Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 2, April 2026
DOI : https://doi.org/10.33395/sinkron.v10i2.15928
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Bidirectional Long Short-Term Memory for Early Detection of Running Injuries in Imbalanced Data

David Naufal Zaki[1)*], Defri Kurniawan[2)]
[1)]Faculty of Computer Science, [2)]Universitas Dian Nuswantoro, Semarang, Indonesia
[1)]111202214773@mhs.dinus.ac.id, [2)]defri.kurniawan@dsn.dinus.ac.id

Submitted :Feb 20, 2026 | **Accepted** : Mar 20, 2026 | **Published** : April 2, 2026

Abstract: Running-related injuries are a common sports health issue that can impair athletic performance and potentially terminate an athlete's career. Early injury detection is therefore critical, as injuries are cumulative in nature and influenced by training load patterns over time. Consequently, data-driven predictive approaches based on time-series analysis are required to support athlete monitoring systems with a safety-oriented focus. This study aims to develop an efficient, accurate, and safety-first injury prediction model for running athletes. The study utilizes daily running activity time-series data obtained from Kaggle. The proposed model is based on a Bi-Directional Long Short-Term Memory (Bi-LSTM) architecture to capture bidirectional temporal dependencies, combined with Focal Loss to address extreme class imbalance. In addition, domain-specific feature engineering is applied through the Acute:Chronic Workload Ratio (ACWR). Model performance is evaluated against tabular-data-based models, namely XGBoost and Balanced Bagging, across multiple experimental configurations. Experimental results indicate that the lightweight Bi-LSTM configuration achieves a Recall of 90.7%, outperforming the benchmark models while maintaining a competitive AUC. These findings demonstrate that sequential modeling is more effective in detecting rare injury events. Overall, this study confirms that Bi-LSTM-based sequential modeling is well suited for early detection of running injuries and suggests its potential applicability in athlete monitoring systems that prioritize safety.

Keywords: Acute:Chronic Workload Ratio; Bi-Directional Long Short-Term Memory; Focal Loss; sports injury prediction; time-series data

INTRODUCTION

Long-distance running has become one of the most popular sports activities worldwide due to its accessibility and significant cardiovascular benefits (Lin et al., 2023; Venckunas et al., 2025). However, this growing participation has been accompanied by a high prevalence of musculoskeletal injuries, particularly overuse injuries, which represent a major challenge in sports health (Kakouris et al., 2021). Running-related injuries not only impair athletic performance but may also lead to long-term consequences, including career disruption, financial burden, and psychological distress (Kalkhoven et al., 2021).

The etiology of running injuries is multifactorial, involving complex interactions among biomechanical, anatomical, and training-related factors. Numerous studies have emphasized that errors in training load management and inadequate recovery are dominant contributors to the accumulation of physical stress that exceeds tissue adaptive capacity (Syauqi et al., 2025; van Poppel et al., 2021). This gradual and cumulative nature of injury development necessitates predictive approaches capable of continuously monitoring training load dynamics over time.

In recent years, injury prevention paradigms have shifted from subjective assessments toward data-driven and predictive modeling approaches (Jimenez & Verhagen, 2025). Systematic reviews report a growing application of machine learning techniques for predicting injury risk in athletes (Van Eetvelde et al., 2021). However, a critical limitation remains in how athlete activity data are modeled. Most existing studies rely on static models that treat each training session as an independent observation, thereby ignoring temporal dependencies that are essential for capturing cumulative fatigue and injury risk progression (Wu et al., 2025). As a result, these models often fail to provide reliable early detection of injury events.

*name of corresponding author

In addition, injury prediction is inherently challenged by extreme class imbalance, where injury events represent only a small fraction of the available data. Conventional models tend to be biased toward the majority class, resulting in deceptively high accuracy but poor sensitivity in detecting actual injury cases (Husain et al., 2025; Leckey et al., 2025). Although more advanced deep learning approaches, such as attention-based architectures and graph neural networks, have demonstrated improved performance, these methods often require complex data transformations and high computational resources, limiting their practical applicability in real-world athlete monitoring systems (Shao et al., 2021; Ye et al., 2023)

Based on these limitations, a clear research gap can be identified. Previous studies have shown that static models fail to capture temporal dependencies in injury prediction (Wu et al., 2025), while more advanced approaches such as graph-based models introduce high computational complexity and limited practical applicability (Fang & Chen, 2025; Ye et al., 2023). Furthermore, existing research rarely integrates sequential modeling with explicit class imbalance handling and domain-specific workload features in a unified framework. This indicates that a practical yet effective approach for injury prediction in imbalanced time-series data remains underexplored.

To address this gap, this study proposes an efficient and safety-oriented injury prediction framework based on a Bi-Directional Long Short-Term Memory (Bi-LSTM) architecture. In this study, Bi-LSTM refers to Bidirectional Long Short-Term Memory, a type of recurrent neural network designed to capture temporal dependencies in sequential data by processing information in both forward and backward directions. This architecture has been shown to effectively capture long-term temporal dependencies in sequential data (Alizadegan et al., 2025). The proposed approach integrates three key components: bidirectional sequential modeling to capture temporal dependencies, a combination of SMOTE-Tomek resampling applied during preprocessing and Focal Loss during model training to address extreme class imbalance (Husain et al., 2025), and domain-specific feature engineering using the Acute:Chronic Workload Ratio (ACWR), which is widely recognized as an important indicator of injury risk (van Poppel et al., 2021).

The primary objective of this study is to investigate whether an efficient sequential modeling approach can improve early injury detection performance, particularly in terms of sensitivity (Recall), without introducing excessive computational complexity. This study is intended to contribute to the development of practical athlete monitoring systems and to support decision-making for coaches, sports scientists, and health practitioners. The main contributions of this research are threefold: proposing an efficient Bi-LSTM-based framework for imbalanced time-series injury prediction, demonstrating the effectiveness of combining Focal Loss and ACWR in improving injury detection sensitivity, and providing empirical evidence that lightweight sequential models can achieve competitive performance while maintaining practical feasibility.

LITERATURE REVIEW

Research on running-related injury prediction has evolved alongside the increasing availability of athlete activity data and advances in machine learning techniques. Epidemiological studies consistently establish that running injuries, particularly overuse injuries, are primarily caused by cumulative training load exceeding the adaptive capacity of biological tissues (van Poppel et al., 2021). This understanding has led to the development of data-driven approaches aimed at predicting injury risk based on training patterns.

Early studies predominantly employed statistical methods and conventional machine learning models such as Decision Trees and Random Forests. While these approaches offer interpretability, they generally treat each training session as an independent observation, thereby failing to capture the progressive and cumulative nature of injury development. This limitation has been widely acknowledged in the literature, with systematic reviews indicating that most existing models do not explicitly incorporate temporal dependencies (Pratama et al., 2024; Van Eetvelde et al., 2021).

More recent studies have recognized that athlete activity data are inherently sequential. Time-series-based approaches have demonstrated improved capability in modeling injury risk by capturing temporal dependencies in training load (Wu et al., 2025). However, despite this advancement, a major challenge remains in the form of extreme class imbalance, where injury events constitute only a small fraction of the dataset. Many models struggle to maintain high sensitivity under such conditions, often resulting in poor detection of actual injury cases.

To overcome these limitations, several studies have explored more complex deep learning approaches, including time-series image encoding and Graph Neural Networks. These methods have achieved high discriminative performance, particularly in terms of AUC (Fang & Chen, 2025; Ye et al., 2023). However, their reliance on complex data transformations and high computational cost limits their practical applicability, especially in real-time athlete monitoring systems.

Recurrent Neural Network architectures, particularly Long Short-Term Memory (LSTM), have been widely adopted for sequential data modeling due to their ability to capture long-term dependencies. The Bi-LSTM extends this capability by processing sequences in both forward and backward directions, resulting in more comprehensive contextual understanding (Alizadegan et al., 2025). Despite its effectiveness in other domains, its application in

*name of corresponding author

SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 2, April 2026
DOI : https://doi.org/10.33395/sinkron.v10i2.15928
e-ISSN : 2541-2019
p-ISSN : 2541-044X

sports injury prediction remains limited, particularly in combination with explicit strategies to address class imbalance.

Based on the existing literature, several key observations can be identified. First, it is well established that injury development is a temporal process influenced by cumulative training load. Second, sequential models such as LSTM have demonstrated advantages over static models in capturing temporal patterns. However, it remains unclear how to effectively integrate sequential modeling with robust class imbalance handling and domain-specific features within a practical and computationally efficient framework.

Therefore, a clear research gap exists in the development of an efficient injury prediction model that simultaneously addresses temporal dependencies, extreme class imbalance, and practical implementation constraints. This study aims to fill this gap by proposing a Bi-LSTM-based framework integrated with Focal Loss and domain-specific workload features such as the Acute:Chronic Workload Ratio (ACWR).

Table 1 Comparative Analysis and Research Gap in Running Injury Prediction

Authors	Year	Data Type	Modeling Approach	Strengths	Key Limitations
(Pratama et al., 2024)	2024	Running activity (tabular)	Decision Tree	Simple and interpretable	Treats sessions as independent; low injury recall
(Van Eetvelde et al., 2021)	2021	Multi-study review	ML-based injury prediction	Comprehensive overview of ML usage	Most studies rely on static models; limited temporal modeling
(van Poppel et al., 2021)	2021	Epidemiological running data	Statistical analysis	Identifies workload-related risk factors	Not predictive; no ML/DL implementation
(Wu et al., 2025)	2025	Running activity (time-series)	Time-sequenced ML	Explicit temporal modeling; improved recall	Performance degrades under extreme class imbalance
(Ye et al., 2023)	2023	Sensor-based sports data	Time-series image encoding + DL	High discriminative performance (AUC)	Complex data transformation; high computational cost
(Fang & Chen, 2025)	2025	Multi-sport injury dataset	Temporal Graph Neural Network	Captures complex spatiotemporal interactions	Overly complex; limited practical scalability
(Alizadegan et al., 2025)	2025	Energy consumption time-series	Bi-LSTM	Stable modeling of long-term dependencies	Not applied to sports injury domain
This Study	2026	Running activity (7-year time-series)	Bi-LSTM + Focal Loss + ACWR	High injury recall; efficient temporal modeling; safety-oriented	Lower precision in general population; unimodal data

Based on the comparative analysis summarized in Table 1, a clear research gap can be identified. Existing approaches either rely on static models that fail to capture temporal injury patterns, or adopt highly complex deep learning architectures that limit practical implementation. Furthermore, the integration of sequential modeling with explicit imbalance-aware learning and domain-specific workload features remains underexplored. This study positions itself at this intersection by proposing an efficient Bi-LSTM framework combined with Focal Loss and Acute:Chronic Workload Ratio, aiming to achieve high injury sensitivity while maintaining computational practicality.

METHOD

Research Workflow

To provide a systematic overview of the research procedure, this study is accompanied by a workflow diagram. The diagram summarizes the entire research process, starting from data collection, preprocessing, and feature engineering, followed by sequential data construction, and concluding with model training and performance evaluation for injury prediction.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 2, April 2026
DOI : https://doi.org/10.33395/sinkron.v10i2.15928
e-ISSN : 2541-2019
p-ISSN : 2541-044X

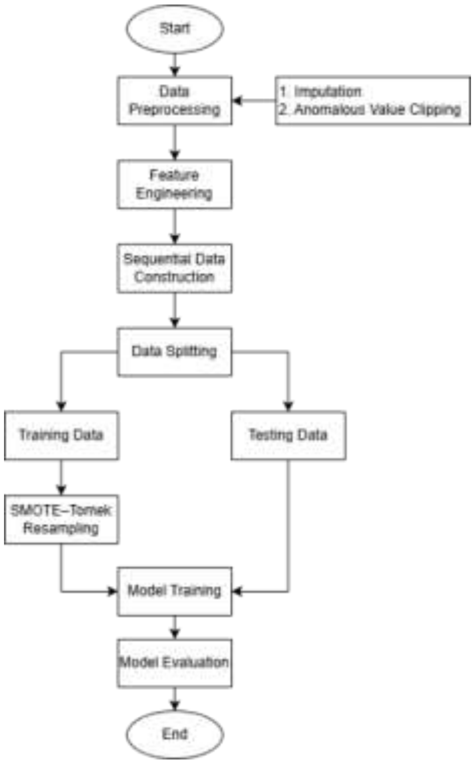

Fig. 1 Research Workflow Diagram

Data Collection

This study utilizes secondary data obtained from a public Kaggle repository (https://www.kaggle.com/datasets/shashwatwork/injury-prediction-for-competitive-runners). The dataset consists of 42,766 daily activity records of long-distance running athletes collected over a seven-year period (2012–2019). Each record represents a single day of training activity for an individual athlete, resulting in hundreds to thousands of observations per athlete depending on the monitoring duration.

The target variable, injury, is binary and indicates the injury status on a given day. Of the total dataset, 583 records (1.36%) correspond to injury events, while 42,183 records (98.64%) represent non-injury conditions. This distribution reflects an extremely imbalanced class ratio, which poses a major challenge for injury prediction modeling and motivates the need for dedicated class imbalance handling strategies in subsequent stages of the analysis. See Table 2.

Table 2 Overview of Features in the Daily Running Activity Dataset

Feature Name	Data Type	Short Description	Injury Relevance
total_km	Numeric	Daily running distance	Training load accumulation
nr_sessions	Numeric	Number of daily sessions	Training frequency
km_z3_z4	Numeric	Moderate-high intensity distance	Submaximal load
km_z5_t1_t2	Numeric	Very high intensity distance	Acute stress
km_sprinting	Numeric	Sprint training distance	Explosive load
strength_training	Binary	Strength training activity	Protective factor
hours_alternative	Numeric	Alternative training duration	Active recovery
perceived_exertion	Numeric	Subjective fatigue (RPE)	Overtraining indicator

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 2, April 2026
DOI : https://doi.org/10.33395/sinkron.v10i2.15928
e-ISSN : 2541-2019
p-ISSN : 2541-044X

perceived_trainingSuccess	Numeric	Training success	Athlete readiness
perceived_recovery	Numeric	Recovery level	Injury risk indicator
ACWR	Numeric	Acute-Chronic Workload Ratio	Training load spike
injury	Binary	Injury status	Target variable

Data Preprocessing, Feature Engineering, and Sequential Data Construction

The data preprocessing stage was conducted to ensure data integrity and consistency prior to model training. Missing values were handled using median-based imputation computed separately for each individual athlete. This approach was adopted to preserve athlete-specific characteristics and to avoid distributional bias that may arise from global imputation. In addition, non-physical negative values in training load variables were addressed using a clipping method, setting such values to zero. This step was implemented to maintain data continuity without removing sequential information that is essential for time-series modeling.

Subsequently, feature engineering was performed based on findings reported by (van Poppel et al., 2021), which emphasize that errors in training load management are a primary predictor of running-related injuries. Guided by this domain knowledge, the study constructed derived features specific to sports science. One of the key features is the Acute:Chronic Workload Ratio (ACWR), which represents the ratio between acute training load (seven-day moving average) and chronic training load (twenty-eight-day moving average).

$$ACWR_t = \frac{\text{Acute Load}_t}{\text{Chronic Load}_t} \quad \textbf{(1)}$$

where $ACWR_t$ represents the Acute:Chronic Workload Ratio at time step t, Acute Load denotes the average training load over the previous seven days, and Chronic Load represents the rolling average training load over the previous twenty-eight days.

In addition to ACWR, a Rolling Volatility feature was calculated as the standard deviation of training load within a seven-day window to quantify day-to-day training inconsistency. High volatility values indicate abrupt fluctuations in training intensity, which have been associated in the literature with an increased risk of injury. All numerical features were subsequently normalized using a Min–Max scaler to the range [0,1] to ensure scale uniformity across variables and to support stable neural network training.

To accommodate the time-series nature of the data and explicitly capture cumulative training load effects, daily records were not treated as independent observations. Instead, the data were transformed into sequential samples using a sliding window approach spanning seven consecutive days. Under this scheme, training load information from day (t−6) to day t was used as input to predict the injury status on day t.

Research Design and Implementation

The selection of the Bi-LSTM architecture in this study is motivated by the sequential nature of athlete activity data, where injury risk develops as a result of cumulative training load over time. Recurrent Neural Network architectures, particularly Long Short-Term Memory (LSTM), have been widely recognized for their ability to capture long-term temporal dependencies in sequential data (Hanafiah et al., 2023; Landi et al., 2021). Furthermore, the bidirectional structure of Bi-LSTM allows the model to incorporate both past and future contextual information within a given time window, resulting in more comprehensive temporal representation (Alizadegan et al., 2025). Alternative approaches such as Convolutional Neural Networks (CNN) are less effective in capturing long-term dependencies, while more advanced architectures such as Transformers generally require larger datasets and higher computational resources, making them less suitable for this study's objective of developing an efficient and practical model.

All data processing and model development in this study were conducted using Python programming language. The implementation utilized several open-source libraries, including Pandas and NumPy for data preprocessing, Scikit-learn for baseline models and evaluation metrics, and TensorFlow/Keras for building and training the Bi-LSTM model. All software tools used in this study are open-source and freely available.

Data analysis in this study was performed through a structured pipeline. First, descriptive statistics were used to summarize the dataset characteristics, including class distribution and feature ranges. Next, data transformation techniques such as normalization and sequence construction were applied to prepare the data for modeling. Model performance was evaluated using quantitative metrics, including Area Under the Curve (AUC) and Recall, to assess both global discriminative ability and sensitivity to injury events. Comparative analysis was then conducted across different model configurations to identify the optimal architecture.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 2, April 2026
DOI : https://doi.org/10.33395/sinkron.v10i2.15928
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Data Splitting and Class Imbalance Handling

The dataset was divided into training (80%) and testing (20%) subsets using a Group Shuffle Split strategy based on athlete identifiers to prevent data leakage across subjects. Prior to data splitting, daily observations were transformed into sequential samples using a seven-day sliding window approach, ensuring that only records with complete historical information were included in the analysis.

As a consequence of sequential sample construction and subject-wise data splitting, the effective number of observations was reduced from the original dataset to approximately 28,000 sequential samples. Although this reduction decreased the sample size, the approach was necessary to preserve temporal consistency and ensure the validity of model evaluation. Given that injury events remained a highly minority class (less than 2% of the data), class imbalance was addressed using a hybrid strategy. SMOTE–Tomek resampling was applied to the training set to balance the data distribution, while Focal Loss was employed during model training to emphasize hard-to-classify minority samples. Previous studies have demonstrated that SMOTE-based resampling can improve classification performance, particularly in scenarios with extremely small minority class proportions (Syukron et al., 2023). Figure 2 illustrates the class distribution before and after the application of the SMOTE–Tomek method.

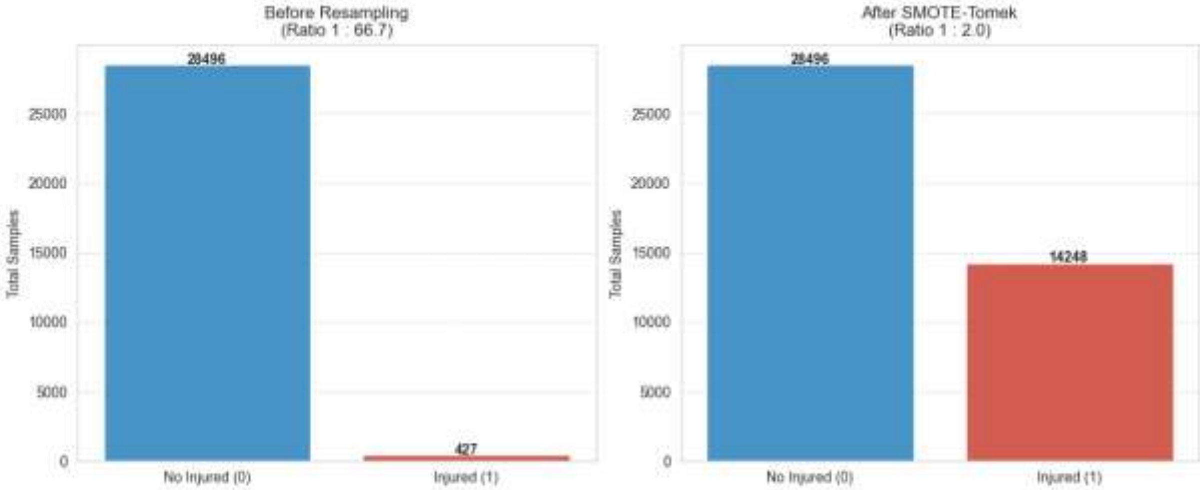

Fig. 2 Class distribution before and after SMOTE–Tomek resampling

Bi-LSTM

Long Short-Term Memory (LSTM) is an extension of the Recurrent Neural Network (RNN) architecture designed to address the vanishing gradient problem commonly encountered in long sequential data (Al-Selwi et al., 2023; Landi et al., 2021). The memory cell mechanism in LSTM enables the model to retain relevant historical information over extended temporal windows, thereby improving its ability to learn long-term dependencies in time-series data (Hanafiah et al., 2023).

Building upon this foundation, Bi-LSTM enhances predictive capability by processing sequential data simultaneously in two temporal directions: forward (from past to future) and backward (from future to past), as proposed by (Alizadegan et al., 2025). This bidirectional processing allows the model to capture a more comprehensive temporal context compared to unidirectional architectures. Mathematically, a Bi-LSTM consists of two separate hidden layers that operate in parallel. For each time step t, the forward hidden state h_t and the backward hidden state $h_t^{\leftarrow}$ are computed as follows:

$$h_t^f = \sigma\left(W_{xh}^f x_t + W_{hh}^f h_{t-1}^f + b_h^f\right) \quad \textbf{(2)}$$

$$h_t^b = \sigma\left(W_{xh}^b x_t + W_{hh}^b h_{t+1}^b + b_h^b\right) \quad \textbf{(3)}$$

where x_t notes the input feature vector at time step t, $h_t^{(f)}$ and $h_t^{(b)}$ represent the forward and backward hidden states, respectively, W_{xh} and W_{hh} denote weight matrices associated with input-to-hidden and hidden-to-hidden connections, b represents bias terms, and σ denotes a nonlinear activation function. The final output y_t at time step t is obtained by concatenating the forward and backward hidden state vectors, followed by a linear transformation:

*name of corresponding author

e-ISSN : 2541-2019
p-ISSN : 2541-044X

$$y_t = W_{hy}^f h_t^f + W_{hy}^b h_t^b + b_y \quad \textbf{(4)}$$

This mechanism ensures that predictions at a given time step are informed by both historical information from past observations and contextual information from future observations within the specified temporal window (window size).

Model Architecture: Bi-LSTM with Focal Loss and ACWR

Based on the theoretical foundations discussed earlier, this study implements a deep learning architecture specifically designed to handle time-series data under extreme class imbalance conditions. The model architecture begins with an input layer that receives data in the form of a three-dimensional tensor (N, T, F), where N denotes the number of samples, T represents the temporal window size of seven days, and F corresponds to the number of input features, including the ACWR variable. The input data are subsequently processed through two stacked Bi-LSTM layers. This design enables the model to extract high-level temporal features by scanning training load sequences simultaneously in both temporal directions (past to future and future to past), thereby capturing a more comprehensive chronological context of injury risk compared to unidirectional models.

The use of stacked Bi-LSTM layers is consistent with previous findings reported by (Mahadevaswamy & Swathi, 2022; Pavlatos et al., 2023), which demonstrate that multi-layer architectures are more effective in extracting higher-level sequential feature representations than single-layer models, particularly for time-series data with complex temporal dependencies.

To ensure training stability and mitigate overfitting, each Bi-LSTM block is followed by regularization layers consisting of Batch Normalization and AlphaDropout. The extracted features are then passed to a fully connected (Dense) layer employing the Scaled Exponential Linear Unit (SELU) activation function (Ko et al., 2021). SELU is selected for its self-normalizing properties, which help maintain stable mean and variance of activations throughout the network. At the final stage, an output layer with a single neuron and a Sigmoid activation function produces injury probability estimates within the range of 0 to 1 (Zul et al., 2024)

As a critical component for addressing class imbalance, the model is trained using the Focal Loss function. Unlike standard cross-entropy loss, which treats all misclassifications equally, Focal Loss applies a modulation factor that down-weights easy-to-classify samples from the majority (non-injury) class while assigning substantially higher penalties to misclassified samples from the minority (injury) class. This mechanism forces the model to focus learning on hard-to-detect injury patterns.

To examine the impact of architectural complexity on model performance, particularly with respect to safety-oriented metrics such as Recall, this study evaluates multiple Bi-LSTM configurations that vary in terms of the number of hidden layers, LSTM units, training epochs, and batch sizes. All model configurations are trained using the Adadelta optimizer in conjunction with Focal Loss to ensure consistent optimization settings across experiments. This design choice ensures that observed performance differences arise solely from architectural and training configuration variations rather than differences in optimization strategies or loss functions.

The detailed specifications of the evaluated Bi-LSTM configurations are summarized in Table 3, covering lightweight, deep, wide, and extended training architectures.

Table 3 Bi-LSTM Model Configuration

Configuration	**Layers**	**Units**	**Window Size**	**Epoch**	**Batch Size**
Baseline	2	64	7	100	256
Lightweight	1	32	7	80	128
Deep	3	64	7	120	256
Wide	2	128	7	100	512
Long Training	2	64	7	200	256

Experimental Evaluation Design

The evaluation framework in this study is designed to assess the effects of architectural configuration variations and training strategies on injury prediction performance. Unlike multi-scenario evaluation approaches that rely on varying data distributions, this study adopts a configuration-based evaluation strategy, in which multiple model

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 2, April 2026
DOI : https://doi.org/10.33395/sinkron.v10i2.15928
e-ISSN : 2541-2019
p-ISSN : 2541-044X

architectures are systematically compared using the same data splitting scheme. All experiments are conducted with a single, consistent training and testing split, as described in Section 2.4, to ensure that observed performance differences are solely attributable to variations in model configurations rather than differences in data distribution.

The evaluated configurations vary in terms of the number of Bi-LSTM layers, the number of LSTM units, batch size, and training duration (number of epochs), as summarized in Table 3. This experimental design enables a controlled analysis of how architectural complexity and training parameters influence model performance.

Each configuration is evaluated using two primary performance metrics: Area Under the Curve (AUC), which reflects global discriminative capability, and Recall (Sensitivity), which serves as the primary safety-oriented metric for detecting injury events. Recall is prioritized because it directly represents the model's ability to minimize false negatives, which pose a higher risk than false positives in medical and sports injury prevention contexts. This evaluation strategy facilitates an explicit analysis of the trade-offs between model complexity, global discrimination performance, and sensitivity to injury events, thereby providing a principled basis for selecting the optimal Bi-LSTM configuration.

RESULT

Bi-LSTM Configuration Experiments

To ensure that model performance was not dependent on a single architectural configuration, this study conducted a series of experiments using multiple Bi-LSTM configurations. The evaluated variations included the number of hidden layers, the number of LSTM units, and different training strategies. Model performance was assessed using two primary metrics: Area Under the Curve (AUC) as an indicator of global discriminative capability and Recall (Sensitivity) as a safety-oriented metric for injury detection.

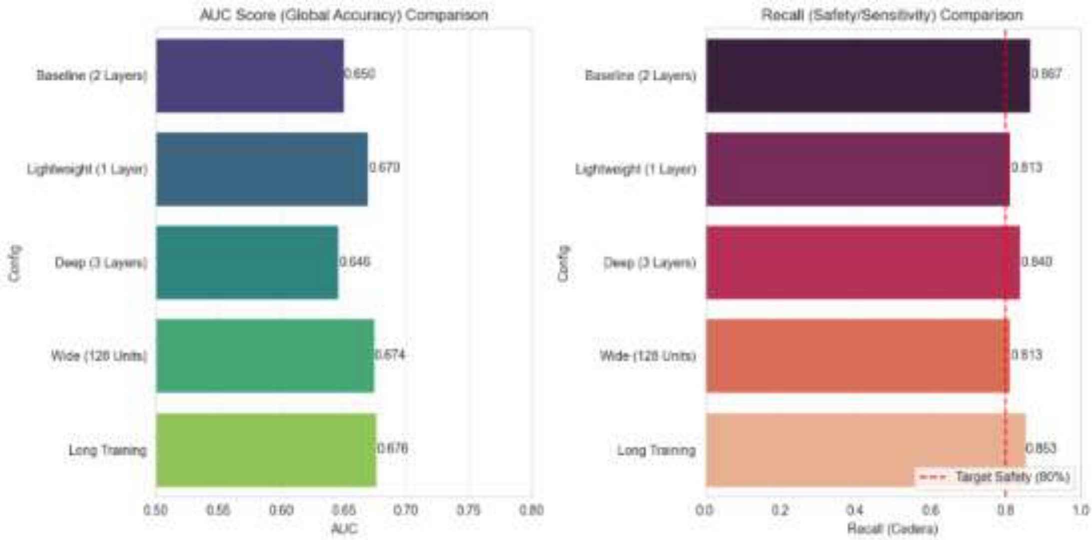

Fig. 3 Evaluation results of different Bi-LSTM configurations

Figure 3 presents the evaluation results across different Bi-LSTM configurations. The Deep configuration achieved the highest AUC value (0.693). In contrast, the Lightweight configuration achieved the highest Recall value (90.7%), despite having a slightly lower AUC compared to the Deep model.

Selection of the Optimal Bi-LSTM Configuration

Based on this criterion, the Lightweight Bi-LSTM configuration was selected as the proposed model. This configuration exceeded the minimum safety threshold of 80% Recall by achieving a Recall score of 90.7%, while still maintaining a competitive AUC value.

Evaluation of Baseline Models (Balanced Bagging and XGBoost)

For comparative purposes, this study evaluated two commonly used non-sequential models for tabular data: the Balanced Bagging and XGBoost. Multiple configurations were tested for each model to ensure a fair comparison.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 2, April 2026 e-ISSN : 2541-2019
DOI : https://doi.org/10.33395/sinkron.v10i2.15928 p-ISSN : 2541-044X

Table 4 Performance Evaluation of the Balanced Bagging Model

Configuration	n_estimators	max_depth	AUC	Recall
Baseline	50	5	0.71	0.55
Deeper Trees	50	10	0.72	0.52
Higher voting	200	5	0.74	0.57

Table 4 summarizes the performance of the Balanced Bagging models. Increasing the number of estimators through massive voting improved the AUC; however, Recall values remained well below the minimum safety threshold of 80%. The configuration labels such as "Baseline", "Deeper Trees", and "Higher Voting" are used solely as naming conventions to distinguish different parameter settings of the Balanced Bagging configuration, and do not represent distinct algorithms.

Table 5 Performance Evaluation of the XGBoost Model

Configuration	n_estimators	Learning Rate	max_depth	AUC	Recall
Baseline	100	0.1	5	0.66	0.35
Conservative	300	0.01	6	0.68	0.39
High Recall	100	0.1	5	0.67	0.62

Table 5 presents the evaluation results for the XGBoost models. The High Recall configuration achieved improved sensitivity compared to other XGBoost variants but still failed to meet the minimum safety requirement. Similarly, the configuration labels "Baseline", "Conservative", and "High Recall" are used only as naming conventions to differentiate parameter settings of the XGBoost model, rather than indicating fundamentally different model types.

Comparison of Best Models Across Approaches

Table 6 compares the best-performing models from each approach. The results demonstrate that the Bi-LSTM model significantly outperforms static ensemble models in terms of Recall, despite having comparable AUC values.

Table 6 Comparison of Evaluation Metrics Across Models

Model	Configuration	Approach	AUC	Recall (Injured)
Balanced Bagging	Higher Voting	Ensemble Static	0.74	0.570
XGBoost	High Recall	Boosting	0.67	0.620
Bi-LSTM	Lightweight	Deep Learning Sequential	0.67	0.907

DISCUSSIONS

Overall, the results indicate that the proposed Bi-LSTM model significantly improves injury detection sensitivity compared to conventional approaches, particularly in highly imbalanced time-series data.

This study aims to develop an efficient and safety-oriented framework for early detection of running-related injuries using sequential modeling. The experimental results demonstrate that the proposed Bi-LSTM-based approach achieves superior performance in terms of Recall compared to static ensemble models such as XGBoost and Balanced Bagging. The Lightweight Bi-LSTM configuration achieves the highest Recall of 90.7%, indicating its strong capability in detecting injury events, which are inherently rare and difficult to identify.

The results confirm that incorporating temporal dependencies plays a critical role in injury prediction. Unlike static models that treat each observation independently, the Bi-LSTM model captures cumulative training load patterns over time, enabling more accurate identification of injury risk. This finding is consistent with previous studies (Wu et al., 2025), which emphasize that injury development is a temporal process rather than an isolated event.

Interestingly, increasing model complexity does not necessarily lead to improved performance. While deeper configurations achieved higher AUC values, they did not consistently improve Recall. In contrast, the Lightweight configuration demonstrated the best sensitivity performance. This suggests that simpler architectures may generalize better in highly imbalanced time-series data, particularly when the primary objective is to detect

*name of corresponding author

SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 2, April 2026
DOI : https://doi.org/10.33395/sinkron.v10i2.15928
e-ISSN : 2541-2019
p-ISSN : 2541-044X

minority-class events. A possible explanation is that deeper models may overfit to dominant patterns in the majority class, thereby reducing their ability to detect rare injury cases.

When compared with previous studies, the proposed approach provides a balance between performance and practicality. While more complex models such as Graph Neural Networks have reported higher AUC values (Fang & Chen, 2025; Ye et al., 2023), they require sophisticated data transformations and higher computational resources. In contrast, this study demonstrates that competitive performance can be achieved using a simpler sequential model combined with domain-specific features, making it more suitable for real-world deployment in athlete monitoring systems.

Another important finding is the contribution of domain-informed feature engineering, particularly the Acute:Chronic Workload Ratio (ACWR). The inclusion of ACWR, combined with Focal Loss, enhances the model's ability to focus on hard-to-classify injury cases by emphasizing sudden spikes in training load. This result supports prior findings (van Poppel et al., 2021) that highlight the importance of workload management in injury prevention.

The strengths of this study lie in its ability to achieve high sensitivity using a relatively simple and computationally efficient model, making it suitable for real-time and large-scale deployment. However, the model also exhibits several limitations. First, the model demonstrates relatively low precision, indicating a high number of false positives. This suggests that daily activity data alone may not be sufficient to clearly distinguish between normal fatigue and injury-prone conditions. An alternative explanation is that the available features may contain overlapping patterns between injured and non-injured states, making precise classification challenging. Second, the study relies on a single dataset with unimodal features, which may limit the generalizability of the findings to other sports or populations.

Despite these limitations, the proposed model offers significant practical implications. In safety-critical applications such as injury prevention, high Recall is more important than overall accuracy, as missing an injury case (false negative) can have serious consequences. Therefore, the model is well suited as an early screening tool in athlete monitoring systems, where it can provide initial risk alerts for further evaluation by coaches or medical professionals.

In a broader context, these findings highlight the importance of adopting time-series modeling approaches in sports injury prediction globally, particularly in the development of scalable and real-time athlete monitoring systems. This study contributes to the advancement of data-driven sports analytics by demonstrating that efficient sequential modeling can provide a practical alternative to complex deep learning architectures. Future research should focus on integrating multimodal data sources, such as physiological signals and biomechanical measurements, to improve model precision. Additionally, exploring hybrid approaches that combine data-driven models with expert knowledge may further enhance the reliability and interpretability of injury prediction systems.

CONCLUSION

This study proposes a sequential modeling framework based on Bi-LSTM for early detection of running-related injuries in highly imbalanced time-series data. The findings indicate that incorporating temporal dependencies can substantially improve injury detection sensitivity compared to conventional static models. In particular, the proposed approach achieves a high Recall performance, suggesting its effectiveness in identifying injury events that are typically rare and difficult to detect.

The integration of domain-specific feature engineering, particularly the Acute:Chronic Workload Ratio (ACWR), together with Focal Loss, contributes to mitigating class imbalance and enhancing the model's ability to focus on minority-class patterns. These results highlight the potential of combining sequential modeling and imbalance-aware learning in developing safety-oriented injury prediction systems.

However, the findings of this study should be interpreted with caution. The model exhibits limitations in terms of precision, indicating a relatively high number of false positives. In addition, the use of a single dataset with unimodal activity features may limit the generalizability of the results across different sports contexts or populations.

Despite these limitations, the proposed framework demonstrates practical potential as an early screening tool in athlete monitoring systems, where high sensitivity is prioritized to minimize missed injury cases. Future research is recommended to incorporate multimodal data sources, such as physiological and biomechanical measurements, and to explore hybrid approaches that integrate data-driven models with expert knowledge.

Overall, this study contributes to the growing body of research in sports injury prediction by demonstrating that efficient and relatively simple sequential models can provide a practical and scalable alternative for real-world applications, particularly in safety-critical contexts. These findings provide practical insights for the development of data-driven, safety-oriented athlete monitoring systems.

*name of corresponding author

REFERENCES

Alizadegan, H., Rashidi Malki, B., Radmehr, A., Karimi, H., & Ilani, M. A. (2025). Comparative study of long short-term memory (LSTM), bidirectional LSTM, and traditional machine learning approaches for energy consumption prediction. *Energy Exploration and Exploitation*, *43*(1), 281–301. https://doi.org/10.1177/01445987241269496

Al-Selwi, S. M., Hassan, M. F., Abdulkadir, S. J., & Muneer, A. (2023). LSTM Inefficiency in Long-Term Dependencies Regression Problems. *Journal of Advanced Research in Applied Sciences and Engineering Technology*, *30*(3), 16–31. https://doi.org/10.37934/araset.30.3.1631

Fang, D., & Chen, C. (2025). Sports injury risk prediction based on temporal graph encoding and graph neural networks: A cross-sport transfer learning framework. *Scientific Reports*, *15*(1). https://doi.org/10.1038/s41598-025-21613-2

Hanafiah, A., Arta, Y., Nasution, H. O., & Lestari, Y. D. (2023). Penerapan Metode Recurrent Neural Network dengan Pendekatan Long Short-Term Memory (LSTM) Untuk Prediksi Harga Saham. *Bulletin of Computer Science Research*, *4*(1), 27–33. https://doi.org/10.47065/bulletincsr.v4i1.321

Husain, G., Nasef, D., Jose, R., Mayer, J., Bekbolatova, M., Devine, T., & Toma, M. (2025). SMOTE vs. SMOTEENN: A Study on the Performance of Resampling Algorithms for Addressing Class Imbalance in Regression Models. *Algorithms*, *18*(1). https://doi.org/10.3390/a18010037

Jimenez, C., & Verhagen, E. (2025). Reimagining athlete monitoring for true indicative injury prevention. In *BMJ Open Sport and Exercise Medicine* (Vol. 11, Number 2). BMJ Publishing Group. https://doi.org/10.1136/bmjsem-2025-002479

Kakouris, N., Yener, N., & Fong, D. T. P. (2021). A systematic review of running-related musculoskeletal injuries in runners. In *Journal of Sport and Health Science* (Vol. 10, Number 5, pp. 513–522). Elsevier B.V. https://doi.org/10.1016/j.jshs.2021.04.001

Kalkhoven, J. T., Watsford, M. L., Coutts, A. J., Edwards, W. B., & Impellizzeri, F. M. (2021). Training Load and Injury: Causal Pathways and Future Directions. In *Sports Medicine* (Vol. 51, Number 6, pp. 1137–1150). Springer Science and Business Media Deutschland GmbH. https://doi.org/10.1007/s40279-020-01413-6

Ko, M. S., Lee, K., Kim, J. K., Hong, C. W., Dong, Z. Y., & Hur, K. (2021). Deep Concatenated Residual Network with Bidirectional LSTM for One-Hour-Ahead Wind Power Forecasting. *IEEE Transactions on Sustainable Energy*, *12*(2), 1321–1335. https://doi.org/10.1109/TSTE.2020.3043884

Landi, F., Baraldi, L., Cornia, M., & Cucchiara, R. (2021). *Working Memory Connections for LSTM*. https://doi.org/10.1016/j.neunet.2021.08.030

Leckey, C., Van Dyk, N., Doherty, C., Lawlor, A., & Delahunt, E. (2025). Machine learning approaches to injury risk prediction in sport: A scoping review with evidence synthesis. In *British Journal of Sports Medicine* (Vol. 59, Number 7, pp. 491–500). BMJ Publishing Group. https://doi.org/10.1136/bjsports-2024-108576

Lin, Y. J., Lee, C. C., Huang, T. W., Hsu, W. C., Wu, L. W., Lin, C. C., & Hsiu, H. (2023). Using Arterial Pulse and Laser Doppler Analyses to Discriminate between the Cardiovascular Effects of Different Running Levels. *Sensors*, *23*(8). https://doi.org/10.3390/s23083855

Mahadevaswamy, U. B., & Swathi, P. (2022). Sentiment Analysis using Bidirectional LSTM Network. *Procedia Computer Science*, *218*, 45–56. https://doi.org/10.1016/j.procs.2022.12.400

Pavlatos, C., Makris, E., Fotis, G., Vita, V., & Mladenov, V. (2023). Enhancing Electrical Load Prediction Using a Bidirectional LSTM Neural Network. *Electronics (Switzerland)*, *12*(22). https://doi.org/10.3390/electronics12224652

Pratama, A. Y., Tiara, S. D., Daniarti, E., Informasi, S., Nusantara, U., & Kediri, P. (2024). Analisis Cedera Atlet Lari Menggunakan Metode Decision Tree Berdasarkan Data Aktivitas Latihan. In *INOTEK* (Vol. 9).

Shao, Y., Li, R. D., Luo, Y. J., & Zhu, M. (2021). Research on Running Data Analysis Method Based on Attention-LSTM. *Proceedings - 2021 International Conference on Intelligent Transportation, Big Data and Smart City, ICITBS 2021*, 446–450. https://doi.org/10.1109/ICITBS53129.2021.00116

Syauqi, M. A., Syauqy, D., & Kurniawan, W. (2025). *Sistem Wearable Deteksi Postur pada Training Lateral Raise menggunakan Sensor MPU6050 dengan Algoritma Random Forest* (Vol. 1, Number 1). http://j-ptiik.ub.ac.id

Syukron, A., Saputro, E., & Widodo, P. (2023). Penerapan Metode Smote Untuk Mengatasi Ketidakseimbangan Kelas Pada Prediksi Gagal Jantung. In *Jurnal Teknologi Informasi dan Terapan (J-TIT* (Vol. 10, Number 1). https://doi.org/10/25047/jtit.v10i1.312

Van Eetvelde, H., Mendonça, L. D., Ley, C., Seil, R., & Tischer, T. (2021). Machine learning methods in sport injury prediction and prevention: a systematic review. In *Journal of Experimental Orthopaedics* (Vol. 8, Number 1). Springer Science and Business Media Deutschland GmbH. https://doi.org/10.1186/s40634-021-00346-x

van Poppel, D., van der Worp, M., Slabbekoorn, A., van den Heuvel, S. S. P., van Middelkoop, M., Koes, B. W., Verhagen, A. P., & Scholten-Peeters, G. G. M. (2021). Risk factors for overuse injuries in short- and long-

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 2, April 2026
DOI : https://doi.org/10.33395/sinkron.v10i2.15928
e-ISSN : 2541-2019
p-ISSN : 2541-044X

distance running: A systematic review. *Journal of Sport and Health Science*, *10*(1), 14–28. https://doi.org/10.1016/j.jshs.2020.06.006

Venckunas, T., Gumauskiene, B., Muanjai, P., Cadefau, J. A., & Kamandulis, S. (2025). High-Intensity Interval Training Improves Cardiovascular Fitness and Induces Left-Ventricular Hypertrophy During Off-Season. *Journal of Functional Morphology and Kinesiology*, *10*(3). https://doi.org/10.3390/jfmk10030271

Wu, H., Brooke-Wavell, K., Barnes, M. R., Awan, Z., Mastana, S., Allen, S., & Blagrove, R. C. (2025). *A time-sequenced approach to machine learning prognostic modelling with implementation on running-related injury prediction*. https://doi.org/10.1101/2025.05.07.25327162

Ye, X., Huang, Y., Bai, Z., & Wang, Y. (2023). A novel approach for sports injury risk prediction: based on time-series image encoding and deep learning. *Frontiers in Physiology*, *14*. https://doi.org/10.3389/fphys.2023.1174525

Zul, M., Khairuddin, F., Sankaranarayanan, S., Hasikin, K., Anuar, N., Razak, A., & Omar, R. (2024). *Contextualizing injury severity from occupational accident reports using an optimized deep learning prediction model*. 20. https://doi.org/10.7717/peerj

*name of corresponding author