IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 21 July 2023, accepted 7 August 2023, date of publication 11 August 2023, date of current version 16 August 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3304328

RESEARCH ARTICLE

Harnessing the Multimodal Data Integration and Deep Learning for Occupational Injury Severity Prediction

MOHAMED ZUL FADHLI KHAIRUDDIN[1,2], KHAIRUNNISA HASIKIN[1,3], NASRUL ANUAR ABD RAZAK[1], SITI AFIFAH MOHSHIM[4], AND SITI SALWA IBRAHIM[5]

[1]Department of Biomedical Engineering, Faculty of Engineering, University Malaya, Kuala Lumpur 50603, Malaysia
[2]Environmental Health Section, Institute of Medical Science Technology, Universiti Kuala Lumpur, Kuala Lumpur, Selangor 43000, Malaysia
[3]Centre of Intelligent Systems for Emerging Technology, Faculty of Engineering, University Malaya, Kuala Lumpur 50603, Malaysia
[4]Medical Engineering Technology Section, British Malaysian Institute, Universiti Kuala Lumpur, Kuala Lumpur, Selangor 53100, Malaysia
[5]Negeri Sembilan State Health Department, Ministry of Health, Seremban, Negeri Sembilan 70300, Malaysia

Corresponding author: Khairunnisa Hasikin (khairunnisa@um.edu.my)

ABSTRACT Most previous studies have neglected the potential of integrating structured data and unstructured workplace injury reports to perform a predictive analysis of occupational injury severity. This study proposes an optimized integrated approach for occupational injury severity prediction using multimodal machine and deep learning techniques. We used 66,405 data points gathered from the US OSHA Severe Injury Reports from January 2015 to July 2021. Structured labeled data are preprocessed and normalized, whereas unstructured injury reports undergo text cleaning using Natural Language Processing techniques and text representation using Term Frequency-Inverse Document Frequency (TF-IDF) and Global Vector (GloVe) to convert them into numerical representations. Both modalities, in the form of vector representations, were concatenated and fed as input features for the proposed models. Seven sets of classifiers, namely Naïve Bayes, Support Vector Machine, Random Forest, Decision Tree, K-Nearest Neighbors, Long Short-Term Memory, and Bidirectional Long Short-Term Memory, were employed to learn the multimodal representations. The algorithm with superior performance was further optimized using the proposed feature importance and hyperparameter optimization techniques. Our findings revealed that the proposed optimized-Bi-LSTM architecture outperformed other classifiers in learning multimodal data to predict the likelihood of hospitalization and amputation with higher accuracies of 0.93 and 0.99, respectively. Consequently, the proposed approach enhances the performance by significantly improving the model processing time. This performance prediction provides a convincing benchmark for the successful execution of multimodal deep learning in occupational injury research. Therefore, the proposed multimodal occupational injury severity prediction model enhances the early screening and identification of at-risk workers with severe occupational injury outcomes, as well as, provides valuable information to improve the workplace safety, health, and well-being of the workers.

INDEX TERMS Artificial intelligence, machine learning, multimodal integration, natural language processing, occupational safety.

I. INTRODUCTION

Occupational injury is defined as ''any physical injury that affects a worker while working''. Similar terms refer to workplace injury, work-related accidents, and occupational accidents [1], [2]. The International Labour Organization (ILO) reported that occupational injury may potentially be declared a 'public health emergency'; as it has killed more than thousands of workers each year [3]. It has substantially contributed to fatality events, reduced work productivity, and

The associate editor coordinating the review of this manuscript and approving it for publication was Gustavo Olague.

worsened the economy at large [4]. In addition, the economic cost of workplace injuries can be measured through the amount of medical and rehabilitation compensation, including the loss-time costs, social security benefits, as well as, the training and re-training of the worker [5]. This may cost up to 4-5% of the global 'Gross National Product' (GNP) [6]. Therefore, workplace injuries may affect workers physically, but they can also manifest detrimental psychological effects, including depression, anxiety, and post-traumatic injury [7]. These psychological effects of workplace injuries result in prolonged injury recovery, thus; increasing compensation expenses. Other research has revealed that workers who experienced occupational injury developed difficulties in physical and physiological health which interrupted their workplace relationships [8]. To some extent, they have reported psychiatric symptoms and incidences of suicide attempts [9]. Evidently, work-related accidents can cause a 'domino effect', contributing not only to the immediate physical health and financial burden but also contribute to the long-term psychological impacts. Therefore, the avoidance of workplace accidents and injuries should be a top priority for occupational health and safety throughout all sectors of the economy.

Occupational injury reports are records of injured worker information consisting of structured information (workers' demographics, type of injury, accident cause, etc.) and unstructured data such as textual injury history or reports. These records are valuable and may provide remarkable opportunities, especially among Artificial Intelligence researchers to extract and analyze records in a more reliable and efficient manner. With recent advances, machine learning, including deep neural network techniques has gained interest as the method of choice for predicting occupational injury outcomes [10].

Some related studies executing those techniques with structured occupational injury information are as follows: (i) Yedla et al. [11] used categorical input from the mining industry to predict the occupational accident outcome of days away from work with Neural Network outperformed other models in their study; (ii) Chadwiya in their study had utilized South African workplace accident-labelled data and revealed the Support Vector Machine (SVM) as the best prediction model in predicting occupational injury based on affected body parts [12], and (iii) Khairuddin et al. [13] investigated the prediction performance of occupational injury severity using categorical variables through a Random Forest optimized model across all industrial sectors. In addition, it is believed that unstructured text data are a valuable source of information; thus, the extraction of insights can be achieved using text-mining techniques [14]. Often, unstructured data contains rich semantic information that can yield insightful insights. In text data, for instance, the choice of words, sentence structure, and tone can convey significant meaning. By incorporating this information, the model is better to comprehend and capture the data's underlying patterns. Unstructured data can provide crucial additional context for comprehending the relationship within a dataset. Several studies have focused on analyzing unstructured injury reports as the input features. For example: (i) Jing [15] proposed Word2Vec and Long Short-Term Memory (LSTM), a recurrent neural network (RNN) variant as a text-mining predictive tool for workplace accidents in the chemical industry, (ii) Baker et al. [16] developed an improved text-mining with model stacking of the XGBoost-Random Forest algorithm to predict the occupational injury outcomes, and (iii) Goldberg [17] analyzed injury narratives to compare the techniques of word embedding, such as Word2Vec and TFIDF, including several machine learning algorithms in predicting the severity of occupational injury in the United States. As most of the preceding works focused on employing structured data or unstructured text separately, it has been verified that the development of the occupational injury severity prediction model by combining both modalities; structured and unstructured data are neglected and restricted in the occupational injury research domain [18].

Consequently, the purpose of this research is to propose an integrated predictive model based on multimodal learning of structured data and unstructured information using machine and deep neural network approaches in predicting occupational injury severity.

To summarize, the main contributions of our study are as follows:

i. The potential of integrating structured data, such as labelled data points and unstructured data, for example workplace injury reports has been neglected by the majority of previous studies in occupational injury severity prediction [18]. This work acknowledges the significance of both modalities and proposes a novel strategy that exploits the power of multimodal data integration. Integrating the unstructured data will enhance the feature representation of the overall dataset. The unstructured data in text narratives contain valuable information that may not be captured in the structured data alone.

ii. Unstructured occupational injury reports are subjected to several preprocessing stages, including text cleaning using Natural Language Processing (NLP) and tokenization, followed by text representation techniques. Our study proposed an innovative approach for integrating Term Frequency-Inverse Document Frequency (TF-IDF) and the Global Vector (GloVe) as text representations. These stages allow unstructured textual data to be converted into numerical representations, making them appropriate for machine and deep learning models.

iii. The vectors representing structured and preprocessed unstructured text data are concatenated and utilized as input features for the proposed predictive models. Because of this integration, the models may learn from both modalities simultaneously, collecting the complementary information available in structured and unstructured data.

iv. The multimodal occupational injury severity prediction model presented herein has practical implications for workplace safety and health. By enhancing the early screening and identification of at-risk employees with severe occupational injury outcomes, the predictive model can contribute to the improvement of workplace safety measures and overall working environment. The model's information can guide interventions and initiatives designed to promote workers' well-being and prevent the occurrence of severe workplace injuries. These insights will assist in identifying potential hazards, implementing proactive measures, and enhancing overall workplace safety.

Therefore, this paper enhances the field of occupational injury research as the findings from the multimodal machine and deep learning will presents the benchmark model performance for occupational injury severity prediction tasks.

This paper is organized into eight sections, including the Introduction. In Section II, a summary of previous related studies is presented. The proposed methodology is explained in Section III and a step-by-step model experiment is summarized in Section IV. The model prediction findings are presented in Section V, followed by the results of model optimization in Section VI. Section VII discusses the overall findings, and Section VIII concludes the paper.

II. RELATED WORKS

Multimodal learning is defined as 'the area of applying machine and deep learning techniques in integrating multiple types of data into a single model to optimum the uniqueness and valuable information in an algorithmic framework' [24]. The ultimate aim of this multimodal learning is to harmonize the diversity of data in improving the data quality, thereby enhancing the performance predictions [25]. The application of multimodal learning has been explored in several fields, especially the use of different modalities extracted from 'Electronic Health Record' (EHR), as it contains categorical/numerical variables, clinical notes, and clinical images. For example, Ross et al. [26] integrated structured data and clinical text to be trained into Logistic Regression (LR) and Random Forest (RF) models in predicting cardiovascular diseases. Then, Lei et al. combined structured data, unstructured text, including audio and clinical images into deep neural network architectures to categorize the events of communicable disease, whereas Zhang et al. [27] executed LR, RF and advanced deep learning methods; Long Short-Term Memory (LSTM) and Convolutional Neural Network (CNN) in structured, unstructured clinical notes, and the integration of both data to predict hospitalization stay and mortality, hence revealing the integration of both data achieved the best performance predictions. Therefore, multimodal learning with machine and deep learning techniques have emerged progressively in those fields including clinical diagnosis prediction [28], [29], pathological screening [30], and business intelligence purposes [31], as it has proven in improving the model predictability.

However, the exploration of multimodal data for occupational injury severity remains limited. Johansson et al. [32] proposed a study protocol to integrate structured occupational injury registry and photovoice inputs in predicting the event of medical leave by the injured workers. They believed that multimodal data could improve model predictability and assist in discovering insights into workplace injuries among Swedish adults. Next, a study by Paraskevopoulos et al. [33] used multimodal dataset of safety reports and workplace images, prepared by Safety Officers and executed an NLP-based neural network as the predictive model to predict the outcomes of workplace safety audits. Their study agreed that multimodal data can discover hidden information, thus providing a better accuracy performance. Recently, Sarkar et al. developed a multimodal deep neural network model by integrating occupational injury narratives and categorical variables [18]. Their study compared the performance prediction of deep neural network with several optimizer and found the model with 'adaptive moment estimation' (adam) optimizer became the best-performing prediction of occupational injury in steel manufacturing industry. Nevertheless, they emphasized that predictive analysis using multimodal data in occupational injury areas is undiscovered, limited, and requires further exploration.

Table 1 summarizes recent literature on occupational injury severity prediction models. On the basis of this overview, one can conclude that the current trend of occupational injury severity predictive analysis is to make use of structured and unstructured text data, and it concentrates exclusively on one industry sector. This causes a lack of generalizability of the existing predictive models, where data from different industries and sectors are neglected. These variations in work environments, hazards, and job tasks are disregarded, and predictive models do not have the capability to learn from diverse scenarios across various occupational settings. In addition, there is a paucity of multimodal data integration to develop an occupational injury prediction model. However, the exploration of advanced neural network architectures is gaining attention in the domain of occupational injury.

III. PROPOSED METHODOLOGY

Our study proposes a multimodal occupational injury severity prediction model that encompasses three main processes; first, the gathering of the data, followed by data pre-processing, and finally, the prediction classifier stage. These two modalities; structured and unstructured data, were preprocessed separately. Subsequently, the feature representations generated by the structured information and text representation were merged. These vectors were concatenated and fed as the input of the proposed classifiers to predict the severity of the occupational injury.

TABLE 1. Related studies in occupational injury prediction.

Author	Year	Industry	Techniques	Data Modality	Limitation/Research Gap
Oyedele, et al. [19]	2021	Utility	Deep neural network and Boosted trees.	Structured	Limited focus only on the construction industry.
Ganguli, et al. [20]	2021	Mining	NLP-based Random Forest	Unstructured	Dependency on terminology and writing styles.
Chadyiwa, et al. [12]	2022	Tourism	Decision Trees, XG Boost, Support Vector Machines, K-nearest neighbours, and deep neural networks.	Structured	Limited features on nature of work and textual injury reports.
Kim, et al. [21]	2022	Construction	Deep neural network	Structured	Limited in model comparison; to further explore other deep neural models such as RNNs.
Jing, et al. [15]	2022	Chemical	Word2Vec with Bi-LSTM-Attention model	Unstructured	Limited accessibility on occupational injuries data.
Goldberg [17]	2022	Multiple	Word-embedding-based machine learning models: Logistic Regression, Decision Trees, Random Forests, Naïve Bayes, Support Vector Machines, Multi-layer perceptron, and Bi-directional Long Short-Term Memory.	Unstructured	Limited on cross-industrial data and high reliability on human-labelled data.
Zhang [22]	2022	Construction	Word2Vec with hybrid structured deep neural network (CNN-LSTM).	Unstructured	Size of corpus is small, computational expensive, and limited exploration on hyperparameter such as use of early stopping function.
Sarkar, et al. [18]	2022	Steel	Latent Dirichlet allocation (LDA)-based topic modeling and Deep Neural Network with Adam optimizer	Multimodal	Limited number of workplace injuries records and required extensive human-assistance in data cleaning.
Zermane, et al. [23]	2023	Construction	Extreme Gradient Boosting, Decision Trees, Random Forest, Multi-Layer Perception, K-Nearest Neighbour, Support Vector Machine, and Logistic Regression.	Structured	Limited to fatal accident due to fall from height at construction industry only.

A. DATASET DESCRIPTION

A publicly accessed dataset was gathered from an established Occupational Safety and Health Administration (OSHA) located in the United States [34]. The dataset can be accessed through (https://www.osha.gov/severeinjury). In this study, the dataset comprises injured worker information with over sixty thousand data between January 2015 and July 2021. This dataset named as 'Occupational Injury Severity Report' includes variable columns such as (i) ID number, (ii) event date, (iii) employer's address, (iv) the state with latitude and longitude, (v) nature of the injury, (vi) affected body parts, (vii) type of exposure, (viii) type of source, as well as, (ix) the injury narratives. Additionally, it contains information on (i) amputation and (ii) hospitalization, as the indicator of occupational injury severity. Table 2 presents the types of variables used in the dataset.

B. STRUCTURED DATA PREPROCESSING

In this research, five categorical variables were considered as the input features: the (i) type of industry, (ii) nature of the injury, (iii) affected body part(s), (iv) type of event, and (v) type of source. These data are pre-coded according to the top label only, as guided by the Occupational Injury and Illness Classification Manual (OIICS). In addition, any rows with non-value or empty columns were removed, whereas other columns, for example, ID number, employer's address, latitude, and longitude were excluded because of their irrelevancy. After data preparation, 66,405 data points were used for predictive analysis.

For the predictive analysis, a data preprocessing step was employed to ensure that the input features had consistent contributions during the machine and deep learning development process [35]. Categorical data were manually encoded by referring to the OIICS manual system, which represents categories as numerical labels.

The next steps involved developing machine and deep learning models using these encoded categorical data along with unstructured feature representations to predict the outcomes of occupational injury severity. This approach leverages information encoded in categorical variables and incorporates it into predictive classifiers for multimodal learning research.

C. UNSTRUCTURED DATA PREPROCESSING

Textual narratives of occupational injuries were included as sequential unstructured input features. This text report was prepared by Safety and Health personnel, assisted by Occupational Health Doctor and we believe it encompasses a large amount of insight that can be extracted for predictive analysis in this study. The conversion of text data into numerical values is essential before they can be processed by machine and deep learning algorithms [36].

In this study, text preprocessing was performed for Natural Language Processing (NLP) tasks through the following steps: (i) removal of punctuation and digit numbers as they did not contribute to the analysis, (ii) removal of extra whitespaces such as tabs and line breaks, (iii) removal of characters that may potentially interfere during the text vectorization step [37], (iv) removal of stop words such as "a" and "the" as they consider as the 'unnecessary words' which did not contribute the classification taks and may create higher dimension of vector [38], and (v) lower case capitalization of the text. Next, the text underwent the tokenization step,

TABLE 2. Types of variables.

Variables	Values
Workers' ID number	Numerical (vary)
Accident Date	
Nature of Injury	10: Traumatic injuries, 11: Bones, Nerves, Spinal Cord, 12: Muscles, Tendons, Ligaments, 13: Open wounds, 14: Surface wounds and bruises, 15: Burns and corrosions, 16: Intracranial injuries, 17: Environmental conditions, 18: Multiple traumatic injuries, 19: Other(s)
Affected Body Parts	1: Head, 2: Neck and throat, 3: Trunk, 4: Upper extremities, 5: Lower extremities, 6: Body systems, 8: Multiple body parts, 9: Other(s)
Type of Event	1: Violence/other injuries (persons/animals), 2: Transportation, 3: Fires/explosions, 4: Falls, slips, trips, 5: Harmful substances/environments, 6: Contact with objects/equipment, 7: Overexertion and bodily reaction, 9: Other(s)
Type of Source	1: Chemicals/chemical products, 2: Containers, furniture, and fixtures, 3: Machinery, 4: Parts and materials, 5: Persons/plants/animals, and minerals, 6: Structures and surfaces, 7: Tools/instruments/ equipment, 8: Vehicles, 9: Other(s)
Type of Industry	11: Agriculture, Forestry, Fishing and Hunting, 21: Mining, 22: Utilities, 23: Construction, 31: Manufacturing, 42: Wholesale trade, 44: Retail trade, 48: Transportation and Warehousing, 51: Information, 52: Finance and Insurance, 53:Real estate, 54:
	Professional, Scientific, and Technical Services, 55: Management of Companies and Enterprises, 56: Waste Management and Remediation, 61: Educational Services, 62: Health Care and Social Assistance, 71: Arts, Entertainment and Recreation, 72: Accommodation and Food Services, 81: Other Services, 92: Public Administration
Employer's Address	Unstructured (Text/Number)
Injury Narratives	

where a string of words is segmented into its component words named 'tokens'. Each tokenized word is numbered to identify a particular word. This is a crucial step in converting words into numerical features [39].

Next, text representations were generated by converting the tokens into numbers to be processed by the classifiers. The text representation methods employed in this study were Term Frequency-Inverse Document Frequency (TF-IDF) and the Global Vector (GloVe). TF-IDF was considered as it commonly appears as a high-performing text vectorization technique [40], [41]. It consists of two elements; the 'term frequency' (TF) and the 'inverse document frequency' (IDF). TF depends on the number of occurrences of the word in each injury narrative, meanwhile, IDF is computed based on how much the word contains throughout the entire injury narrative dataset. TF is measured as $tf_{i,w}$, in which i represent a given keyword in each given document, w. The data in this experiment consists of D documents and df_i represents the number of occurrences of a word across all the documents. The IDF of a keyword, idf_i, is calculated by taking the logarithmic inverse, as shown in the following equation (1). Next, the final score of TF-IDF is derived from the equation (2).

$$idf_i = \log\left(\frac{1+D}{1+df_i}\right) + 1 \tag{1}$$

$$tfidf_{i,w} = tf_{i,w} \times idf_i \tag{2}$$

The vectorized word was then followed by a word-embedding of GloVe. It is a word-embedding method in which words are represented as vectors in a high-dimensional space that uses word2vec-word representation to learn word embeddings from textual materials efficiently [42]. To generate the word vector representation in this study, a pre-trained GloVe model from Stanford NLP labeled "Glove.6B" was used. This pre-trained GloVe is a 100-dimensional vector that has been trained on six billion tokens from Wikipedia articles and the Gigaword dataset. It is freely accessible to the public under a Public Domain Dedication and License [43]. Figure 1 is the schematic diagram of the unstructured text preprocessing used in this study.

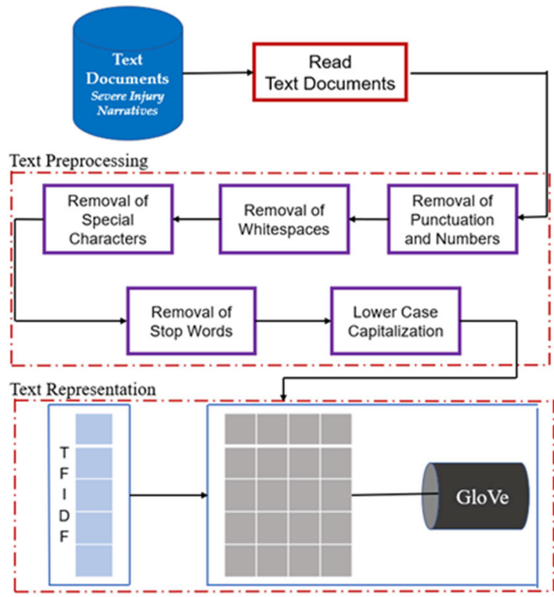

FIGURE 1. Schematic diagram of unstructured text processing.

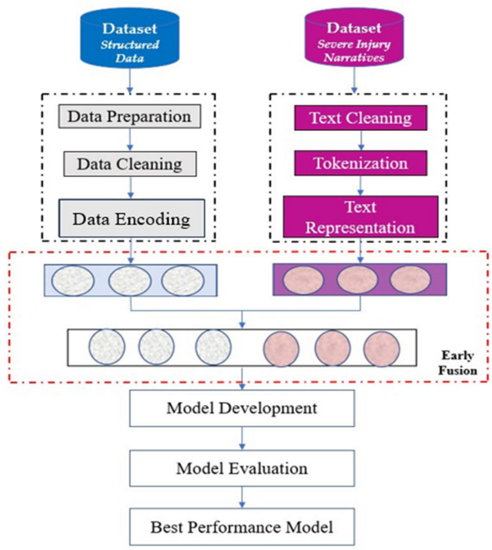

FIGURE 2. Flowchart of multimodal data learning.

D. MULTIMODAL DATA FUSION

Structured and unstructured data were combined as input representations to predict occupational injury severity, in terms of (i) hospitalization and (ii) amputation for multimodal learning. The structured data were prepared and normalized as explained in subsection B, whereas the unstructured text data underwent text preprocessing, tokenization, and text representation, as described in subsection C. Subsequently, both preprocessed representations were concatenated into a single input representation vector using an early fusion strategy. The early fusion strategy integrates both data modalities after their preprocessing steps and fed them as input representations to the sets of classifiers. The application of an early fusion strategy is preferred in multimodal learning owing to its practicality and simplicity [44]. Moreover, this strategy tends to generate better performance predictions than the unimodal versions [45], [46]. Figure 2 illustrates the flowchart of multimodal learning in this study.

E. PREDICTION MODELLING

Both prediction outcomes are composed of a binary classification problem, where the label indicates Yes (1) or No (0) for the occurrence of hospitalization, as well as, the likelihood of an amputation event. Prior to the model development, the data were partitioned into two sets using stratified sampling; 80% of the data were used as the training set, and the remaining 20% of the data were applied as the testing set.

Five sets of machine learning algorithms and two deep neural architectures were proposed to analyze the multimodal data in predicting the severity of occupational injuries. The five sets of machine learning predictive models were (i) Naïve Bayes (NB), (ii) K-Nearest Neighbors (KNN), (iii) Decision Tree (DT), (iv) Random Forest (RF), and (v) Support Vector Machine (SVM). These ML models were selected because of their consistency in occupational injury prediction studies [11]; therefore, a comprehensive comparative analysis is required to assess model effectiveness [47]. Because several previous studies have recommended the exploration of RNN variants in the multimodal occupational injury domain [48], [49], this study executes two types of RNN variants as the proposed deep learning architectures: (i) Long Short-Term Memory (LSTM) and (ii) Bidirectional Long Short-Term Memory (Bi-LSTM).

NB is a supervised learning algorithm that is based on Bayes' theorem and is preferred due to its simplicity and ability to predict performance accurately. It is named ''naive'' because it assumes that given the class label, the features in the data are conditionally independent of each other. Despite the ''naïve'' assumption between input variables, the NB algorithm performs well in a variety of classification problems [50].

KNN is an algorithm that implies on similar information exists nearby or in close proximity to one another. The model calculates the distance between data points and then categorizes them based on their proximity. Normally, it is based on a distance metric, such as Euclidean distance among the training samples, and then making a judgement based on the majority vote or average of its k neighbors' labels [51].

Next, DT algorithm uses the training data to generate a tree-like decision structure, with the starting point being a 'root node' and the ending point being certain leaves. The classification strategy in DT begins with the division of the root node into the leaf node. The splitting process makes use of the input variables. The splitting will progress until it reaches the leaf node. Then, the leaf nodes, also known as end nodes in certain literature, indicate the final outputs, which is the classification problem [52].

SVM uses a hyperplane to maximize the margin between two classes in binary classification tasks. The training data points that are nearer to the boundary will impact the creation of the hyperplane, and they are referred to as ''support vectors''. An interesting feature of SVM is it can accommodate both linearly and non-linearly separable datasets by mapping the data into a higher dimensional space where it can be separated by a hyperplane using a kernel function. A kernel function is introduced to aid in the separation of different classes. The linear, sigmoid, and radial basic function (RBF) kernels are the most commonly employed kernels, and this design of SVM produces the best generalization of decision boundaries for data categorization [53]. This study employed the radial basic kernel function.

RF combines numerous decision trees to provide a more accurate and stable prediction. Each tree in the RF model produces a classification and accounts it as a 'vote'. The final classification is then established based on the majority of these votes, with the category with the most votes chosen as the final prediction. Because the RF model adheres to the 'majority votes decision rule,' the aggregation of these outcomes will provide a good generalization, resulting in improved accuracy. Additionally, each tree in the forest is trained using a randomly chosen portion of the training data, known as the 'bootstrap sample', and a randomly chosen subset of the features, known as the feature subset. This reduces variance and the risk of overfitting [54].

LSTM is an improved technique to solve a well-known drawback in training Recurrent Neural Network, which is the vanishing gradient issue. The LSTM method overcomes the problem by adding a gate mechanism and a memory unit. Three gates of LSTM are: (i) the input gate, controls which information is stored in memory cells, (ii) the output gate determines which information is used in prediction, and (iii) the forget gate controls which information is ignored. The configuration of these gates in LSTM enables information control, which is the primary rationale for reducing the vanishing gradient problem in standard RNNs [55].

Bi-LSTM is an advanced architecture of LSTM, which composed of forward and backward LSTMs. The key idea behind this bidirectional structure is the capacity to collect information patterns that may be overlooked by unidirectional LSTM [56]. Because the Bi-LSTM network is constructed of two LSTMs, their outputs are concatenated and utilized as inputs to the final prediction output layer. This enables the network to generate predictions while considering the complete sequence. The strength of the 'forward-backward' in Bi-LSTM leads to improve learning long-term sequences, as a result, improves the model's performance prediction.

In this study, both deep learning models; LSTM and Bi-LSTM, were structured as follows: the hidden unit was chosen as 256, trained with a batch size of 64, and the maximum epoch number was set to 25. In addition, the models were implemented with an 'Adam' optimizer, 'ReLU' activation with a dropout rate of 0.2. ReLU activation was proposed based on recent studies that demonstrated the modified approach of ReLU with an LSTMs network has empirically improved model performance in terms of comparison with other activation functions [57] and existing deep learning tools [58], [59], [60]. In addition, an early stopping function was introduced to prevent overfitting and improve model generalization, including assisting in determining the optimal stopping point for training [61]. The metric criterion used was validation loss with a patience of 5, in which the models were trained for a maximum of 25 epochs but stopped the training earlier if the validation loss did not improve for consecutive 5 epochs. As this study is a binary classification task, a dense output layer functions as a sigmoid and loss function used was binary cross-entropy. The customized parameters of each classifier are summarized in Table 3.

TABLE 3. Classifier's parameter.

Model	Parameters	
NB	Preset:	GaussianNB
KNN	Metric:	Gower
	Number of neighbors:	20
	Power parameter:	1
DT	Split criterion:	Entropy
	Estimator randomness:	None
RF	Split criterion:	Entropy
	Number of learners:	50
	Estimator randomness:	None
SVM	Kernel function:	Radial basis (rbf)
	Randomness:	None
LSTM	Hidden unit:	256
	Batch size:	64
	Epoch:	25
	Dense unit:	10
	Dropout:	0.2
	Activation:	ReLU
	Optimizer:	Adam
	Output activation:	Sigmoid
	Early stopping	
	Metric:	Validation loss
	Patience:	5
Bi-LSTM	Hidden unit:	256
	Batch size:	64
	Epoch:	25
	Dense unit:	10
	Dropout:	0.2
	Activation:	ReLU
	Optimizer:	Adam
	Output activation:	Sigmoid
	Early stopping	
	Metric:	Validation loss
	Patience:	5

All of the multimodal classifiers were imported and developed in the Python environment with the execution of suitable libraries and packages: NumPy (np), pandas (pd), matploblit (plt), sklearn, natural language toolkit (nltk), Keras and TensorFlow. The development of the prediction models was performed on a laptop with the following specifications: (i) CPU: AMD Ryzen 7 3700U @ 2.30GHz with 12GB RAM and (ii) GPU: Radeon^{TM} RX Vega 10 Graphics 1400 MHz.

F. MODEL EVALUATION METRICS

A confusion matrix is a foundation for computing the predictability performance of classifiers. It is in the form of a ''contingency table'' that visualizes how the findings are disseminated over the actual class (represented in rows) and predicted class (represented in columns) [62]. The matrix comprises four instances: ''True Positive'' (TP) and ''False Positive'' (FP) are observations of correct and incorrect predictions per actual classes, accordingly, while, ''True Negative'' (TN) and ''False Negative'' (FN) are instances of right and wrong rejections per actual classes, respectively. Based on this matrix, the following evaluation metrics for classification tasks were used to assess the prediction performance of all models proposed in this study: (i) precision, (ii) recall, (iii) F1-score, (iv) accuracy, and (v) AUC values. Figure 3 summarized the model evaluation metrics employed in this study.

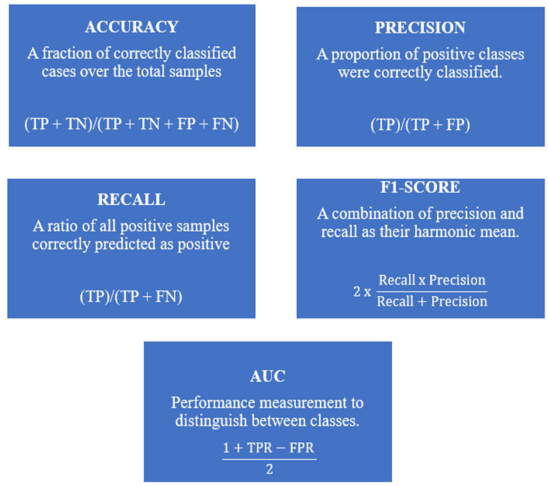

FIGURE 3. Model evaluation metrics.

IV. MODEL EXPERIMENTATION

This subsection summarizes and simplifies the overall implementation of the occupational injury severity prediction model used in this study. The step-by-step implementation is explained as follows:

i. The experiment began with the preprocessing of structured data, involving data cleaning and preparation. Five categorical variables were selected from the occupational injury dataset: (i) type of industry, (ii) nature of injury, (iii) affected body part(s), (iv) type of event, and (v) type of source. These categorical data are encoded using a reference system.
ii. Next, unstructured injury narratives were extracted through NLP, followed by text representation using the integration of TFIDF and pretrained GloVe word embedding to convert the textual data into numerical representation for further analysis.
iii. The experiment was then resumed with the integration of both preprocessed data modalities–structured and unstructured–as multimodal representations.
iv. The data were split into two sets using stratified sampling at an 80:20 ratio; 80% of the data were used as the training set, whereas the remaining 20% were used as the testing set.
v. Several candidate models were explored, consisting of five sets of ML models (NB, KNN, DT, RF, and SVM)

and two deep learning models (LSTM and Bi-LSTM), for predicting the severity outcomes of hospitalization and amputation.

vi. Comparative analyses were performed by utilizing established model evaluation metrics, including accuracy, precision, recall, F1-score, AUC, and prediction time. These metrics are used to assess the performance of each candidate model.

vii. Based on the evaluation, the model exhibiting superior performance in terms of accuracy, F1-score, AUC, precision, and recall was selected as the best-performing prediction model.

viii. Then, the experiment was resumed with model optimization step. Firstly, both data modalities were assessed using the RF feature importance algorithm to determine the most important features and predictive keywords.

ix. The selected best-performing model in (vii) was redeveloped using only the important features and keywords, followed by hyperparameter optimization using random search cross-validation.

x. Finally, the optimized model was compared with the initially developed model through model evaluation metrics to determine the best occupational injury severity prediction model for final model deployment. Additionally, its compatibility with computational efficiency in terms of processing time (training and testing) was a significant factor in its selection.

Overall, this careful and systematic model selection process ensures that the proposed approach represents the execution of rigorous experimentation and thorough evaluation, thereby generating an accurate interpretable practical occupational injury severity prediction model. The pseudocode of the proposed model experimentation is presented in Figure 4.

V. RESULTS

The prediction outcomes investigated in this study was the likelihood of severity, in terms of hospitalization and amputation. Each classifier was evaluated on all performance metrics and comprehensively compared to determine the best-performing model.

A. HOSPITALIZATION PREDICTION

Table 4 presents the performance prediction of all proposed models on the hospitalization prediction task. From this table, it shows that the Bi-LSTM outperformed other models with a slightly higher accuracy of 0.93 as compared to RF and LSTM, both achieved 0.92, respectively. Also, Bi-LSTM achieved the best F1-score at 0.95, meanwhile, the AUC value is slightly better than the SVM model at 0.93.

B. AMPUTATION PREDICTION

Table 5 summarizes the findings of each prediction algorithm for the amputation prediction task. From the table, it can be seen that Bi-LSTM is the best-performing model as it

TABLE 4. Performance prediction of hospitalization.

Classifiers	Accuracy	Precision	Recall	F1-score	AUC
Machine learning					
NB	0.88	0.97	0.89	0.92	0.90
KNN	0.91	0.97	0.93	0.94	0.90
SVM	0.91	0.98	0.90	0.94	0.92
DT	0.90	0.96	0.92	0.94	0.89
RF	0.92	0.97	0.94	0.94	0.91
Deep learning					
LSTM	0.92	0.98	0.90	0.93	0.91
Bi-LSTM	**0.93**	**0.98**	**0.92**	**0.95**	**0.93**

TABLE 5. Performance prediction of amputation.

Classifiers	Accuracy	Precision	Recall	F1-score	AUC
Machine learning					
NB	0.93	0.87	0.93	0.89	0.93
KNN	0.95	0.95	0.94	0.94	0.96
SVM	0.96	0.90	0.97	0.93	0.97
DT	0.95	0.92	0.95	0.94	0.95
RF	0.97	0.93	0.96	0.96	0.97
Deep learning					
LSTM	0.95	0.85	0.97	0.92	0.95
Bi-LSTM	**0.99**	**0.98**	**0.98**	**0.97**	**0.98**

achieved higher accuracy (0.99), F1-score (0.97), and AUC of 0.98, compared to SVM, KNN, LSTM, DT, and NB. Meanwhile, the RF model ranked second, with accuracy and an AUC value of 0.97.

Based on both tables, the Bi-LSTM models have been discovered as the best-performing prediction model in predicting occupational injury severity, as this model performed significantly in each model evaluation metric, specifically in accuracy, recall, and F1-score, compared to other models.

C. PREDICTION TIME

Additionally, the prediction or testing time for each model was investigated in this study. Although the Bi-LSTM model achieved higher accuracy, F1-score, and AUC values for both prediction tasks, the prediction time for Bi-LSTM may be longer than those for RF, SVM, and LSTM. The Bi-LSTM model required up to 67s to predict the hospitalization and at 66s to predict the amputation outcomes. However, it is presumed that the testing time of the Bi-LSTM model is still acceptable. A comparison of each model's prediction time is depicted in the line graph in Figure 5 for both the predictions.

```
Development of Multimodal Occupational Injury Severity Prediction Model
Input: Structured Data, S and Unstructured Injury Text, U
Output: Occupational Injury Severity, H or A
Begin
Preprocessing structured data
preprocess_structured_data(data)
    cleaned_data = data_cleaning(data)
    encoded_data = encode_data(cleaned_data)
    return encoded_data
Extract unstructured injury text
extract_injury_narratives(data)
    injury_narratives = extract_nlp(data)
    preprocessed_text = preprocess_text(injury_text)
    tfidf_representation = apply_tfidf(preprocessed_text)
    text_representation = apply_glove_embedding(tfidf_representation)
    return text_representation
Integration of preprocessed data modalities
integrate_data(normalized_data, text_representation)
    multimodal_data = early_fusion(encoded_data, text_representation)
    return multimodal_data
Split data into training and testing sets
split_data(data, train_ratio)
    train_data, test_data = stratified_sampling(data, train_ratio=0.8)
    return train_data, test_data
Explore candidate models
explore_candidate_models(train_data)
    candidate_models = [NB, KNN, SVM, DT, RF, LSTM, Bi-LSTM]
    trained_models = train_models(candidate_models, train_data)
    return trained_models
Perform comparative analyses
perform_comparative_analyses(trained_models, test_data)
    evaluation_metrics = evaluate_models(trained_models, test_data)
    evaluation_metrics[model] = {
        'accuracy': accuracy,
        'precision': precision,
        'recall': recall,
         'f1_score': f1_score,
        'auc': auc
        }
    return evaluation_metrics
Select the best-performing model
select_best_model(evaluation_metrics)
    best_model = choose_best_model(evaluation_metrics)
    return best_model
Feature importance analysis
feature_optimization(structured_data, textual_data)
    important_features = rf_feature_importance(structured_data)
    important_keywords = rf_feature_importance(textual_data)
    return important_features, important_keywords
Optimization of the best-performing model
redevelop_best_model(best_model, important_features, important_keywords)
    multimodal_model = early_fusion(best_model, important_features, important_keywords)
    optimized_model = hyperparameter_optimization(multimodal_model, random_search, k=10)
    return optimized_model
Compare the optimized model with the initial model
compare_models(initial_model, optimized_model)
    initial_model_metrics = evaluate_model(initial_model, test_data, prediction_time=True)
    optimized_model_metrics = evaluate_model(optimized_model, test_data, prediction_time=True)
    return initial_model_metrics, optimized_model_metrics
Deploy the selected best model for final use
if optimized_model_metrics is superior to initial_model_metrics:
    final_model = optimized_model
else:
    final_model = initial_model
deploy_model(final_model)
End
```

FIGURE 4. Pseudocode of prediction model development.

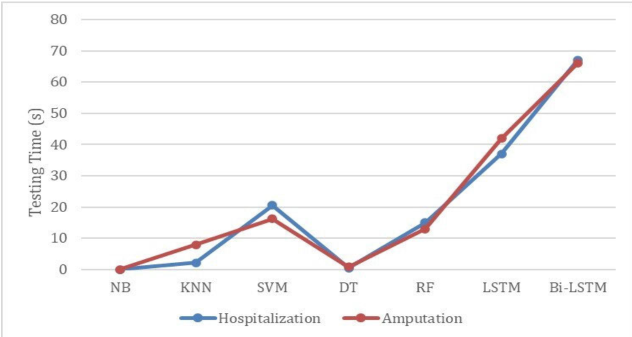

FIGURE 5. Prediction time of classification tasks.

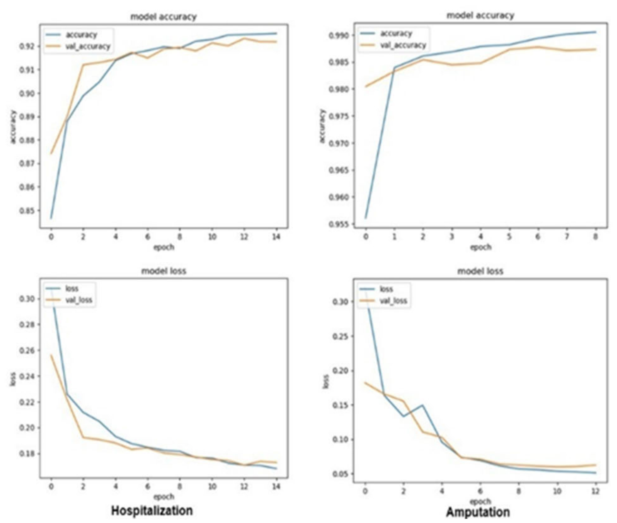

FIGURE 6. Learning curves of Bi-LSTM.

(a) Hospitalization

hospitalized	0.0628
finger	0.0595
caught	0.0138
tip	0.0127
ring	0.0124
amputated	0.0114
fell	0.0114
machine	0.0113
partial	0.0112
thumb	0.0106
hand	0.0093
pinched	0.0077
knuckle	0.0066
employee	0.0057
leg	0.0051
blade	0.0045
fractured	0.0044
broken	0.0041
slip	0.0024
ladder	0.0015

(b) Amputation

amputated	0.0941
finger	0.0707
amputation	0.0644
fingertip	0.0526
fell	0.0177
caught	0.0174
tip	0.0138
partial	0.0133
thumb	0.0125
hospitalized	0.0123
hand	0.0106
machine	0.0104
right	0.0097
left	0.0087
pinched	0.0083
knuckle	0.0080
saw	0.0058
leg	0.0055
blade	0.0046
fractured	0.0045

FIGURE 7. Top predictive words.

D. LEARNING CURVES OF BI-LSTM MODEL

Another metric to measure the performance of our proposed multimodal Bi-LSTM model is the visualization of the accuracy and loss of progress in the training and validation sets [63], [64]. Figure 6 shows the performance of the Bi-LSTM model based on the learning curves of accuracy and loss on the training and validation sets for both prediction tasks. From the figure, by implementing early stopping function, it appears that the validation loss (model loss) for the hospitalization task stopped decreasing and stabilized at epoch 14, whereas for the amputation prediction, it halted at epoch 12. This indicates that the Bi-LSTM model achieved the best performance in terms of minimizing the validation loss within these epochs. Additionally, it can be observed that the accuracy of the training and validation sets shows an upward trend and gradually becomes flat. Overall, as the epochs advanced, the learning curve for both the loss and accuracy values became more stable, with less fluctuation in the training and validation sets. This demonstrates that the models reached a point of convergence, and training them beyond these epochs may result in overfitting.

VI. MODEL OPTIMIZATION

We conducted a comprehensive methodology that involved feature importance analysis and hyperparameter optimization to achieve the optimum performance of the proposed multimodal Bi-LSTM predictive model. Firstly, a Random Forest Feature Importance algorithm was employed to determine the most relevant features of structured variables and unstructured text, respectively. Subsequently, the proposed Bi-LSTM model was optimized with these significant features and underwent hyperparameter tuning using Randomized Search Cross Validation. The optimized Bi-LSTM was then compared with the initial developed model to determine the best performing model for final model deployment.

A. FEATURE IMPORTANCE ANALYSIS

This step was introduced to assess the most important variables of the occupational injury severity prediction model, thereby providing valuable insights for classification tasks [57]. The RF algorithm determines the feature importance by measuring the reduction in Gini impurity in the model. The impurity values for all variables were added and standardized across the trees. The final value is organized in decreasing order, with the most important attribute at the top; the greater the value, the more important the feature [13], [58].

The findings revealed the top three most significant features of structured data were similar for both prediction tasks, which were ‘nature of injury’, ‘type of event’, and ‘affected body part’. These results were consistent with other related studies that measured ‘nature of injury’ and ‘affected body part’ [11], [65], [66], as well as, ‘type of event’ [18] as the essential predictors for occupational injury outcomes.

Moreover, the feature importance of unstructured text data was determined based on the importance of keywords. In this study, we proposed the top 20 keywords to be measured

TABLE 6. The multiple architectures of Bi-LSTM.

	Arch I	Arch 2	Arch 3
Batch Size	32	64	128
Epochs	20	25	30
LSTM Units	128	256	512
Dense		10	
Dropout		0.2	
Optimizer		Adam	

TABLE 7. The performance of multiple architectures.

	Hospitalization			Amputation		
	Accuracy	F1-score	AUC	Accuracy	F1-score	AUC
Arch 1	0.90	0.90	0.88	0.92	0.89	0.90
Arch 2	**0.93**	**0.95**	**0.93**	**0.99**	**0.97**	**0.98**
Archi3	0.89	0.88	0.89	0.88	0.86	0.87

and ranked for both the prediction tasks. Figure 7 illustrates the top 20 keywords for the prediction of hospitalization and amputation. Based on the figure, it is observed that the extracted predictive words were closely related to the 'type of event' of injury severity, such as 'fell', 'pinched', 'slip', 'caught', 'burns', 'tripped', 'broken', and 'fractured'. In addition, some keywords were identified as common objects that may cause the workplace injuries, such as 'machine', 'blade', 'ladder', saw', and 'floor'. With this interpretability analysis, it can be concluded that the content of occupational injury narratives comprised keywords that indicated the event and source of workplace injuries, including the severity outcomes, such as 'hospitalized/hospitalization' and 'amputation/amputated'.

B. HYPERPARAMETER OPTIMIZATION

Initially, the Bi-LSTM model was configured through multiple experiments based on the number of epochs, batch size, and LSTM units. Three architectures were proposed to determine the best accuracy, F1-score, and AUC for both the predictions. The proposed architectures are presented in Table 6.

Based on the experiments, it was found that the Bi-LSTMs configured with Arch 2 (epochs 25, batch size 64 and LSTM units 256) were the highest as presented in Table 7. Therefore, to further verify the Bi-LSTM configurations, hyperparameter optimization is introduced.

In this study, a Random Search algorithm was employed with cross-validation method (k-fold=10). This method allows a thorough exploration of the hyperparameter space to identify the optimal configuration for the Bi-LSTM predictive model. The dataset was divided into 10 equal-sized folds, in which each fold acted as a testing set, whereas the remaining folds served as the training set. This process was resumed until each fold was used once for testing [67]. For each iteration, a combination of hyperparameters for the defined search space is randomly sampled and trained on the training set. The performance of each configuration was then evaluated on the respective testing set using the assigned model evaluation metrics; such as accuracy, F1-score, and AUC. By repeating this process 10 times, a comprehensive review of the model's performance across different hyperparameter combinations was obtained. It is believed that the Random Search algorithm provided better performance prediction and efficient approach to tune the model's hyperparameter [68], thus ensuring the model predictability was reliable and generalized on the unseen data [69]. The corresponding hyperparameters are listed in Table 8.

TABLE 8. Corresponding hyperparameters for Bi-LSTM.

Hyperparameters	Ranges	Optimal
Lstm units	[128, 256, 512]	256
Dense units	[10, 20, 30]	10
Dropout rate	[0.2, 0.3, 0.40]	0.2
Batch size	[32, 64, 128]	64
Epochs	[20, 25, 30]	25
Activation	[relu, tanh, sigmoid]	ReLu
Optimizer	[Adam, RMSprop, SGD]	Adam
Initializer	[GlorotUniform, HeUniform]	HeUniform

Based on the table, HeUniform was determined as the preferred weight initialization method, as it is specifically designed to work well with ReLU activation [70]. This is in agreement with Huimin et al. [71], who concluded that the weight initializers corresponded to the activation function. The other identified best hyperparameter values were similar to the initial configurations.

This step was followed by optimizing the Bi-LSTM model with important features as input representations. For each prediction, the model was fed with three important features (structured data) and the most predictive keywords (unstructured data), and the hyperparameters were adjusted based on the optimal parameters. Next, the performance prediction of the optimized Bi-LSTM model was compared with the following architectures: (i) Bi-LSTM I, which was initially developed with all features and without hyperparameter tuning; (ii) Bi-LSTM II, a model with important features without hyperparameter tuning; and (iii) Optimized Bi-LSTM I (OPTIM Bi-LSTM I), a model developed with all features with hyperparameter tuning, whereas the optimized Bi-LSTM was labeled as OPTIM Bi-LSTM II composed of important features with hyperparameter tuning. The findings are presented in Table 9 for hospitalization prediction and Table 10 for amputation outcomes. Based on both tables, it was observed that the performance of the model evaluation metrics for each proposed model was consistent. Although the models using all features may produce slightly better

TABLE 9. Comparison of Bi-LSTMs performance for hospitalization.

Metrics	Bi-LSTM I (All Features)	Bi-LSTM II (Important Features)	OPTIM Bi-LSTM I (All Features)	OPTIM Bi-LSTM II (Important Feature)
	Without Hyperparameter Tuning		With Hyperparameter Tuning	
Accuracy	0.93	0.89	0.91	0.92
F1-Score	0.95	0.91	0.93	0.95
AUC	0.93	0.90	0.90	0.90
Training Time (s)	1625	1552	1049	**781**
Testing Time (s)	67	88	65	**49**

TABLE 10. Comparison of Bi-LSTMs performance for amputation.

Metrics	Bi-LSTM I (All Features)	Bi-LSTM II (Important Features)	OPTIM Bi-LSTM I (All Features)	OPTIM Bi-LSTM II (Important Features)
	Without Hyperparameter Tuning		With Hyperparameter Tuning	
Accuracy	0.99	0.94	0.97	0.98
F1-Score	0.97	0.95	0.95	0.96
AUC	0.98	0.94	0.96	0.98
Training Time (s)	1939	1667	1159	**428**
Testing Time (s)	66	79	6	**42**

performance metrics, the possibility of utilizing those features without feature importance or hyperparameter tuning may introduce noise and irrelevant information for model development.

Additionally, it was found that the OPTIM Bi-LSTM II, a proposed model with important features and hyperparameter tuning managed to generate prediction outputs in a timely manner. It is believed that the optimized hyperparameters can lead to faster convergence during model training, thereby allowing the model to reach its optimal performance more quickly. The optimized Bi-LSTM can predict hospitalization outcome at 49s and amputation severity at only 42s. Therefore, the feature optimization algorithms conducted in this study optimize the multimodal Bi-LSTM occupational injury severity prediction model and excellently accelerate the model prediction time, making it suitable for occupational injury decision support systems in real field applications.

Consequently, this study is in agreement with a recent study by [72], which preferred the execution of an effective and optimum prediction model that utilizes fewer important features than numerous features as input representations. Some arguments in the existing literature highlighted the 'impracticality' of using larger set of variables in developing the machine and deep learning classifiers; (i) numerous features may increase the complexity of the model and (ii) training process suffers with overfitting problems [73], including (iii) a complex model may generates higher computational tasks, making it cost expensive and less efficient [74]. Therefore, any technique that promotes the reduction of data dimensionality is recommended to improve model performance prediction [75].

Accordingly, this study emphasizes this feature optimization approach to assist Safety and Health Practitioners in executing an accurate interpretable practical and time-efficient occupational injury severity prediction model, thus guiding practitioners and policymakers to improve workplace injury intervention strategies. Figure 8 depicts the overall proposed framework of our multimodal Bi-LSTM occupational injury severity prediction model. This proposed framework concludes the innovative approaches developed from our study; multimodal learning with Bi-LSTM predictive model integrates with model optimization techniques to enhance model interpretability, practicality, and predictability.

VII. DISCUSSION

A. UNIQUENESS OF THE PROPOSED BI-LSTM MODEL

From the findings, the recurrent neural network variant; the Bi-LSTM model showed promising prediction performances for both prediction tasks using the multimodal where the; structured and unstructured notes used as the input features. The new aspect of this study is the optimization of the proposed Bi-LSTM model in predicting the outcomes of occupational injuries by making use of both structured and unstructured data as input features. In addition, the proposed model of the optimized Bi-LSTM has two LSTMs applied to the input features. Firstly, an LSTM is executed on the input sequence ("forward layer") and followed the training in reverse order with another LSTM ("backward layer") [76]. Because of its innovative architecture, which includes both forward and backward LSTM layers, the proposed model is able to do an analysis on each and every component of the input sequences. As a result, the model's accuracy is improved, and the results are more relevant. We believed that the architectures of the proposed model, in which the desired algorithm is trained, not only from the 'input to output' but also from the 'output to input' leads to its high model performances. This nature of architecture gives additional advantages as the proposed model is able to analyze every component of the input sequences, thus, providing more meaningful outputs and enhancing the model's accuracy [77].

Additionally, the recurrent layers in the proposed Bi-LSTM model have been assumed as the reason for the capability of this deep learning algorithm to learn better the feature representations as the networks and layers grow deeper [27]. The results highlighted the remarkable performance of our

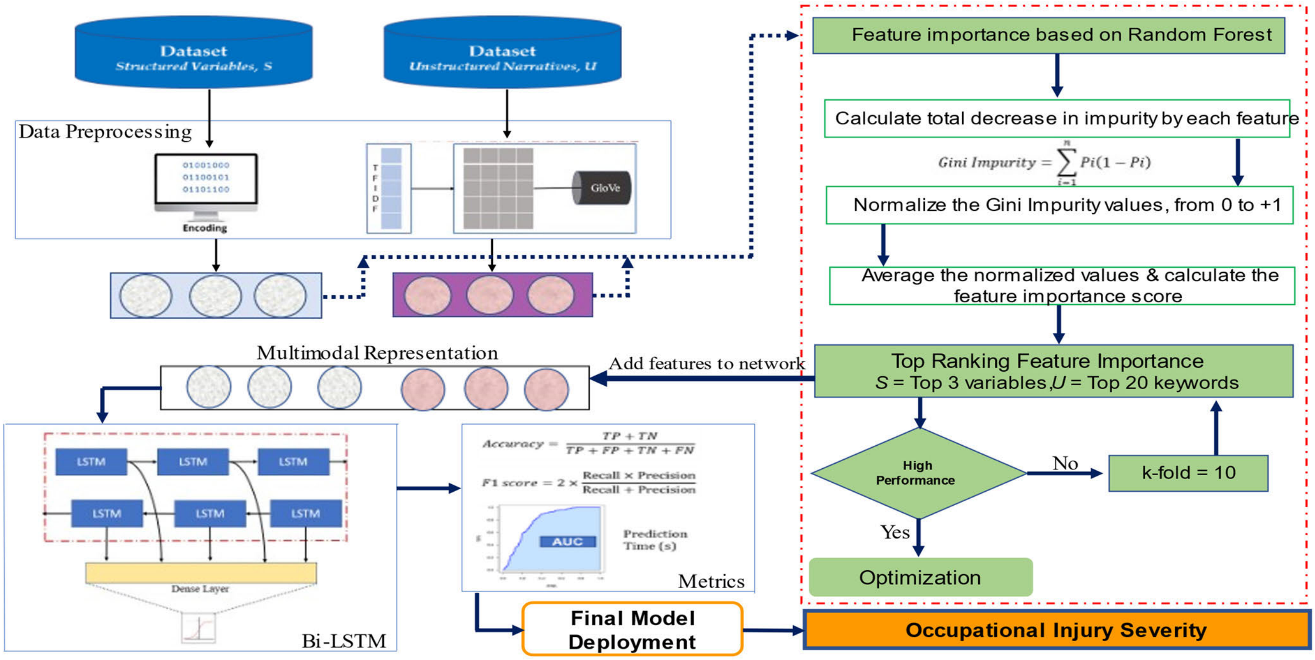

FIGURE 8. The overall proposed framework of multimodal Bi-LSTM occupational injury severity prediction model.

proposed Bi-LSTM optimized model, in terms of the uniqueness of its architectures that are well-suited in handling 'massive-length' data sequences in multimodal learning.

Moreover, the execution of the LSTMs from both orders or directions justified more time required for prediction, in terms of training and testing time using this model. It is known that more complex the model, more time required to train and test the prediction outcomes [78]. Therefore, it is fair to mention the Bi-LSTM model in this multimodal learning required a bit longer training and testing time, compared to other algorithms.

Next, we observed the prediction accuracy, F1-score and AUC values of RF and SVM are quite close to the best-performing model for both prediction tasks. In RF, the classifier is a combination of prediction trees and acts as an "ensemble" [19]. This capability of integrating the prediction ability of multiple learners into a single RF model leads RF to perform quite well in this study. Not to mention, the SVM algorithm has the ability to map the input representation in a high dimensional space by using the 'kernel function'. The 'kernel function' used in this study is a radial-based function ("rbf") and this function significantly increased the accuracy performance [65] making SVM one of the most effective machine learning classifiers [53].

B. COMPLEMENTARY NATURE OF MULTIMODAL DATA

This study emphasizes the use of multimodal data sources to develop an occupational injury severity model. Structured input highlighted the injured worker's information, whereas the sequential injury narratives contained the workplace injury history. Based on the findings, the performances of the proposed machine and deep learning models are satisfactory, as they ranged from 0.8 to 0.9 above in each metric, for both prediction tasks. This performance prediction implicitly clarifies the harmonious nature of multimodal data to complement one another, thus, generating good-quality of predictive models. Additionally, multimodal learning can enhance model robustness by reducing the impact of noisy or incomplete data in a single modality. If one modality, such as structured data, is ambiguous, the presence of other modalities, such as unstructured text, can compensate for it, thus providing a more realistic prediction performance. Moreover, the predictive model generated from multimodal learning can generalize better to hidden data because it learns and trains from multiple sources, thus capturing a broader range of patterns and data relationships

In the context of occupational injury, the integrated analysis of multimodal learning permits the extraction of rich useful information from occupational injury records prepared by Safety and Health Practitioners and Occupational Health Doctors. This kind of integration of field experts in the occupational injury domain and technical aspects of workplace safety resulted in dependable stable successful occupational injury severity prediction outcomes [79], [80], [81]. It is believed that by integrating both modalities, it can provide a more comprehensive data representations as each modality contains unique and complementary information, thereby integrating them lead to a more holistic understanding of the underlying occupational injury severity events.

Consequently, the combination of the structured tabular data and unstructured injury notes in this study justified the successful execution of multimodal deep neural architectures as it appears as a convincing strategy to improve the prediction performance of occupational injury severity.

C. LONG-TERM BENEFIT OF THE PROPOSED SYSTEM

The detection of occupational injury severity is important to address the post-injury consequences to the injured worker, as well as, the organization [82]. In the case of an injured worker being hospitalized, they have to take days off for recovery, and those amputated may face longer treatment as it may involve physical and emotional rehabilitation. The absence of work owing to the severity of workplace injuries may affect the organization's lost-time injuries (LTI). LTIs are indicators of the effectiveness of workplace safety and health in preventing work-related accidents and injuries. A high rate of LTIs indicates poor safety and health monitoring in the organization [83]. Moreover, workplace injury severity is related to the chances of injured workers returning to work [8]. Amputation due to workplace injury may result in permanent disability, and functional deterioration are the main reason for not returning to work. This may generate long-term psychological effects, especially on mental health, thus, prolonging injury recovery [84].

As a result, this study offers a number of contributions to the actual applications used in industry. The ultimate objective of this research was to develop an accurate forecasting system that would be of assistance to safety and health professionals, particularly in the area of estimating the severity of occupational injuries. Intervention techniques for workplace safety can be applied as first preventive steps to lessen the severity of injuries based on the severity outcomes that have been forecasted for them. This newly created multimodal prediction model has the capability of successfully identifying high-risk regions and activities within the workplace. This is accomplished by precisely predicting the possibility of occupational injury severity. Using this information to guide, the adoption of targeted safety interventions and remedial strategies to lower the likelihood of future injuries is beneficial.

Additionally, this predictive system can aid in the early screening and identification of at-risk workers with severe occupational injury outcomes, thereby allowing the prioritization of safety interventions and support systems for those workers, such as providing specialize safety training and offer support for physical and mental well-being. The proactive approach of this predictive system can lead to improved workplace safety, health, and overall well-being. Furthermore, this predictive system potential to be useful to industry practitioners in the field of resource management. When workers are incapable of working due to hospitalization and rehabilitation, this can cause significant productivity losses to the company. Therefore, management may reallocate resources, such as assigning additional manpower or redesigning work assignments to ensure that job productivity is still underway. Next, additional ongoing support, such as physical and counselling support, including job retraining can be allocated in assisting those injured workers to recover safely and timely manner.

Consequently, occupational injury severity predictive analytics utilizing multimodal learning are essential for early screening, anticipation, and identification tools for at-risk workers with severe occupational injury outcomes. Correspondingly, the information obtained from multimodal dataset analysis is beneficial in addressing the compelling concerns among Safety and Health Practitioners to foresee effective intervention strategies for preventing the severity of workplace accidents [85], [86], thereby, promoting workplace environment that is safer and healthier for employees. Worker safety, health, and well-being are of the greatest priority in occupational safety and health; thus, it is vital to employ the latest advanced Artificial Intelligence (AI) approach in constructing an accurate and robust occupational injury severity prediction model.

D. COMPARISON WITH RECENT SIMILAR APPROACH

Sarkar et al. [18] employed a multimodal dataset (structured and unstructured injury reports) from the steel manufacturing industry to develop an occupational injury prediction model. They developed a simple DNN model and tuned it using three optimizers: Adam, RMSprop, and SGD, with a 10-fold cross-validation method. The findings revealed that the DNN with Adam-optimizer (ADNN) achieved the best classifier with 0.79 accuracy compared to SGD-DNN, RMSprop-DNN, KNN, SVM, and RF. Our study was in agreement with their study in terms of using the Adam-optimizer, 10-fold cross validation scheme, and compared with similar state-of-the-art ML models.

In other related studies, Mahajan et al. [28] developed a baseline multimodal predictive model using the LR algorithm to predict 30-day readmission for heart failure. They used structured data from EHR and combined them with unstructured clinical notes. The multimodal LR prediction model achieved 0.65 AUC values, compared to using only structured (0.64) or unstructured notes (0.52). Moreover, Zhang et al. [27] developed a multimodal DNN named 'Fusion-LSTM' to predict mortality, hospitalization stay and hospital readmission using the 'MIMIC-III' health records. They utilized unstructured clinical notes and static information, and the model produced more accurate predictions and outperformed the baseline methods: LR and RF models with AUC scores of 0.87. Compared with our proposed approach, the multimodal Bi-LSTM model achieved a higher AUC of 0.90.

Recent approaches have become more advanced in terms of the diverse integration of data modalities with advanced DNN architectures. For example, Saleh and Murab [87] developed a Convolutional Neural Network (CNN) to predict fall injuries. In their study, they combined images and sensor data into multimodal representations. The findings revealed that the multimodal CNN model outperformed other conventional ML methods: SVM, KNN, DT, and RF with an accuracy of 0.97. The latest work by Jujjavarapu et al. [88] integrated structured and unstructured health data, consisting of patients' personal information, diagnosis codes, drug names, and diagnostic imaging reports, to predict decompression surgery for low back pain due to occupational back injury. They proposed a multimodal deep learning architecture composed of (i) a layer of CNN, (ii) a layer of Gated Recurrent Unit (GRU) model, an advanced simpler architecture of LSTM, and (iii) 2-layer-Fully Connected, compared to baseline LASSO Logistic Regression. The findings revealed that multimodal deep learning achieved a better AUC value of 0.73.

Our findings and reviewed studies consistently indicate that multimodal deep learning architectures generate better predictive performance than traditional ML models. Further exploration in terms of data accessibility and advancement in adopting standardized multimodal deep learning methodologies in the occupational injury domain is required, with the potential to assist decision-making, resource allocation, and enhance workplace injury intervention strategies in real-industry applications. In the following, we highlight the limitations of study and potential opportunities for future research.

E. LIMITATIONS OF STUDY

Although this study utilized occupational injury data across broad industrial sectors, we noticed the necessity of performing additional transfer learning to evaluate the generalizability of the developed model. However, most occupational injury datasets are restricted and did not reveal sufficient features to indicate the severity of hospitalization and amputation [15], thus limiting the accessibility and data quality to undergo the predictive analysis process. Besides, most of the dataset are 'domain-specific' and difficult to be transferred to other settings [27], such as the 'technical-language' related to workplace safety, including 'manner of injury' may vary between industries [17], [89].

Additionally, the dataset relies on 'human-labelled' data; as each Safety and Health Practitioner may have diverse interpretation of workplace injury severities due to their experience and training level, thereby impacting the consistency of data labelling and categorization. This limitation requires an extensive human assistance to clean and sanitize the data labelling before it can be proceeded for further analysis.

The scalability of our machine learning models is another limitation to consider. Given a large dataset comprising multiple modalities, our computational resources were constrained, thereby limiting the exploration of a wider range of model architectures and parameters. Although our study provides valuable insights into multimodal data analysis, the impact of the parameter choices of each machine-learning classifier cannot be overlooked. More in-depth investigations of the effects of specific parameters on different aspects of the analysis would enrich the understanding of our findings.

F. FUTURE RESEARCH TREND

There are multiple sources of occupational safety data that can be used to develop occupational injury prediction models, such as workplace safety audit reports, hazard evaluation reports, and injured workers' compensation records. Integrating these data sources could improve the comprehensiveness, generalizability, and transferability of the model. As a way forward, we anticipate that future research should integrate injured worker information from occupational injury reports and worker compensation documents to further analyze the pattern of workplace injury severity and the accurate cost implication, thereby enhancing the model interpretability for efficient utilization as an intelligent occupational injury decision support system for real industrial applications.

Another direction is to improve multimodal learning in occupational injury research by exploring other types of multimodal data fusion strategies, such as joint fusion and late fusion, including hybrid fusion in generating more robust occupational injury prediction model. Moreover, multimodal learning using workplace injury images with structured data and unstructured text is recommended; thus, alternative neural architectures such as convolutional neural networks (CNN) have been proposed.

Finally, future research could benefit from extensive feature optimization to assess the robustness of our findings with respect to the parameter variations. Additionally, conducting multiple experiments or iterative refinement approaches to explore a wide range of parameter settings and leveraging more advanced computational resources would help enhance the efficiency and effectiveness of our multimodal analysis.

VIII. CONCLUSION

In conclusion, our study highlights the need to utilize all modalities in occupational injury records to determine the risk of occupational injury severity, such as hospitalization and amputation. The proposed model has been proven to work well in this multimodal learning for both prediction tasks. These findings are significant in practicing workplace accidents and injury analytics because the model shows a high predictive and accurate classification performance.

To the best of our knowledge, this study is the first to propose multimodal integration learning with traditional machine learning algorithms and recurrent neural network variants; hence, our study serves as a crucial foundation and benchmark for further advancements in multimodal deep learning for occupational injury prediction. In addition, we merged a large historical workplace injury-specific dataset to classify the severity of occupational injuries across broad industrial sectors.

CODE AVAILABILITY STATEMENT

The sample code used in this study is available at https://github.com/mzf23/oshinjury. Any updates or improvements to the code will be made available in the repository to ensure accessibility and sustainability of the research findings.

REFERENCES

[1] Z. Aderaw, D. Engdaw, and T. Tadesse, ''Determinants of occupational injury: A case control study among textile factory workers in Amhara regional state, Ethiopia,'' *J. Tropical Med.*, vol. 2011, pp. 1–8, Oct. 2011.

[2] G. Mulu, K. Abera, and G. Gebrekiros, ''Magnitude of occupational injury and associated factors among factory workers in Ethiopia: The case of Mugher Cement Factory,'' *J. Public Health Epidemiol.*, vol. 9, no. 12, pp. 318–331, Dec. 2017.

[3] E. Wadsworth and D. Walters, ''Safety and health at the heart of the future of work: Building on 100 years of experience,'' Int. Labour Org., Geneva, Switzerland, 2019.

[4] C. Atombo, C. Wu, E. O. Tettehfio, G. Y. Nyamuame, and A. A. Agbo, ''Safety and health perceptions in work-related transport activities in Ghanaian industries,'' *Saf. Health Work*, vol. 8, no. 2, pp. 175–182, Jun. 2017.

[5] Y. M. Alamneh, A. Z. Wondifraw, A. Negesse, D. B. Ketema, and T. Y. Akalu, ''The prevalence of occupational injury and its associated factors in Ethiopia: A systematic review and meta-analysis,'' *J. Occupat. Med. Toxicol.*, vol. 15, no. 1, pp. 1–11, Dec. 2020.

[6] K. Boczkowska, K. Niziołek, and E. Roszko-Wójtowicz, ''A multivariate approach towards the measurement of active employee participation in the area of occupational health and safety in different sectors of the economy,'' *Equilibrium. Quart. J. Econ. Econ. Policy*, vol. 17, no. 4, pp. 1051–1085, 2022.

[7] D. Kendrick, P. Dhiman, B. Kellezi, C. Coupland, J. Whitehead, K. Beckett, N. Christie, J. Sleney, J. Barnes, S. Joseph, and R. Morriss, ''Psychological morbidity and return to work after injury: Multicentre cohort study,'' *Brit. J. Gen. Pract.*, vol. 67, no. 661, pp. e555–e564, Aug. 2017.

[8] P.-C. Chu, W.-S. Chin, Y. L. Guo, and J. S.-C. Shiao, ''Long-term effects of psychological symptoms after occupational injury on return to work: A 6-year follow-up,'' *Int. J. Environ. Res. Public Health*, vol. 16, no. 2, p. 235, Jan. 2019. [Online]. Available: https://www.mdpi.com/1660-4601/16/2/235

[9] W.-S. Chin, Y. L. Guo, S.-C. Liao, K.-H. Lin, C.-Y. Kuo, C.-C. Chen, and J. S. C. Shiao, ''Suicidality 6 years after occupational injury,'' *J. Clin. Psychiatry*, vol. 79, no. 5, p. 20333, Sep. 2018.

[10] A. O. Oyedele, A. O. Ajayi, and L. O. Oyedele, ''Machine learning predictions for lost time injuries in power transmission and distribution projects,'' *Mach. Learn. Appl.*, vol. 6, Dec. 2021, Art. no. 100158.

[11] A. Yedla, F. D. Kakhki, and A. Jannesari, ''Predictive modeling for occupational safety outcomes and days away from work analysis in mining operations,'' *Int. J. Environ. Res. Public Health*, vol. 17, no. 19, p. 7054, Sep. 2020. [Online]. Available: https://www.mdpi.com/1660-4601/17/19/7054

[12] M. Chadyiwa, J. Kagura, and A. Stewart, ''Investigating machine learning applications in the prediction of occupational injuries in South African National Parks,'' *Mach. Learn. Knowl. Extraction*, vol. 4, no. 3, pp. 768–778, Aug. 2022. [Online]. Available: https://www.mdpi.com/2504-4990/4/3/37

[13] M. Z. F. Khairuddin, P. L. Hui, K. Hasikin, N. A. A. Razak, K. W. Lai, A. S. M. Saudi, and S. S. Ibrahim, ''Occupational injury risk mitigation: Machine learning approach and feature optimization for smart workplace surveillance,'' *Int. J. Environ. Res. Public Health*, vol. 19, no. 21, p. 13962, Oct. 2022, doi: 10.3390/ijerph192113962.

[14] S. B. Rao, ''A comparative approach of text mining: Classification, clustering and extraction techniques,'' *J. Mech. Continua Math. Sci.*, vol. 5, no. 1, pp. 120–131, Jan. 2020, doi: 10.26782/jmcms.spl.5/2020.01.00010.

[15] S. Jing, X. Liu, X. Gong, Y. Tang, G. Xiong, S. Liu, S. Xiang, and R. Bi, ''Correlation analysis and text classification of chemical accident cases based on word embedding,'' *Process Saf. Environ. Protection*, vol. 158, pp. 698–710, Feb. 2022, doi: 10.1016/j.psep.2021.12.038.

[16] H. Baker, M. R. Hallowell, and A. J.-P. Tixier, ''AI-based prediction of independent construction safety outcomes from universal attributes,'' *Autom. Construct.*, vol. 118, Oct. 2020, Art. no. 103146, doi: 10.1016/j.autcon.2020.103146.

[17] D. M. Goldberg, ''Characterizing accident narratives with word embeddings: Improving accuracy, richness, and generalizability,'' *J. Saf. Res.*, vol. 80, pp. 441–455, Feb. 2022, doi: 10.1016/j.jsr.2021.12.024.

[18] S. Sarkar, S. Vinay, C. Djeddi, and J. Maiti, ''Classification and pattern extraction of incidents: A deep learning-based approach,'' *Neural Comput. Appl.*, vol. 34, no. 17, pp. 14253–14274, Sep. 2022, doi: 10.1007/s00521-021-06780-3.

[19] A. Oyedele, A. Ajayi, L. O. Oyedele, J. M. D. Delgado, L. Akanbi, O. Akinade, H. Owolabi, and M. Bilal, ''Deep learning and boosted trees for injuries prediction in power infrastructure projects,'' *Appl. Soft Comput.*, vol. 110, Oct. 2021, Art. no. 107587, doi: 10.1016/j.asoc.2021.107587.

[20] R. Ganguli, P. Miller, and R. Pothina, ''Effectiveness of natural language processing based machine learning in analyzing incident narratives at a mine,'' *Minerals*, vol. 11, no. 7, p. 776, Jul. 2021. [Online]. Available: https://www.mdpi.com/2075-163X/11/7/776

[21] J.-M. Kim, K.-K. Lim, S.-G. Yum, and S. Son, ''A deep learning model development to predict safety accidents for sustainable construction: A case study of fall accidents in South Korea,'' *Sustainability*, vol. 14, no. 3, p. 1583, Jan. 2022. [Online]. Available: https://www.mdpi.com/2071-1050/14/3/1583

[22] F. Zhang, ''A hybrid structured deep neural network with Word2Vec for construction accident causes classification,'' *Int. J. Construct. Manage.*, vol. 22, no. 6, pp. 1120–1140, Apr. 2022, doi: 10.1080/15623599.2019.1683692.

[23] A. Zermane, M. Z. M. Tohir, H. Zermane, M. R. Baharudin, and H. M. Yusoff, ''Predicting fatal fall from heights accidents using random forest classification machine learning model,'' *Saf. Sci.*, vol. 159, Mar. 2023, Art. no. 106023, doi: 10.1016/j.ssci.2022.106023.

[24] A. Kline, H. Wang, Y. Li, S. Dennis, M. Hutch, Z. Xu, F. Wang, F. Cheng, and Y. Luo, ''Multimodal machine learning in precision health: A scoping review,'' *NPJ Digit. Med.*, vol. 5, no. 1, p. 171, Nov. 2022, doi: 10.1038/s41746-022-00712-8.

[25] S.-C. Huang, A. Pareek, S. Seyyedi, I. Banerjee, and M. P. Lungren, ''Fusion of medical imaging and electronic health records using deep learning: A systematic review and implementation guidelines,'' *NPJ Digit. Med.*, vol. 3, no. 1, p. 136, Oct. 2020, doi: 10.1038/s41746-020-00341-z.

[26] E. G. Ross, K. Jung, J. T. Dudley, L. Li, N. J. Leeper, and N. H. Shah, ''Predicting future cardiovascular events in patients with peripheral artery disease using electronic health record data,'' *Circulat., Cardiovascular Qual. Outcomes*, vol. 12, no. 3, Mar. 2019, Art. no. e004741, doi: 10.1161/circoutcomes.118.004741.

[27] D. Zhang, C. Yin, J. Zeng, X. Yuan, and P. Zhang, ''Combining structured and unstructured data for predictive models: A deep learning approach,'' *BMC Med. Informat. Decis. Making*, vol. 20, no. 1, pp. 1–11, Dec. 2020.

[28] S. M. Mahajan and R. Ghani, ''Combining structured and unstructured data for predicting risk of readmission for heart failure patients,'' in *Proc. MedInfo*, 2019, pp. 238–242.

[29] A. Khan and S. Zubair, ''An improved multi-modal based machine learning approach for the prognosis of Alzheimer's disease,'' *J. King Saud Univ., Comput. Inf. Sci.*, vol. 34, no. 6, pp. 2688–2706, Jun. 2022.

[30] S. Kayikci and T. Khoshgoftaar, ''A stack based multimodal machine learning model for breast cancer diagnosis,'' in *Proc. Int. Congr. Hum.-Comput. Interact., Optim. Robotic Appl. (HORA)*, Jun. 2022, pp. 1–5.

[31] L. Dey, H. Meisheri, and I. Verma, ''Predictive analytics with structured and unstructured data—A deep learning based approach,'' *IEEE Intell. Inform. Bull.*, vol. 18, no. 2, pp. 27–34, Dec. 2017.

[32] M. K. Johansson, M. Hasselberg, and R. Rissanen, ''Return to work and sick leave patterns following a work injury among young adults: A study protocol of a Swedish multimodal study,'' *BMJ Open*, vol. 11, no. 6, Jun. 2021, Art. no. e045143, doi: 10.1136/bmjopen-2020-045143.

[33] G. Paraskevopoulos, P. Pistofidis, G. Banoutsos, E. Georgiou, and V. Katsouros, ''Multimodal classification of safety-report observations,'' *Appl. Sci.*, vol. 12, no. 12, p. 5781, Jun. 2022. [Online]. Available: https://www.mdpi.com/2076-3417/12/12/5781

[34] OSHA. *OSHA Severe Injury Report*. Accessed: Jan. 25, 2023. [Online]. Available: https://www.osha.gov/severeinjury

[35] D. Singh and B. Singh, ''Investigating the impact of data normalization on classification performance,'' *Appl. Soft Comput.*, vol. 97, Dec. 2020, Art. no. 105524, doi: 10.1016/j.asoc.2019.105524.

[36] J. P. Mueller and L. Massaron, *Deep Learning for Dummies*. Hoboken, NJ, USA: Wiley, 2019.

[37] F. Zhang, H. Fleyeh, X. Wang, and M. Lu, ''Construction site accident analysis using text mining and natural language processing techniques,'' *Autom. Construct.*, vol. 99, pp. 238–248, Mar. 2019, doi: 10.1016/j.autcon.2018.12.016.

[38] S. R. Basha and J. K. Rani, "A comparative approach of dimensionality reduction techniques in text classification," *Eng., Technol. Appl. Sci. Res.*, vol. 9, no. 6, pp. 4974–4979, Dec. 2019, doi: 10.48084/etasr.3146.

[39] B. Song and Y. Suh, "Narrative texts-based anomaly detection using accident report documents: The case of chemical process safety," *J. Loss Prevention Process Industries*, vol. 57, pp. 47–54, Jan. 2019.

[40] Y. M. Goh and C. U. Ubeynarayana, "Construction accident narrative classification: An evaluation of text mining techniques," *Accident Anal. Prevention*, vol. 108, pp. 122–130, Nov. 2017, doi: 10.1016/j.aap.2017.08.026.

[41] X. Pan, H. Wang, W. You, M. Zhang, and Y. Yang, "Assessing the reliability of electronic products using customer knowledge discovery," *Rel. Eng. Syst. Saf.*, vol. 199, Jul. 2020, Art. no. 106925.

[42] L.-C. Yu, J. Wang, K. R. Lai, and X. Zhang, "Refining word embeddings using intensity scores for sentiment analysis," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 26, no. 3, pp. 671–681, Mar. 2018, doi: 10.1109/TASLP.2017.2788182.

[43] J. Pennington, R. Socher, and C. Manning, "Glove: Global vectors for word representation," in *Proc. Conf. Empirical Methods Natural Lang. Process. (EMNLP)*, 2014, pp. 1532–1543.

[44] D. Ramachandram and G. W. Taylor, "Deep multimodal learning: A survey on recent advances and trends," *IEEE Signal Process. Mag.*, vol. 34, no. 6, pp. 96–108, Nov. 2017.

[45] H. Chai, X. Zhou, Z. Zhang, J. Rao, H. Zhao, and Y. Yang, "Integrating multi-omics data through deep learning for accurate cancer prognosis prediction," *Comput. Biol. Med.*, vol. 134, Jul. 2021, Art. no. 104481, doi: 10.1016/j.compbiomed.2021.104481.

[46] G. Lee, K. Nho, B. Kang, K.-A. Sohn, and D. Kim, "Predicting Alzheimer's disease progression using multi-modal deep learning approach," *Sci. Rep.*, vol. 9, no. 1, p. 1952, Feb. 2019, doi: 10.1038/s41598-018-37769-z.

[47] E. Scott, L. Hirabayashi, A. Levenstein, N. Krupa, and P. Jenkins, "The development of a machine learning algorithm to identify occupational injuries in agriculture using pre-hospital care reports," *Health Inf. Sci. Syst.*, vol. 9, no. 1, p. 31, Dec. 2021, doi: 10.1007/s13755-021-00161-9.

[48] H. Baker, M. R. Hallowell, and A. J.-P. Tixier, "Automatically learning construction injury precursors from text," *Autom. Construct.*, vol. 118, Oct. 2020, Art. no. 103145, doi: 10.1016/j.autcon.2020.103145.

[49] M.-Y. Cheng, D. Kusoemo, and R. A. Gosno, "Text mining-based construction site accident classification using hybrid supervised machine learning," *Autom. Construct.*, vol. 118, Oct. 2020, Art. no. 103265, doi: 10.1016/j.autcon.2020.103265.

[50] S. C. J. Lim and S. Kang, "A classification model for predicting the injured body part in construction accidents in Korea," in *Proc. 9th Int. Conf. Construct. Eng. Project Manag.*, Las Vegas, NV, USA, 2022. [Online]. Available: http://www.iccepm2022.com/index.php

[51] I. Häkkä, "Application of machine learning to predict occurrence of accidents at Finnish construction sites," M.S. thesis, Aalto Univ., School Bus., Espoo, Finland, 2022. [Online]. Available: http://urn.fi/URN:NBN:fi:aalto-202209115534

[52] Y. Y. Song and Y. Lu, "Decision tree methods: Applications for classification and prediction," *Shanghai Arch. Psychiatry*, vol. 27, no. 2, pp. 130–135, Apr. 2015, doi: 10.11919/j.issn.1002-0829.215044.

[53] A. Afifi, "Improving the classification accuracy using support vector machines (SVMS) with new kernel," *J. Global Res. Comput. Sci.*, vol. 4, no. 2, pp. 1–7, 2013.

[54] G. Colmenarejo, "Machine learning models to predict childhood and adolescent obesity: A review," *Nutrients*, vol. 12, no. 8, p. 2466, Aug. 2020. [Online]. Available: https://www.mdpi.com/2072-6643/12/8/2466

[55] H. Zhao, S. Sun, and B. Jin, "Sequential fault diagnosis based on LSTM neural network," *IEEE Access*, vol. 6, pp. 12929–12939, 2018.

[56] K. Wu, J. Wu, L. Feng, B. Yang, R. Liang, S. Yang, and R. Zhao, "An attention-based CNN-LSTM-BiLSTM model for short-term electric load forecasting in integrated energy system," *Int. Trans. Electr. Energy Syst.*, vol. 31, no. 1, Jan. 2021, Art. no. e12637, doi: 10.1002/2050-7038.12637.

[57] A. Singh, R. Thapliyal, R. Vanave, R. Shedge, and S. Mumbaikar, "Analysis of hyperparameters in sentiment analysis of movie reviews using Bi-LSTM," in *Proc. ITM Web Conf.*, vol. 44, 2022, p. 3012, doi: 10.1051/itmconf/20224403012.

[58] R. A. Hameed, W. J. Abed, and A. T. Sadiq, "Evaluation of hotel performance with sentiment analysis by deep learning techniques," *Int. J. Interact. Mobile Technol.*, vol. 17, no. 9, pp. 70–87, May 2023.

[59] S. Ansari, F. Naghdy, H. Du, and Y. N. Pahnwar, "Driver mental fatigue detection based on head posture using new modified ReLU-BiLSTM deep neural network," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 8, pp. 10957–10969, Aug. 2022, doi: 10.1109/TITS.2021.3098309.

[60] M. S. Islam, M. N. Islam, N. Hashim, M. Rashid, B. S. Bari, and F. A. Farid, "New hybrid deep learning approach using BiGRU-BiLSTM and multilayered dilated CNN to detect arrhythmia," *IEEE Access*, vol. 10, pp. 58081–58096, 2022, doi: 10.1109/ACCESS.2022.3178710.

[61] M. V. Ferro, Y. D. Mosquera, F. J. R. Pena, and V. M. D. Bilbao, "Early stopping by correlating online indicators in neural networks," *Neural Netw.*, vol. 159, pp. 109–124, Feb. 2023, doi: 10.1016/j.neunet.2022.11.035.

[62] R. Guns, C. Lioma, and B. Larsen, "The tipping point: F-score as a function of the number of retrieved items," *Inf. Process. Manage.*, vol. 48, no. 6, pp. 1171–1180, Nov. 2012.

[63] J. Huan, B. Chen, X. G. Xu, H. Li, M. B. Li, and H. Zhang, "River dissolved oxygen prediction based on random forest and LSTM," *Appl. Eng. Agricult.*, vol. 37, no. 5, pp. 901–910, 2021.

[64] S. N. Njimbouom, K. Lee, and J.-D. Kim, "MMDCP: Multi-modal dental caries prediction for decision support system using deep learning," *Int. J. Environ. Res. Public Health*, vol. 19, no. 17, p. 10928, Sep. 2022.

[65] F. D. Kakhki, S. A. Freeman, and G. A. Mosher, "Evaluating machine learning performance in predicting injury severity in agribusiness industries," *Saf. Sci.*, vol. 117, pp. 257–262, Aug. 2019, doi: 10.1016/j.ssci.2019.04.026.

[66] K.-S. Kang, C. Koo, and H.-G. Ryu, "An interpretable machine learning approach for evaluating the feature importance affecting lost workdays at construction sites," *J. Building Eng.*, vol. 53, Aug. 2022, Art. no. 104534, doi: 10.1016/j.jobe.2022.104534.

[67] Y. Jung and J. Hu, "A K-fold averaging cross-validation procedure," *J. Nonparametric Statist.*, vol. 27, no. 2, pp. 167–179, Apr. 2015, doi: 10.1080/10485252.2015.1010532.

[68] J. Bergstra and Y. Bengio, "Random search for hyper-parameter optimization," *J. Mach. Learn. Res.*, vol. 13, no. 2, pp. 281–305, 2012.

[69] S. Wang, C. Ma, Y. Xu, J. Wang, and W. Wu, "A hyperparameter optimization algorithm for the LSTM temperature prediction model in data center," *Sci. Program.*, vol. 2022, Dec. 2022, Art. no. 6519909, doi: 10.1155/2022/6519909.

[70] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 1026–1034, doi: 10.1109/ICCV.2015.123.

[71] H. Li, M. Krček, and G. Perin, "A comparison of weight initializers in deep learning-based side-channel analysis," in *Proc. Appl. Cryptogr. Netw. Secur. Workshops*. Cham, Switzerland: Springer, 2020, pp. 126–143.

[72] M. Z. I. Chowdhury and T. C. Turin, "Variable selection strategies and its importance in clinical prediction modelling," *Family Med. Community Health*, vol. 8, no. 1, Feb. 2020, Art. no. e000262, doi: 10.1136/fmch-2019-000262.

[73] D. Rengasamy, J. M. Mase, A. Kumar, B. Rothwell, M. T. Torres, M. R. Alexander, D. A. Winkler, and G. P. Figueredo, "Feature importance in machine learning models: A fuzzy information fusion approach," *Neurocomputing*, vol. 511, pp. 163–174, Oct. 2022, doi: 10.1016/j.neucom.2022.09.053.

[74] Y. A. Majeed, S. S. Awadalla, and J. L. Patton, "Regression techniques employing feature selection to predict clinical outcomes in stroke," *PLoS ONE*, vol. 13, no. 10, Oct. 2018, Art. no. e0205639.

[75] K. Maharana, S. Mondal, and B. Nemade, "A review: Data pre-processing and data augmentation techniques," *Global Transitions Proc.*, vol. 3, no. 1, pp. 91–99, Jun. 2022, doi: 10.1016/j.gltp.2022.04.020.

[76] S. Siami-Namini, N. Tavakoli, and A. S. Namin, "The performance of LSTM and BiLSTM in forecasting time series," in *Proc. IEEE Int. Conf. Big Data (Big Data)*, Dec. 2019, pp. 3285–3292.

[77] N. Tavakoli, "Modeling genome data using bidirectional LSTM," in *Proc. IEEE 43rd Annu. Comput. Softw. Appl. Conf. (COMPSAC)*, vol. 2, Jul. 2019, pp. 183–188.

[78] V. Dogra, S. Verma, P. Chatterjee, J. Shafi, J. Choi, and M. F. Ijaz, "A complete process of text classification system using state-of-the-art NLP models," *Comput. Intell. Neurosci.*, vol. 2022, Jun. 2022, Art. no. 1883698, doi: 10.1155/2022/1883698.

[79] K. Vallmuur, H. R. Marucci-Wellman, J. A. Taylor, M. Lehto, H. L. Corns, and G. S. Smith, ''Harnessing information from injury narratives in the 'big data' era: Understanding and applying machine learning for injury surveillance,'' *Injury Prevention*, vol. 22, pp. i34–i42, Apr. 2016, doi: 10.1136/injuryprev-2015-041813.

[80] S. Amal, L. Safarnejad, J. A. Omiye, I. Ghanzouri, J. H. Cabot, and E. G. Ross, ''Use of multi-modal data and machine learning to improve cardiovascular disease care,'' *Frontiers Cardiovascular Med.*, vol. 9, Apr. 2022, Art. no. 840262, doi: 10.3389/fcvm.2022.840262.

[81] M. Z. F. Khairuddin, K. Hasikin, N. A. A. Razak, K. W. Lai, M. Z. Osman, M. F. Aslan, K. Sabanci, M. M. Azizan, S. C. Satapathy, and X. Wu, ''Predicting occupational injury causal factors using text-based analytics: A systematic review,'' *Frontiers Public Health*, vol. 10, Sep. 2022, Art. no. 984099.

[82] T. Yosef, E. Sineshaw, and N. Shifera, ''Occupational injuries and contributing factors among industry park construction workers in Northwest Ethiopia,'' *Frontiers Public Health*, vol. 10, Jan. 2023, Art. no. 1060755, doi: 10.3389/fpubh.2022.1060755.

[83] S. Winge, E. Albrechtsen, and J. Arnesen, ''A comparative analysis of safety management and safety performance in twelve construction projects,'' *J. Saf. Res.*, vol. 71, pp. 139–152, Dec. 2019.

[84] W. S. Journeay, T. Pauley, M. Kowgier, and M. Devlin, ''Return to work after occupational and non-occupational lower extremity amputation,'' *Occupat. Med.*, vol. 68, no. 7, pp. 438–443, Sep. 2018.

[85] G. A. Shirali, M. V. Noroozi, and A. S. Malehi, ''Predicting the outcome of occupational accidents by CART and CHAID methods at a steel factory in Iran,'' *J. Public Health Res.*, vol. 7, no. 2, p. 1361, Oct. 2018, doi: 10.4081/jphr.2018.1361.

[86] V. Gallego, A. Sánchez, I. Martón, and S. Martorell, ''Analysis of occupational accidents in Spain using shrinkage regression methods,'' *Saf. Sci.*, vol. 133, Jan. 2021, Art. no. 105000, doi: 10.1016/j.ssci.2020.105000.

[87] S. A. Meraikhi and M. Al-Rajab, ''A multimodal approach of machine and deep learnings to enhance the fall of elderly people,'' *J. Inf. Technol. Manage.*, vol. 14, no. 3, pp. 168–184, 2022, doi: 10.22059/jitm.2022.88290.

[88] C. Jujjavarapu, P. Suri, V. Pejaver, J. Friedly, L. S. Gold, E. Meier, T. Cohen, S. D. Mooney, P. J. Heagerty, and J. G. Jarvik, ''Predicting decompression surgery by applying multimodal deep learning to patients' structured and unstructured health data,'' *BMC Med. Informat. Decis. Making*, vol. 23, no. 1, p. 2, Jan. 2023, doi: 10.1186/s12911-022-02096-x.

[89] N. Zaman, D. M. Goldberg, R. J. Gruss, A. S. Abrahams, S. Srisawas, P. Ractham, and M. M. H. Seref, ''Cross-category defect discovery from online reviews: Supplementing sentiment with category-specific semantics,'' *Inf. Syst. Frontiers*, vol. 24, pp. 1265–1285, Aug. 2022.

MOHAMED ZUL FADHLI KHAIRUDDIN received the B.Sc. and M.Eng. degrees in safety, health, and environment from Universiti Malaya, Malaysia, where he is currently pursuing the Ph.D. degree in engineering science. His research interests include occupational safety and health, ergonomics and integrated occupational, and environmental management systems. He is also exploring the utilization of artificial intelligence (AI) approaches, namely machine learning and deep learning, in the occupational safety and health research domain, specifically predictive analytics for occupational accidents and injuries.

KHAIRUNNISA HASIKIN received the B.Eng. and M.Eng. (Science) degrees in biomedical engineering from Universiti Malaya, Malaysia, and the Ph.D. degree from Universiti Sains Malaysia. She is currently a Senior Lecturer with the Department of Biomedical Engineering, Faculty of Engineering, Universiti Malaya. She is also a Professional Engineer (P.Eng.) with the Electrical Engineering Discipline, Board of Engineer Malaysia (BEM). Her research interests include artificial intelligence (AI), machine learning, and image processing in various fields especially in medical and healthcare applications. She is also expanding her expertise in AI for sustainability management projects in intelligent environmental decision support systems.

NASRUL ANUAR ABD RAZAK received the B.Eng. degree in mechatronics from International Islamic University Malaysia, Selangor, Malaysia, and the M.Eng. and Ph.D. degrees from the Department of Biomedical Engineering, Universiti Malaya, Malaysia. Since 2015, he has been a Senior Lecturer with the Department of Biomedical Engineering, Faculty of Engineering, Universiti Malaya, Kuala Lumpur, Malaysia. His research interests include healthcare technology management, assistive device managements, biomechanical engineering (manufacturing in prosthetics and bio sculptor CAD/CAM systems), prosthetics and orthotics engineering, and rehabilitation engineering (upper limb prosthetic, assistive devices and technology in prosthetic and orthotics, and biomechatronic in prosthetic and orthotics).

SITI AFIFAH MOHSHIM received the bachelor's degree in applied science, majoring in medical physics and the M.Sc. degree in medical physics from Universiti Sains Malaysia. She is currently a Senior Lecturer with the Medical Engineering Technology Section, British Malaysian Institute, Universiti Kuala Lumpur, Malaysia. Her research interests include the medical physics applications, medical radiation technology, and mathematical statistics.

SITI SALWA IBRAHIM received the Bachelor of Medicine and Bachelor of Surgery (M.B.B.S.) degrees from Universiti Malaya, Malaysia. She is currently a Medical Officer by profession and also with the Negeri Sembilan State Health Department, Ministry of Health, Malaysia. Her job scope was related to managing daily data monitoring, surveillance, and analysis of public health data.