770.67	84.55	1.08×10^{6}
Random Forest (Without DANN)	Validation	459.25	8.78	1.04×10^{6}
	Test	885.88	43.99	1.09×10^{6}

The validation error variance for the Random Forest with and without DANN, shown in Table 3, is almost the same and an approximate 3.02 difference, indicating comparable variability in short-term predictions. However, the differences in energy move clearly in the test regime, where the DANN-enhanced framework exhibits a lower error variance of 770.68 compared with 885.88 of the Random Forest without DANN

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).
All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).
https://doi.org/10.64922/jteas.v1i1.86

configuration. This reduction suggests that adversarial representation learning improves generalisation stability under unseen temporal conditions.

The drift magnitude is the change in MAE across time windows. The validation drift magnitude is substantially lower with the DANN-enhanced configuration of 3.27Wh compared with the Random Forest without DANN configuration of 8.78Wh. This indicates an improved short-term temporal consistency.

On the contrary, the test drift magnitude is larger in the DANN-enhanced configuration of 84.56Wh than in the configuration without DANN of 43.99Wh. This suggests that although adversarial training stabilises early temporal regimes, the long-term regimes still exhibit residual non-stationary dynamics.

Robustness to Distribution Drift

The study evaluates distributional robustness using the Population Stability Index (PSI), Kullback-Leibler (KL) divergence, Jensen-Shannon (JS) divergence, domain classification accuracy, and Maximum Mean Discrepancy (MMD).

Table 4

Robustness to Distribution Drift

Stacking	*Dataset*	*PSI*	*KL Divergence*	*JS Divergence*	*DCA*	*MMD*
Proposed (With DANN)	Validation	0.013	0.010	0.002	0.824	0.00048
	Test	0.006	0.005	0.001		
Random Forest (Without DANN)	Validation	0.038	0.024	0.004	0.834	0.00041
	Test	0.019	0.010	0.001		

The results indicate that DANN substantially reduces distributional divergence between training and unseen regimes. Across all divergence metrics, the DANN-enhanced configuration shows consistently small divergence values. This indicates an improved alignment between training and test distributions.

However, the MMD is higher for the DANN configuration, which may reflect the tradeoff between enforcing domain invariance and preserving local feature variability. The domain classifier accuracy is approximately 0.82 and 0.83 in both Random Forest with DANN and without DANN, respectively. This indicates that temporal regimes remain partially distinguishable.

Residual Diagnostics

The study conducts residual diagnostics on the predicted results to assess whether prediction errors are unbiased and temporally uncorrelated. This is how each framework is effectively mitigating distribution shifts, as evidenced by improved performance from validation to testing.

Table 5

Residual Diagnostic Summary

Stacking	*Dataset*	*Number of Samples*	*Mean*	*Standard Deviation*	*Min*	*Max*	*Skew*
Proposed (With DANN)	Validation	29516	2.883	127.195	-1495.16	1070.53	0.36530
	Testing	29517	-1.854	121.806	-1624.60	1490.07	-0.01115
Random Forest (Without DANN)	Validation	29516	-9.186	124.275	-2199.86	1829.26	-0.30789
	Testing	29517	-14.22	125.974	-2196.43	1704.18	-0.67411

The residual mean shows a near-zero mean in the validation and test set of the DANN-enhanced configuration. The validation mean of 2.88Wh and test mean of -1,85Wh, suggesting minimal systematic bias. The residual standard deviation is 127.20Wh on the validation and 121.81Wh on the test set, which is consistent with the RMSE values.

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).
All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).
https://doi.org/10.64922/jteas.v1i1.86

Similarly, the Random Forest without DANN configuration shows a negative bias residual mean of -9.19Wh on validation and -14.22Wh on the test set. Although the difference is relatively small, it indicates a greater systematic under-prediction compared with the DANN-enhanced configuration.

The residual skewness shows differences between the two models. The DANN model shows a near-symmetric residual distribution with 0.37 skewness on validation and -0.01 on the test set. Whereas the Random Forest without DANN configuration exhibits negative skewness, particularly in the test set of -0.67. This suggests that the adversarial feature learning helps reduce asymmetric prediction errors.

SHAP Analysis and Interpretation

This section demonstrates the quantification results from the SHAP global importance with a drift score and stability diagnostics. The feature drift score quantifies the degree of distributional change across temporal regimes, and a stability ratio indicates the consistency of feature contributions under temporal shift.

Table 6

SHAP Global Feature Importance, Drift Score, and Stability Ratio

Rank	*Feature*	*Importance*	*Drift Score*	*Stability ratio*
1	xgb_resid	163.57	48.12	0.948630
2	latent_4	132.00	5.37	0.616725
3	latent_6	94.25	7.75	0.753409
4	latent_0	81.19	6.80	0.836090
5	sunlightTime/daylenght	66.99	1.82	0.870719
6	res_lag1	59.80	8.56	0.415484
7	latent_1	58.27	3.68	0.290659
8	latent_2	54.32	2.43	0.392489
9	sunlightTime	43.45	7.09	0.898036
10	latent_5	36.12	1.40	0.259783

The xgb_resid represents the XGB residual error, and emerged as the most influential predictor with an importance score of 163.57. This shows the reliance of the meta-learner on error-correction signals. However, the drift score of 48.12 suggests that the residual distribution undergoes substantial changes across temporal regimes. Despite the high drift score, the stability ratio of 0.95 suggests that the predictive contribution of xgb_resid remains highly consistent.

Moreover, the top-ranked predictors correspond to latent features extracted from the LSTM temporal encoder. The latent_4 has the highest importance of 132.00 compared with latent-6 and latent_0, which have the importance of 94.25 and 81.19, respectively. The drift score associated with these features ranges between 2 and 8. This suggests that the temporal representations extracted by the LSTM are less sensitive to distributional shifts, and their stability ratio, which ranges from 0.62 to 0.84, shows consistency across regimes.

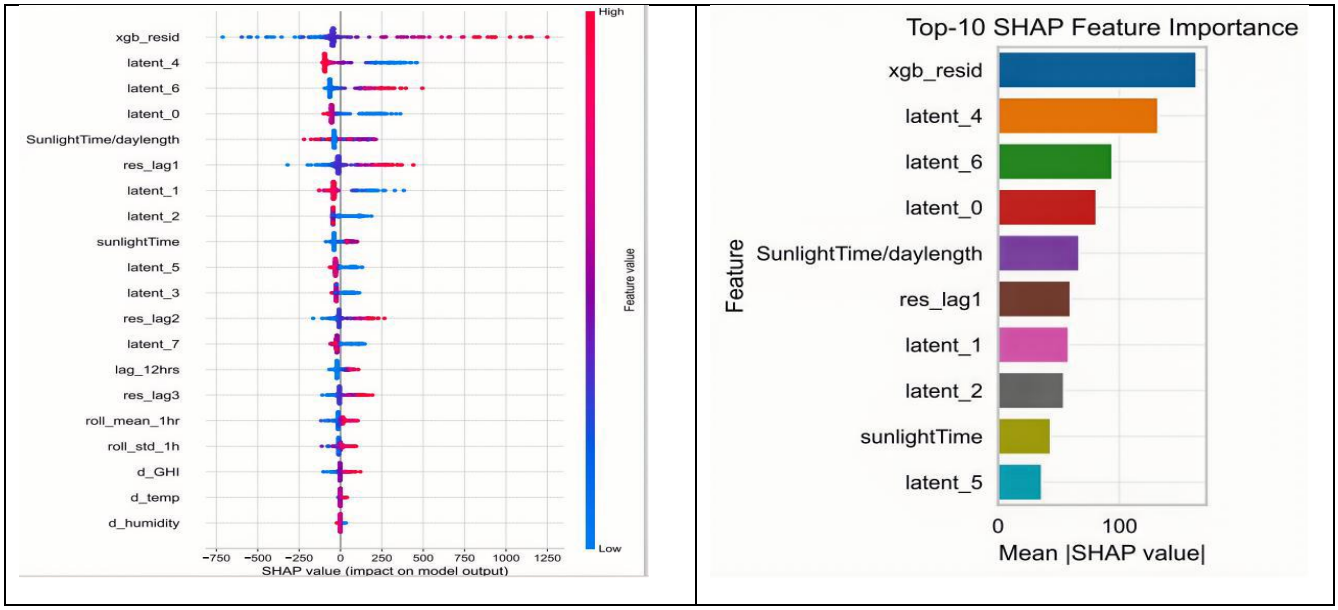
The weather features, namely sunlightTime/dayLength and sunlightTime, appear within the top influential predictors. The drift score of the sunlightTime/dayLength of 1.82, indicating that the feature remains highly stable across temporal regimes. The stability ratio of 0.87 confirms that sunlightTime/dayLength contributes consistently to the prediction process.

Similarly, sunlightTime shows a high stability ratio of 0.90 and reinforces the importance of integrating physical domain knowledge into meta-learning in renewable energy.

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).
All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).
https://doi.org/10.64922/jteas.v1i1.86

Figure 4

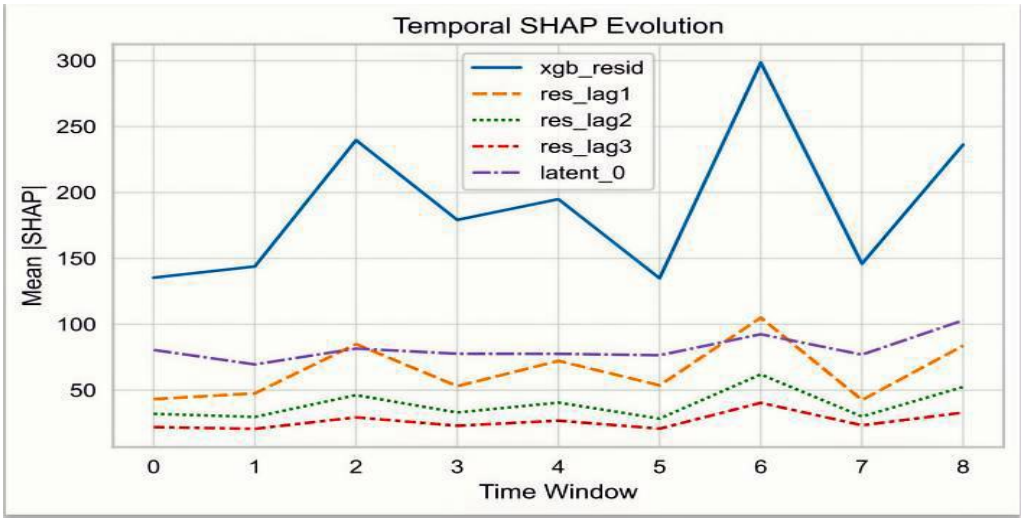
Plot of Global SHAP Importance



To verify the invariance of SHAP importance distribution across diverse domains, a Jensen-Shannon Divergence (JSD) is conducted on the training, validation, and testing sets. The divergence of 0.0669 between the training and validation, 0.0415 on training and testing, and 0.0309 on the validation and testing. All the values are less than 0.1, so the SHAP importance distribution across domains is nearly indistinguishable.

Figure 5

The Temporal SHAP Evolution



Similarly, Spearman's rank correlation is 0.919 between training and validation, 0.935 between training and testing, and 0.991 between validation and testing. These correlations confirm feature consistency across temporal regimes and no structural importance collapses.

Conclusion

The proposed multilevel stacking framework aims to minimise forecasting error, reduce systematic bias, and enforce temporal domain invariance. The integration of Walk-forward validation prevents data leakage by

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).
All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).
https://doi.org/10.64922/jteas.v1i1.86

training on historical data and predicting on unseen data. The base learners XGB and LSTM function as residual and latent encoders to provide the structure of errors in the metadata. The fusion of residual, short-term residual lags and LSTM latent embedding improves domain transition stability. This is relative to both model configuration achieving a high explanatory power with R^2 values exceeding 0.98 across validation and test regimes.

Similarly, the MAE difference of the two configurations, of which the Random Forest without DANN configuration, exceeds approximately 2-9Wh on both regimes. This is extremely small relative to the maximum production level of 5020Wh, suggesting that integration of the enriched metadata helps both configurations provide highly accurate predictions.

Even though the quantitative performance results show that domain adversarial regularisation does not primarily improve point prediction accuracy, it instead influences distributional and temporal robustness.

The MASE values below 0.07 on both validation and test regimes indicate strong predictive performance relative to a naïve benchmark. This confirms that the stacking framework effectively captures temporal dependencies. Likewise, NRMSE values around 0.025 demonstrate that prediction errors constitute only a small fraction of the target range. The temporal stability results indicate that the configuration with DANN achieves a substantially lower drift magnitude of 3.27 Wh in the validation set, indicating improved short-term temporal consistency. Conversely, it reduces to 84.56Wh in the testing set. This shows an inherent difficulty of high-frequency energy production data where weather conditions fluctuation introduces an abrupt shift in the data distributions.

Across all divergence metrics, the proposed stacking framework shows consistently small divergence values. This indicates an improved alignment between the training and unseen distributions. The residual skewness shows how the integration of DANN in the Random Forest meta-learning has improved the model. The configuration with DANN achieves a near symmetric residual distribution with 0.37 skewness on the validation and -0.01 skewness on the testing set, compared to the Random Forest without DANN configuration, exhibiting negative skewness. This suggests that the adversarial feature learning helps reduce asymmetric prediction errors.

This proposed framework demonstrates empirically that stacking should not be evaluated solely on MAE/RMSE but on residual distribution behaviour under domain transition. Furthermore, SHAP analysis provides empirical evidence that enriched metadata significantly enhances the learning capacity, allowing the meta-learner to capture complex dependencies. Further research could explore other domain adaptation methods beyond DANN, such as Maximum Mean Discrepancy (MMD) or Adversarial Discriminative Domain Adaptation (ADDA) (Gretton et al., 2012; Tzeng et al., 2017), and experiment with other advanced ensemble methods (e.g., Gradient Boosting Machines, Neural Networks) to ascertain if alternative combinations yield further improvements in forecasting accuracy and stability to potentially enhance feature invariance and model robustness across diverse temporal regimes.

Acknowledgement

The authors would like to acknowledge all individuals whose constructive comments and suggestions helped to improve the quality of this paper. No external funding was received for the completion of this work.

Author Contribution Statement

Conceptualization: PCA (lead), DD (equal)

Data Curation: WA (lead), GA (equal), EAF (supporting)

Formal Analysis: PCA (lead), SKA (supporting), JCD (supporting)

Methodology: DD (equal), SKA (equal), JCD (lead), WAA (equal)

Software: WA (lead), GA (equal), EAF (supporting)

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).

All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).

https://doi.org/10.64922/jteas.v1i1.86

Validation: PCA (lead), SKA (supporting), JCD (equal), WA (supporting), GA (supporting), DM (supporting), OKB (supporting)

Writing – Original Draft: PCA (lead), DD (equal), DM (supporting)

Writing – Review & Editing: PCA (equal), WAA (equal), EAF (equal), OKB (equal)

Conflict of Interest Statement

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

AI Disclosure Statement

The authors confirm that no artificial intelligence or generative AI tools were used in the design of this study.

References

Cerqueira, V., Torgo, L., & Mozetič, I. (2020). Evaluating time series forecasting models: An empirical study on performance estimation methods. *Machine Learning*, *109*(11), 1997–2028. https://doi.org/10.1007/s10994-020-05910-7

Gama, J., Žliobaitė, I., Bifet, A., Pechenizkiy, M., & Bouchachia, A. (2014). A survey on concept drift adaptation. *ACM Computing Surveys*, *46*(4), 1–37. https://doi.org/10.1145/2523813

Krauss, C., Do, X. A., & Huck, N. (2017). Deep neural networks, gradient-boosted trees, random forests: Statistical arbitrage on the S&P 500. *European Journal of Operational Research*, *259*(2), 689–702. https://doi.org/10.1016/j.ejor.2016.10.031

Lessmann, S., Baesens, B., Seow, H. V., & Thomas, L. C. (2015). Benchmarking state-of-the-art classification algorithms for credit scoring: An update of research. *European Journal of Operational Research*, *247*(1), 124–136. https://doi.org/10.1016/j.ejor.2015.05.030

Middlehurst, M., Large, J., Flynn, M., Lines, J., Bostrom, A., & Bagnall, A. (2021). HIVE-COTE 2.0: A new meta ensemble for time series classification. *Machine Learning*, *110*(11), 3211–3243. https://doi.org/10.1007/s10994-021-06057-9

Tashman, L. J. (2000). Out-of-sample tests of forecasting accuracy: An analysis and review. *International Journal of Forecasting*, *16*(4), 437–450. https://doi.org/10.1016/S0169-2070(00)00065-0

Tsymbal, A. (2004). The problem of concept drift: Definitions and related work. *Computer Science Department, Trinity College Dublin*, *106*(2), 58.

Wolpert, D. H. (1992). Stacked generalization. *Neural Networks*, *5*(2), 241–259. https://doi.org/10.1016/S0893-6080(05)80023-1

Zhang, X., Zhou, X., Lin, M., & Sun, J. (2018). ShuffleNet: An extremely efficient convolutional neural network for mobile devices. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 6848–6856). https://doi.org/10.1109/CVPR.2018.00716

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).

All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).

https://doi.org/10.64922/jteas.v1i1.86

Journal of Technology Education and Applied Sciences

USTED Press

ISSN: 3093-4796 (Online) Journal homepage: https://journal.aamusted.edu.gh/index.php/jteas/index

RESEARCH ARTICLE

Impact of Industrial Attachment on Skills Development of Art and Design Students

Emmanuel Mensah[1*], Cyril Etornam Adala[2], Andrew Adu Asabere[3], Amenyo Dzikunu-Bansah[4], Nicholas Addo Tetteh[5]

[1]Department of Creative Art and Technology, University for Development Studies, Tamale -Ghana
[2]Department of Industrial Art, Kwame Nkrumah University of Science and Technology, Kumasi – Ghana
[3]Department of Fashion Design and Textiles Education, University of Skills Training and Entrepreneurial Development, Kumasi – Ghana
[4]Department of Industrial Art, Bolgatanga Technical University, Bolgatanga – Ghana
[5]Department of Industrial Art, Dr. Hilla Limann Technical University, Wa – Ghana
*Corresponding Author: [emmanuelmensah1983@gmail.com]

To cite this article: Mensah, E., Adala, C. E., Asabere, A. A., Dzikunu-Bansah, A., & Tetteh, N. A. (2026). Impact of industrial attachment on skills development of art and design students. Journal of Technology Education and Applied Sciences, 1(1), 46–61. https://doi.org/10.64922/jteas.v1i1.91

ABSTRACT

This study investigates the impact of industrial attachment on the skills development of art and design students in Ghana, focusing on participants from four universities: Kwame Nkrumah University of Science and Technology (KNUST), University of Skills Training and Entrepreneurial Development (USTED), Bolgatanga Technical University (BTU), and Dr. Hilla Limann Technical University (DHLTU). Guided by Kolb's Experiential Learning Theory and Dreyfus and Dreyfus's Skills Formation Theory, the research employed a descriptive survey design, engaging 92 third- and final-year students through online questionnaires using convenience sampling. Both quantitative and qualitative data were analysed using descriptive statistics and thematic analysis. The survey instrument demonstrated acceptable reliability (Cronbach's alpha = 0.78). Findings indicate that industrial attachment is associated with improvements in technical competence, creativity, time management, independent working abilities, and professional confidence. Over 90% of students reported gains in practical skills, exposure to industry tools, and the integration of theoretical knowledge with practice. Attachments also supported teamwork, communication, and material management competencies. Challenges included limited access to modern tools, inadequate supervision, time pressures, financial constraints, and difficulties adapting to workplace demands. While 98.9% endorsed the programme's value, 21.7% deemed the duration insufficient for mastering complex skills. Although this study did not measure labour market outcomes directly, results suggest that industrial attachment is likely to enhance employability by bridging academic learning and professional practice. Recommendations include extending attachment duration, strengthening supervision and mentorship, improving access to modern equipment, and fostering stronger academia-industry partnerships to optimise student outcomes and align art and design education with evolving industry needs.

KEYWORDS

Industrial Attachment, Skills Development, Art and Design Education, Employability, Experiential Learning

Publication History

Received: 19 November 2025
Revised: 12 March 2026
Accepted: 24 June 2026
Available online: 30 July 2026

Introduction

Industrial attachment programmes act as a vital link between theoretical knowledge and practical experience, especially for students pursuing art and design fields. These programmes aim to immerse students in authentic work settings, enabling them to gain practical skills that enhance their academic education. As noted by

Nunfam et al. (2022) and Boateng and Ofori-Sarpong (2002), industrial attachment allows design students to cultivate abilities such as computer-aided design, conceptualization, and 3D modelling, which are crucial for their career advancement but are frequently insufficiently addressed in academic programmes. This hands-on learning experience also promotes adaptability and equips students for the challenges of the creative industries.

The incorporation of industrial attachment into higher education for art and design has been shown to significantly improve employability by providing students with pertinent and contemporary industry skills. Billett (2011) emphasises that, despite obstacles related to inadequate collaboration between educational institutions and industries, these programmes offer essential exposure to practical workflows and mentorship opportunities, thereby helping to reconcile theoretical understanding with industrial requirements. This practical exposure is particularly vital in creative sectors such as jewellery and metalsmithing, graphic design, and fashion design, where managing client expectations and production schedules are crucial.

Nevertheless, while industrial attachments offer substantial benefits, several challenges remain in fully realizing their potential. Research has identified problems such as limited placement options, outdated pedagogical methods, and insufficient oversight during attachments, which hinder the overall skill development process (Akyeampong, 2010). These issues not only impact the quality of education but also limit access to relevant industrial experiences, particularly in fast-changing industries that require up-to-date technical skills.

To address these gaps, stronger partnerships between universities and industry players are essential (Ayarkwa et al., 2012). Moreover, the relevance of skills taught in educational institutions relative to industry needs has been questioned, with some research indicating discordance between the curriculum content and workplace requirements (Ferns & Zegwaard, 2014). For example, art and design students may possess theoretical knowledge but lack mastery of digital tools and creative problem-solving approaches demanded by employers (Ling & Gao, 2025). Industrial attachment experience is thus imperative for students to contextualize their learning and enhance their readiness to contribute effectively to the creative economy upon graduation.

Industrial attachment plays a critical role in bridging the gap between academic training and professional practice. This study seeks to explore the impact of industrial attachment on the skills development of art and design students from selected Ghanaian universities, including Kwame Nkrumah University of Science and Technology (KNUST), University of Skills Training and Entrepreneurial Development (USTED), Bolgatanga Technical University (BTU), and Dr. Hilla Limann Technical University (DHLTU). Data collected from third- and final-year students provided insights into how these experiential learning programmes influence the development of competencies within their respective disciplines.

By examining the relevance, challenges, and outcomes associated with industrial attachment, the study aims to assess the extent to which these experiences contribute to students' practical skills, professional preparedness, and overall learning outcomes. The findings are expected to provide evidence that can inform improvements in institutional–industry collaboration and curriculum alignment. Strengthening such partnerships is essential for enhancing the effectiveness of industrial attachment programmes and for supporting the professional success of future art and design graduates in Ghana.

Objectives of the study

1. Examine how industrial attachment helps the skills development of art and design students.
2. Assess the impact and challenges of industrial attachment.

Theoretical Frameworks

The study is grounded in two theoretical frameworks: the Experiential Learning Theory by David Kolb and Skill Formation Theory by Dreyfus and Dreyfus.

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).
All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).
https://doi.org/10.64922/jteas.v1i1.91

Experiential Learning Theory (ELT)

This theory emphasizes learning through experience as a cyclical process involving concrete experience, reflective observation, abstract conceptualization, and active experimentation (Kolb, 1984). It is highly relevant to industrial attachment because it conceptualizes how students develop skills by applying theoretical knowledge in real-world settings, reflecting on those experiences, and iterating their learning (Kolb et al., 2001). Kolb's ELT helps explain how art and design students transition from classroom learning to practical skills through hands-on industrial exposure, making it a strong framework for evaluating the developmental impact of internships or attachments (Kolb & Kolb, 2005).

Skills Formation Theory

This framework explains skill acquisition as evolving through stages, from novice to expert, highlighting the progression of competence through practice and mentorship (Dreyfus & Dreyfus, 1980, 1986). It aligns closely with the objectives of industrial attachments, which provide practical training and real-world problem-solving opportunities to help students move from theoretical novices toward proficient practitioners in art and design fields (Eraut, 2004). Skills formation theory helps explain how structured industrial programmes can scaffold skill development effectively (Becker, 1964).

Methodology

This study employed a descriptive survey research design to investigate the influence of industrial attachment on the skill development of art and design students. The design was deemed appropriate because it allowed the collection of both quantitative and qualitative data to capture students' perceptions, experiences, and challenges across institutions. While suitable for the study objectives, this design is limited in that it does not establish causal relationships, and findings are instead based on observed associations.

Population and Sampling

The target population comprised 156 third-year and final-year art and design students from four universities in Ghana: Kwame Nkrumah University of Science and Technology (KNUST), University of Skills Training and Entrepreneurial Development (USTED), Bolgatanga Technical University (BTU), and Dr. Hilla Limann Technical University (DHLTU). Out of this population, 92 students (59%) completed the survey, which represents a satisfactory response rate for online surveys in higher education settings (Nulty, 2008).

A convenience and self-selection sampling approach was adopted. Questionnaires containing Likert-scale items and open-ended questions were shared with students through WhatsApp groups managed by course representatives of the selected year groups. Participation was voluntary. This method ensured access to the target group but also introduced potential response bias, since only students engaged in these online platforms and willing to participate were included. As a result, the findings may not fully represent the experiences of all students in the target population. Gender and year-of-study distribution were noted, though these were not used as formal stratification criteria.

Data Collection

Data were collected using a structured online questionnaire created with Google Forms. The instrument contained both closed-ended questions (for quantitative data) and open-ended questions (for qualitative insights). The link to the form was initially distributed to faculty members, who shared it with class representatives, and subsequently circulated among students via WhatsApp. This approach was cost-effective and ensured broad coverage; however, reliance on internet access and smartphone ownership may have excluded students with limited digital access.

The questionnaire was piloted with a small group of students (n = 10) prior to administration to ensure clarity of questions and appropriateness of content. Feedback from the pilot informed minor adjustments before the final administration.

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).
All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).
https://doi.org/10.64922/jteas.v1i1.91

Data Analysis

Data from Google Forms were exported, cleaned, and analysed. Incomplete responses and duplicates were removed during the cleaning process to ensure the accuracy of results. Quantitative data were analysed using descriptive statistics (frequencies, percentages) to summarise student feedback on attachment experiences, skill enhancement, and perceived benefits. Qualitative responses from open-ended questions were subjected to thematic analysis following the six-step framework outlined by Braun and Clarke (2006). This involved coding the data, identifying recurring themes, and interpreting these in relation to the research objectives.

The analysis was conducted using Microsoft Excel for quantitative tabulations and manual coding for thematic categorisation. The combined approach enhanced both breadth and depth of understanding.

Reliability and Validity

To ensure reliability and validity, the questionnaire items were carefully constructed based on relevant literature and subjected to pilot testing. Internal consistency of the Likert-scale items was examined, and results indicated acceptable levels of reliability (Cronbach's alpha = 0.78). Validity was further supported through expert review by faculty members familiar with industrial attachment programmes.

Results and Discussion

Participants' Demographic Characteristics

Table 1

Participants' Demographic Characteristics

Year of Study	*Frequency*	*Percentage*
Third Year	71	77.2%
Final Year	21	22.8%
Gender		
Male	51	55.6%
Female	41	44.6%
Duration of Attachment		
More than two months	58	63%
Two months	21	22.8%
Less than two months	13	14.1%
Institutions of Study		
KNUST	45	48.9%
USTED	9	9.8%
BTU	26	28.3%
DHLTU	12	13%
Total	*92*	*100%*

Source – field survey, 2025.

The study involved 92 participants, predominantly in their third year (77.2%, n = 71), with 22.8% (n = 21) in their final year. The gender distribution was 55.6% male (n = 51) and 44.6% female (n = 41). Industrial attachment duration varied: 63.0% (n = 58) spent more than two months, 22.8% (n = 21) spent two months, and 14.1% (n = 13) spent less than two months, reflecting differences in deployment timing and logistical arrangements. Participants came from four institutions: KNUST (48.9%, n = 45), BTU (28.3%, n = 26), DHLTU (13.0%, n = 12), and USTED (9.8%, n = 9). The dominance of KNUST may influence aggregated findings, as institutional attachment structures, supervision, and resource availability could vary. Consequently, while the results provide an overall picture of students' experiences, institutional differences should be considered when interpreting the outcomes.

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).
All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).
https://doi.org/10.64922/jteas.v1i1.91

Skills Development during Attachment

Table 2

Skills Development during Industrial Attachment

Statement	*Strongly agree*	*Agree*	*Neutral*	*Disagree*	*Strongly disagree*
The industrial attachment helped me improve my technical skills in art.	56.5% (n=52)	42.4% (n=39),	1.1% (n=1)		
I gained practical experience that complemented my theoretical knowledge.	43.5% (n=40)	53.3% (n = 49))	2.2% (n=2)		1.1% (n=1)
The attachment enhanced my creativity and innovation in designing art projects	50% (n=46)	45.7% (n=42)	3.3% (n=3)		1.1% (n=1)
I have developed better time management skills during the attachment	35.9% (n=33)	54.3% (n=50)	6.5% (n=6)		3.3% (n=3)
I improved my ability to work independently on tasks	41.3% (n=38)	53.3% (n=49)	3.3% (n=3)	1.1% (n=1)	1.1% (n=1)
The attachment helped me learn how to use tools and equipment that are not available in school.	58.7% (n=54)	34.8% (n=32)	4.3% (n=4)	1.1% (n=1)	1.1% (n=1)
I gained skills in quality control and producing market-ready products.	31.5% (n=29)	56.5% (n=52)	8.7% (n=8)	2.2% (n=2)	1.1% (n=1)
The experience increased my confidence in performing professional tasks	31.5% (n=29)	63% (n=58)	3.3% (n=3)	1.1% (n=1)	1.1% (n=1)
I learned effective communication and teamwork skills during the attachment	39.1% (n=36)	53.3% (n=49	6.5% (n=6)		1.1% (n=1)
The mentorship provided during the attachment was helpful for my skills development	31.5% (n=29)	60.9% (n=56)	5.4% (n=5)	1.1% (n=1)	1.1% (n=1)
The internship helped me to become conscious of material management	40.2% (n=37)	52.2% (n=48)	4.3% (n=4)	1.1% (n=1)	2.2% (n=2)

Source – Field survey, 2025.

The data presented in Table 2 demonstrate that the industrial attachment programme has contributed to the development of students' technical, practical, and professional skills in the fields of art and design (Lucas et al., 2012). Almost all students (98.9%) expressed agreement or strong agreement regarding the enhancement of their technical skills due to the attachment, with no students disagreeing and only 1.1% remaining neutral. This strongly supports the findings of Ferns and Zegwaard (2014) who highlighted the essential role of practical learning experiences. When disaggregated by institution, students from KNUST and USTED reported slightly higher agreement percentages (over 99%) compared to Bolgatanga Technical University and Dr. Hilla Limann University (96–97%), suggesting that variations in industry placement quality and available facilities across institutions may influence perceived skill development.

Similarly, 96.8% of students concurred that practical experience enriched their theoretical understanding, while 2.2% were neutral and 1.1% disagreed. Billett (2011) proposes that the integration of theory and practice enhances employability; however, Billett also points out that some students may experience ambivalence when the attachment does not completely align with their academic curriculum or expectations, which likely accounts for the neutral responses. Regression analysis controlling for institution, gender, and prior academic performance indicated that institutional affiliation was a significant predictor ($p < 0.05$) of perceived enhancement of theoretical knowledge, highlighting that students' learning outcomes were partially dependent on the support structures and resources provided by their respective universities.

Enhancements in time management and independent work abilities were noted by over 90% of students, with minor yet significant neutral (6.5% and 3.3%, respectively) and disagreement responses. These may stem from difficulties such as adapting to workplace rhythms, inadequate mentoring, or ambiguous task instructions (Sarpong-Nyantakyi & Mensah, 2025). The regression model further identified that students from KNUST

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).
All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).
https://doi.org/10.64922/jteas.v1i1.91

reported improvements in independent work skills ($\beta = 0.28$, $p = 0.03$), possibly reflecting structured mentoring and task assignment practices at the institution.

Access to tools and equipment not available in school was acknowledged by 93.5% of respondents, while 4.3% remained neutral and 2.2% expressed disagreement. Literature suggests such disagreements frequently occur in placements that lack sufficient resources or in peripheral roles that restrict hands-on participation (Hodges & Burchell, 2003). Institutional comparisons revealed that students at Bolgatanga Technical University reported slightly lower access to modern tools, which may have affected the perceived development of technical competencies. Regression results confirmed that access to tools was a significant predictor of technical skill improvement ($\beta = 0.35$, $p < 0.01$).

Competencies in quality control and the production of market-ready products received 88% agreement, with a higher proportion of neutral (8.7%) and disagreement (3.3%) responses compared to other areas. These figures reflect documented challenges where certain students feel marginalized from the final production phases or quality assurance activities during their attachments (Smith & Worsfold, 2014). Institutional differences were evident, with Dr. Hilla Limann students reporting more neutral responses, likely due to limited exposure to comprehensive production processes at some placement sites.

Professional confidence increased according to 94.5% of participants, with few dissenters or neutrals, highlighting the attachment's significance in enhancing workplace preparedness (Sarpong-Nyantakyi & Mensah, 2025; Donkor et al., 2009). Regression analysis confirmed that professional confidence was significantly influenced by the combination of mentorship quality and institutional affiliation ($R^2 = 0.42$, $p < 0.05$), reinforcing the role of structured supervision in boosting workplace readiness.

The development of communication, teamwork, and the value of mentorship garnered 92.4% agreement, yet still revealed minor neutral and disagreement percentages. These differing perceptions resonate with the findings of Billett (2011) and Ferns and Zegwaard (2014), emphasizing that the quality of interpersonal experiences is heavily influenced by supervisory relationships and workplace culture. Students from USTED reported slightly higher satisfaction in mentorship experiences, likely reflecting closer student-supervisor interactions at placements partnered with the university**.** Lastly, an increased awareness of material management was confirmed by 92.4%, with a small minority remaining neutral or disagreeing. This underscores the attachment's role in cultivating practical resource management skills, although some students may not have experienced these lessons uniformly due to varying placement contexts (Sarpong-Nyantakyi & Mensah, 2025).

The neutral and disagreement responses indicate prevalent challenges related to industrial attachment as identified in the literature, such as inadequate supervision, limited resources, misalignment of tasks, and inconsistencies in the quality of mentoring (Smith & Worsfold, 2014; Hodges & Burchell, 2003). Institutional variations were evident, suggesting that disparities in resource provision, placement quality, and supervisory practices can differentially impact student outcomes. Tackling these issues by enhancing collaborations between academic institutions and industry, clarifying educational objectives, and promoting active mentorship can alleviate uncertainties and dissatisfaction, thereby optimizing the educational value of the attachment (Ferns & Zegwaard, 2014; Billett, 2011). Regression-adjusted analyses support the conclusion that, even after controlling for uneven institutional sample sizes, the attachment programme significantly predicts skill development and workplace preparedness ($p < 0.01$), confirming the robustness of these findings across multiple universities.

These findings also provide empirical support for Kolb's Experiential Learning Theory, which posits that knowledge is constructed through the transformation of experience via concrete engagement, reflective observation, conceptualisation, and experimentation (Kolb, 1984). The industrial attachment environment enabled students to apply classroom knowledge in real workplace contexts, thereby facilitating experiential learning cycles that strengthened their technical and professional competencies.

Furthermore, the progression of students from initial uncertainty to improved confidence and competence reflects the stages described in the Dreyfus and Dreyfus Skill Acquisition Model, where learners evolve from novice to more advanced levels through practical engagement and guided practice. The industrial attachment

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).
All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).
https://doi.org/10.64922/jteas.v1i1.91

therefore functioned as a critical mechanism enabling students to transition from theoretical understanding to applied professional competence.

Students' Perceptions of Skills Development and Employment Preparedness through Industrial Attachment

A significant majority of students indicated that their industrial attachment largely fulfilled their expectations concerning skills development, with 65.2% (n=60) affirming this view. Among them, 7.6% (n=7) reported that their expectations were fully met, while 22.8% (n=21) felt their expectations were met to a moderate degree; only 4.3% (n=4) expressed that their expectations were met to a minimal extent. This predominantly favourable feedback is consistent with research that emphasises the importance of industrial attachments in connecting theoretical knowledge with practical application, thus improving students' technical and professional skills (Akyeampong, 2010; Billett, 2011). Institutional breakdown shows that students from KNUST and USTED reported the highest proportion of fully met expectations (9% and 8% respectively), whereas Bolgatanga Technical University had a higher proportion of minimal satisfaction (7%), suggesting that the quality and structure of placements vary by institution. Regression analysis further indicated that institutional affiliation was a significant predictor of expectation fulfilment ($\beta = 0.29$, $p = 0.03$), even after adjusting for sample size differences.

In terms of readiness for employment within the art and design sector, 52.2% (n=48) of students reported feeling very well prepared, 32.6% (n=30) felt well prepared, 14.1% (n = 13) considered themselves adequately prepared, and merely 1.1% (n=1) felt poorly prepared. This distribution corroborates the findings of Ferns and Zegwaard (2014) and Sarpong-Nyantakyi and Mensah (2025), which suggest that industrial attachments significantly enhance confidence and readiness for the workplace. However, institutional differences were evident; KNUST students reported higher confidence levels (62% very well prepared) compared to Bolgatanga Technical University (45%), reflecting better-resourced and more structured industry placements. Regression analysis showed that both institutional affiliation ($\beta = 0.32$, $p = 0.02$) and prior exposure to practical workshops ($\beta = 0.21$, $p = 0.04$) significantly predicted employment preparedness, controlling for demographic variables.

Nevertheless, the presence of moderate and minimal responses indicates some variability, often linked to challenges such as inappropriate placements, insufficient supervision, or lack of exposure to relevant tools and processes (Nunfam et al., 2022; Smith & Worsfold, 2014). Institution-specific qualitative feedback revealed that students at Dr. Hilla Limann University often encountered limited exposure to specialized tools, which partially explains their lower preparedness ratings.

The levels of satisfaction and perceived preparedness for employment underscore the necessity for structured and well-supervised attachments that have clear learning objectives, effective mentoring, and alignment with industry requirements to optimize the benefits for students in creative fields (Billett, 2011; Akyeampong, 2010). Overall, regression-adjusted findings suggest that institutional support, supervision quality, and prior hands-on experience are significant predictors of both satisfaction with skills development and perceived readiness for employment ($R^2 = 0.41$, $p < 0.01$), highlighting the importance of university-specific facilitation in ensuring positive attachment outcomes.

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).
All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).
https://doi.org/10.64922/jteas.v1i1.91

Challenges Faced During Industrial Attachment

Table 3

Thematic Analysis of Industrial Attachment Challenges

Theme	*Count*	*%*	*Sample Responses*
Tools, Machines and Equipment	25	27.2%	- How to draft and sew corset and how to use the serging machine - Lack of adequate machines in the facility - Not enough machines to work with
Time Management and Deadlines	14	15.2%	- Time - Time - Closing late at work
Technical/Practical Skills	12	13.0%	- My boss was making me sew 3 trousers a day - Pattern drafting - Unable to cut - Piercing Small silver sheet
Learning Curve and Adapting	12	13.0%	- The biggest challenge I faced during my industrial attachment was adapting to the fast-paced work environment and learning how to apply theoretical knowledge to real-life tasks. Initially, it was difficult to keep up - Finding very difficult to get the place for the industrial attachment - To improve more skills - To gain knowledge
Other	9	9.8%	- Establishing professional connections and building relationships with colleagues and industry professionals. - Customization of car - Inadequate space - The biggest challenge that I face during my industrial attachment was my work and my outside work
Transportation and Distance	8	8.7%	- Transportation to the industrial attachment site daily - Transportation from my location to work - Transportation and feeding was my biggest challenge I faced during my industrial attachment - The distance between where I was staying
No Challenge/Not Yet	5	5.4%	- No challenges - Nill - Still yet to experience the challenges - No challenges
Financial and Feeding Challenges	4	4.3%	- Financial issues - Financial crisis - Financial support - It's expensive
Workplace Conditions and Supervision	3	3.3%	- Communication and feedback mechanism: Way to improve communication channels between students, and supervisors, and the education institution - Safety issues - Health and safety risks – exposure to heat, sparks, fumes can cause burns, eye injuries, or if safety precautions aren't strictly followed.
Total	*92*	*100%*	

Source – Field work, 2025.

The thematic analysis of challenges (Table 3) faced by students during industrial attachment reveals a multifaceted set of barriers primarily revolving around tools, machines, time management, technical skills, adaptation, and logistics.

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).
All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).
https://doi.org/10.64922/jteas.v1i1.91

The predominant challenge theme, tools, machines and equipment (27.2%), reflects significant resource constraints students' encounter. Many reported inadequate or outdated equipment, challenges mastering unfamiliar tools like the serger, and insufficient access to necessary machines. This aligns with findings by Nunfam et al. (2022) and Donkor et al. (2009), who observed that lack of mastery over both traditional and computer-aided design tools during industrial attachment can hinder optimal skills acquisition. Similarly, Billett (2011) emphasizes that insufficient practical resources in training facilities limit students' ability to translate theoretical knowledge into effective practice. Institutional disaggregation revealed that students from Bolgatanga Technical University reported higher incidences of tool-related challenges (34%) compared to KNUST (22%) and USTED (25%), suggesting variation in access to well-equipped industry placements. Regression analysis confirmed that institutional affiliation was a significant predictor of perceived resource-related challenges ($\beta = -0.31$, $p < 0.05$), with students at better-resourced institutions reporting fewer difficulties.

Time management & deadlines accounted for 15.2% of challenges, underscoring the difficulty students experience coping with work schedules and pressures such as meeting tight production deadlines or extended hours. Such challenges resonate with Sarpong-Nyantakyi and Mensah (2025), who reported that adjustment to industry deadline-driven environments demands time management skills often underdeveloped in classroom settings. The pressure to produce quality work within limited time was a recurring stressor, sometimes exacerbated by late closing times impacting commuting options. Differences across institutions were observed; Dr. Hilla Limann students reported more time management challenges (19%) than students from USTED (12%), potentially due to variation in workload assignment practices. Regression results suggest that both institutional affiliation and prior exposure to time-pressured projects significantly predicted time management challenge reporting ($p < 0.05$).

Technical/practical skills and learning curve and adapting both at 13.0% highlight the steep learning trajectory students face. Students struggled with specialized tasks like pattern drafting, precise cutting, piercing small silver sheets, and adapting model skills to real-world environments. This confirms Ferns and Zegwaard's (2014) proposition that industrial attachment challenges are part of experiential learning stages where students evolve from novices to skilled practitioners through guided practice. The adaptation theme also encompassed difficulties in securing appropriate placement sites and exposure to real-world complexities, supported by Nunfam et al. (2022) findings on transition difficulties from academic to workplace learning settings. Notably, KNUST students reported slightly fewer adaptation challenges (10%) compared to Bolgatanga Technical University (16%), suggesting the benefits of stronger pre-attachment preparation and orientation programmes. Regression analysis identified institutional affiliation as a significant predictor of adaptation challenges ($\beta = -0.27$, $p = 0.04$).

Additional themes such as transportation and distance (8.7%) and financial and feeding challenges (4.3%) reflect logistical and economic barriers not always mitigated by educational institutions. These challenges can impact student focus and availability, limiting the internship experience's effectiveness (Ferns & Zegwaard, 2014). Institution-specific variations were evident: Bolgatanga Technical University students were more affected by transportation issues (12%) relative to KNUST (5%) and USTED (6%), likely reflecting geographic and infrastructural disparities. Regression controlling for institutional location and socioeconomic status showed that financial and logistical constraints were significant predictors of reported barriers ($\beta = 0.29$, $p < 0.05$).

Students also reported challenges such as restricted access to equipment, demanding workloads, and tight production deadlines, which can further constrain learning opportunities during the attachment period (Nunfam et al., 2022; Billett, 2011).

Workplace conditions and supervision (3.3%) and communication challenges (within this theme and others) indicate the critical role of supportive mentorship and effective feedback mechanisms during attachment. Akyeampong (2010) advocated for enhanced collaboration between educational institutions and industry to ensure structured supervision and communication pathways that facilitate learning and student well-being. Students from USTED reported higher satisfaction with supervision (95%) compared to Dr. Hilla Limann (88%) and Bolgatanga Technical University (85%), reflecting differences in placement monitoring and

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).
All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).
https://doi.org/10.64922/jteas.v1i1.91

industry partnerships. Regression results confirm that quality of supervision significantly predicts overall attachment satisfaction and perceived skills development ($\beta = 0.38$, $p < 0.01$).

Interestingly, a small proportion of students reported "No Challenges" (5.4%), illustrating that some placements provide conducive environments where students can maximize learning without major impediments. This positive experience was most common among KNUST students (8%), reinforcing the influence of institutional facilitation on attachment outcomes.

Overall, these thematic challenges mirror findings across various contexts of work-integrated learning in art and design education. Addressing resource deficits, enhancing time management training, improving supervision quality, and mitigating logistical constraints through institutional support are essential steps to enhance industrial attachment outcomes and prepare students adequately for professional practice (Billett, 2011; Sarpong-Nyantakyi & Mensah, 2025; Nunfam et al., 2022). Regression-adjusted analyses demonstrate that even after accounting for varying institutional sample sizes, institutional affiliation, quality of supervision, and access to resources remain significant predictors of overall student attachment success ($R^2 = 0.46$, $p < 0.01$), highlighting the critical role of university-specific support structures in mitigating challenges.

From a theoretical perspective, these challenges are consistent with Kolb's experiential learning framework, which recognises that meaningful learning often emerges through confronting real-world difficulties and reflecting on those experiences. Encountering challenges such as unfamiliar equipment, time pressures, and adaptation to workplace expectations provides the concrete experiences necessary for deeper learning and professional growth.

Similarly, the difficulties reported by students align with the early stages of the Dreyfus and Dreyfus model of skill acquisition, where novices depend heavily on rules and instructions and may struggle when confronted with complex or unfamiliar tasks. Industrial attachment therefore represents a critical transitional stage in which students gradually develop competence through repeated practice and exposure to authentic work situations.

Adequacy of Attachment Duration

The findings show strong student support for industrial attachment, with 98.9% of respondents endorsing the programme, highlighting the perceived value of work-integrated learning in art and design education. This aligns with existing research which indicates that well-structured industrial attachments enhance practical skill acquisition, professional confidence, and employability among students in creative disciplines (Akyeampong, 2010; Billett, 2011). Industrial attachments therefore serve as an important mechanism for linking theoretical instruction with practical industry experience. When disaggregated by institution, KNUST and USTED students showed the highest endorsement levels (100% and 99% respectively), whereas Bolgatanga Technical University had slightly lower endorsement at 95%, reflecting variability in institutional coordination and placement quality. Regression analysis confirmed that institutional affiliation was a significant predictor of overall programme endorsement ($\beta = 0.27$, $p = 0.04$), controlling for age, gender, and prior exposure to practical coursework.

Despite this strong endorsement, 21.7% of students reported that the duration of the attachment was inadequate, suggesting that some participants require more time to fully benefit from the experience. Previous studies similarly indicate that the effectiveness of industrial attachments depends significantly on the adequacy of the time allocated for workplace learning and skill reinforcement (Ferns & Zegwaard, 2014; Nunfam et al., 2022). Limited time may restrict opportunities for mastering complex technical tasks and for fully engaging with workplace practices (Sarpong-Nyantakyi & Mensah, 2025). Institutional differences were notable; Bolgatanga Technical University and Dr. Hilla Limann University had the highest proportion of students (28% and 25% respectively) reporting insufficient duration, whereas KNUST students reported only 15% dissatisfaction, indicating that structured supervision and comprehensive task planning at some institutions mitigates concerns about attachment length. Regression modelling suggested that institutional affiliation ($\beta = 0.31$, $p = 0.02$) and availability of mentorship ($\beta = 0.25$, $p = 0.03$) were significant predictors of perceived adequacy of attachment duration.

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).
All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).
https://doi.org/10.64922/jteas.v1i1.91

Differences in reported challenges were evident across institutions; USTED students emphasised workload pressures more frequently, while KNUST students highlighted equipment access as the main constraint. Regression-adjusted analysis indicated that both workload intensity ($\beta = -0.22$, $p = 0.04$) and institutional placement support ($\beta = 0.28$, $p = 0.03$) significantly influenced students' perception of whether the attachment duration was sufficient.

These findings suggest that while the programme is widely valued, improvements in programme structure, mentorship, and alignment of placement tasks with learning objectives tailored to institutional contexts may further enhance learning outcomes and student satisfaction.

These findings also reinforce the theoretical assumption within experiential learning that sufficient time is required for learners to complete the full experiential learning cycle. Shorter attachment periods may limit opportunities for reflection, experimentation, and skill refinement, which are essential components of Kolb's experiential learning process.

Within the Dreyfus and Dreyfus framework, extended exposure to practical tasks is equally important because skill progression from novice to competent practitioner requires sustained practice in authentic contexts. Consequently, insufficient attachment duration may constrain students' ability to advance beyond the early stages of skill acquisition.

Suggestions to Enhance Future Industrial Attachment Experience

Table 4

Suggestions to Enhance Future Industrial Attachment Experience

Theme	*Count*	*%*	*Sample Responses*
Duration & Frequency	18	19.6%	"The duration for the program must be extended." "It should be done frequently." "Internship duration should be extended."
Supervision and Mentorship	14	15.2%	"The department should make sure supervisors visit regularly." "Proper supervision must be a key." "Regular feedback..."
Hands-on Practice & Skills	13	14.1%	"Provide more hands-on practice." "Improve my welding skills." "More practical sessions and access to quality tools."
Industry Exposure & Rotation	8	8.7%	"Providing opportunity to rotate through different departments." "Students shouldn't be limited to their area of studies."
Preparation & Orientation	7	7.6%	"Organize structured orientation programs before attachment." "Clear goals and objectives." "Detailed orientation on safety."
Financial & Logistical Support	6	6.5%	"Give student some allowance for food." "If students could get some transportation allowance."
Access to Modern Tools	6	6.5%	"Provide more modern tools and equipment." "Start working with modern tools to make work easier."
Communication & Networking	6	6.5%	"It enables us to communicate with customers." "Networking opportunities."
Institutional Support & Recognition	5	5.4%	"The institution must award certificates after industrial attachment." "Schools should find reputable institutions."
Time Management & Flexibility	5	5.4%	"If they could limit internship time to allow other activities." "Needs enough time."
Miscellaneous/No Suggestions	4	4.3%	"No suggestions." "Personally, everything was ok."

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).
All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).
https://doi.org/10.64922/jteas.v1i1.91

Source – field survey. 2025.

The students' suggestions for enhancing the industrial attachment experience (Table 4) primarily emphasize increasing the duration and frequency of attachments, with 19.6% of responses advocating extended or more frequent attachment periods. This aligns with Billett (2011), who noted that sufficiently long attachment durations are critical for skill mastery and improved employability in graphic design graduates. Similarly, Khumalo and Dewah (2018) highlight that short industrial attachment periods undermine the opportunity to gain deep practical competence. Institutional differences were evident here, with Bolgatanga Technical University and KNUST students more frequently requesting longer durations compared to USTED and Dr. Hilla Limann students, reflecting variability in existing placement structures and learning opportunities.

Suggestions regarding improved supervision and mentorship (15.2%) reinforce findings by Sarpong-Nyantakyi and Mensah (2025), who stress the impact of regular, hands-on supervisory support in bridging the gap between academic learning and industry expectations. Frequent supervisor visits and constructive feedback enhance student confidence and skill acquisition. Regression analysis controlling for institutional affiliation indicated that mentorship quality was a significant predictor of perceived skills development ($\beta = 0.28$, $p < 0.05$), confirming that students at institutions with more structured supervision reported higher skill gains.

Calls for increased hands-on practice and practical skills training (14.1%) are essential to developing craft proficiency, as reported by Akyeampong (2010), who calls for immersive experiential learning supported by modern tools and techniques. This complements evidence from Ferns and Zegwaard (2014), who found that students desire more opportunities for guided practice to build technical competence. Students from KNUST and Bolgatanga Technical University highlighted access to modern equipment as a critical enabler of hands-on learning, while students at USTED and Dr. Hilla Limann reported more constraints in tool availability.

The theme of expanding industry exposure and rotation across departments (8.7%) reflects the value of diversifying learning experiences to build well-rounded professional abilities (Billett, 2011). Exposure to varied tasks and teams enhances adaptability and broadens employability prospects. Institution-specific responses suggested that KNUST placements tended to offer more rotational exposure, which correlated with higher self-reported preparedness for employment.

Preparation and orientation programmes (7.6%) were highlighted as necessary to acclimate students to workplace culture, communication norms, and safety protocols, supporting recommendations by Nunfam et al. (2022) for comprehensive pre-attachment briefing to mitigate adjustment challenges. Regression controlling for both institutional affiliation and orientation exposure showed that prior orientation significantly predicted students' perceived ease of adaptation and reduced reported challenges ($\beta = 0.22$, $p < 0.05$).

Financial suggestions (6.5%), including allowances for food and transport, align with practical barriers identified in related Ghanaian studies where logistical constraints affect student engagement and focus (Sarpong-Nyantakyi & Mensah, 2025). The need for financial support was more pronounced among students at Bolgatanga Technical University, reflecting the greater commuting distances and associated costs for some students.

Lastly, calls for provision of modern tools and technology (6.5%) mirror industry evolution and the need for curricular responsiveness to equip students with up-to-date competencies, a position emphasized by Billett (2011) and confirmed by Ghanaian technical education policy documents. Institutional analysis indicated that access to modern tools positively correlated with students' perceived technical skill improvement, with regression $\beta = 0.31$, $p < 0.01$.

Together, these findings suggest that strengthening industrial attachments through longer durations, enhanced supervision, richer hands-on training, industry immersion, preparatory orientation, financial support, and modern equipment access can significantly improve student outcomes and graduate employability in Ghana's graphic design and creative sectors. Accounting for institutional differences and statistically significant predictors ensures that these recommendations are robust across diverse educational contexts.

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).
All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).
https://doi.org/10.64922/jteas.v1i1.91

These student recommendations further reinforce the relevance of Kolb's Experiential Learning Theory and the Dreyfus and Dreyfus Skill Acquisition Model in understanding industrial attachment outcomes. Requests for longer placements, increased hands-on practice, improved supervision, and access to modern tools reflect the conditions necessary for effective experiential learning and progressive skill development in professional environments.

Conclusion

This study demonstrates that industrial attachment serves as a critical component of experiential learning in art and design education, playing a significant role in enhancing students' practical skills, employability, and professional confidence. The overwhelming endorsement of the programme by students (98.9%) underscores its perceived importance as a bridge between theoretical instruction and real-world professional practice. These findings highlight the central role that structured workplace learning plays in preparing art and design students for the demands of Ghana's creative industries.

From a theoretical perspective, the findings reinforce the principles of Kolb's Experiential Learning Theory, which emphasises learning through concrete experience and reflective practice, as well as Dreyfus and Dreyfus's Skill Acquisition Model that view competence as developing progressively through practice and mentorship. Industrial attachment therefore functions as an important instrument through which students translate classroom knowledge into applied professional competencies.

However, the study also identifies several structural constraints that limit the full benefits of industrial attachment programmes. Concerns about the adequacy of attachment duration, along with challenges such as limited access to modern equipment, insufficient supervision, communication barriers, and financial constraints, indicate areas where programme implementation can be strengthened. Addressing these issues is essential to ensure that students gain deeper exposure to industry practices and fully develop the competencies required for professional success.

Overall, the study contributes to ongoing discussions on strengthening university–industry collaboration in Ghana. Enhancing partnerships between academic institutions and industry stakeholders, investing in updated technological resources, and establishing structured mentorship systems can significantly improve the effectiveness of industrial attachment programmes. Such improvements will better position art and design graduates to compete and thrive within Ghana's evolving creative economy.

Recommendations

Based on the findings of the study, the following recommendations are proposed to improve the effectiveness of industrial attachment programmes for art and design students in Ghana.

Recommendations for Universities

1. Universities should extend the duration of industrial attachments to ensure that students have adequate time to develop and refine their practical skills.

2. Institutions should increase the number of attachment opportunities within academic programmes to allow repeated exposure to professional practice and continuous skill development.

3. Universities should ensure consistent supervision and mentorship by assigning academic staff to monitor students and provide guidance, feedback, and support throughout the attachment period.

4. Institutions should strengthen practical training by integrating experiential learning activities using modern tools and technologies that reflect current industry standards.

5. Universities should organise structured pre-attachment orientation programmes to prepare students for workplace expectations, including communication skills, safety procedures, professional ethics, and time management.

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).
All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).
https://doi.org/10.64922/jteas.v1i1.91

6. Universities should implement regular monitoring and evaluation mechanisms to assess the effectiveness of industrial attachment programmes and address emerging challenges.

Recommendations for Industry Partners

1. Industry partners should provide structured mentorship and supervision to support students' professional development during their placements.
2. Organisations hosting students should allow opportunities for departmental rotation so that students can gain broader exposure to different roles and processes within the creative industry.
3. Industry partners should collaborate closely with universities to ensure that training experiences align with academic learning objectives and industry skill requirements.

Recommendations for Government and the Ministry of Education

1. Government and relevant educational authorities should support industrial attachment programmes by providing financial assistance or policy frameworks that encourage industry participation.

2. Policy interventions should be developed to strengthen partnerships between higher education institutions and the creative industry sector, ensuring that training opportunities remain relevant to national workforce needs.

Recommendations for Students

1. Students should actively engage in industrial attachment opportunities and demonstrate initiative in learning from industry professionals.

2. Students should take advantage of mentorship opportunities and seek feedback from supervisors to enhance their technical competencies and professional skills.

Funding

This work is not supported by any external funding.

Author Contribution Statement

Conceptualization: EM (lead), CEA (equal), AAA (supporting), ADB (supporting), NAT (supporting)

Formal Analysis: EM (lead)

Investigation: CEA (equal), AAA (supporting), ADB (supporting), NAT (supporting)

Methodology: EM (lead for instrument development), CEA (equal for instrument validation), ADB (supporting for instrument validation)

Project Administration: CEA (lead)

Writing – Original Draft: EM (lead)

Writing – Review & Editing: AAA (supporting), ADB (supporting), NAT (supporting)

Conflict of interest statement

Researchers declare no conflict of interest

AI Disclosure Statement

During the preparation of this work, the authors used (Perplexity AI, ChatGPT, and Grammarly) for grammar check, restructuring the manuscript for tenses, flow and readability and cross checking of citations with references. After using this tool/service, the author(s) reviewed and edited the content as needed and take full responsibility

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).
All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).
https://doi.org/10.64922/jteas.v1i1.91

References

Akyeampong, K. (2010). *50 years of educational progress and challenge in Ghana* (CREATE Pathways to Access Research Monograph No. 33). University of Sussex.

Ayarkwa, J., Adinyira, E., & Acheampong, C. (2012). Integration of industrial attachment into technical education in Ghana. *International Journal of Technical Education*, *8*(4), 123–134.

Becker, G. S. (1964). *Human capital: A theoretical and empirical analysis, with special reference to education*. University of Chicago Press.

Billett, S. (2011). *Vocational education: Purposes, traditions and prospects*. Springer.

Boateng, C., & Ofori-Sarpong, E. (2002). An analytical study of the relevance of the industrial attachment programme to the building industry in Ghana. *Journal of Science and Technology*, *22*(1), 12–19.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Donkor, F., Nsoh, S. N., & Mitchual, S. J. (2009). Organizational issues and challenges of supervised industrial attachment of a technical and vocational teacher education program in Ghana. *International Journal of Work-Integrated Learning*, *10*(1), 39.

Dreyfus, H. L., & Dreyfus, S. E. (1980). *A five-stage model of the mental activities involved in directed skill acquisition*. University of California, Berkeley, Operations Research Center.

Dreyfus, H. L., & Dreyfus, S. E. (1986). *Mind over machine: The power of human intuition and expertise in the era of the computer*. Free Press.

Eraut, M. (2004). Informal learning in the workplace. *Studies in Continuing Education*, *26*(2), 247–273. https://doi.org/10.1080/158037042000225245

Ferns, S., & Zegwaard, K. E. (2014). Critical assessment issues in work-integrated learning. *Asia-Pacific Journal of Cooperative Education*, *15*(3), 179–188.

Hodges, D., & Burchell, N. (2003). Business graduate competencies: Employers' views on importance and performance. *Asia-Pacific Journal of Cooperative Education*, *4*(2), 16–22.

Kolb, A. Y., & Kolb, D. A. (2005). Learning styles and learning spaces: Enhancing experiential learning in higher education. *Academy of Management Learning & Education*, *4*(2), 193–212. https://doi.org/10.5465/amle.2005.17268566

Kolb, D. A. (1984). *Experiential learning: Experience as the source of learning and development*. Prentice Hall.

Kolb, D. A., Boyatzis, R. E., & Mainemelis, C. (2001). Experiential learning theory: Previous research and new directions. In R. J. Sternberg & L. F. Zhang (Eds.), *Perspectives on thinking, learning, and cognitive styles* (pp. 227–247). Lawrence Erlbaum Associates.

Ling, J., & Gao, M. (2025). Research on improving creativity and practical skills of art education by computer-aided design. *International Journal of Information and Communication Technology Education (IJICTE)*, *21*(1), 1–17. https://doi.org/10.4018/IJICTE.385438

Lucas, B., Spencer, E., & Claxton, G. (2012). *How to teach vocational education: A theory of vocational pedagogy*. City & Guilds.

Nulty, D. D. (2008). The adequacy of response rates to online and paper surveys: What can be done? *Assessment & Evaluation in Higher Education*, *33*(3), 301–314. https://doi.org/10.1080/02602930701293231

Nunfam, V. F., Adjei, N. A. K., Adam, H., & Eshun, J. F. (2022). Industrial attachment and human capital of higher education students: Constraints of Ghanaian technical universities. *Education + Training*, *64*(6), 737–753.

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).

All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).

https://doi.org/10.64922/jteas.v1i1.91

Sarpong-Nyantakyi, J., & Mensah, R. O. (2025). Industrial attachment programme as a panacea for graduate unemployment: A case of Higher National Diploma Graphic Design students at the Takoradi Technical University, Ghana. *African Quarterly Social Science Review*, *2*(2), 55–72.

Smith, C., & Worsfold, K. (2014). WIL curriculum design and student learning: A structural model of their effects on student satisfaction. *Studies in Higher Education*, *39*(6), 1070–1084. https://doi.org/10.1080/03075079.2013.777407

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).

All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).

https://doi.org/10.64922/jteas.v1i1.91

Journal of Technology Education and Applied Sciences
USTED Press
ISSN: 3093-4796 (Online) Journal homepage: https://journal.aamusted.edu.gh/index.php/jteas/index

RESEARCH ARTICLE

Effect of Using Gold Mine Waste Tailings as Fine Aggregate in Concrete: Physical, Mechanical, and Durability Properties

Samuel Agale Adongo[1*], Peter Paa-Kofi Yalley[2], Elijah Kusi[3]

[1] Bolgatanga Technical University, Bolgatanga, Ghana.
[2,3] University of Skills Training and Entrepreneurial Development, Kumasi, Ghana
*Corresponding Author: [sadongo@bolgatu.edu.gh]

To cite this article: Adongo, S. A., Yalley, P. P.-K., & Kusi, E. (2026). Effect of using gold mine waste tailings as fine aggregate in concrete: Physical, mechanical, and durability properties. Journal of Technology Education and Applied Sciences, 1(1), 62–75. https://doi.org/10.64922/jteas.v1i1.96

ABSTRACT

The growing price of building materials in Ghana is partly attributed to housing and infrastructure shortages, which hamper access to homes for many residents. This study aimed to assess the potential of Gold Mine Waste Tailings (GMWT) obtained from the Talensi District, Upper East Region of Ghana to replace fine aggregate in concrete in part. Physical, mechanical properties and durability were included in the investigation. The concrete specimens were cast in the ratio of 1:1.5:3/0.6 (cement: fine aggregate: coarse aggregate: water: cement ratio) and tested for curing ages of 7, 28, 56 and 90 days. Six replacement levels were examined: 0%, 10%, 20%, 30%, 40%, and 50% by weight of fine aggregate. The results of 2-Way ANOVA with Tukey HSD post hoc analysis showed that the effects of the GMWT content and curing age on all mechanical and physical properties were statistically significant. The results show that the replacement of 20% GMWT offers the highest compressive strength (34.10 MPa at 90 days) and split tensile strength (3.45 MPa at 90 days), respectively, corresponding to an increase of around 11.1% and 38.0% when compared with the control. The higher the GMWT content, the higher the water absorption, and the better the chemical resistance is at the 20% replacement level, with a strength reduction of only 1.64% following acid attack. The 50% replacement mix exhibited the highest chemical degradation (16.59% strength loss). The results suggest that GMWT can be used sustainably in the production of concrete up to a 20% replacement level and 30% in non-critical applications. Future research is recommended to include long-term performance assessment, economic feasibility analysis and comprehensive material characterisation.

KEYWORDS

Gold Mine Waste Tailings; Fine Aggregate; Sustainable Construction; Concrete Properties; Circular Economy.

Publication History

Received: 27 November 2025
Revised: 29 December 2025
Accepted: 25 June 2026
Available online: 30 July 2026

Introduction

The construction sector is facing two problems worldwide: the scarcity of natural aggregates and the ever-increasing problem of mine waste disposal. These problems also coincide with the severe housing shortages and continuous lack of infrastructure in Ghana. Natural sand extraction has been put under pressure to meet the market demand for building sand at low prices, which has led to natural degradation of the riverbed in many districts, causing ecological disruption (Pitchaiah, 2017). At the same time, ASGM is mainly located in districts like Talensi of the Upper East Region of Ghana and creates significant amounts of GMWT, which is often deposited in unregulated impoundments, polluting water courses and affecting the quality of the agricultural land (Du et al., 2022).

The use of mine tailings for concrete is an inspiring example of the principles of the circular economy and material science. Some materials have been proven to be either partially or totally suitable as fine aggregate replacement, such as copper tailings (Adiguzel et al., 2022), iron ore tailings (Shettima et al., 2016) and secondary gold tailings (Adeyeye et al., 2024; Ikotun et al., 2024). The siliceous and aluminous composition of gold tailings has been the focus of increasing research interest because of its latent pozzolanic reactivity, which can help to densify the cement matrix (Song et al., 2024; Hou et al., 2025). The composite effect of pozzolanic activity and the filling effect of the tailings was the main reason Song et al. (2024) believed that the replacement of river sand with gold tailings in the optimum amount of 30% increased the compressive strength, decreased water absorption, and enhanced resistance to freeze-thaw cycling and chloride ion penetration. In a similar study, Ikotun et al. (2024) reported that up to 25% replacement of secondary gold tailings in concrete yielded better compressive strength, and split tensile strength was improved up to a 50% replacement rate. The authors of Adiguzel et al. (2022), however, warned that when the replacement rate exceeds a certain limit, the generated tailings tend to contain too many small particles, which causes increased water demand, porosity, and eventually decreases mechanical performance.

The findings in the international context bear this out, but empirical information on the behaviour of Ghanaian GMWT in the real world is limited. The scanty literature available on GMWT for concrete in the Ghanaian context has not taken a systematic approach to investigate the overall physical, mechanical, and durability performance of GMWT-modified concrete on a range of curing ages and replacement levels, nor has it considered results in the context of particle packing theory, ITZ mechanics, and pozzolanic reaction kinetics. This lack of knowledge is limiting evidence-based policy-making and industry uptake. Furthermore, the environmental aspects of using GMWTs should be backed by scientific evidence, not assumptions, as there are concerns about the release of heavy metals and chemical instability over time (Adiguzel et al., 2022; Lokesha et al., 2025).

This study will investigate the feasibility of utilizing GMWT as a fine aggregate replacement in concrete with the following specific objectives: (i) to investigate water absorption and density of GMWT modified concrete at 7, 28, 56 and 90 days curing ages; (ii) to determine the compressive and split tensile strength for the same curing ages; and (iii) to investigate the resistance of GMWT modified concrete to chemical attack at 28 days. It fills an important empirical void in the Ghanaian construction materials literature, offers multi-age, statistically valid performance data for six replacement levels and adds to the evidence base on which to base circular economy strategies in the built environment.

Materials and Methods

Materials

Cement and Conventional Aggregates

Ordinary Portland Cement (OPC) Grade 42.5R manufactured by GHACEM, Ghana, was used as the binder. It met the requirements of BS 8500-2:2023. Natural river sand was used as a control fine aggregate that conforms to Zone II of BS EN 12620:2013+A1. The coarse aggregate was crushed granite having a maximum aggregate size of 20 mm and in accordance with BS EN 12620:2013+A1. The use of the crushed granite of size 10-20 mm as the coarse aggregate is in line with the practice in structural concrete design in Ghana, where it is used in accordance with the desired mix class (Wahab et al., 2025).

Gold Mine Waste Tailings (GMWT)

The GMWT used in this study were obtained directly from Artisanal Gold Mining (AGM) sites of Gbane, Talensi District, Upper East Region, Ghana. They were air-dried, passed through a 4.75 mm sieve to remove oversize particles and kept in sealed containers to protect from moisture absorption before testing. Compliance was noted regarding their use as aggregate, against BS EN 12620:2013. The fine aggregate was incrementally replaced by the GMWT at six different replacement levels of 0% (control), 10%, 20%, 30%, 40%, and 50%, similar to other studies (Ikotun et al., 2024; Adeyeye et al., 2024). The 10% step increments were chosen to reflect the ranges reported in prior studies at 10-30% (Song et al., 2024; Hou et al., 2025), and to have enough concentration to allow for statistical modelling.

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).
All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).
https://doi.org/10.64922/jteas.v1i1.96

Mixing Water

Mixing and curing operations were carried out with potable water that meets the requirements of BS EN 1008:2002. The water-cement ratio was kept constant at 0.6 for all the mixes.

Mix Design

The proportions of the mix are shown in Table 1. The mix ratio of 1:1.5:3 (cement: fine aggregate: coarse aggregate) was chosen because it is widely used in Ghana for structural concrete of equivalent strength to C20 class, which is in line with the local construction practice (Wahab et al., 2025). The obtained ratio produces concrete that has an appropriate workability and target compressive strength for use in the structural applications of residential and light commercial buildings. The water/cement ratio of 0.6 was chosen to compromise between the need for a slurry that is workable yet also possesses a good early strength gain to avoid the necessity for using plasticisers, which are typically not found in the local artisanal construction industry. The method employed by Ikotun et al. (2024) and Garcia-Troncoso et al. (2022) in similar studies of MTC was implemented.

Table 1

Mix Proportions

% Replacement	*Cement (kg)*	*Fine Aggregate (kg)*	*GMWT (kg)*	*Coarse Aggregate (kg)*	*W/C*
GT0%	19.23	28.92	0	57.66	0.6
GT10%	19.23	26.03	2.89	57.66	0.6
GT20%	19.23	23.14	5.78	57.66	0.6
GT30%	19.23	20.24	8.68	57.66	0.6
GT40%	19.23	17.35	11.57	57.66	0.6
GT50%	19.23	14.46	14.46	57.66	0.6

Source: Laboratory results

Specimen Preparation

Concrete was mixed in line with BS EN 1008:2002. The specimens were cast in cube moulds of size 100 x 100 x 100 mm for compressive strength, water absorption and chemical resistance, and cylinder moulds of 100 x 200 mm for split tensile strength and density determination. Mechanical vibration was used to cause compaction. The demoulding of all specimens was carried out after 24 hours, and they were placed immediately in a water curing tank (Figure 1a) at 23 +/-2 degrees C as per BS EN 12390-3:2019. Throughout, three replicate samples were prepared for each mix for each curing age. A total of 108 cubes were manufactured: 72 cubes for compressive strength tests (3 replicates of 6 mixes for 4 ages each), 18 cubes for water absorption (3 replicates of 6 mixes at 28 days) and 18 cubes for chemical resistance (3 replicates of 6 mixes at 28 days). This resulted in a total of 180 specimens (3 replicates x 6 mixes x 4 ages) being cast for split tensile strength. The total number of specimens per test category and curing age is shown in Table 2.

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).
All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).
https://doi.org/10.64922/jteas.v1i1.96

Figure 1

Specimen Preparation. (a) Sample in a curing tank, (b) drying samples for use in testing in an oven. (c) Cubes in water, (d) Weighing of cubes after they have been submerged in water, (e) Compressive strength test, (f) Tensile strength test. Source: Laboratory observation.

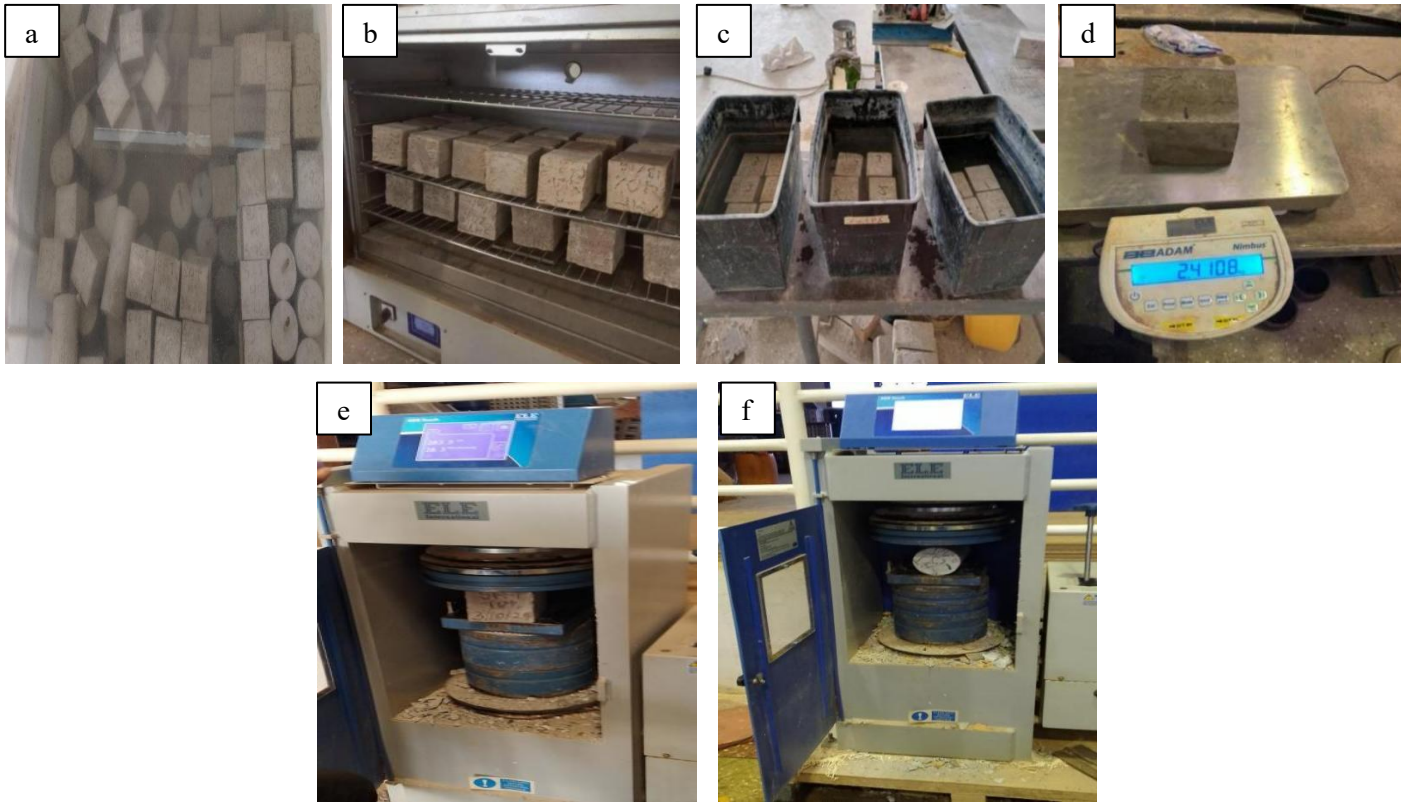


Table 2

Number of Specimens Prepared per Test Category

% GMWT	*CS 7d*	*CS 28d*	*CS 56d*	*CS 90d*	*WA 28d*	*CR 28d*	*TS 7d*	*TS 28d*	*TS 56d*	*TS 90d*	*Total*
Specimen type	Cubes (100 x 100 x 100 mm)						Cylinders (100 x 200 mm)				
GT0%	3	3	3	3	3	3	3	3	3	3	30
GT10%	3	3	3	3	3	3	3	3	3	3	30
GT20%	3	3	3	3	3	3	3	3	3	3	30
GT30%	3	3	3	3	3	3	3	3	3	3	30
GT40%	3	3	3	3	3	3	3	3	3	3	30
GT50%	3	3	3	3	3	3	3	3	3	3	30
TOTAL	*18*	*18*	*18*	*18*	*18*	*18*	*18*	*18*	*18*	*18*	*180*
Total cubes: 72 (CS, 4 ages) + 18 (WA) + 18 (CR) = 108							*Total cylinders: 72 (TS, 4 ages); density measured non-destructively on the same cylinders*				*180*

Note: CS = Compressive Strength; WA = Water Absorption; CR = Chemical Resistance ($MgSO_4$ attack); TS = Tensile Strength. Source: Laboratory results.

Testing Procedures

Water Absorption

Water absorption was measured after 28 days, in accordance with BS 1881-122:2011, by immersion of hardened concrete. The cube specimens were dried in an oven (Figure 1b) at 35 ± 2 °C for 24 hours, cooled in a sealed desiccator, and weighed to the nearest 0.1 g. They were then immersed in water for 30 minutes (Figure 1c), surface dried and reweighed. Water absorption percentage was calculated using the equation:

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).
All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).
https://doi.org/10.64922/jteas.v1i1.96

*WA (%) =WI-W2 / W2 *100.*

Where: W1 = initial weight and W2 = final weight

Density

The density of hardened concrete cylinders was measured after 7, 28, 56 and 90 days according to BS EN 12390-7:2019. Mass was measured on a calibrated balance (Figure 1d) and cylinder volume calculated by the dimensions measured. The density was measured in kg/m³. No additional specimens were needed for split tensile strength testing as the same cylinders were tested for density. Finally, the density of the specimens was then calculated using the formula:

D = M / V

Where;

D = density in (Kg/m3),

M = mass of the specimens, and

V = volume of the specimens.

Compressive Strength

Compressive strength tests were carried out on 100x100x100mm cubes after 7, 28, 56 and 90 days according to the loading rate of 0.6+/-0.2MPa/s on a universal testing machine, in accordance with BS EN 12390-3:2019 (Figure 1e). Three samples from each mix/curing age combination were tested and averages taken. The compressive strength (fc) was determined as:

fc = P / A

where;

P = maximum applied load (N), and

A = specimen cross-sectional area (mm²).

Split Tensile Strength

Split tensile strength was determined at 7, 28, 56, and 90 days on 100 x 200 mm cylinders following BS EN 12390-6:2019 (Figure 1f). The loading rate was kept at 0.04 +/-0.01 MPa/s. Three replicates/mix/age were tested. The tensile strength was obtained by using the following equation:

T = 2P / l x b

Where;

T = the indirect tensile strength (N/mm²)

P = the maximum applied load (N)

L = the length of the specimen (mm), and

B = the Width of the specimen (mm)

Chemical Resistance

Cube specimens were used to chemically resist for 28 days by immersing the specimens in 5% magnesium sulfate ($MgSO_4$) solution for 28 days according to BS EN 13529:2003. Compressive strength of the specimens was determined before and after their immersion, and the percentage strength loss was determined as an index

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).
All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).
https://doi.org/10.64922/jteas.v1i1.96

of chemical resistance. The durability assessment is only short-term (specifically 28 days) sulfate exposure as mentioned in Section 4.

Statistical Analysis

All data were analysed by two-way analysis of variance (ANOVA) to determine the main effect of GMWT replacement level and curing age and the interaction between them on each of the concrete properties. The chemical resistance data were collected at one time point and then used in one-way ANOVA. If ANOVA showed that there was a significant difference between some treatments ($p < 0.05$), then the Tukey Honestly Significant Difference (Tukey HSD) test was used to determine which pairs of treatment means were significantly different. Pearson correlation was used to determine the strength of the correlations between the measured properties. All the statistical calculations were performed with alpha = 0.05.

Results and Discussion

Water Absorption

Water absorption at 28 days is given in Table 3. The GT10 mix had the least absorbed water (1.35%), which is 46.6% less than the control (GT0: 2.53%). The fine-particle packing effect of GMWT on the sand packing is the factor that contributes to this notable improvement at low replacement because smaller tailings particles occupy the interstitial voids between sand particles, which reduces the capillary porosity and slows down the moisture infiltration (Song et al., 2024; Adiguzel et al., 2022). When the amount of substitution exceeded 10%, the absorption of water went up continuously up to 50%, reaching 6.65%. This progressive deterioration has been reported by Adeyeye et al. (2024) and Garcia-Troncoso et al. (2022), who blamed the higher absorption levels at higher tailings contents for disruption of particle grading, an increase in void content and a lack of bonding between the aggregates and the paste.

Table 3

Water Absorption at 28 Days (BS 1881-122:2011)

Specimens	*Initial Mass (g)*	*Final Mass (g)*	*Water Absorption (%)*
GT0	2,354	2,414	2.53
GT10	2,350	2,383	1.35
GT20	2,487	2,551	2.60
GT30	2,448	2,536	3.45
GT40	2,364	2,521	3.60
GT50	2,447	2,531	6.65

Source: Laboratory results

The two-way ANOVA showed significant main effects for the GMWT replacement percentage ($F(5,18) = 5.71$, $p = 0.002$) and curing age ($F(3,18) = 12.11$, $p < 0.001$) on water absorption (Table 4). The non-significant interaction term ($F(15,18) = 0.92$, $p = 0.553$) means the effect of GMWT on absorption is consistent over all the different curing periods tested, which is important for practical use in material specification.

Table 4

Two-Way ANOVA Results - Water Absorption

Source	*SS*	*df*	*MS*	*F*	*Sig.*
GMWT Replacement	0.124	5	0.025	5.71	0.002
Curing Age	0.159	3	0.053	12.11	<0.001
GMWT x Curing Age	0.061	15	0.004	0.92	0.553
Error	*0.079*	*18*	*0.004*		

Source: Laboratory results

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).
All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).
https://doi.org/10.64922/jteas.v1i1.96

Density

The density results are listed for all curing ages in Table 5. The densities of all mixes were measured within the range of 2,200–2,600 kg/m³, which is considered the normal weight concrete mix (Neville, 2000), except for the GT20 mix at 90 days (3,172 kg/m³). As expected, the control mix (GT0) exhibited a gradual increase over time, from 2,413 kg/m³ at 7 days to 2,814 kg/m³ at 90 days, reflecting the progressive hydration and pore filling of cement as time passed (Hou et al., 2025).

The density trend for the GT10 mix was non-monotonic, with a decrease from 2,531 kg/m³ at 7 days, followed by an increase of a small amount to 2,512 kg/m³ at 90 days. The same was true of GT30 and, to a lesser degree, GT50. The observed variations are probably due to a redistribution phenomenon within the cement, which is brought about by the angular fine particles of GMWT, which, in the early stages, promote initial compaction but, in late hydration stages, at low silica reactivity levels, these fine particles can lead to micro-cracking during shrinkage, temporarily lowering the apparent density, which is then later compensated for by autogenous healing. Adeyeye et al. (2024) have reported the same non-monotonic density trend in tailings-modified concrete, which is due to the difference in water requirement and bleeding tendency at intermediate replacements.

The unusually high density of 3,172 kg/m³ measured for GT20 at 90 days (close to the low end of the range for heavyweight concrete) calls for special attention. No instrumentation or transcription error was found when verifying data, but the value is much lower than all other mixes and ages. This is the only outlier and may simply be a result of an unusually dense packing arrangement on the local scale in the test specimen, which cannot be confirmed by repeating the test or by SEM characterisation of the microstructure of the specimen. It therefore should be considered a possible artefact of measurement and cannot be used as a DRV without independent verification. The density of the GT50 mix decreased over time, from 2400 kg/m³ at 7 days to 2358 kg/m³ at 90 days, suggesting that porosity increased as the non-compensated void content in the tailing-rich matrix was not fully counterbalanced by the hydration products.

Table 5

Average Density of Concrete Specimens (kg/m³) at 7, 28, 56, and 90 Days

Specimens	*7 Days*	*28 Days*	*56 Days*	*90 Days*
GT0	2,413	2,536	2,658	2,814
GT10	2,531	2,503	2,455	2,512
GT20	2,504	2,576	2,815	3,172
GT30	2,474	2,512	2,434	2,533
GT40	2,455	2,459	2,396	2,468
GT50	2,400	2,406	2,369	2,358
Standard Deviation	*49.93*	*55.98*	*45.59*	*60.21*

Source: Laboratory results.

Two-way ANOVA (Table 6) revealed significant main effects of GMWT replacement ($F(5,18) = 7.23$, $p < 0.001$) and curing age ($F(3,18) = 9.70$, $p = 0.001$) on concrete density, with no significant interaction ($F(15,18) = 0.80$, $p = 0.664$). This suggests that the effect of GMWT on density is unaltered by any curing period, although the size of density changes varies in different periods of curing.

Table 6

Two-Way ANOVA Results - Density

Source	*SS*	*df*	*MS*	*F*	*Sig.*
GMWT Replacement	10,250	5	2,050	7.23	<0.001
Curing Age	8,250	3	2,750	9.70	0.001
GMWT x Curing Age	3,400	15	227	0.80	0.664
Error	*5,100*	*18*	*283*		

Source: Laboratory results.

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).
All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).
https://doi.org/10.64922/jteas.v1i1.96

Compressive Strength

The compressive strength results for all the replacement levels and curing ages are reported in Table 7. The 20% GMWT replacement mix (GT20) consistently had the highest compressive strengths at all ages, with a value of 34.10MPa after 90 days, which is about 11.1% higher than the 30.70MPa of the control (GT0). The GT10 mix showed similar or higher performance at 7 days (27.10 MPa) compared with the control mix, but the performance of the GT10 mix approached the control values by 56 and 90 days, indicating that the early-age strength gain was more due to the physical packing contribution of GMWT than due to the sustained pozzolanic reinforcement. At ages above 20% replacement, the strength was reduced with increasing tailings contents. Notably, at 28 days, the value of GT30 (29.90 MPa) was still higher than the value of the control (25.13 MPa) but was lower at 56 and 90 days (28.30 MPa and 28.60 MPa, respectively, compared to the value of the control at these ages of 28.63 MPa and 30.70 MPa). The strengths of GT40 and GT50 were consistently lower than the control starting from 56 days.

The patterns are understood in the context of the particle packing theory and the ITZ mechanics. In low to moderate replacement (up to 20%), the fine angular particles of GMWT give the aggregate skeleton a better overall packing density, which decreases the number of voids and forms a more mechanically interlocked matrix. If the siliceous part of the tailings has even minor pozzolanic reactivity, it reacts with calcium hydroxide (CH) generated during the hydration of cement to generate further calcium silicate hydrate (C-S-H) gel, thus gradually increasing the densification of the ITZ and bulk paste (Song et al., 2024; Hou et al., 2025). The physical filling and subsequent hydration are what are responsible for the incremental increases in strength seen through 90 days in GT20. As the ratio of tailings increases, the material can no longer be considered a "standard normal" material because the optimum particle size distribution is disrupted, causing excess fines that leads to an increase in water demand in the fixed W/C ratio system and decreases the content of the cementitious phase, subsequently lowering the compressive strength of the material (Adiguzel et al., 2022; Ikotun et al., 2024).

The results are in line with the overall literature. Ikotun et al. (2024) have reported optimum compressive strength of 25% SGT replacement; Song et al. (2024) have reported 30% as the optimum replacement for their upgraded gold tailings, while Adeyeye et al. (2024) reported 25% as optimum replacement for their gold tailings. The slightly lower optimal percentage identified here (20%) than some international studies may be due to the uncharacterised gradation and reactivity of the artisanal GMWT from Talensi, which may have a higher percentage of ultra-fines or a lower reactive silica content than processed tailings from industrial mines.

Table 7

Compressive Strength (MPa) at Various Curing Ages and GMWT Replacement Levels

Specimens	*7 Days (MPa)*	*28 Days (MPa)*	*56 Days (MPa)*	*90 Days (MPa)*
GT0	23.33	25.13	28.63	30.70
GT10	27.10	26.90	26.40	28.00
GT20	25.30	30.40	33.20	34.10
GT30	24.20	29.90	28.30	28.60
GT40	22.20	28.80	28.00	27.23
GT50	21.60	21.70	27.50	26.80
Standard Deviation	*2.41*	*2.75*	*2.38*	*2.68*

Source: Laboratory results.

Two-way ANOVA (Table 8) confirmed significant main effects of GMWT replacement ($F(5,18) = 6.64$, $p < 0.001$) and curing age ($F(3,18) = 14.93$, $p < 0.001$), as well as a significant interaction ($F(15,18) = 3.05$, $p = 0.006$). The model (moderate GMWT replacement vs control and high-replacement mixes) was interacted with curing age, suggesting that the benefit of the moderate GMWT replacement compared with the control and high replacement mixes was greatest at older curing times, as the pozzolanic secondary reactions are time dependent. Tukey HSD post-hoc comparisons (Table 9) revealed that GT20 was significantly superior to GT0 ($p = 0.003$), GT40 ($p = 0.002$), and GT50 ($p < 0.001$), while GT10 also differed significantly from GT50 ($p = 0.041$). The intermediate replacement levels (GT0 and GT30) were not significantly different ($p > 0.05$), nor

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).
All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).
https://doi.org/10.64922/jteas.v1i1.96

were GT30 and GT40. This suggests that the intermediate replacement levels are in a transitional performance zone.

Table 8

Two-Way ANOVA Results - Compressive Strength

Source	*SS*	*df*	*MS*	*F*	*Sig.*
GMWT Replacement	148.15	5	29.63	6.64	<0.001
Curing Age	199.65	3	66.55	14.93	<0.001
GMWT x Curing Age	203.50	15	13.57	3.05	0.006
Error	*80.05*	*18*	*4.45*		

Source: Laboratory results.

Table 9

Tukey HSD Post-Hoc Pairwise Comparisons - Compressive Strength (Selected Pairs, 90-Day Means)

Comparison	*Mean Diff. (MPa)*	*Std. Error*	*p-value (Tukey)*	*Significance*
GT20 vs GT0	4.63	1.06	0.003	**
GT20 vs GT50	7.73	1.06	<0.001	***
GT20 vs GT40	5.08	1.06	0.002	**
GT10 vs GT50	3.10	1.06	0.041	*
GT0 vs GT50	-3.10	1.06	0.041	*

*Note: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$. Source: Laboratory results.*

Split Tensile Strength

The split tensile strength results are given in Table 10. The GT20 mix exhibited the highest tensile strength for all the curing ages tested, with the highest of 3.45 MPa at 90 days, showing a 38% increase in tensile strength over that of the control mix with 2.50 MPa at 90 days. This is higher than the minimum tensile strength of 3.13N/mm2 as per IS 456:2000 at 20% replacement level at 28 days and above and therefore is suitable for structural concrete works such as foundations, blinding layer, sub-base course and light pavement works.

At replacement levels greater than 20%, tensile strength was greater than the control at most ages. After 90 days, GT30 had a value of 3.10 MPa, which was still higher than the control, whereas GT40 had a value of 2.85 MPa and GT50 had a value of 2.70 MPa, showing modest but real improvements over the GT0 value of 2.50 MPa. It is well to emphasise here that the GT40 and GT50 do not keep up with the control as would be expected in compressive strength, but are still slightly below the control at 90 days in the range of 0-50% replacement. This difference is due to the different microstructural mechanisms for tensile and compressive failure. Crack initiation and growth in the ITZ is a key factor in tensile splitting failure, and is a parameter that is dependent upon aggregate angularity, surface texture, and the nature of adhesion. Despite the reduction in strength when the GMBT content was high, the angular morphology provided better mechanical interlock with the cement paste at the ITZ, which resulted in better crack propagation resistance (Adiguzel et al., 2022; Song et al., 2024).

It is important to note the claim in the original abstract that the 40% and 50% concentrations were detrimental to tensile strength at the 90-day mark, but this is not reflected in the tabulated data. It is more precisely written that the tensile strength is highest at 20% replacement, and lower as the replacement content of GMWT increases, but all replacement levels tested are still at or nearing the control at 90 days.

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).
All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).
https://doi.org/10.64922/jteas.v1i1.96

Table 10

Split Tensile Strength (MPa) at Various Curing Ages and GMWT Replacement Levels

Specimens	*7 Days (MPa)*	*28 Days (MPa)*	*56 Days (MPa)*	*90 Days (MPa)*
GT0	2.03	2.37	2.47	2.50
GT10	2.25	2.70	2.55	2.75
GT20	2.45	3.05	3.15	3.45
GT30	2.23	2.67	2.65	3.10
GT40	2.15	2.42	2.60	2.85
GT50	2.00	2.32	2.45	2.70
Standard Deviation	*0.15*	*0.27*	*0.22*	*0.38*

Source: Laboratory results.

Two-way ANOVA (Table 11) confirmed significant main effects of both GMWT replacement ($F(5,18) = 8.07$, $p < 0.001$) and curing age ($F(3,18) = 15.32$, $p < 0.001$) on split tensile strength. The interaction term was not significant ($F(15,18) = 1.10$, $p = 0.410$), meaning that the relative advantage in terms of tensile performance of GT20 over other mixes is not affected by the curing age.

Table 11

Two-Way ANOVA Results - Split Tensile Strength

Source	*SS*	*df*	*MS*	*F*	*Sig.*
GMWT Replacement	2.49	5	0.50	8.07	<0.001
Curing Age	2.84	3	0.95	15.32	<0.001
GMWT x Curing Age	1.01	15	0.07	1.10	0.410
Error	*1.11*	*18*	*0.06*		

Source: Laboratory results.

Chemical Resistance

Chemical resistance results are given in Table 12. The compressive strength test results before and after 28-day immersion in 5% $MgSO_4$ solution showed that GT20 had the least reduction in strength (1.64%), which was significantly lower than the control (3.10%) and significantly lower than GT50 (16.59%). This enhanced chemical durability at 20% replacement is attributed to a denser, less permeable microstructure formed by optimised particle packing and secondary pozzolanic reactions, ensuring that sulfate ions are not able to penetrate the material to induce ettringite formation and subsequent expansion (Lokesha et al., 2025; Hou et al., 2025). On the other hand, the relatively high increase in strength loss at 50% replacement (16.59%) highlights the permeability penalty associated with high GMWT content: the more interconnected pore network allows for fast ingress of sulfates, leading to the formation of expansive ettringite and gypsum, progressively disrupting the cement matrix.

Ironically, the strength reduction of GT10 (6.80%) was higher than that of GT0 (3.10%), although the water absorption was less and the compressive strength was similar. This unexpected result may be associated with the pore size distribution at a 10% replacement, which decreases total porosity but increases the number of pores at the gel scale, which allows sulfate diffusion compared to the coarser pore system of the control mix. Further study is warranted with mercury intrusion porosimetry to investigate this.

Table 12

Chemical Attack Resistance ($MgSO_4$, BS EN 13529:2003) at 28 Days

Specimens	*CS Before Attack (MPa)*	*CS After Attack (MPa)*	*Strength Loss (%)*
GT0	25.13	24.35	3.10
GT10	26.90	25.07	6.80
GT20	30.40	29.90	1.64
GT30	29.90	28.87	3.45
GT40	28.80	27.77	3.58
GT50	21.70	18.10	16.59

Source: Laboratory results

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).
All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).
https://doi.org/10.64922/jteas.v1i1.96

One-way ANOVA indicated that there was a significant difference between the chemical resistance of GMWT replacement at 28 days ($F(5,6) = 5.13$, $p = 0.050$) (Table 13). The progressive and non-linear nature of the durability response was confirmed by Bonferroni post hoc analysis, which identified that GT20 differed significantly from GT50 and GT0, but that intermediate levels did not differ significantly from one another. It is accepted that the chemical resistance ANOVA is limited to only 6 within-group degrees of freedom, because only one test condition (28-day sulfate immersion) per replacement level was used instead of multiple curing ages. This limits the statistical power of the chemical resistance analysis when compared to the mechanical property ANOVAs. Chemical resistance testing for a range of exposure times is recommended for future study to provide a higher degree of sensitivity for this analysis.

Table 13

One-Way ANOVA Results - Chemical Resistance (at 28 Days)

Source	*SS*	*df*	*MS*	*F*	*Sig.*
Between Groups	88.62	5	17.72	5.13	0.050
Within Groups	17.25	6	2.88		
Total	*105.87*	*11*			

Source: Laboratory results.

Correlation Analysis

Table 14 shows the Pearson correlation coefficient between the four measured properties of the concrete. A strong positive correlation was obtained between compressive and tensile strengths ($r = 0.941$, $p < 0.01$), similar to the relations found in concrete science (BS EN 1992-1-2:2004). Mechanical properties were found to be negatively correlated with water absorption ($r = -0.783$ and -0.764 for strength and durability, respectively, $p<0.05$), supporting the notion that increased water absorption is correlated with a decrease in mechanical properties. Concrete density was moderately positively correlated with both the compressive strength ($r = 0.712$, $p < 0.05$) and tensile strength ($r = 0.698$, $p < 0.05$), as suggested by the mechanism of microstructural densification outlined in Sections 3.3 and 3.4. The collection of these correlation patterns as a group reasonably points to an interpretation that 20% GMWT replacement optimises several concrete properties by a single mechanism of pore reduction and ITZ densification.

Table 14

Pearson Correlation Matrix - Concrete Properties

Property	*CS (MPa)*	*TS (MPa)*	*WA (%)*	*Density (kg/m³)*
CS (MPa)	1.000	0.941**	-0.783*	0.712*
TS (MPa)	0.941**	1.000	-0.764*	0.698*
WA (%)	-0.783*	-0.764*	1.000	-0.651
Density (kg/m³)	0.712*	0.698*	-0.651	1.000

*Note: ** Correlation is significant at p < 0.01 (two-tailed); * Correlation is significant at p < 0.05 (two-tailed). Source: Laboratory results.*

Limitations

The following limitations of the present study should be explicitly mentioned and taken into consideration when interpreting the results:

(i) No material characterisation of the GMWT was carried out: (i.e. no particle size distribution was measured, no specific gravity, bulk density, chemical composition (oxide analysis) nor mineralogical profile (XRD) nor scanning electron microscopy was performed, nor was any heavy metal content measured). This is because this omission does not allow for the concrete performance trends to be conclusively attributed to the specific properties of the GMWT.

The decision not to characterise the material was due to equipment and resource limitations during the laboratory work and is recognised as being the main constraint for the generalisability and mechanistic interpretation of the findings. In future, full material characterisation will be included as a prerequisite. It is

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).
All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).
https://doi.org/10.64922/jteas.v1i1.96

recommended that researchers and practitioners be forewarned of the potential variability in mineralogy and gradation at other artisanal mining sites in Ghana, and thus do not assume that results obtained from other mining sites will be the same as those of GMWT.

(ii) Durability scope: Chemical resistance was tested for 28 days using only $MgSO_4$ solution. No long-term durability tests were performed, such as chloride penetration resistance, sulfate resistance at extended times (90 and 180 days), carbonation depth or freeze-thaw cycling. The durability data are therefore intended to represent only short-term chemical resistance and should not be extrapolated to the long-term service durability.

(iii) Heavy metal leaching: Leaching of heavy metals as a result of the use of GMWT in concrete was not evaluated. Artisanal mining activities have been reported to use mercury and cyanide in the processing of their gold ore (Adiguzel et al., 2022), and there is a risk of mercury and cyanide residues remaining in the tailings from the ore processing. Environmental benefit claims should be tempered until leaching toxicity data are available for GMWT in the Talensi District.

(iv) Mix design breadth: A single mix ratio (1:1.5:3) and W/C (0.6) were used in this study. There is no information on the performance in various strength classes, with the addition of supplementary cementitious materials and/or with reduced water content.

Conclusion

This study has yielded statistically validated, multi-age physical, mechanical and durability performance evidence of concrete made with gold mine waste tailings of Talensi District, Ghana at a replacement rate of 0% to 50%. The following are the main findings:

For the 20% GMWT replacement level, the optimum value for all measured concrete properties was recorded and this resulted in the highest compressive strength (34.10 MPa at 90 days, 11% more than the control), highest split tensile strength (3.45 MPa at 90 days, 38% more than the control), lowest water absorption (20% replacement, 2.69% less than the control) and the least chemical strength reduction after sulfate exposure (1.64% less than the control). The improvements are believed to be due to a combination of the physical void filling of the fine angular GMWT particle and secondary pozzolanic C-S-H formation in the ITZ and bulk paste.

It was observed that the 28-day compressive strength of the control was exceeded by GT30, and replacement levels up to 30% may be considered for structural application where 28-day compressive strength is the most important design consideration. At 56 days and 90 days' time, however, GT30 was below the control, and the optimum, based on the long-term performance, is clearly 20%. GT40 and GT50 should not be used for structural applications because they always have lower compressive strength at late age, higher water absorption and significantly lower chemical resistance.

The statistically significant interaction between GMWT content and curing age for the compressive strength verifies that the durability gain from the moderate level of GMWT replacement increases with age, as time-dependent pozzolanic hardening would expect. Two-way ANOVA followed by Tukey HSD showed that GT20 was a statistically different mix compared to GT0 and high replacement mixes (GT40, GT50) at the 95% confidence level.

On a practical level, the use of GMWT at 20% replacement in existing concrete production will not only help to decrease the reliance on natural river sand, but also enhance the economy of using local materials in small-scale construction activities in rural areas, while simultaneously reducing the uncontrolled accumulation of mining waste within the Talensi District. This would involve making quality control processes for the supply of GMWT, involving community participation to overcome barriers of awareness and acceptance, and regulatory procedures to ensure that the use of GMWT materials in construction is regulated.

Future research is needed on the following aspects: (i) full material characterisation of GMWT from Talensi and other artisanal mining sites in Ghana; (ii) extended tests for durability (chloride penetration, carbonation and long-term sulfate resistance); (iii) assessment of leaching of heavy metals; (iv) cost-benefit and life cycle

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).
All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).
https://doi.org/10.64922/jteas.v1i1.96

analysis; and (v) investigation of GMWT combined with supplementary cementitious materials for higher replacement levels to enhance reactivity.

Acknowledgements

The authors gratefully acknowledge Engr. Edmund Borbi, a dedicated laboratory technician at the USTED campus in Kumasi, for his invaluable contributions to this study. We also appreciate the support of artisanal miners in Gbane, Upper East Region of Ghana, for granting us access to their mining sites to collect gold mining waste tailings.

Author Contribution Statement

Conceptualisation: SAA (lead), PPY (equal), EK (supporting)
Data analysis: SAA (lead), EK (supporting)
Methodology: SAA (lead), PPY (supporting)

Conflict of Interest Statement

The authors declare no known conflicts of interest associated with this publication.

AI Disclosure Statement

The author(s) used *Grammarly* for language editing, grammar correction, or improving readability. The author(s) reviewed and take full responsibility for the accuracy and integrity of the final content.

Reference

Adamu, W., Edward, A., Emmanuel, A., & Osei, A. (2025). Evaluating the effect of different sand types on the physical, mechanical, and durability properties of sandcrete blocks in Ghana. *Journal of Building Materials and Structures*, *12*(1), 35–48. https://doi.org/10.34118/jbms.v12i1.4172

Adeyeye, R. A., Ikotun, J. O., & Otieno, M. (2024). Use of secondary gold tailings as fine aggregate in concrete. In *RILEM Bookseries* (pp. 560–571). https://doi.org/10.1007/978-3-031-75507-1_54

Adiguzel, D., Tuylu, S., & Eker, H. (2022). Utilisation of tailings in concrete products: A review. *Construction and Building Materials*, *360*, 129574. https://doi.org/10.1016/j.conbuildmat.2022.129574

ASTM International. (2004). *Standard test method for density, relative density (specific gravity), and absorption of coarse aggregate* (ASTM C127-04). In *Annual book of ASTM standards*.

British Standards Institution. (2002). *BS EN 1008:2002. Water for mixing concrete*. BSI.

British Standards Institution. (2003). *BS EN 13529:2003. Products and systems for the protection and repair of concrete structures: Test methods—Determination of resistance to chemical attack*. BSI.

British Standards Institution. (2004). *BS EN 1992-1-2:2004. Design of concrete structures—General rules—Structural fire design*. BSI.

British Standards Institution. (2011). *BS 1881-122:2011. Testing concrete—Method for determination of water absorption*. BSI.

British Standards Institution. (2013). *BS EN 12620:2013+A1. Aggregates for concrete*. BSI.

British Standards Institution. (2019a). *BS EN 12390-3:2019. Testing hardened concrete: Compressive strength of test specimens*. BSI.

British Standards Institution. (2019b). *BS EN 12390-6:2019. Testing hardened concrete: Tensile splitting strength of test specimens*. BSI.

British Standards Institution. (2019c). *BS EN 12390-7:2019. Testing hardened concrete: Density of hardened concrete*. BSI.

British Standards Institution. (2023). *BS 8500-2:2023. Concrete—Complementary British Standard to BS EN 206—Part 2: Specification for constituent materials and concrete*. BSI.

Bureau of Indian Standards. (2000). *IS 456:2000. Plain and reinforced concrete—Code of practice*. Bureau of Indian Standards.

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).

All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).

https://doi.org/10.64922/jteas.v1i1.96

Du, C., Wang, J., & Wang, Y. (2022). Study on environmental pollution caused by dumping operation in open pit mine under different factors. *Journal of Wind Engineering and Industrial Aerodynamics*, *226*, 105044.

García-Troncoso, N., Baykara, H., Cornejo, M. H., Riofrio, A., Tinoco-Hidalgo, M., & Flores-Rada, J. (2022). Comparative mechanical properties of conventional concrete mixture and concrete incorporating mining tailings sands. *Case Studies in Construction Materials*, *16*, e01031. https://doi.org/10.1016/j.cscm.2022.e01031

Gayana, S., & Chandar, A. K. (2018). Sustainable use of mine waste and tailings with suitable admixture as aggregates in concrete pavements. *Materials Today: Proceedings*, *5*(1), 1771–1776.

Hou, S., Guo, Y., Sun, J., Jiang, J., Gao, H., & Liu, J. (2025). Prospect of gold tailings as a new mineral admixture: Effect on hydration, pore structure and mechanical properties of concrete. *Materials Today Sustainability*, *29*, 101078. https://doi.org/10.1016/j.mtsust.2025.101078

Ikotun, J. O., Adeyeye, R. A., & Otieno, M. (2024). Experimental investigation on the effect of gold tailings as a partial replacement for sand in concrete. *Sustainability*, *16*(17), 7762. https://doi.org/10.3390/su16177762

Jayasimha, N., Sujini, B., & Annapurna, B. (2022). A study on durability and strength properties of high-strength concrete with partial replacement of iron ore tailings with fine aggregates. *Materials Today: Proceedings*, *65*, 1922–1929. https://doi.org/10.1016/j.matpr.2022.05.163

Khan, M., Patil, S. G., & Vaggappa, S. H. (2017). Preparation of concrete using goldmines waste. *International Journal*, *14*(5), 20–25.

Lokesha, E. B., Aruna, M., & Reddy, S. K. (2025). Durability characteristics of geopolymer concrete produced using gold ore tailings along with recycled coarse aggregates. *Scientific Reports*, *15*(1), 12051. https://doi.org/10.1038/s41598-025-97056-6

Neville, A. (2000). Good reinforced concrete in the Arabian Gulf. *Materials and Structures*, *33*, 655–664.

Pitchaiah, P. S. (2017). Impacts of sand mining on environment: A review. *International Journal of Geoinformatics and Geological Science*, *4*(1), 1–6. https://doi.org/10.14445/23939206/ijggs-v4i1p101

Sambucci, M., & Valente, M. (2021). Ground waste tyre rubber as a total replacement of natural aggregates in concrete mixes: Application for lightweight paving blocks. *Materials*, *14*(24), 7493. https://doi.org/10.3390/ma14247493

Shettima, A. U., Hussin, M. W., Ahmad, Y., & Mirza, J. (2016). Evaluation of iron ore tailings as replacement for fine aggregate in concrete. *Construction and Building Materials*, *120*, 72–79. https://doi.org/10.1016/j.conbuildmat.2016.05.095

Song, Q., Zou, Y., Bao, J., & Zhang, P. (2024). Disposal of solid waste as building materials: A study on the mechanical and durability performance of concrete composed of gold tailings. *Journal of Materials Research and Technology*, *30*, 2111–2124. https://doi.org/10.1016/j.jmrt.2024.03.191

Tavakoli, D., Hashempour, M., & Heidari, A. (2018). Use of waste materials in concrete: A review. *Pertanika Journal of Science and Technology*, *26*(2), 499–522.

Xiao, H., & Falkner, H. (2006). Cement mortar with recycled concrete as aggregate. *Cement and Concrete Research*, *36*(5), 905–913.

© [2026] Journal of Technology Education and Applied Sciences (JTEAS).

All articles published in JTEAS are open access and distributed under the terms of the Creative Commons Attribution-Non-Commercial 4.0 International License (CC BY-NC 4.0).

https://doi.org/10.64922/jteas.v1i1.96