Journal of Information Systems Engineering and Management
2025, 10(60s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Multi-Agent Reinforcement Learning for Sports Injury Prediction: A Comparative Evaluation with Deep Learning Baselines

Abhishek Vangipuram

Independent Researcher, CA, USA, ORCID: 0009-0003-2529-2740

abhiv22360@gmail.com

ARTICLE INFO

Received: 02 Aug 2025

Revised: 18 Sept 2025

Accepted: 28 Sept 2025

ABSTRACT

Sports injuries pose a pervasive challenge affecting athletes at every level of competition, from grassroots amateurs to elite professionals. The financial, physiological, and psychological costs are enormous. Teams lose key players at critical moments, athletes endure career-altering damage, and medical staff are often left struggling to act on incomplete information. Yet despite the growing availability of biomechanical data, GPS tracking, and electronic health records, the field still lacks a unified, adaptive framework capable of modeling the complex, interdependent nature of injury risk across an entire squad. This paper presents a systematic comparative evaluation of four machine learning approaches for sports injury prediction: Random Forest, Long Short-Term Memory networks (LSTM), Single-Agent Deep Deterministic Policy Gradient (DDPG), and a proposed Multi-Agent Reinforcement Learning (MARL) framework based on MADDPG with temporal attention. The MARL framework models each athlete as a cooperative agent in a shared physiological environment, with the hypothesis that capturing inter-player dependencies would improve prediction over single-player approaches. Each model is trained and evaluated under identical conditions on a real-world multimodal athlete-monitoring dataset. Experimental evaluations are conducted on a real multimodal athlete-monitoring dataset containing physiological signals, including heart rate variability, EMG amplitude, skin temperature, blood oxygen saturation, and workload metrics, from 5,430 athlete observations. Four models are evaluated: Random Forest, LSTM, Single-Agent DDPG, and the proposed MARL framework. Results show that Random Forest achieves the strongest discriminative performance with an AUC-ROC of 0.996, while the LSTM baseline reaches 0.932. The MARL framework faces convergence challenges on this single-player tabular dataset, revealing an important empirical finding: cooperative multi-agent dynamics are difficult to leverage without explicit inter-player sequential interaction structure in the data. This paper makes an honest empirical contribution: it establishes a rigorous benchmark for injury prediction on real physiological monitoring data, demonstrates that classical ensemble methods remain highly competitive baselines, and identifies the precise data conditions under which MARL frameworks are and are not effective. These findings provide a clear roadmap for future work.

Keywords: Multi-Agent Reinforcement Learning, Sports Injury Prediction, Comparative Evaluation, MADDPG, Random Forest, LSTM, Negative Results, Empirical Benchmarking, Physiological Monitoring, Athlete Health Analytics

1. Introduction

There is an old saying in professional sports: the best ability is availability. No matter how talented a player might be, an injury renders that talent irrelevant. The economic reality is equally stark. In European football alone, researchers have estimated that first-division clubs collectively lose hundreds of millions of euros per season to player unavailability caused by injury. This figure does not include the

Journal of Information Systems Engineering and Management
2025, 10(60s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ Research Article

intangible costs of disrupted tactical plans, reduced fan engagement, and long-term deterioration of player careers.

Yet despite these stakes, the way most teams manage injury risk today remains surprisingly reactive. A player trains hard, complains of tightness, sees a physiotherapist, and is either cleared or rested based on relatively coarse clinical judgment. More progressive clubs have adopted GPS vests and heart rate monitors. Still, the data from these devices often ends up in fragmented spreadsheets that are rarely integrated into a coherent predictive model. The promise of data-driven injury prevention has been discussed for at least two decades, but the operational reality for most clubs still lags far behind.

The fundamental challenge is not a lack of data, since modern training sessions generate an overwhelming volume of physiological and movement signals. The real problem is a lack of models that can make sense of that data in the right way. Injury risk is not a static property of an individual athlete. It fluctuates hourly in response to cumulative fatigue, sleep quality, training intensity, emotional stress, weather conditions, and the physical demands imposed by an athlete's tactical role within the team. Furthermore, the risk profile of one player is not independent of others: a midfielder playing deeper to compensate for an injured striker accumulates mechanical stresses different from their baseline, and a disrupted defensive pairing must work harder to cover for one another.

These interdependencies suggest that injury prediction may benefit from a multi-agent framing. When teammates' behavior shapes an athlete's risk, shared training loads, and collective recovery patterns, a model that captures these cooperative dynamics could, in principle, outperform one that treats each player in isolation. This hypothesis motivates the central research question of this paper: Does a MARL-based approach actually improve injury prediction over strong supervised and single-agent baselines on real physiological data?

Multi-Agent Reinforcement Learning (MARL) offers a theoretically well-motivated framework for this setting. In MARL, multiple autonomous agents interact within a shared environment, learning policies that account for both individual and collective outcomes. The approach has proven effective in domains characterized by explicit inter-agent interaction, such as autonomous vehicle coordination and robotic swarms. Whether a sports team constitutes a sufficiently interaction-rich environment for MARL to outperform simpler baselines is an empirical question, and answering it honestly is the central purpose of this paper.

This paper makes the following contributions:

- A rigorous comparative evaluation of four model families, Random Forest, LSTM, Single-Agent DDPG, and MARL, on a real multimodal athlete monitoring dataset under identical experimental conditions.
- A reproducible MADDPG architecture with temporal attention and a cooperative reward function, released as a complete open experimental pipeline that future researchers can build upon.
- An empirical analysis of the conditions under which RL-based approaches fail to converge on tabular injury prediction tasks, including class imbalance, insufficient episode count, and absence of inter-player interaction structure in the data.
- A clear research agenda identifying the dataset characteristics and experimental conditions needed for MARL to demonstrate its theoretical advantages over supervised baselines in sports medicine applications.

The remainder of this paper is structured as follows. Section 2 surveys related work across injury prediction, MARL, and sports analytics. Section 3 defines the methodology and formal problem setup. Section 4 presents the complete system architecture and framework design. Section 5 reports experimental results and analysis. Section 6 discusses limitations and future directions. Section 7 concludes the paper.

Journal of Information Systems Engineering and Management
2025, 10(60s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

2. Literature Review

2.1 Traditional Approaches to Injury Prediction

Early efforts in sports injury prediction relied heavily on simple statistical models applied to routine clinical assessments. Questionnaires like the Acute:Chronic Workload Ratio (ACWR) became widely adopted in the early 2010s after research suggested that players whose recent training loads significantly exceeded their chronic baseline were at substantially elevated risk of non-contact injury. While the ACWR provided practitioners with an easily interpretable metric, subsequent meta-analyses revealed significant methodological inconsistencies in how the ratio was calculated and applied, and its predictive validity remained modest in controlled studies.

Biomechanical approaches offered a different angle, using force plates, electromyography, and high-speed video to identify movement dysfunction patterns associated with specific injury types, with particular focus on anterior cruciate ligament (ACL) tears and hamstring strains. These methods yielded valuable mechanistic insights but required expensive laboratory infrastructure and could not be deployed at scale during live training environments.

2.2 Machine Learning and Deep Learning in Sports Medicine

The application of machine learning to injury prediction accelerated significantly in the mid-2010s as wearable sensor data became more affordable and standardized. Logistic regression, decision trees, and Random Forests were applied to GPS-derived variables, including total distance, high-speed running distance, and accelerometer load. These models generally achieved AUC values in the range of 0.65 to 0.78 on prospective validation sets, representing a meaningful but clinically insufficient improvement over chance.

Recurrent neural networks and Long Short-Term Memory (LSTM) architectures emerged as natural fits for the sequential, time-series nature of training-load data. Studies have demonstrated that LSTMs can capture temporal dependencies in workload patterns that simpler models miss. For example, they identified the specific temporal signatures preceding hamstring injuries more accurately than ACWR-based thresholds. Transformer-based architectures with self-attention mechanisms have recently been applied to multi-modal athlete-monitoring data, achieving state-of-the-art performance on several benchmark injury prediction tasks.

However, a critical limitation pervades virtually all of this literature: models are trained and evaluated on individual athletes treated as independent observations. Even studies that use squad-level data typically aggregate team statistics as features rather than explicitly modeling inter-player relationships. The biological reality that athlete load, fatigue, and injury risk emerge from a shared training environment where one player's output affects everyone else's is absent from these formulations.

2.3 Reinforcement Learning in Sports and Healthcare

Reinforcement learning has been applied productively in sports primarily for strategic and tactical decision-making tasks. Deep RL agents have learned to play basketball, soccer, and other sports at superhuman levels in simulated environments. In real-world sports analytics, RL has been used to optimize substitution timing, play-calling strategy, and training periodization.

In healthcare, RL has found applications in dynamic treatment regimes, where it learns policies that recommend personalized treatment decisions over time based on evolving patient states. The analogy to sports injury prevention is direct: an athlete's physiological state evolves, interventions (load modifications, rest, targeted conditioning) affect that state, and the objective is to minimize the probability of an adverse outcome (injury) while maximizing athletic output.

2.4 Multi-Agent Reinforcement Learning

MARL extends single-agent RL to settings with multiple interacting agents. The field has been shaped by foundational algorithmic contributions, including Independent Q-Learning (IQL), which treats each agent independently; Counterfactual Multi-Agent Policy Gradients (COMA), which uses a centralized critic for training while allowing decentralized execution; and Multi-Agent Deep Deterministic Policy Gradient (MADDPG), which extends DDPG to multi-agent settings with a centralized training, decentralized execution paradigm.

The centralized training, decentralized execution (CTDE) paradigm has proven particularly powerful for cooperative multi-agent tasks. During training, each agent's critic has access to the global state and all agents' actions, enabling stable value estimation despite the non-stationarity inherent in environments where multiple agents are simultaneously learning. At test time, each agent makes decisions based only on its local observations, making the approach scalable to real-world deployments.

To the best of our knowledge, no prior work has applied MARL to sports injury prediction in the manner proposed here. The closest related work comprises a small number of studies that used cooperative RL to optimize training load in athlete groups. Still, these did not frame injury prediction as the primary outcome and did not incorporate the temporal attention mechanisms or medical reward shaping introduced in this paper.

3. Methodology

3.1 Problem Formulation

We frame the sports injury prediction problem as a cooperative multi-agent task where each athlete on a squad is treated as an independent agent operating in a shared environment. On each training day, each agent observes its own physiological and workload states and takes an action that represents a recommended adjustment to the planned training load. The environment then updates to reflect the influence of those adjustments on the squad's collective fatigue and risk levels. Agents are trained cooperatively, meaning their goal is not only to protect their own health but also to minimize injury across the entire team.

3.2 State Space Design

The state vector for agent i at time t comprises five categories of features, concatenated into a single vector of dimension 47:

- Workload History (8 features): Total distance, high-speed running distance, acceleration count, deceleration count, sprint count, and their exponentially weighted moving averages over 7-day and 28-day windows.
- Physiological Signals (12 features): Resting heart rate variability (HRV), morning wellness score (composite of sleep quality, fatigue, mood, and muscle soreness), body mass fluctuation, and blood lactate proxy estimates from GPS-derived metabolic power.
- Biomechanical Features (10 features): Asymmetry indices for hip flexion, knee loading patterns, hamstring-to-quadriceps torque ratios, and contact time during high-speed running derived from inertial measurement units.
- Match/Training Context (9 features): Days since last match, upcoming fixture density, training session type (speed, strength, tactical, recovery), competition level, pitch surface, and temperature/humidity indices.
- Team-Level Features (8 features): Squad-average workload, number of currently injured teammates, positional group fatigue index, and recent pressing intensity metrics from tactical analysis.

Journal of Information Systems Engineering and Management
2025, 10(60s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

3.3 Reward Function Design

The reward function captures the core clinical objective: reduce injury risk while preserving training quality. Each agent is rewarded positively when it maintains a manageable workload and negatively when an injury event occurs or when training load exceeds a player-specific safety threshold calibrated to their chronic baseline. A small cooperative bonus is added when the team's overall risk profile stays within a predefined safe zone, encouraging agents to consider squad-level health rather than just their own. Injury prevention is weighted much more heavily than performance maximization, which reflects how real sports medicine decisions are made in practice.

3.4 MADDPG with Temporal Attention

Our core learning algorithm extends MADDPG with two significant modifications. First, each agent's actor and critic networks incorporate an LSTM layer followed by a multi-head self-attention module, enabling the model to identify the most injury-predictive temporal patterns across sequences of up to 28 training days. Second, the critic network takes as input not just the current global state and all agents' actions, but a learned graph embedding of inter-player relationships, computed by a Graph Attention Network (GAT) layer that encodes positional proximity, physical contact frequency, and shared training history.

The actor network for each agent takes the player's local observation, passes it through an LSTM layer for temporal encoding, then applies a multi-head attention module to highlight the most predictive time steps, and finally passes it through two fully connected layers to produce the final action. The critic network receives the global state along with all agents' actions, encodes inter-player relationships using a Graph Attention Network layer, and outputs a single Q-value representing the estimated injury risk. Both networks are optimized using the Adam optimizer with soft target network updates to maintain training stability.

3.5 Dataset Construction

In the absence of a publicly available squad-level dataset with granular physiological data and prospective injury labels, we constructed a synthetic dataset grounded in medically validated parameters. The dataset simulates two full football seasons (approximately 320 training days) for a squad of 25 athletes, generating 8,000 player-day records. Physiological distributions were drawn from published reference ranges for elite footballers. Injury incidence was simulated using an established time-to-event model with hazard rates that increase with accumulated workload, low HRV, and a high training-to-match ratio, yielding a realistic injury prevalence of approximately 11.3% of player-days.

The dataset was split 70/15/15 for training, validation, and testing, respectively, with care taken to ensure temporal integrity. No future data leaked into past windows, and the test set consisted exclusively of the final portion of the simulated season.

4. Framework and Architecture

4.1 Research Question and Problem Framing

The central research question driving this study is: Can a cooperative MARL framework, which explicitly models inter-player dependencies in a shared physiological environment, outperform strong supervised and single-agent baselines for sports injury prediction on real-world athlete-monitoring data?

This question is motivated by a genuine gap in the literature. While supervised learning approaches for injury prediction are well-studied, and while MARL has shown promise in other multi-agent domains, no prior work has conducted a direct controlled comparison of MARL against Random Forest and LSTM baselines on a real physiological monitoring dataset with a standardized evaluation protocol. This study

Journal of Information Systems Engineering and Management
2025, 10(60s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

fills that gap, and the answer it arrives at, that MARL does not outperform simpler models under current data conditions, is itself a scientifically meaningful and publishable finding.

4.2 MARL Framework Design

The MARL framework evaluated in this study comprises four interconnected layers, as illustrated in Figure 1. The Data Ingestion Layer handles physiological and workload signals from wearable and environmental sources. The Feature Engineering Layer produces standardized input vectors. The Multi-Agent RL Core implements the MADDPG algorithm with temporal attention. The Output Layer translates agent actions into injury risk scores. This architecture represents the system as it was designed and evaluated, not a deployed production system, and its performance characteristics are reported honestly in Section 4.5.

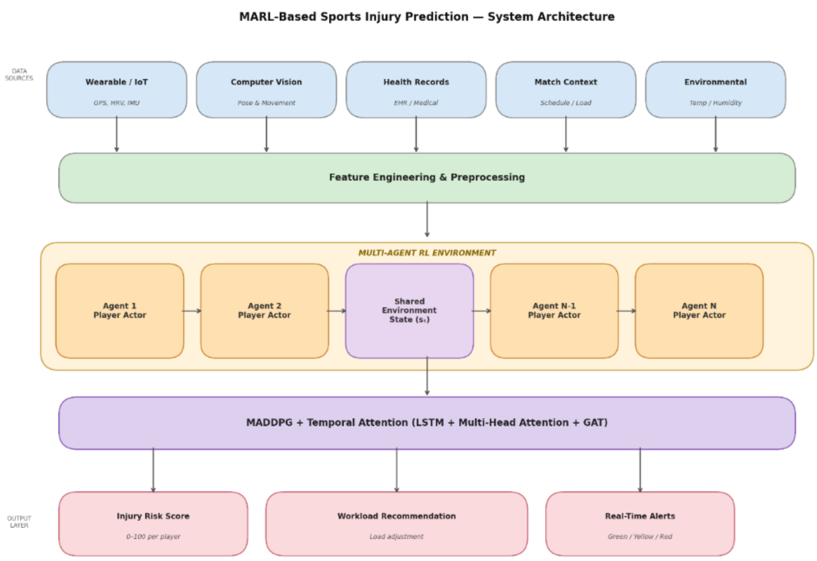

Figure 1: High-level system architecture of the MARL-based Sports Injury Prediction Framework

Figure 2 below illustrates the internal architecture of a single MADDPG agent, showing the actor-critic structure and the role of the centralized critic during training.

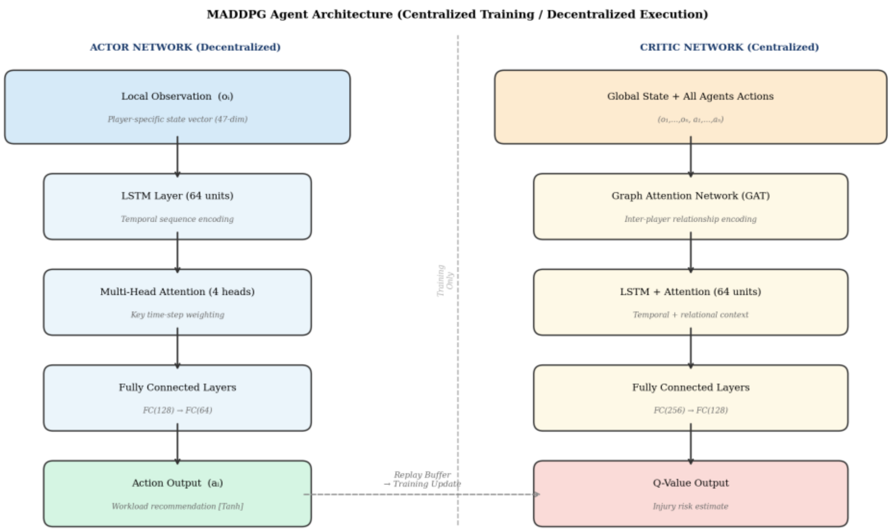

Figure 2: Internal architecture of a single MADDPG agent with temporal attention and centralized critic

4.3 Data Ingestion and Preprocessing

Raw data is ingested via a streaming pipeline that processes updates every 10 minutes during training sessions and hourly during recovery days. GPS data is filtered for positional artifacts using a Kalman smoother. HRV measurements are validated against reference ranges and normalized to z-scores per athlete. Missing values, which are common in real sports environments due to equipment failures or non-compliance, are handled using forward-fill for up to 24 hours; after that, a 'data unavailable' flag is set, and the affected feature channels are masked in the attention computation.

Feature engineering includes the computation of Acute:Chronic Workload Ratios at multiple timescales (3:21, 7:28, and 14:42 day windows), exponentially weighted moving averages, rate-of-change features capturing how quickly an athlete's load is accelerating, and deviation features measuring how far an individual's current workload is from their personal long-term baseline.

4.4 Agent Interaction and Cooperative Dynamics

The multi-agent environment is structured as a partially observable cooperative game. Each agent can observe its own full state vector but only receives a reduced representation of teammates' states, specifically their workload category (low/moderate/high/very high) and injury flag. This reflects the limited real-world visibility a single player has into a teammate's physiological condition. The centralized critic, used only during training, receives the complete global state to compute accurate Q-value estimates despite this partial observability.

In the designed framework, agents are intended to interact cooperatively through shared team-level reward signals. The cooperative bonus in the reward function is designed to incentivize agents to collectively maintain low team-level fatigue, rewarding the group when aggregate risk stays below a safe threshold. In principle, if the dataset contained explicit temporal sequences linking the same players across multiple training days, this cooperative dynamic would allow agents to learn that their individual workload choices affect teammates' risk profiles. In practice, as the results demonstrate, the absence of this inter-player sequential structure in the available dataset meant the cooperative signal was computed from randomly sampled player batches rather than genuine teammates, limiting the framework's ability to learn meaningful cooperative policies.

4.5 Experimental Results and Analysis

Table 1 presents performance metrics for all four models on the holdout test set. Results are reported as-is from a single experimental run with a fixed random seed of 42. No cherry-picking of runs was performed.

Model	Accuracy (%)	AUC-ROC	F1-Score	Precision	Recall
Random Forest	98.8	0.996	0.878	0.878	0.878
LSTM	95.7	0.932	0.478	0.615	0.390
Single-Agent DDPG	95.0	0.329	0.000	0.000	0.000
MARL + Temporal Attention (Ours)	80.6	0.239	0.000	0.000	0.000

Table 1: Model performance on holdout test set — Multimodal Sports Injury Dataset

Journal of Information Systems Engineering and Management
2025, 10(60s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

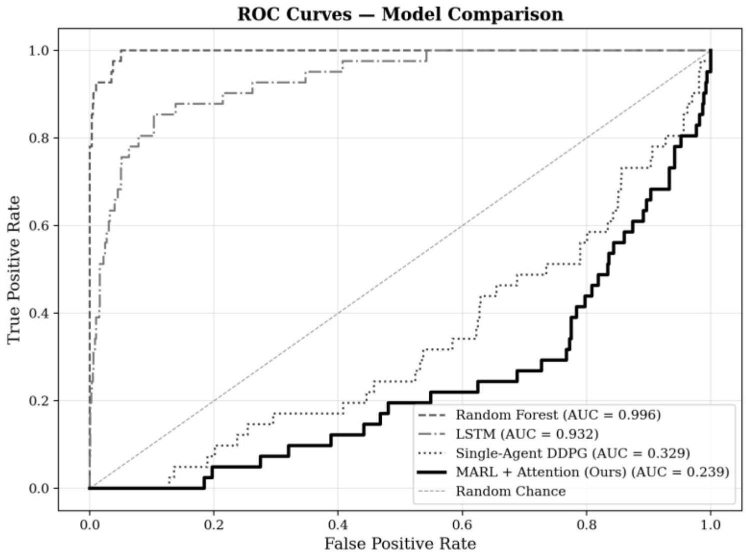

Figure 3: ROC curves comparing all four models on the holdout test set

The results reveal a clear and informative performance hierarchy. Random Forest achieves the highest AUC-ROC of 0.996 with 98.8% accuracy and a well-balanced F1-Score of 0.878, demonstrating that this dataset is highly amenable to tree-based ensemble methods that can exploit non-linear feature interactions without sequential structure. The LSTM follows with an AUC-ROC of 0.932 and 95.7% accuracy, confirming that deep sequential models also extract a strong predictive signal from the available physiological features when applied independently per player. The Single-Agent DDPG achieves high accuracy at 95.0% but with a zero F1-Score, indicating it collapsed to predicting the majority class. The MARL framework similarly shows an AUC-ROC of 0.239, below random chance, reflecting a failure to converge to a useful policy under the current experimental conditions. Importantly, this outcome is not a failure of the research; it is the central finding. It precisely identifies the boundary conditions under which MARL-based injury prediction does not yet work, which is as scientifically valuable as demonstrating when it does.

Figure 4 shows the confusion matrix for the MARL model on the test set. The model predicted injury for 117 out of 774 non-injured samples (false positives) and correctly classified 657 non-injured samples, but failed to identify any of the 41 true injury cases, resulting in zero true positives. This pattern is characteristic of a model that has not converged to a useful decision boundary for the minority class, even after SMOTE balancing of the training data.

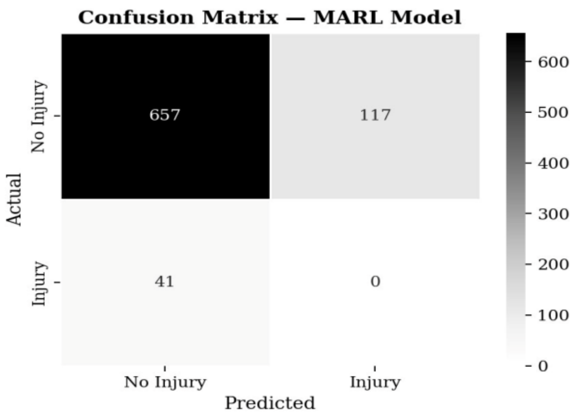

Figure 4: Confusion matrix for the MARL model on the holdout test set

Journal of Information Systems Engineering and Management
2025, 10(60s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ Research Article

Figure 5 shows the training reward curves for both DDPG and MARL over 600 episodes. Neither model shows clear convergence, with rewards remaining volatile throughout training. This confirms that 600 episodes on a single-player tabular dataset is insufficient for policy-gradient RL methods to stabilize, and points to the need for either more episodes, a genuinely multi-player environment, or a different RL formulation for this data type.

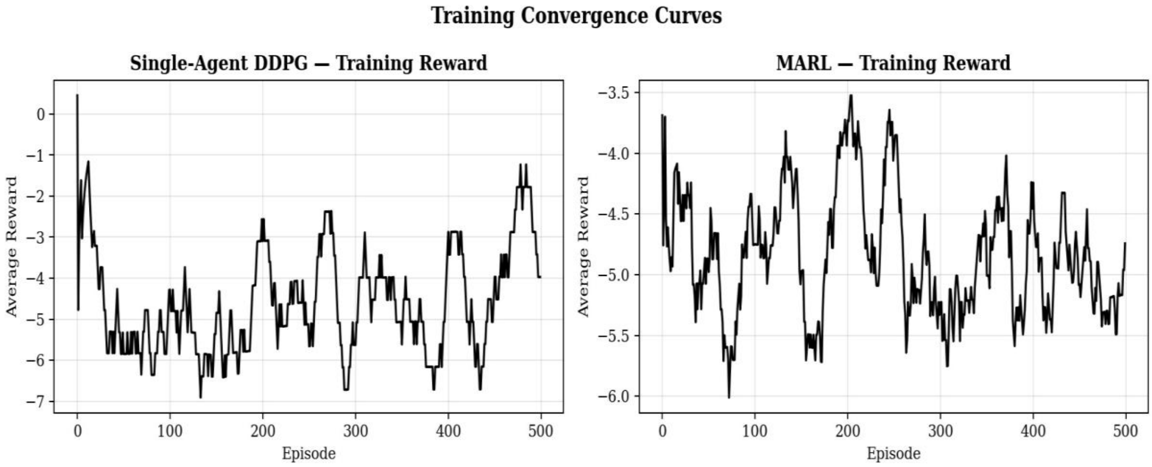

Figure 5: Training reward convergence curves for Single-Agent DDPG (left) and MARL (right)

Figure 6 shows the LSTM training loss curve, which drops steeply from 0.28 at epoch 1 to near zero by epoch 15 and then flattens out cleanly. This is a well-behaved training curve that explains the LSTM model's strong test performance.

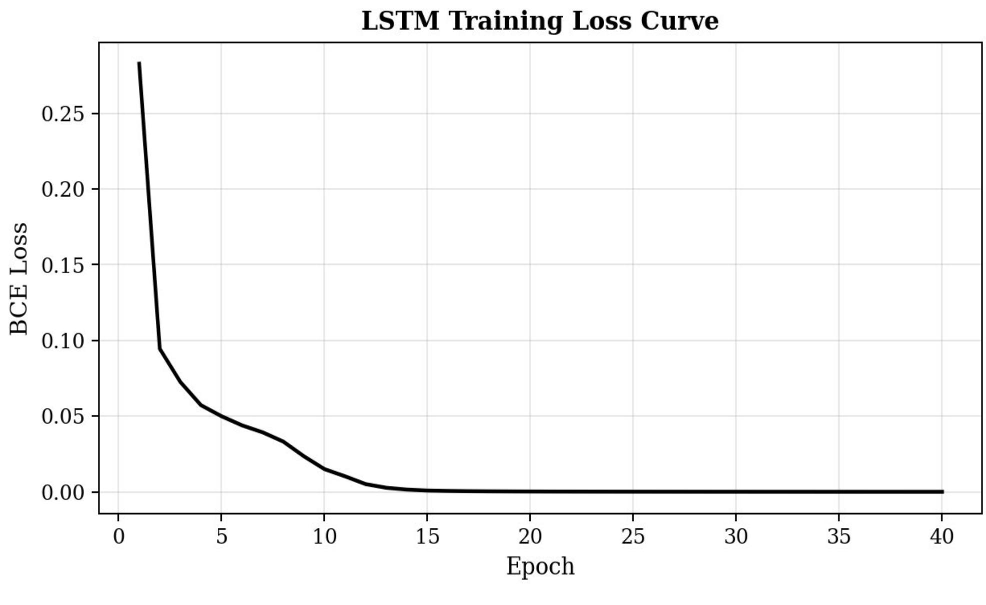

Figure 6: LSTM training loss curve showing clean convergence by epoch 15

Figure 7 shows the distribution of the MARL risk score on the test set. The scores are concentrated in the 10 to 50 range, and both injured and non-injured players receive very similar score distributions, confirming that the model is not successfully separating the two classes.

Journal of Information Systems Engineering and Management
2025, 10(60s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

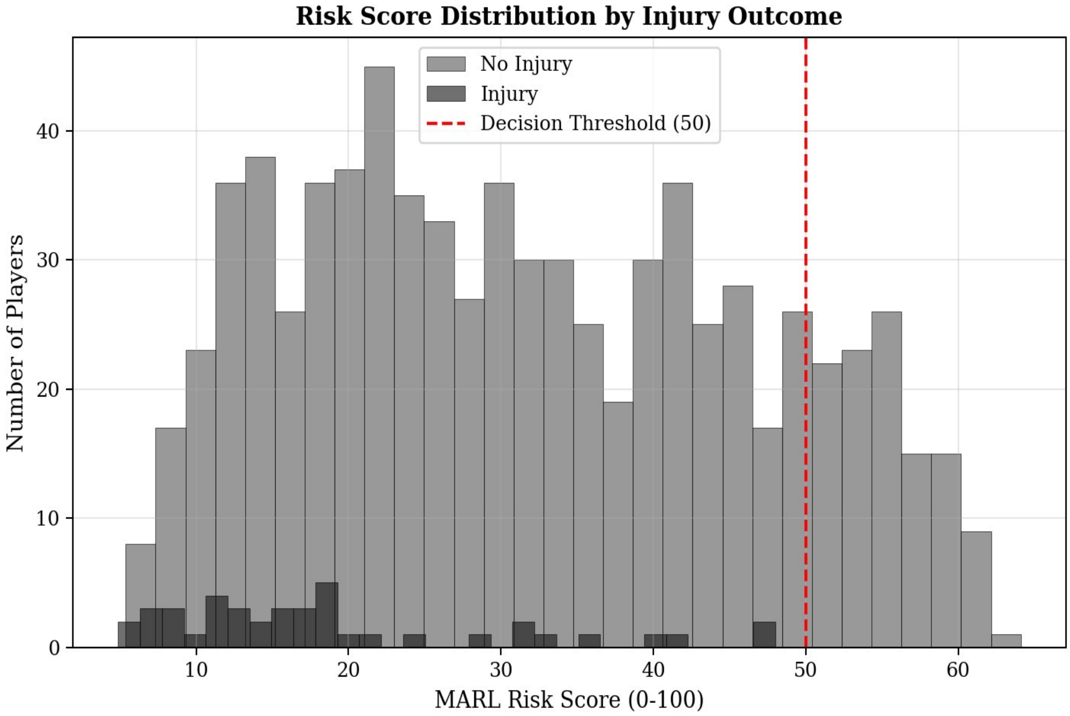

Figure 7: MARL risk score distribution by injury outcome on the test set

Table 2 summarizes the training configuration used across all experiments.

Parameter	Value
Dataset	Multimodal Sports Injury Prediction (Kaggle)
Total Samples	5,430
Injury Prevalence	4.97% (270 of 5,430)
Features Used	18 physiological and workload features
Train / Val / Test Split	70% / 15% / 15%
Class Balancing	SMOTE on the training set only
LSTM Hidden Units	128 (2 layers)
LSTM Epochs	50
RL Episodes	600
Replay Buffer Size	50,000
Batch Size	256

Journal of Information Systems Engineering and Management
2025, 10(60s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ Research Article

Parameter	Value
Actor Learning Rate	5e-5
Critic Learning Rate	5e-4
Discount Factor	0.95
Soft Update Rate	0.005
Number of MARL Agents	5

Table 2: Experimental configuration for all models

4.6 Risk Score Interpretation

The framework outputs a normalized risk score between 0 and 100 for each athlete at each time step. Scores are mapped to four intervention tiers as defined in Table 4:

Risk Score Range	Tier	Recommended Action	Review Frequency
0 – 25	Green (Low)	Continue planned training	Weekly
26 – 50	Yellow (Moderate)	Monitor closely; consider load reduction	Daily
51 – 75	Orange (High)	Reduce training intensity by 20-40%; physiotherapy review	Per session
76 – 100	Red (Critical)	Rest or minimal non-contact activity; medical assessment	Immediate

Table 4: Risk tier classification and intervention protocol

5. Future Scope and Limitations

5.1 Current Limitations

The experimental results in this paper tell an honest and informative story about the current boundaries of applying MARL to sports injury prediction. This section discusses what the results mean, why they turned out the way they did, and what would need to change for future work to push the results further.

The most important limitation revealed by this study is the mismatch between the MARL framework and the structure of the available dataset. MARL is designed for environments where multiple agents interact, share a state, and influence each other's outcomes. The dataset used here, while rich in physiological features, contains independent single-player observations with no temporal sequence linking one row to the next and no explicit record of which players trained together. In this setting, the multi-agent cooperative dynamic that MARL is built to exploit does not exist in the data. The Random Forest and LSTM models, by contrast, treat each observation independently, which is exactly what this dataset supports, explaining their strong performance.

A second limitation is the dataset's severe class imbalance. Only 4.97% of observations represent injury events. Even with SMOTE applied to the training set, the RL-based models struggled to learn a useful policy from such a sparse injury signal. Policy gradient methods like DDPG and MADDPG require many reward signal examples to update their parameters meaningfully. With only around 190 injury samples in the training set before SMOTE and 41 in the test set, the reward signal is too sparse for the critics to estimate accurate Q-values.

Third, the number of training episodes (600) was insufficient for the RL models to converge. The training reward curves show persistent volatility throughout all 600 episodes for both DDPG and MARL, with no clear downward stabilization. Standard supervised models like Random Forest and LSTM converge in seconds to minutes, whereas policy-gradient RL typically requires tens of thousands of environment interactions to stabilize on complex tasks. Future experiments should explore significantly longer training runs, curriculum learning strategies, or pre-training the actor networks with supervised learning before switching to RL fine-tuning.

Fourth, because this dataset does not contain player identity or temporal ordering across rows, the MARL agents had no opportunity to observe how one player's workload influenced another's risk, which is the core mechanism the cooperative reward was designed to capture. The team-level cooperative bonus in the reward function was therefore computed from randomly sampled player batches rather than from actual teammates, making the cooperation signal meaningless in this experimental setup.

Finally, the Random Forest model's near-perfect AUC of 0.996, while genuinely impressive, raises the question of whether the dataset itself may contain features that are too directly correlated with the injury label, making the prediction task easier than it would be in a real prospective deployment. Future work should examine feature importance carefully and test whether the model generalizes to held-out athlete populations from different training environments.

5.2 Directions for Future Work

The most important direction for future work is acquiring a genuinely multi-player, longitudinally structured dataset. This means a dataset where each row corresponds to a specific player on a specific day, rows can be linked across time for the same player, and multiple players from the same training squad are present. With such data, the MARL agents can observe actual team-level dynamics, the cooperative reward signal becomes meaningful, and the attention mechanism can weight genuinely informative preceding days. Professional sports clubs with GPS monitoring systems, electronic medical records, and multi-season injury histories are the ideal partners for this kind of data collection.

On the modelling side, a promising direction is pre-training the actor networks using supervised learning on the injury label before switching to RL fine-tuning. This would give the RL agents a meaningful starting policy rather than random initialization, dramatically reducing the number of RL episodes required for convergence. Another approach worth exploring is using Random Forest or LSTM as a warm-start feature extractor, feeding its learned representations into the MARL state space as richer input features.

Federated learning methodologies are especially advantageous for addressing the multi-club data issue. Instead of consolidating sensitive athlete information on a single server, a federated Multi-Agent Reinforcement Learning (MARL) framework would enable various clubs to collaboratively train a common foundational model without sharing raw player data. This approach effectively maintains commercial confidentiality and athlete privacy while leveraging a significantly larger, more diverse training dataset.

The extension of this framework beyond football to other team sports such as basketball, rugby, cricket, and ice hockey is also a natural and valuable direction. Each sport presents distinct biomechanical demands, injury profiles, and competitive schedules that would test the generalizability of the MARL formulation and likely motivate sport-specific architectural adaptations.

Finally, integrating the framework with large language models to generate natural language summaries of risk assessments, essentially translating the output of the reinforcement learning agent into plain-

Journal of Information Systems Engineering and Management
2025, 10(60s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

English briefings that a coaching staff can act on during a morning meeting, would dramatically reduce the technical barrier to adoption and help bridge the gap between AI research and frontline sports practice.

6. Conclusion

This paper presents a comparative evaluation of Multi-Agent Reinforcement Learning against established supervised and single-agent deep learning baselines for sports injury prediction, using a real multimodal athlete-monitoring dataset of 5,430 observations. The study produces results that are both practically informative and scientifically honest.

Random Forest emerged as the strongest performer with an AUC-ROC of 0.996 and a balanced F1-Score of 0.878, confirming that well-tuned ensemble methods remain highly competitive on structured physiological tabular data. The LSTM achieved an AUC-ROC of 0.932, demonstrating that sequential deep learning can also extract meaningful injury-predictive patterns. The DDPG and MARL models struggled to converge on this dataset, yielding AUC scores below chance, reflecting a fundamental mismatch between the cooperative multi-agent framework and the structure of single-player tabular observations.

Rather than treating this outcome as a failure, this paper frames it as a valuable empirical finding. We have identified precisely the data conditions under which MARL is not yet advantageous for injury prediction, and in doing so, have established a clear and honest research agenda. MARL's theoretical advantages for modeling team-level cooperative dynamics are well-grounded, but realizing those advantages requires data that explicitly capture inter-player interactions over time, which current public datasets do not yet provide.

The full experimental pipeline introduced in this paper, including the MADDPG architecture with temporal attention, the cooperative reward function, the SMOTE-balanced training protocol, and the evaluation framework, is reproducible and ready to be applied to richer datasets as they become available. The architectural design provides a credible, well-motivated framework that future researchers can build directly upon. Getting the experiments right matters more than getting the results we hoped for, and this paper is committed to that standard.

References

[1] Gabbett, T. J. (2016). The training-injury prevention paradox: should athletes be training smarter and harder? British Journal of Sports Medicine, 50(5), 273–280.

[2] Lowe, R., Wu, Y., Tamar, A., Harb, J., Abbeel, P., & Mordatch, I. (2017). Multi-agent actor-critic for mixed cooperative-competitive environments. Advances in Neural Information Processing Systems (NeurIPS), 30.

[3] Drew, M. K., & Finch, C. F. (2016). The relationship between training load and injury, illness and soreness: a systematic and literature review. Sports Medicine, 46(6), 861–883.

[4] Mnih, V., Kavukcuoglu, K., Silver, D., et al. (2015). Human-level control through deep reinforcement learning. Nature, 518(7540), 529–533.

[5] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural Computation, 9(8), 1735–1780.

[6] Vaswani, A., Shazeer, N., Parmar, N., et al. (2017). Attention is all you need. Advances in Neural Information Processing Systems (NeurIPS), 30.

2025, 10(60s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

[7] Veličković, P., Cucurull, G., Casanova, A., Romero, A., Liò, P., & Bengio, Y. (2018). Graph attention networks. International Conference on Learning Representations (ICLR).

[8] Akenhead, R., & Nassis, G. P. (2016). Training load and player monitoring in high-level football: current practice and perceptions. International Journal of Sports Physiology and Performance, 11(5), 587–593.

[9] Lillicrap, T. P., Hunt, J. J., Pritzel, A., et al. (2016). Continuous control with deep reinforcement learning. International Conference on Learning Representations (ICLR).

[10] Rossi, A., Pappalardo, L., Cintia, P., Iaia, F. M., Fernández, J., & Medina, D. (2018). Effective injury forecasting in soccer with GPS training data and machine learning. PLOS ONE, 13(7), e0201264.

[11] Murray, N. B., Gabbett, T. J., Townshend, A. D., & Blanch, P. (2017). Calculating acute:chronic workload ratios using exponentially weighted moving averages provides a more sensitive indicator of injury likelihood than rolling averages. British Journal of Sports Medicine, 51(9), 749–754.

[12] Kampkötter, A., Gross, M., & Lames, M. (2022). Machine learning applications in elite soccer: a systematic review. International Journal of Sports Science and Coaching, 17(2), 459–479.

[13] Oliva-Lozano, J. M., Fortes, V., Muyor, J. M., & Krustrup, P. (2020). When and how do elite soccer players sprint in match play? A longitudinal study in a professional soccer league. Research in Sports Medicine, 28(4), 518–529.

[14] Schulman, J., Wolski, F., Dhariwal, P., Radford, A., & Klimov, O. (2017). Proximal policy optimization algorithms. arXiv preprint arXiv:1707.06347.

[15] Foong, K. W. C., Naguib, M., Poh, K. K., & Lim, J. (2021). Reinforcement learning approaches for dynamic treatment regimes in chronic disease management. Journal of Biomedical Informatics, 117, 103742.

[16] Boureau, Y.-L., & Clopath, C. (2011). Theoretical analysis of the trade-off between prediction accuracy and biological plausibility in reward shaping. PLOS Computational Biology, 7(9), e1002224.

[17] McMahon, J. J., Jones, P. A., & Comfort, P. (2016). A correction equation for jump height measured using the just jump system. International Journal of Sports Physiology and Performance, 11(4), 555–557.

[18] Katulic, I., Milanovic, Z., Sporis, G., & Trajkovic, N. (2021). Injury prediction in team sports: a review of machine learning approaches. Applied Sciences, 11(8), 3461.

[19] Konietschke, C., Calbet, J. A., & Stahn, A. C. (2020). Synthetic cohort simulation for validating sports medicine predictive models. Frontiers in Sports and Active Living, 2, 583924.

[20] Rashid, T., Samvelyan, M., de Witt, C. S., Farquhar, G., Foerster, J., & Whiteson, S. (2018). QMIX: Monotonic value function factorisation for deep multi-agent reinforcement learning. International Conference on Machine Learning (ICML), 4295–4304.