NIPES - Journal of Science and Technology Research
www.nipes.org

Enhanced Scorch Occurrence Prediction in Foam Production via a Fusion SMOTE-Tomek Balanced Deep Learning Scheme

Tabitha Chukwudi Aghaunor[1], Edith Ugochi Omede[2], Eferire Val Ugbotu[3], Joy Agboi[2], Chris Chukwudi Onochie[4], Asuobite ThankGod Max-Egba[5], Victor Ochuko Geteloma[5], Paul Awerosuo Onoma[5], Andrew Okonji Eboka[4], Arnold Adim Ojugo[5], Chris Chukwufunaya Odiakaose[6], and Amaka Patience Binitie[4]

[1]School of Data Intelligence and Tech., Robert Morrison University, Pittsburgh, PA15108, United States. tabitha.aghaunor@gmail.com
[2]Department of Computer Science, Delta State University, Abraka, Delta State 330105, Nigeria. edithomede@delsu.edu.ng
[6]Department of Data Science, University of Salford, Manchester M5 4WT, United Kingdom. eferhire.ugbotu@gmail.com
[4]Federal College of Education (Technical), Asaba, Delta State 320212, Nigeria. xtoline2@gmail.com; andrew.eboka@fcetasaba.edu.ng; amaka.binitie@fcetasaba.edu.ng
[5]Department of Computer Science, Federal University of Petroleum Resources Effurun, Delta State 330102, Nigeria. max-egba.asuobite@fupre.edu.ng, geteloma.victor@fupre.edu.ng; kenbridge14@gmail.com; ojugo.arnold@fupre.edu.ng,
[6]Department of Computer Science, Dennis Osadebey University, Asaba, Delta State 320212, Nigeria. osegalaxy@gmail.com

Article Info

Keywords: *Enhanced, Scorch, Prediction, Fusion SMOTE-Tomek Balanced Deep Learning, Scheme*

Received 08 February 2025
Revised 28 March 2025
Accepted 29 May 2025
Available online 29 May 2025

https://doi.org/10.37933/nipes/7.2.2025.25

eISSN-2682-5821, pISSN-2734-2352

Abstract

Scorch is a common occurrence in the production of flexible polyurethane with significant negative impacts on its resilience, compactness, and integrity. With the increased likelihood of the scorch menace on foam production – conventional remedies have become of great concern due to the increased emission of these chemical constituents against an eco-friendly environment, and the consequent rise in operational costs to provide such remedies. Tackling scorch occurrence requires skilled professionals to efficiently navigate the flexible proportioning of chemical additives, which can also be achieved via the utilization of a cost-effective machine learning scheme that also provisions early warning to predict the occurrence of scorch prior the physical processing via the thermodynamic profile of polyurethane foam. Previous works observe the impact of an imbalanced dataset. Our study investigates the impact of the data balancing scheme and feature selection via the utilization of a SMOTE-Tomek-based chi-squared fused BiLSTM model for a scorch dataset, recorded during the production of the polyurethane foam. The result shows the BiLSTM outperformed benchmark models yielding an Accuracy of 0.9895, F1 of 0.9892, Precision of 0.9817, Recall of 0.9901, AUC of 0.98, and a smooth, monotonic decrease without fluctuations in model loss respectively. Thus, implies that the proposed BiLSTM accurately handles the minority class without the instability due to the vanishing gradients problem. In addition, its benchmark models (i.e. Decision Tree, Logistic Regression, Random Forest, and XGBoost) yield F1 of [0.8145, 0.9105, 0.9210, and 0.9125] with Accuracy of [0.8032, 0.9105, 0.9228, and 0.9574] respectively. Results show the proposed BiLSTM accurately predicts 2-distinct cases of non(occurrence) of scorch. The model demonstrates its capability to effectively predict the occurrence of scorch.

1. Introduction

Polyurethane foam is a porous structured, synthetic created with a mixture of diisocyanates and polyols [1], [2]. It consists of blowing agents (gases) mixed with polyurethane elastomer material [3] – and its use is attributed to its beneficial physical characteristics. Their low density [4], [5] and thermal conductivity [6] allow their flexible use in applications such as thermal/sound insulation, bedding, furniture, encapsulating components, underlays, and other packaging forms [7], [8]. Water as a replacement for chlorofluorocarbons [9], tends to increase the occurrence of scorch in polyurethane foam due to high exothermic reaction that ripples across the constituents. This temperature rise often requires fire retardants as formulating ingredients [10], [11]. However, some fire retardants have also been known to introduce scorch [12]; while others during the interaction with the constituents formulation dissipate aniline from ammonia (-NH2) found to be present in water-blown polyurethane foam [13]. This interaction as in have necessitated the experts' quest to mitigate scorch occurrence via the addition of antioxidants [14], [15] to lessen the effects of fire retardants. However, reports have introduced the use of secondary antioxidants [16] in extreme conditions for low-density foam productions with constituents that yield high moisture content. Nabata et al. [17] studied heating polyurethane foam to observe its cure using diisocyanates. Its drawbacks include longer train time [18], more resources and energy with unsustainable laborious production process as convention. Thus, experts explore the utilization of simulation models to predict ahead with the same variables and chemical constituents – and was found to yield improved result at shorter time [19], [20]. These, use evolutionary modeling approaches, and contributes to searched (domain) target-class amongst others underlying features via its utilization of the trial-n-error approach [21], with several trial iterations in its quest to yield the desired result. This, has been found by many researchers and experts – to be quite cost-effective with reduced waste of resources [22], [23].

With chlorofluorocarbons previously used – scorch was greatly reduced [24]. However, its consequent emissions of harmful materials [25] into the environment led to its ban with water reinvented as its substitute [26]. With experts seeking to evaluate the right proportion of constituents to prevent scorching via the use of forecast knowledge models [27] – scorch has become a menace in polyurethane foam production. Scorch can be visually recognized during the cure stage with the foam's exposure to air [28] to cool. Scorch is a slight yellowish-brown coloration found in the foam slab during its cure stage [29]. When fully formed – scorching reduces a foam's degree of compactness to yield decreased durability, so that such scorched foams quickly suppress [30]. Major causes of scorch include the oxidation of phenols and amines [31], and the use of non-polymeric components as responsible for discoloration [32]. Thus, scorching can be explained as heat-induced changes in the polyurethane foam production process, due to inadequate exposure that prevents proper dissipation of trapped heat at its cure phase [33]. Known impact of scorching includes [34]: (a) porous cellular structure, (b) low load-carrying capacity, (c) low resilience and elasticity, (d) reduced life span with foam's durability, (d) high-volume wastage, (e) poor utility of resources, and (f) reduced profitability.

Conventional remedies include: (a) the use of temperature suppressants, (b) antioxidants as additives with salts as anti-scorch essence, (c) the addition of free isocyanate to moderate the fast exothermic reaction that reduces temperature rise, (d) use of scorch inhibitors like halogenated phosphate ester additives in its proper ratio with diphenylamine derivative and hydroquinone [35]. While these yield their benefits [36], the increased concern of chemical emissions and the required high operational cost of environmental clean-up – renders such solutions unsuitable; And the use of inhibitors has been found to exert morphological and chemical changes to the foam [37]. While scorch inhibitors are known to contribute to foam discoloration (not necessarily from scorch) – they can yield such discoloration arising from foam slab's exposure to sunlight and warehousing fumes [38]. Thus, it is evident that an available option to tackle scorch requires expert skillsets to properly navigate [39] via careful proportions of materials using cost-efficient solution(s) that simulate the process, as it responds to scorch occurrence before the physical manufacturing. In the quest for ground truth [40], performance generalization must adequately account for various specifications in the polyurethane materials including (a) diisocyanates and polyols materials, and (b) environmental conditions within production plants and scenarios where scorch is typically promoted as a result of mechanical defects [41]. While these adjustments if made during production are not a smart choice and decision – experts can only make adjustments before scorch occurrence. There is also the issue with the difficulty of generalizing these recommendations for all production processes as they may not adequately account for various specifications of the supplied (raw) polyurethane materials from a variety of suppliers. This case is especially true for polyols, and diisocyanates vis-à-vis other conditions inherent in the production plants (in the case where scorch is typically promoted as a result of mechanical defects) [42], [43].

We model these dynamics and complex production processes as variables using machine learning (ML) schemes – so as to provision adequate insight with trial-and-error simulations for a variety of scenarios in the chemical manufacturing phases. Learning aggregates the learned intrinsic feats using a classifier. Many ML schemes utilize the recursive top-down mode that partitions its dataset via binary approach resulting in a k-fold split for its predictors with a distribution of dependent variable y that are successively homogeneous [44]. Each scheme is constructed and trained to aggregate their results into a stronger classifier. This is achieved via: (a) bagging which recursively generates a train-set that sums results to reduce variance and bias via voting, and (b) boost reduces bias by sequentially pooling together the performance of many weak learners into a stronger learner to yield enhanced accuracy by correcting previous mistakes. Both approaches enhance accuracy by mitigating bias that reduce errors in misclassified outcomes. A variety of ML schemes successfully implemented includes: Genetic Algorithm [45], SVM [46], Deep Learning [47], Random Forest [48], etc. While these MLs have their inherent drawbacks – the use of feature selection and data balancing in their quest for prediction accuracy has remained a crucial feature.

Existing knowledge gaps that have motivated this study include:

1. **Performance Generalization with Previous Models**: Each classification of scorch from previous studies has explored varying classification cum identification methods developed specifically for the scenario used. This complexity in design has also responded with a set of performance generalizations for each dataset as explored. Thus, the accuracies of previous works ranged from 0.69-to-0.858 respectively. This study hopes to reach improved performance accuracies and generalization [49], [50].
2. **Imbalanced Dataset** – Many domain studies explore datasets which by nature at the collection and cursory look – are imbalanced with scorched (minority-class) records often found to lag during the production of the polyurethane foam production. ML approaches have been known to classify effectively records cum data labels for the majority class; And are often poised to ignore data labels in the minority class. Thus, the study seeks to assess the impact of data balancing using the SMOTE-Tomek links scheme [51], [52].
3. **Greater Dynamics Complexity in Adopted Heuristics**: Many studies have shown that complex models – though quite intricate to understand – have often proffered improved optimal fit solutions [53], [54]. While some models such as Random Forest, XGB, and Logistic Regression are easier to implement with robust classification accuracies [55], [56]; their inherent performance also suffers setback with complex datasets (as in scorch occurrence classification), especially for its utilization of non-linear boundaries. Thus, deep learning models are best suited in this guise [57].
4. **Increased Dimensionality with Vanishing Gradient Problems**: While Deep learning models are based on recurrent and convolution neural networks – their utilization in identification and classification is often hampered by: (a) their requirement for a larger dataset, (b) their inability to effectively handle categorical dataset, and (c) their requisition of longer training time. However, RNN methods are well-suited for tasks with temporal data such as scorch occurrence – save for its vanishing gradient problem. To the rescue thus, is the exploration of the Long-Short Term Memory (LSTM), which is more sophisticated in structure as well as computationally more efficient at learning long-term dependencies within domain datasets [58].

Thus, we adopt a deep learning Bi-directional Long-Short Term Memory (BiLSTM) learning scheme as studies have shown that deep learning schemes outperform traditional ML approaches. A major issue with the adoption of deep learning schemes based on recurrent neural networks (RNN) such as the LSTM includes: (a) the gradient vanishing problem since data flow is single-direction based, (b) their requirement of longer training time, and (c) their requisition of larger dataset. Thus, we adopt the BiLSTM as a method to curb and address the issue in LSTM, especially with data flow in both directions [59]. Our choice of the SMOTE-Tomek to handle the imbalanced dataset is based on its approach as a hybrid (SMOTE) oversampler + (Tomek-links)-under-sampler technique. We used the SMOTE-Tomek-based BiLSTM deep learning scheme on the scorch dataset from Winco Foam company in Benin City Nigeria for this study. Our choice is hinged on its capability to greatly improve performance generalization, reduce model overfit, explore the SMOTE-Tomek approach to address the imbalanced dataset and yield enhanced prediction accuracy.

2.0. Materials and Methods

2.1. Proposed Methodology and Framework

Our proposed methodology is as thus:

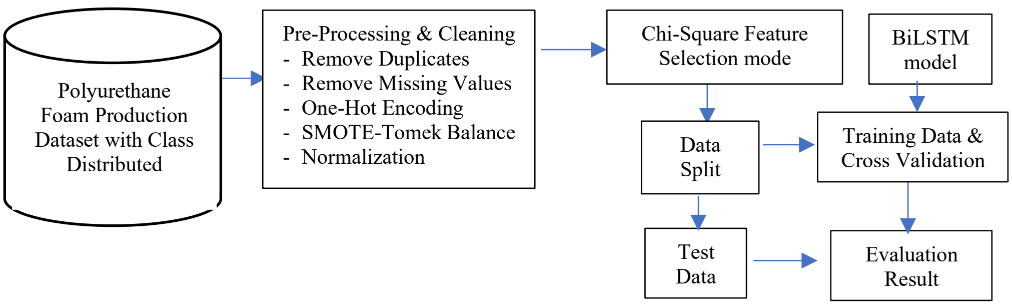

Figure 1. Proposed BiLSTM Deep Learning with SMOTE-Tomek balancing for Scorch Prediction

1. **Step 1 – Data Collection**: Dataset was retrieved from Winco Foams in Edo State. It consists 8540-records with 15-feats: poly_truPut, calc_truPut, TDI_truPut, water_truPut, cal_dialSet, poly_dialSet, TDI_dialSet, water_dialSet, qnty_cal, qnty_poly, qnty_TDI, qnty_water, poly_water_content, prod_time, and scorch as in Table 1 with heatmap as in Figure 2. The dataset was retrieved via the Google Play Scrapper Library. Figure 3a shows the class distribution for the scorched (minority) and unscorched (majority) classes.

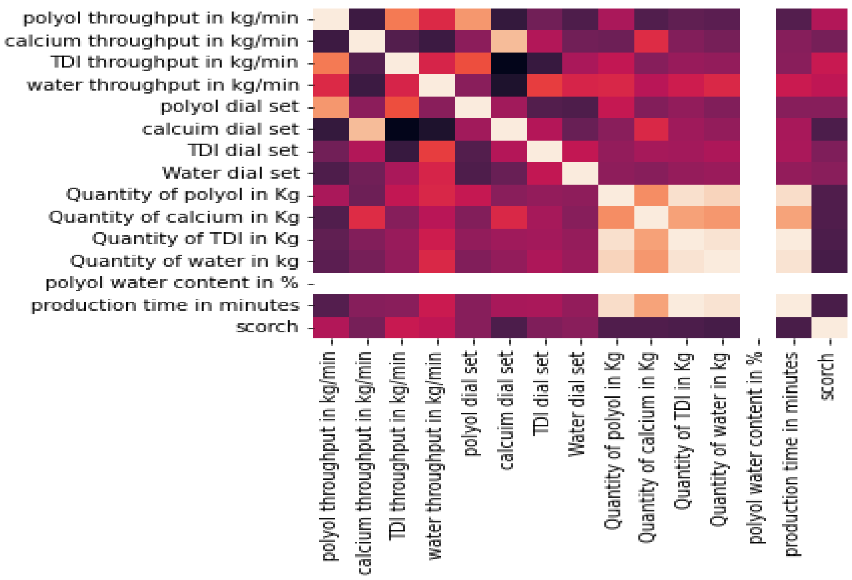

Figure 2. Heatmap plot for the Winco Foam Limited Dataset

Table 1. The Winco Foam Company Dataset Description

Items	**Poly thru**	**Calc thru**	**TDI thru**	**Water thru**	**Poly dial**	**Calc dial**	**TDI dial**	**Water dial**	**Qnty Poly**	**Qnty Calc**	**Qnty TDI**	**Qnty Water**	**Prod Time**	**Scorch**
Mean	71.996	13.142	54.643	4.4988	13.469	18.280	68.111	278.28	1485.5	293.86	1141.5	96.276	0.0800	20.908
Std	9.3113	1.9250	1.4917	0.0865	3.6109	2.1598	0.8544	10.053	729.86	175.89	573.91	46.816	0.0004	10.501
25%	75.000	11.250	54.903	4.4010	11.300	16.100	68.000	270.00	1042.4	194.06	756.23	65.965	0.0008	14.690
50%	75.000	14.005	55.420	4.5600	14.050	18.800	68.000	280.00	1350.0	262.02	1042.5	89.645	0.0008	20.002
75%	75.000	15.000	55.462	4.5640	14.950	20.500	68.000	280.00	1989.0	383.43	1470.3	125.25	0.0008	28.193
Max	80.000	16.000	55.610	4.7000	23.610	20.800	71.000	318.00	3000.0	923.85	2625.0	228.40	0.0008	50.000

2. **Step 2 – Preprocessing** cleans up the dataset by removing duplicates to improve data quality, and remove missing values to avoid redundancy. We utilize the one-hot encoding [60] method that converts categorical values into a suitable form for the ML models; Since ML schemes cannot handle category data directly, it creates a binary equivalence of the dataset by converting categorical variables into their binary form.

3. **Step 3 – Recursive Elimination Feature Selection:** Feature selection selects and extracts what data is input (X), and determines what label will yield ensemble output (Y). It removes all irrelevant features with no importance to the quest for ground truth. This, in turn, reduces the dimensionality of the chosen dataset [61] and fastens the model's construction for improved performance [62], especially in cases where cost is a critical factor. The efficiency of a selected feature is evaluated on how well the model fits to ground truth (i.e. target class). We use the wrapper-based recursive feat elimination [63] mode to unveil how relevant, and ascertain how its occurrence fits with the target class. With the original dataset consisting of 13 features, we categorized the correlation of parameters to the (scorch) target class. With a computed threshold of 9.32 – a total of seven (7) parameters were selected: (a) polyurethane throughput, (b) calcium throughput, (c) water throughput, (d) quantity of water, (e) production time, (f) quantity of polyurethane, and (g) scorch as in Table 2 as examined concerning their correlated contribution to ground truth.

Table 2. Ranking of Attributes score using the Chi-Square

Features	**Selected (Yes/No)**	**X^2-Value**
Polyurethane_throughput	Yes	13.364
calcium_throughput	Yes	15.419
TDI_throughput	No	0.9562
Water_throughput	Yes	20.012
Polyurethane_dial	No	0.2489
calcium_dial	No	2.4701
TDI_dial	No	8.4920
water_dial	No	8.3721

quantity_of_polyurethane	Yes	88.222
quantity_of_calcium	No	0.2589
quantity_of_TDI	No	3.0298
quantity_of_water	Yes	18.006
production_time	Yes	23.092
Scorch	Yes	16.0929

4. **Step 4 – Data Balancing** seeks to redistribute the data points to ensure an almost equitable distribution between major and minor classes. Here, we adopt SMOTE-Tomek [53]: (a) identifies the majority class, (b) interpolates to create synthetic data points via Tomek-link undersample for the majority class, (c) adjusts data points to those of its closest neighbors so that new points overlaps, and (d) adds generated synthetic data to the original dataset to yield a balanced dataset as in Figure 3a and Figure 3b. Afterward, the dataset is split into 75% for the training dataset and 25% remainder for the test dataset [64].

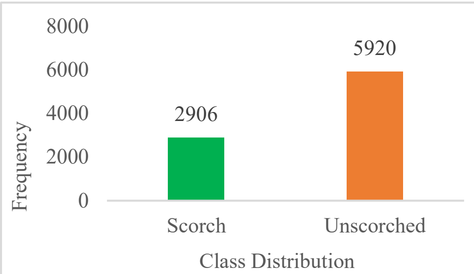

Figure 3a. Original Dataset plot

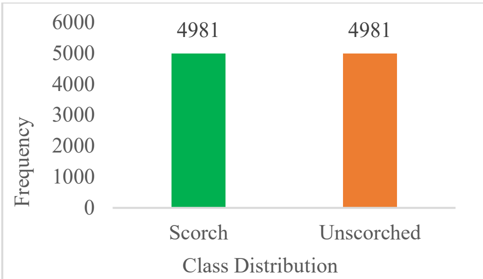

Figure 3b. SMOTE-Tomek Links applied

5. **Step 5 – ML Initialization:** This is further explained as thus:
The Bidirectional Long Short-Term Memory (BiLSTM) based on the RNN, is useful in handling large datasets [50]. The RNN yields a gradient vanishing problem such that its gradient for the learning process becomes quite small. This slows down or eventually stops all forms of learning within the model. LSTM overcomes this challenge via the utilization of (input, forget, and output) gates that effectively allow the network to learn when to 'recall', and 'forget' irrelevant knowledge. In addition, its cell state update function (C_t) maintains all important knowledge over the period and is not impaired or degraded by the vanishing gradient problem. The gates are constructed using the Equation (1)-(3) respectively as [65], [66]:

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad (1a)$$
$$f_t = \sigma\left(W_f[h_{t-1}, x_t] + b_f\right) \quad (1b)$$
$$o_t = \sigma\left(W_f[h_{t-1}, x_t] + b_o\right) \quad (1c)$$
$$\bar{C_t} = tanh(W_c[h_{t-1}, x_t] + b_c) \quad (2a)$$
$$C_t = f_t * C_{t-1} + i_t * C_t \quad (2b)$$
$$h_t = o_t * \tanh(C_t) \quad (3)$$

With i_t as activation of the input gate, o_t is the activation of the output gate, σ is the sigmoid function, W_f is the weight of the forget gate, h_{t-1} is hidden state of the previous timestamp, x_t is input at t, b_f is bias for forget gate, $\bar{C_t}$ is candidate value for the memory cell, and h_{t-1} is the hidden state at t. BiLSTM as a variant of LSTM can process data via forward/backward formations. Its first layer allows data flow in a direction (source to destination); while, the second layer reverses data flow (destination to source) so that the network possesses the past and future context of the dataset [67], [68]. BiLSTM offers greater flexibility via the fusion of knowledge from both directions. It carefully utilizes hyperparameters that tune the model to avoid slow convergence, model overfit, memory efficiency, and task distribution as seen in Table 3.

Table 3. BiLSTM Design and Model configuration with Hyper-predictor tuning

Predictor Settings	**Value(s)**	**Description**
RNN_layer	Bidirectional (LSTM(64))	Bidirectional RNN: 64 LSTM (first layer) and 32 LSTM (second layer)
retun_sequence	True (for the first layer)	Returns the entire output sequence for the first layer
input_shape	x_train_scaled.shape[1], 1	Same length as the number of predictors in x_train_scaled, one feat per timestep

dense_layer	y_train_resampled_max( ) + 1	Layer has the same units as classes in y_train_resampled / output_layer
activation_dense_layer	Softmax	Activation function used in the output for multi-class classification
optimizer	Adam	learning_rate=0.001, beta_1=0.9, beta_2=0.999, epsilon-1e-07
loss_function	categorical_crossentry	The loss function for multi-class classification
metrics	accuracy	Metrics upon which the model is evaluated during training and retraining

6. **Step 6 – Training** as applied estimates learned skills on unseen data. It evaluates the model's performance about its accuracy on how well it has learned the feats of interest via the resampling method. We use a stratified k-fold that rearranges the data to ensure that each fold is a good representation of the dataset. Our ensemble learns from scratch via a pre-designated training dataset and iteratively constructs the decision trees for the RF model [69]. Each tree is trained via bootstrap resampling on the enhanced train dataset; And each tree's collective knowledge is enhanced by this, to identify intricate patterns present in each dataset. Training dataset blends actual examples that guarantee the tree's comprehensive learning experience; And thus, improve its flexibility to the various settings inside the dataset [70]–[72].

3.0. Findings and Discussion

3.1. Training Evaluation and Hyper-Parameters Tuning

Table 4 shows the proposed BiLSTM and benchmark models with hyper-parameters tuned [73], [74].

Table 4. Performance with and without hyper-parameter tuning

Ensembles	Without Hyper-parameter Tuning				With Hyper-parameter Tuning			
	F1	Accuracy	Precision	Recall	F1	Accuracy	Precision	Recall
Decision Tree	0.5263	0.5193	0.5036	0.5199	0.5583	0.5520	0.5501	0.5503
Logistic Regression	0.5361	0.5278	0.5304	0.5301	0.5596	0.5601	0.5589	0.5582
Random Forest	0.5819	0.5738	0.5729	0.5775	0.5987	0.5948	0.5898	0.5899
XGBoost	0.5898	0.5810	0.5801	0.5829	0.5998	0.5985	0.5978	0.5972
BiLSTM	0.6081	0.6090	0.6001	0.6092	0.6285	0.6249	0.6293	0.6292

The result shows that without hyper-parameter tuning – the proposed BiLSTM model yields F1 of 0.6081 with an Accuracy of 0.6090, Precision of 0.6001, and Recall of 0.6092 respectively; While, the benchmark models (i.e. Decision Tree, Logistic Regression, Random Forest and XGBoost) yield as F1 of [0.5263, 0.5361, 0.5819 and 0.5898] with Accuracy of [0.5193, 0.5278, 0.5738 and 0.5810] respectively. Their corresponding Recall and Precision value ranges are seen in Table 4. Conversely, with the tuning application of the hyper-parameters – the proposed BiLSTM proffered F1 of 0.6285 with Accuracy of 0.6249, Precision of 0.6293, and Recall of 0.6292 respectively; while the benchmark models (i.e. Decision Tree, Logistic Regression, Random Forest and XGBoost) yield as F1 of [0.5583, 0.5596, 0.5987 and 0.5998] with Accuracy of [0.5520, 0.5601, 0.5948 and 0.5998] respectively. Their corresponding Recall and Precision values are as in Table 4, and the result affirms that the BiLSTM model outperforms all the benchmarked models explored [75].

Table 5 shows the evaluation report for the proposed and benchmark models with(out) the chi-square feature selection technique explored [76] on the outlier effects of data not previously present from the outset as contained in the domain task dataset.

Table 5. Performance with and without the chi-square feature selection technique applied

Ensembles	Hyper-parameter tuned without chi-square.				Hyper-parameter tuned with chi-square.			
	F1	Accuracy	Precision	Recall	F1	Accuracy	Precision	Recall
Decision Tree	0.5583	0.5520	0.5501	0.5503	0.7208	0.7100	0.7457	0.7398
Logistic Regression	0.5596	0.5601	0.5589	0.5582	0.8219	0.7747	0.7747	0.7506
Random Forest	0.5987	0.5948	0.5898	0.5899	0.8435	0.8318	0.8357	0.8245
XGBoost	0.5998	0.5985	0.5978	0.5972	0.8508	0.8403	0.8562	0.8582
BiLSTM	0.6285	0.6249	0.6293	0.6292	0.8832	0.8502	0.8689	0.8901

The result shows a hyperparameter-tuned proposed BiLSTM model with chi-square feature selection applied yields F1 of 0.8832 with Accuracy of 0.8502, Precision of 0.8689, and Recall of 0.8901 respectively; While, the benchmark models (i.e. Decision Tree, Logistic Regression, Random Forest and XGBoost) yield as F1 of [0.7208, 0.8219, 0.8435 and 0.8508] with Accuracy of [0.71, 0.7747, 0.8318 and 0.8403] respectively. Again, their corresponding Recall and Precision values are as in Table 5 – wherein the results further affirm that the proposed BiLSTM model outperforms all the benchmarked models explored [77], [78].

Table 6 shows the utilization of the SMOTE-Tomek data balancing scheme on the hyperparameter-tuned with chi-square feature selection for both, the proposed and benchmark models on the explored dataset. The result shows that the proposed BiLSTM outperformed benchmark models with an Accuracy of 0.9895 with F1 of 0.9892, Precision of 0.9817, and Recall of 0.9901 respectively; While, the benchmark models (i.e. Decision Tree, Logistic Regression, Random Forest and XGBoost) yield as F1 of [0.8145, 0.9105, 0.9210 and 0.9125] with Accuracy of [0.8032, 0.9105, 0.9228, and 0.9574] respectively [79], [80]. Their corresponding Precision and Recall values are as in Table 6.

Table 6. Performance with and without the SMOTE-Tomek Data Balancing Approach

Ensembles	Hyper-parameter tuned without chi-square.				Hyper-parameter tuned with chi-square.			
	F1	**Accuracy**	**Precision**	**Recall**	**F1**	**Accuracy**	**Precision**	**Recall**
Decision Tree	0.7208	0.7100	0.7457	0.7398	0.8145	0.8032	0.8541	0.8528
Logistic Regression	0.8219	0.7747	0.7747	0.7506	0.9105	0.9105	0.9105	0.9114
Random Forest	0.8435	0.8318	0.8357	0.8245	0.9210	0.9228	0.9480	0.9500
XGBoost	0.8508	0.8403	0.8562	0.8582	0.9125	0.9574	0.9616	0.9609
BiLSTM	0.8832	0.8502	0.8689	0.8901	0.9892	0.9895	0.9817	0.9901

Figure 4a shows the training-and-validation accuracy with a consistent rise from 0.69 in the second epoch to 0.98 by the tenth epoch. This implies the BiLSTM minimizes errors, and that it does not overfit; Rather, it captures the intricate features in the dataset and generalizes well with new data. Its trend demonstrates healthy learning to reliably predict scorch occurrence with increased accuracy. In addition, Figure 4b shows a smooth, monotonic decrease in loss without a sudden burst. It implies the BiLSTM learns well even with the instability pg the vanishing gradients, to accurately and consistently handle the minority class with balancing.

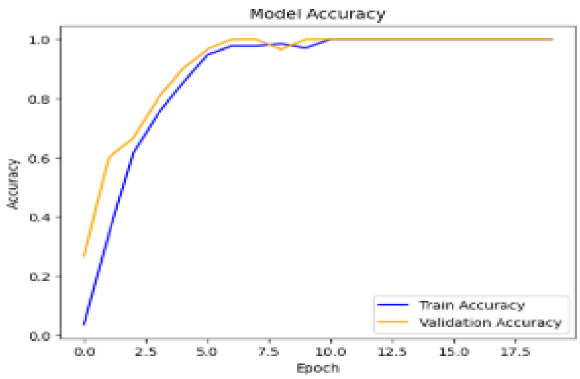

Figure 4a. Training and Validation Accuracy

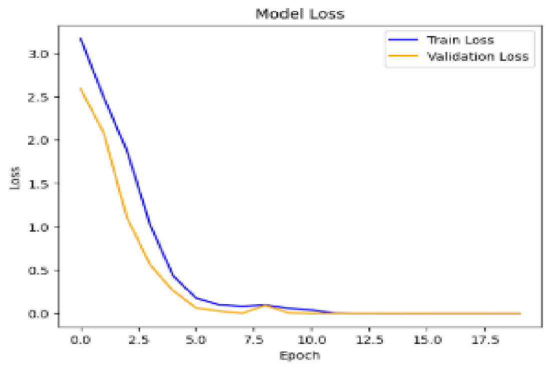

Figure 4b. Training and Validation Loss

The proposed model accurately identifies scorch in the adopted dataset, and has proven to efficiently reduce bias and variance indicative as in Figure 5 [81] to yield a robust model for new data or hidden underlying parameters of interest within a domain's training dataset being considered. The proposed model with an accuracy of 98.95% accurately classified 2475 instances with only 15 incorrectly classified instances to yield 2-distinct predictions for scorch occurrence as in Table 7. Our model successfully predicts the occurrence of the scorch, And its practical implementation saw a significant prediction of scorch occurrence bin before the mixture that unveils a variety of intertwined relations between its constituent features and water as in Table 7.

854	5
10	1,621

Figure 4. Confusion Matrix for the Proposed Ensemble

Table 7. Predicted Values with the value ‘1’ indicate the presence of scorch

Poly thru	Calc thru	TDI thru	Water thru	Poly dial	Calc dial	TDI dial	Water dial	Qnty Poly	Qnty Calc	Qnty TDI	Qnty Water	Prod Time	Scorch
75	11.25	55.50	4.564	11.3	75	68	280	1500	225	91.28	1110	20	0(No)
75	11.25	55.42	4.564	11.3	75	68	280	478.5	71.77	101.1	353.58	6.38	1(Yes)

Our study supports the evidence that SMOTE-Tomek data balancing improves the quest for ground truth and overall performance [82]. It also enhanced model efficiency between true-positives-and-negative vis-à-vis false-positive-and-negative [34], [83] to successfully yield a resultant prediction on the presence of scorch. Data access was severely constrained as facilities yield large amounts of production data daily [84]. Its documentation is rather antiquated and quickly lost if not properly used. The dataset used was retrieved and restricted to only a short duration [85], [86]. The plot in Figure 2 ascertains the skewness of the most important column head ‘scorch’ about density [87], [88]. Its data is fairly distributed across the upper quartiles from the mean. Minimizing the outlier effects as agreed by [89], [90]. The addition of water as a universal solvent has consequently yielded a correlation therein, which is already known feat. It further suggests that water unveils a variety of intertwined relationships between other features, the polyurethane, and water in the dataset. All negative values on the correlation also yield a corresponding negative relationship between both variables and vice versa. And is visually confirmed from the pair plots that agree with [58], [91].

3.2. Comparison/Benchmark

Table 8 yields a benchmark against previous methods that have utilized the same dataset:

Table 8. Benchmarking and Comparative Testing of Proposed Stacking Ensemble

Methods	F1	Accuracy	Precision	Recall
Ref [56]	0.7902	0.7815	0.7372	0.7025
Ref [92]	0.9881	0,9968	0.9848	0.9318
Ref [93]	0.9831	0.9881	0.9783	0.9326
Our Method (SMOTE-Tomek BiLSTM)	0.9892	0.9885	0.9689	0.9901

Results show model by [93] utilized a tree-based XGB ensemble with performance generalization that can also be found to almost be as good as the SMOTE-Tomek-based BiLSTM. This can also be attributed to their utilization of the chi-square feature selection approach and a normalization scaler to ensure a more balanced dataset. However, some task(s) require that the explored ensemble design metric is strongly impacted by the consequence of errors within the captured dataset. Thus, the measure of both specificity and sensitivity becomes 2 critical feats to be evaluated since they are directly related to the patient clinical outcomes.

4.0 Conclusion

Advances in technological development and the widespread adoption of technology-driven business strategies, businesses to operate more efficiently, productively, and profitably. Despite the enormous amount of data generated daily, the polyurethane foam production industry still lags in developing data analytics tools. This study is a positive step and should be improved upon. With the dataset used, NaN (not a number) values imply no relationship; As such, a preprocessing scheme should be adopted to prevent model overfitting, and column flattening prior to the deployment of the ML scheme, which agrees with. These correlation numbers are illogical at first glance. But ML can successfully glean insightful knowledge therein. These algorithms can trace the entangled connections and interpret how each variable interacts with the others and influences the column (variable) that we want to predict.

References

[1] A. Okewale and F. Omoruwuo, “Neem Leaf Extract as a Corrosion Inhibitor on Mild Steel in Acidic Solution,” *Int. J. Eng. Res. Africa*, vol. 35, pp. 208–220, Mar. 2018, doi: 10.4028/www.scientific.net/JERA.35.208.

[2] D. R. I. M. Setiadi *et al.*, “Single Qubit Quantum Logistic-Sine XYZ-Rotation Maps: An Ultra-Wide Range Dynamics for Image Encryption,” *Comput. Mater. Contin.*, vol. 83, no. 2, pp. 1–28, 2025, doi: 10.32604/cmc.2025.063729.

[3] N. V. Gama, A. Ferreira, and A. Barros-Timmons, “Polyurethane foams: Past, present, and future,” *Materials (Basel).*, vol. 11, no. 10, 2018, doi: 10.3390/ma11101841.

[4] R. E. Yoro, F. O. Aghware, M. I. Akazue, A. E. Ibor, and A. A. Ojugo, “Evidence of personality traits on phishing attack menace among selected university undergraduates in Nigerian,” *Int. J. Electr. Comput. Eng.*, vol. 13, no. 2, p. 1943, Apr. 2023, doi: 10.11591/ijece.v13i2.pp1943-1953.

[5] M. I. Akazue, A. A. Ojugo, R. E. Yoro, B. O. Malasowe, and O. Nwankwo, “Empirical evidence of phishing menace among undergraduate smartphone users in selected universities in Nigeria,” *Indones. J. Electr. Eng. Comput. Sci.*, vol. 28, no. 3, pp. 1756–1765, Dec. 2022, doi: 10.11591/ijeecs.v28.i3.pp1756-1765.

[6] A. A. Ojugo and A. O. Eboka, “Empirical Bayesian network to improve service delivery and performance dependability on a campus network,” *IAES Int. J. Artif. Intell.*, vol. 10, no. 3, p. 623, Sep. 2021, doi: 10.11591/ijai.v10.i3.pp623-635.

[7] S. . Okperigho, B. Nwozor, and V. . Geteloma, "Deployment of an IoT Storage Tank Gauge and Monitor," *FUPRE J. Sci. Ind. Res.*, vol. 8, no. 1, pp. 55–68, 2024.

[8] M. Iqbal, J. Putra, M. M. Anugerah, and A. Akbar, "Sentinel 1 (SAR) and Sentinel 2 (MSI) Imagery Capabilities for Oil Spill Detection in Balikpapan Bay," *J. Hydrol.*, vol. 1, no. July 2019, pp. 1–8, 2020.

[9] S. Levchik, M. P. Luda, P. Bracco, P. Nada, and L. Costa, "Discoloration in fire-retardant flexible polyurethane foams," *J. Cell. Plast.*, vol. 41, no. 3, pp. 235–248, 2005, doi: 10.1177/0021955X05053523.

[10] V. O. Geteloma *et al.*, "Enhanced data augmentation for predicting consumer churn rate with monetization and retention strategies : a pilot study," *Appl. Eng. Technol.*, vol. 3, no. 1, pp. 35–51, 2024, doi: 10.31763/aet.v3i1.1408.

[11] B. O. Malasowe, F. O. Aghware, M. D. Okpor, B. E. Edim, R. E. Ako, and A. A. Ojugo, "Techniques and Best Practices for Handling Cybersecurity Risks in Educational Technology Environment (EdTech)," *J. Sci. Technol. Res.*, vol. 6, no. 2, pp. 293–311, 2024, doi: 10.5281/zenodo.12617068.

[12] F. M. A. Hossain, Y. M. Zhang, and M. A. Tonima, "Forest fire flame and smoke detection from UAV-captured images using fire-specific color features and multi-color space local binary pattern," *J. Unmanned Veh. Syst.*, vol. 8, no. 4, pp. 285–309, Dec. 2020, doi: 10.1139/juvs-2020-0009.

[13] A. A. Ojugo, P. O. Ejeh, C. C. Odiakaose, A. O. Eboka, and F. U. Emordi, "Predicting rainfall runoff in Southern Nigeria using a fused hybrid deep learning ensemble," *Int. J. Informatics Commun. Technol.*, vol. 13, no. 1, p. 108, Apr. 2024, doi: 10.11591/ijict.v13i1.pp108-115.

[14] J. Demassa, "Polyol Stabilization and the Introduction of a New PUR Slabstock Foam Antioxidant," *Cell. Plast.*, vol. 41, no. December, pp. 15–24, 2013.

[15] R. E. Yoro, F. O. Aghware, B. O. Malasowe, O. Nwankwo, and A. A. Ojugo, "Assessing contributor features to phishing susceptibility amongst students of petroleum resources varsity in Nigeria," *Int. J. Electr. Comput. Eng.*, vol. 13, no. 2, p. 1922, Apr. 2023, doi: 10.11591/ijece.v13i2.pp1922-1931.

[16] S. Umar, G. O. Adejo, N. U. Imam, A. S. Zaharaddeen, Z. S. Abubakar, and A. Abdullahi, "Diet Fortification with Curcuma longa and Allium cepa Ameliorates 2 , 3 , 7 , 8-Tetracholorodibenzo- p-dioxin-induced Dyslipideamia and Oxidative Stress in Wistar Rats Length Article," *Niger. J. Basic Appl. Sci.*, vol. 30, no. 2, pp. 178–183, 2022, doi: 10.4314/njbas.v30i2.23.

[17] Y. Nabata, A. Mamada, and H. Yamasaki, "Study of the curing process giving the rigid polyurethane foam by dynamic viscoelastic method," *J. Appl. Polym. Sci.*, vol. 35, no. 1, pp. 155–166, Jan. 1988, doi: 10.1002/app.1988.070350114.

[18] S. N. Okofu *et al.*, "Pilot Study on Consumer Preference, Intentions and Trust on Purchasing-Pattern for Online Virtual Shops," *Int. J. Adv. Comput. Sci. Appl.*, vol. 15, no. 7, pp. 804–811, 2024, doi: 10.14569/IJACSA.2024.0150780.

[19] A. O. Eboka and A. A. Ojugo, "Mitigating technical challenges via redesigning campus network for greater efficiency, scalability and robustness: A logical view," *Int. J. Mod. Educ. Comput. Sci.*, vol. 12, no. 6, pp. 29–45, 2020, doi: 10.5815/ijmecs.2020.06.03.

[20] V. O. Geteloma *et al.*, "AQuaMoAS: unmasking a wireless sensor-based ensemble for air quality monitor and alert system," *Appl. Eng. Technol.*, vol. 3, no. 2, pp. 86–101, 2024, doi: 10.31763/aet.v3i2.1409.

[21] A. E. Ibor, B. E. Edim, and A. A. Ojugo, "Secure Health Information System with Blockchain Technology," *J. Niger. Soc. Phys. Sci.*, vol. 5, no. 992, p. 992, Apr. 2023, doi: 10.46481/jnsps.2023.992.

[22] A. O. Eboka *et al.*, "Resolving Data Imbalance Using a Bi-Directional Long-Short Term Memory for Enhanced Diabetes Mellitus Detection," *J. Futur. Artif. Intell. Technol.*, vol. 2, no. 1, pp. 95–109, 2025, doi: 10.62411/faith.3048-3719-73.

[23] D. R. I. M. Setiadi *et al.*, "Integrating Hybrid Statistical and Unsupervised LSTM-Guided Feature Extraction for Breast Cancer Detection," *J. Comput. Theor. Appl.*, vol. 2, no. 4, pp. 536–550, 2025, doi: 10.62411/jcta.12698.

[24] R. L. Gray and R. E. Lee, "Scorch inhibitors for flexible polyurethanes," 1998, pp. 567–575. doi: 10.1007/978-94-011-5862-6_63.

[25] M. Szycher, *Szycher's Handbook of Polyurethanes*, no. December 2021. CRC Press, 2012. doi: 10.1201/b12343.

[26] J. K. Oladele *et al.*, "BEHeDaS: A Blockchain Electronic Health Data System for Secure Medical Records Exchange," *J. Comput. Theor. Appl.*, vol. 1, no. 3, pp. 231–242, 2024, doi: 10.62411/jcta.9509.

[27] R. Vinayakumar, M. Alazab, K. P. Soman, P. Poornachandran, A. Al-Nemrat, and S. Venkatraman, "Deep Learning Approach for Intelligent Intrusion Detection System," *IEEE Access*, vol. 7, pp. 41525–41550, 2019, doi: 10.1109/ACCESS.2019.2895334.

[28] A. Okewale, F. Omoruwuo, and O. A. Adesina, "Comparative Studies of Response Surface Methodology (RSM) and Predictive Capacity of Artificial Neural Network (ANN) on Mild Steel Corrosion Inhibition using Water Hyacinth as an Inhibitor," *J. Phys. Conf. Ser.*, vol. 1378, no. 2, p. 022002, Dec. 2019, doi: 10.1088/1742-6596/1378/2/022002.

[29] I. Tubert-Brohman, W. Sherman, M. Repasky, and T. Beuming, "Improved docking of polypeptides with glide," *J. Chem. Inf. Model.*, vol. 53, no. 7, pp. 1689–1699, 2013, doi: 10.1021/ci400128m.

[30] A. A. Ojugo and D. A. Oyemade, "Boyer moore string-match framework for a hybrid short message service spam filtering technique," *IAES Int. J. Artif. Intell.*, vol. 10, no. 3, pp. 519–527, 2021, doi: 10.11591/ijai.v10.i3.pp519-527.

[31] F. Omoruwou, A. Okewale, and C. N. Owabor, "Statistical Analysis of Corrosion Inhibition of Water Hyacinth on Mild Steel in an Acidic Medium," *J. Environ. Anal. Toxicol.*, vol. 07, no. 04, 2017, doi: 10.4172/2161-0525.1000481.

[32] B. O. Malasowe, M. I. Akazue, A. E. Okpako, F. O. Aghware, D. V. Ojie, and A. A. Ojugo, "Adaptive Learner-CBT with Secured Fault-Tolerant and Resumption Capability for Nigerian Universities," *Int. J. Adv. Comput. Sci. Appl.*, vol. 14, no. 8, pp. 135–142, 2023, doi: 10.14569/IJACSA.2023.0140816.

[33] A. A. Ojugo and O. D. Otakore, "Intelligent cluster connectionist recommender system using implicit graph friendship algorithm for social networks," *IAES Int. J. Artif. Intell.*, vol. 9, no. 3, p. 497~506, 2020, doi: 10.11591/ijai.v9.i3.pp497-506.

[34] S. T. Seydi, M. Hasanlou, M. Amani, and W. Huang, "Oil Spill Detection Based on Multiscale Multidimensional Residual CNN for Optical Remote Sensing Imagery," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 14, no. November, pp. 10941–10952, 2021, doi: 10.1109/JSTARS.2021.3123163.

[35] M. I. Akazue, R. E. Yoro, B. O. Malasowe, O. Nwankwo, and A. A. Ojugo, "Improved services traceability and management of a food value chain using block-chain network : a case of Nigeria," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 29, no. 3, pp. 1623–1633, 2023, doi: 10.11591/ijeecs.v29.i3.pp1623-1633.

[36] M. I. Akazue *et al.*, "Handling Transactional Data Features via Associative Rule Mining for Mobile Online Shopping Platforms," *Int. J. Adv. Comput. Sci. Appl.*, vol. 15, no. 3, pp. 530–538, 2024, doi: 10.14569/IJACSA.2024.0150354.

[37] D. R. Houck *et al.*, "NYX-2925, A Novel N-methyl-D-aspartate Receptor Modulator: A First-in-Human, Randomized, Double-blind Study of Safety and Pharmacokinetics in Adults," *Clin. Transl. Sci.*, vol. 12, no. 2, pp. 164–171, 2019, doi: 10.1111/cts.12584.

[38] B. Kizilkaya, E. Ever, H. Y. Yatbaz, and A. Yazici, "An Effective Forest Fire Detection Framework Using Heterogeneous Wireless Multimedia Sensor Networks," *ACM Trans. Multimed. Comput. Commun. Appl.*, vol. 18, no. 2, pp. 1–21, May 2022, doi: 10.1145/3473037.

[39] A. A. Ojugo and R. E. Yoro, "Forging a deep learning neural network intrusion detection framework to curb the distributed denial of service attack," *Int. J. Electr. Comput. Eng.*, vol. 11, no. 2, pp. 1498–1509, 2021, doi: 10.11591/ijece.v11i2.pp1498-1509.

[40] A. A. Ojugo, A. O. Eboka, R. E. Yoro, M. O. Yerokun, and F. N. Efozia, "Hybrid Model for Early Diabetes Diagnosis," in *2015 Second*

International Conference on Mathematics and Computers in Sciences and in Industry (MCSI), IEEE, Aug. 2015, pp. 55–65. doi: 10.1109/MCSI.2015.35.

[41] H. Zhang, B. He, J. Xing, and M. Lu, "Deep spatial and temporal graph convolutional network for rice planthopper population dynamic forecasting," *Comput. Electron. Agric.*, vol. 210, no. 3, p. 107868, Jul. 2023, doi: 10.1016/j.compag.2023.107868.

[42] E. U. Omede, A. E. Edje, M. I. Akazue, H. Utomwen, and A. A. Ojugo, "IMANoBAS: An Improved Multi-Mode Alert Notification IoT-based Anti-Burglar Defense System," *J. Comput. Theor. Appl.*, vol. 1, no. 3, pp. 273–283, Feb. 2024, doi: 10.62411/jcta.9541.

[43] A. A. Ojugo, P. O. Ejeh, C. C. Odiakaose, A. O. Eboka, and F. U. Emordi, "Improved distribution and food safety for beef processing and management using a blockchain-tracer support framework," *Int. J. Informatics Commun. Technol.*, vol. 12, no. 3, p. 205, Dec. 2023, doi: 10.11591/ijict.v12i3.pp205-213.

[44] F. O. Aghware *et al.*, "Enhancing the Random Forest Model via Synthetic Minority Oversampling Technique for Credit-Card Fraud Detection," *J. Comput. Theor. Appl.*, vol. 1, no. 4, pp. 407–420, Mar. 2024, doi: 10.62411/jcta.10323.

[45] E. Ileberi, Y. Sun, and Z. Wang, "A machine learning based credit card fraud detection using GA algorithm for feature selection," *J. Big Data*, vol. 9, no. 1, p. 24, Dec. 2022, doi: 10.1186/s40537-022-00573-8.

[46] M. D. Okpor *et al.*, "Comparative Data Resample to Predict Subscription Services Attrition Using Tree-based Ensembles," *J. Fuzzy Syst. Control*, vol. 2, no. 2, pp. 117–128, 2024, doi: 10.59247/jfsc.v2i2.213.

[47] I. Benchaji, S. Douzi, B. El Ouahidi, and J. Jaafari, "Enhanced credit card fraud detection based on attention mechanism and LSTM deep model," *J. Big Data*, vol. 8, no. 1, p. 151, Dec. 2021, doi: 10.1186/s40537-021-00541-8.

[48] S. Xuan, G. Liu, Z. Li, L. Zheng, S. Wang, and C. Jiang, "Random forest for credit card fraud detection," in *2018 IEEE 15th International Conference on Networking, Sensing and Control (ICNSC)*, IEEE, Mar. 2018, pp. 1–6. doi: 10.1109/ICNSC.2018.8361343.

[49] O. Plonsky *et al.*, "Predicting human decisions with behavioral theories and machine learning," *ICT Express*, vol. 12, no. 2, pp. 30–41, Apr. 2019, [Online]. Available: http://arxiv.org/abs/1904.06866

[50] L. R. Zuama, D. R. I. M. Setiadi, A. Susanto, S. Santosa, and A. A. Ojugo, "High-Performance Face Spoofing Detection using Feature Fusion of FaceNet and Tuned DenseNet201," *J. Futur. Artif. Intell. Technol.*, vol. 1, no. 4, pp. 385–400, 2025, doi: 10.62411/faith.3048-3719-62.

[51] M. A. Rahman, G. A. Francia, and H. Shahriar, "Leveraging GANs for Synthetic Data Generation to Improve Intrusion Detection Systems," *J. Futur. Artif. Intell. Technol.*, vol. 1, no. 4, pp. 429–439, Feb. 2025, doi: 10.62411/faith.3048-3719-52.

[52] M. F. Ijaz, M. Attique, and Y. Son, "Data-Driven Cervical Cancer Prediction Model with Outlier Detection and Over-Sampling Methods," *Sensors*, vol. 20, no. 10, p. 2809, May 2020, doi: 10.3390/s20102809.

[53] D. R. I. M. Setiadi, K. Nugroho, A. R. Muslikh, S. W. Iriananda, and A. A. Ojugo, "Integrating SMOTE-Tomek and Fusion Learning with XGBoost Meta-Learner for Robust Diabetes Recognition," *J. Futur. Artif. Intell. Technol.*, vol. 1, no. 1, pp. 23–38, May 2024, doi: 10.62411/faith.2024-11.

[54] W.-L. Hsu, J.-Y. Jhuang, C.-S. Huang, C.-K. Liang, and Y.-C. Shiau, "Application of Internet of Things in a Kitchen Fire Prevention System," *Appl. Sci.*, vol. 9, no. 17, p. 3520, Aug. 2019, doi: 10.3390/app9173520.

[55] H. A. Ibrahim and H. S. Syed, "Hazard Analysis of Crude Oil Storage Tank Farm," *Int. J. ChemTech Res.*, vol. 11, no. 11, pp. 300–308, 2018, doi: 10.20902/IJCTR.2018.111132.

[56] U. A. Amran *et al.*, "Production of Rigid Polyurethane Foams Using Polyol from Liquefied Oil Palm Biomass: Variation of Isocyanate Indexes," *Polymers (Basel).*, vol. 13, no. 18, p. 3072, Sep. 2021, doi: 10.3390/polym13183072.

[57] J. P. Ntayagabiri, Y. Bentaleb, J. Ndikumagenge, and H. El Makhtoum, "A Comparative Analysis of Supervised Machine Learning Algorithms for IoT Attack Detection and Classification," *J. Comput. Theor. Appl.*, vol. 2, no. 3, pp. 395–409, Feb. 2025, doi: 10.62411/jcta.11901.

[58] A. A. Ojugo and R. E. Yoro, "Extending the three-tier constructivist learning model for alternative delivery: ahead the COVID-19 pandemic in Nigeria," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 21, no. 3, p. 1673, Mar. 2021, doi: 10.11591/ijeecs.v21.i3.pp1673-1682.

[59] S. Adamu, A. Iorliam, and Ö. Asilkan, "Exploring Explainability in Multi-Category Electronic Markets: A Comparison of Machine Learning and Deep Learning Approaches," *J. Futur. Artif. Intell. Technol.*, vol. 1, no. 4, pp. 440–454, Mar. 2025, doi: 10.62411/faith.3048-3719-58.

[60] C. C. Odiakaose *et al.*, "Hypertension Detection via Tree-Based Stack Ensemble with SMOTE-Tomek Data Balance and XGBoost Meta-Learner," *J. Futur. Artif. Intell. Technol.*, vol. 1, no. 3, pp. 269–283, Dec. 2024, doi: 10.62411/faith.3048-3719-43.

[61] H. Lu and C. Rakovski, "The Effect of Text Data Augmentation Methods and Strategies in Classification Tasks of Unstructured Medical Notes," *Res. Sq.*, vol. 1, no. 1, pp. 1–29, 2022, [Online]. Available: https://doi.org/10.21203/rs.3.rs-2039417/v1

[62] D. R. I. M. Setiadi, A. R. Muslikh, S. W. Iriananda, W. Warto, J. Gondohanindijo, and A. A. Ojugo, "Outlier Detection Using Gaussian Mixture Model Clustering to Optimize XGBoost for Credit Approval Prediction," *J. Comput. Theor. Appl.*, vol. 2, no. 2, pp. 244–255, Nov. 2024, doi: 10.62411/jcta.11638.

[63] A. Bahl *et al.*, "Recursive feature elimination in random forest classification supports nanomaterial grouping," *NanoImpact*, vol. 15, p. 100179, Mar. 2019, doi: 10.1016/j.impact.2019.100179.

[64] M. D. Okpor *et al.*, "Pilot Study on Enhanced Detection of Cues over Malicious Sites Using Data Balancing on the Random Forest Ensemble," *J. Futur. Artif. Intell. Technol.*, vol. 1, no. 2, pp. 109–123, 2024, doi: 10.62411/faith.2024-14.

[65] M. F. Aslan, M. F. Unlersen, K. Sabanci, and A. Durdu, "CNN-based transfer learning–BiLSTM network: A novel approach for COVID-19 infection detection," *Appl. Soft Comput.*, vol. 98, p. 106912, Jan. 2021, doi: 10.1016/j.asoc.2020.106912.

[66] J. Dong *et al.*, "Estimating reference crop evapotranspiration using improved convolutional bidirectional long short-term memory network by multi-head attention mechanism in the four climatic zones of China," *Agric. Water Manag.*, vol. 292, no. December 2023, p. 108665, Mar. 2024, doi: 10.1016/j.agwat.2023.108665.

[67] M. I. Akazue, I. A. Debekeme, A. E. Edje, C. Asuai, and U. J. Osame, "UNMASKING FRAUDSTERS : Ensemble Features Selection to Enhance Random Forest Fraud Detection," *J. Comput. Theor. Appl.*, vol. 1, no. 2, pp. 201–212, 2023, doi: 10.33633/jcta.v1i2.9462.

[68] B. O. Malasowe *et al.*, "Quest for Empirical Solution to Runoff Prediction in Nigeria via Random Forest Ensemble: Pilot Study," *Adv. Multidiscip. Sci. Res. J. Publ.*, vol. 10, no. 1, pp. 73–90, Mar. 2024, doi: 10.22624/AIMS/BHI/V10N1P8.

[69] F. O. Aghware *et al.*, "Effects of Data Balancing in Diabetes Mellitus Detection : A Comparative XGBoost and Random Forest Learning Approach," *NIPES - J. Sci. Technol. Res.*, vol. 7, no. 1, pp. 1–11, 2025, doi: 10.37933/nipes/7.1.2025.1.

[70] E. A. Otorokpo *et al.*, "DaBO-BoostE: Enhanced Data Balancing via Oversampling Technique for a Boosting Ensemble in Card-Fraud Detection," *Adv. Multidiscip. Sci. Res. J. Publ.*, vol. 12, no. 2, pp. 45–66, 2024, doi: 10.22624/AIMS/MATHS/V12N2P4.

[71] H. Huang, Y. Song, Z. Fan, G. Xu, R. Yuan, and J. Zhao, "Estimation of walnut crop evapotranspiration under different micro-irrigation techniques in arid zones based on deep learning sequence models," *Results Appl. Math.*, vol. 20, no. September, p. 100412, Nov. 2023, doi: 10.1016/j.rinam.2023.100412.

[72] H. Said, B. B. S. Tawfik, and M. A. Makhlouf, "A Deep Learning Approach for Sentiment Classification of COVID-19 Vaccination Tweets," *Int. J. Adv. Comput. Sci. Appl.*, vol. 14, no. 4, pp. 530–538, 2023, doi: 10.14569/IJACSA.2023.0140458.

[73] C. L. Kumar *et al.*, "Metaparameter optimized hybrid deep learning model for next generation cybersecurity in software defined networking environment," *Sci. Rep.*, vol. 15, no. 1, 2025, doi: 10.1038/s41598-025-96153-w.
[74] J. Yao, C. Wang, C. Hu, and X. Huang, "Chinese Spam Detection Using a Hybrid BiGRU-CNN Network with Joint Textual and Phonetic Embedding," *Electronics*, vol. 11, no. 15, p. 2418, Aug. 2022, doi: 10.3390/electronics11152418.
[75] D. R. I. M. Setiadi, A. Susanto, K. Nugroho, A. R. Muslikh, A. A. Ojugo, and H. Gan, "Rice yield forecasting using hybrid quantum deep learning model," *MDPI Comput.*, vol. 13, no. 191, pp. 1–18, 2024, doi: 10.3390/computers13080191.
[76] G. TekalignTujo, G. Dileep Kumar, D. ElifeneshYitagesu, and B. MeseretGirma, "Predictive Model to Predict Seed Classes using Machine Learning," *Int. J. Eng. Res. \& Technol.*, vol. 6, no. 08, pp. 334–344, 2017.
[77] R. E. Ako *et al.*, "Pilot Study on Fibromyalgia Disorder Detection via XGBoosted Stacked-Learning with SMOTE-Tomek Data Balancing Approach," *NIPES - J. Sci. Technol. Res.*, vol. 7, no. 1, pp. 12–22, 2025, doi: 10.37933/nipes/7.1.2025.2.
[78] C. Ren *et al.*, "Short-Term Traffic Flow Prediction: A Method of Combined Deep Learnings," *J. Adv. Transp.*, vol. 2021, pp. 1–15, Jul. 2021, doi: 10.1155/2021/9928073.
[79] H. Shin, "XGBoost Regression of the Most Significant Photoplethysmogram Features for Assessing Vascular Aging," *IEEE J. Biomed. Heal. Informatics*, vol. 26, no. 7, pp. 3354–3361, Jul. 2022, doi: 10.1109/JBHI.2022.3151091.
[80] A. Raza, K. Munir, M. S. Almutairi, and R. Sehar, "Novel Transfer Learning Based Deep Features for Diagnosis of Down Syndrome in Children Using Facial Images," *IEEE Access*, vol. 12, no. January, pp. 16386–16396, 2024, doi: 10.1109/ACCESS.2024.3359235.
[81] M. D. Okpor *et al.*, "Unmasking effects of feature selection and SMOTE-Tomek in tree-based random forest for scorch occurrence detection," *Bull. Electr. Eng. Informatics*, vol. 14, no. 3, pp. 1–12, 2025, doi: 10.11591/eei.v14i3.8901.
[82] F. O. Aghware *et al.*, "BloFoPASS: A blockchain food palliatives tracer support system for resolving welfare distribution crisis in Nigeria," *Int. J. Informatics Commun. Technol.*, vol. 13, no. 2, p. 178, Aug. 2024, doi: 10.11591/ijict.v13i2.pp178-187.
[83] A. Taravat and F. Del Frate, "Weibull Multiplicative Model and Machine Learning Models for Full-Automatic Dark-Spot Detection From Sar Images," *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.*, vol. XL-1/W3, no. September 2013, pp. 421–424, 2013, doi: 10.5194/isprsarchives-xl-1-w3-421-2013.
[84] A. O. Eboka *et al.*, "Pilot study on deploying a wireless sensor-based virtual-key access and lock system for home and industrial frontiers," *Int. J. Informatics Commun. Technol.*, vol. 14, no. 1, p. 287, Apr. 2025, doi: 10.11591/ijict.v14i1.pp287-297.
[85] M. I. Akazue *et al.*, "FiMoDeAL: pilot study on shortest path heuristics in wireless sensor network for fire detection and alert ensemble," *Bull. Electr. Eng. Informatics*, vol. 13, no. 5, pp. 3534–3543, Oct. 2024, doi: 10.11591/eei.v13i5.8084.
[86] D. K. Packham, H. S. Rasmussen, B. Singh, D. J. Mccann, and P. Skolnick, "New Agents for Hyperkalemia," *N. Engl. J. Med.*, vol. 372, no. 16, pp. 1569–1572, 2015, doi: 10.1056/nejmc1501933.
[87] K. W. Brown and T. J. Armstrong, *Hydrocarbon Inhalation*. 2023. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/24911841
[88] F. O. Aghware, R. E. Yoro, P. O. Ejeh, C. C. Odiakaose, F. U. Emordi, and A. A. Ojugo, "DeLClustE: Protecting Users from Credit-Card Fraud Transaction via the Deep-Learning Cluster Ensemble," *Int. J. Adv. Comput. Sci. Appl.*, vol. 14, no. 6, pp. 94–100, 2023, doi: 10.14569/IJACSA.2023.0140610.
[89] N. R. Pratama, D. R. I. M. Setiadi, I. Harkespan, and A. A. Ojugo, "Feature Fusion with Albumentation for Enhancing Monkeypox Detection Using Deep Learning Models," *J. Comput. Theor. Appl.*, vol. 2, no. 3, pp. 427–440, 2025, doi: 10.62411/jcta.12255.
[90] E. Yakub, S. E. Agarry, F. Omoruwou, and C. N. Owabor, "Comparative study of the batch adsorption kinetics and mass transfer in phenol-sand and phenol-clay adsorption systems," *Part. Sci. Technol.*, vol. 38, no. 7, pp. 801–811, Oct. 2020, doi: 10.1080/02726351.2019.1616862.
[91] A. P. Binitie *et al.*, "Stacked Learning Anomaly Detection Scheme with Data Augmentation for Spatiotemporal Traffic Flow," *J. Fuzzy Syst. Control*, vol. 2, no. 3, pp. 203–214, 2024, doi: 10.59247/jfsc.v2i3.267.
[92] S.-X. Wang *et al.*, "Inherently flame-retardant rigid polyurethane foams with excellent thermal insulation and mechanical properties," *Polymer (Guildf).*, vol. 153, pp. 616–625, Sep. 2018, doi: 10.1016/j.polymer.2018.08.068.
[93] F. Omoruwou, A. A. Ojugo, and S. E. Ilodigwe, "Strategic Feature Selection for Enhanced Scorch Prediction in Flexible Polyurethane Form Manufacturing," *J. Comput. Theor. Appl.*, vol. 1, no. 3, pp. 346–357, Feb. 2024, doi: 10.62411/jcta.9539.