International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 13s, 2025
https://www.theaspd.com/ijes.php

AI And Deep Learning In Sports For Revolutionizing Performance And Strategic Planning

Dr. S. Santiago[1], Dr. P. Gomathi Sathish[2], Dr. Priya Stella Mary[3]
[1]Assistant Professor, Department of Computer Science, St. Joseph`s College (Autonomous), Affiliated to Bharathidasan University, Tiruchirappalli, Tamilnadu, India, ssantiagosj@gmail.com
[2]Physical Education Directress Grade 1, Dindigul District Government Model School, Dindigul, Tamilnadu,India,gomathidgl@gmail.com
[3]Assistant Professor, Department of Computer Science, CHRIST (Deemed to be University), Yeshwanthpur Campus, Bangalore-560073, Karnataka | India, priya.stella@christuniversity.in

Abstract– *The sports industry is experiencing a profound evolution, driven by cutting-edge technologies that are reshaping how athletes train, compete, and strategize. These innovations are transforming performance analysis, game strategy, and fan engagement, ushering in a new era of sports science and management. Artificial Intelligence strategic planning. Initially the data were collected from Collective Sports [Sensor] DB of Practice Sessions. This data is then meticulously cleaned and pre-processed to ensure accuracy and consistency, involving Z- score normalization. For feature selection using the Random Forest and Genetic Algorithm (RF-GA) and that features were extracted using the Kernel Principle Component Analysis (K-PCA) and finally the models using the Novel Convolutional Long Short-Term Memory with Improved Quantum Particle Swarm Optimization (Conv-LSTM-IQPSO), are trained. These models are designed to recognize complex patterns and correlations that might be overlooked by human analysts. With these trained models, performance analysis is conducted to uncover insights into players' strengths and weaknesses, leading to tailored training programs and strategic recommendations. The AI-driven analysis extends to strategic planning, where simulations and predictive analytics play a crucial role. Coaches and managers can leverage these tools to anticipate opponent tactics, optimize game strategies, and make data-informed decisions during matches. This iterative process of incorporating new data and feedback ensures continuous improvement, keeping strategies relevant and effective. In summary, the integration of AI and deep learning in sports provides a sophisticated approach to enhancing performance and devising strategies, marking a transformative advancement in the field.*

Keywords– *Artificial Intelligence, Deep learning, Kernel Principle Component Analysis, Novel Convolutional Long Short-Term Memory.*

I. INTRODUCTION

The sports industry is growing globally, making it crucial to acquire accurate, rapid, and reliable data. This data is essential for managing information and developing long-term sports development goals, which will speed up sports data organization. While governments construct or grow sports information centers, international and regional entities have increased their sports information operations. Computers optimize and specialize technology, reducing the gap between industrialized and developing nations [1-3]. The introduction of new technologies has advanced athletes and sports, sparking new inventions and technology. Information technology has changed people's perspectives since its introduction and is effective in any situation, so this study focuses on it. Technology makes sports smarter and more exciting for millions of viewers worldwide. In modern clubs and stadiums around the world, trainers see scientific and technology advances as allies and assistants, improving sports training performance and raising the bar for training worldwide. Sports clubs can boost earnings by selling promotional goods and gear worldwide. Due to the internet, the latest sports education and achievements can be obtained without leaving home. The internet provides information from several reputable sports universities worldwide. The internet affects business and marketing. Nearly all major American sports teams have websites that update fans on moves, club stock prices, and revenue. Sports lovers can also get the latest team news anytime. National sports bodies' development of ICT in sports has been proved to be one of their most helpful functions [4,5]. These technologies help sports evolve. These technologies may need to be expanded to advance sports and succeed. This study uses "technology" to refer to sports-related digital and technical advances. Technology has impacted several facets of the sports industry and player experience, according to the report. Wearable technology, analytics tools, virtual platforms, adaptive gear, and other engineering and computer science applications have transformed sports data collection, analysis, and utilization. Even while "technology" covers numerous areas, defining it specifically contextualizes the following discussion of sports digital transformation consequences and tactics. Because many factors affect sports growth and promotion, more experts believe that macro and micro sports development dimensions must be addressed. Sports development involves providing opportunities, procedures, methods, and efficient structures enabling people of all backgrounds and specialized groups to participate in sports and leisure activities and improve their performance. Research calls this the information or knowledge age. Information technology has changed every element of modern life, introducing new ideas and perspectives. Information technology has had a major impact on companies, making them unrecognizable without it. Research reveals that information technology reduces information acquisition and analysis expenses. It reduces administrative and information gathering, distribution, and handling costs for firms. Organizational rigidity is also altered. Sports organizations must establish and use IT like other organizations. Modern companies, especially sports organizations, must improve and emphasize information technology. More advanced technology are employed to improve sports events and training. These cutting-edge technology have changed sporting events, training, and athlete performance. IT adoption affects organizational performance. Several businesses launch concentrated IT development and application initiatives. An information development program begins with identifying barriers to not adopting this technology. Information technology boosts corporate efficiency and productivity. There are also challenges. Sports organisations must employ IT like other organisations [6,7]. Implementation makes organizational procedures faster, more accurate, and cheaper. Therefore, organizations that adapt to information technology will succeed. All sports stakeholders must employ the right information technologies for their field and endeavor to enhance their operations. These systems develop sports quantitatively and qualitatively across all aspects (management, equipment, etc.). Public understanding of sports science's health advantages has driven recent advances in the subject. Studies on the commercialization of sports management reveal that today's society values sports' economic and health benefits equally. Professional sports organizations are also affected by economic considerations, therefore economic development is important. These sports organisations need new sciences and methods to do this. According to research from multiple sources, standard sports equipment, sophisticated training aids, sports performance evaluation, massive sports arenas, performance analysis software, the sports supplement industry, anti-doping initiatives, mathematical modeling, computer simulation of sports movements, etc. highlight sports technologies. A study on the futurism of information technology infrastructure highlighted sports organizations and model presentation and found that emerging information technologies affect sports and leisure. These technologies affect consumer and employee communication, sports and leisure management, and analysis. IT knowledge and utilization will define professional program production and effectiveness in the future. Sports and entertainment enterprises must employ this technology to coordinate all communication, services, programs, and human resources to succeed [8-10].

A. Research aim

This study's main goal is to investigate and illustrate how artificial intelligence (AI) and deep learning may transform sports performance analysis and strategic planning. This includes raising the standard of the spectator experience overall, raising the performance of athletes, and refining game plans.

B. Research Contributions

The research contributions were mentioned as follows:

- introduced a methodical process for gathering and preparing data from various sources, using Z-score normalization to guarantee correctness and consistency.
- To improve the prediction power of the model, Kernel Principle Component Analysis (K-PCA) was used for feature extraction and Random Forest and Genetic Algorithm (RF-GA) for feature selection.

- created models for identifying intricate patterns and correlations in sports data utilizing Novel Convolutional Long Short-Term Memory with Improved Quantum Particle Swarm Optimization (Conv-LSTM-IQPSO).
- enabled AI-driven analysis for strategic recommendations and performance insights, assisting management and coaches in making data-driven decisions and enhancing game plans.

II. LITERATURE SURVEY

The Cognitive Reasoning Sports Teaching Model (CRSTM) [11] is a problem-based technique based on research ideas and findings from teaching computer-based sports learning approaches to temporarily solve dynamic trails. Most computational mathematics training begins with identifying expected performance or behavior. Educational psychology has traditionally supported educational goals. Motor skill training emphasizes incorporating important task elements in various ways. An SVM model created in a lab was tested for real-world shot and pass detection in [12]. Unlike the existing full-instep and side-kick classification, we identified shoots and passes independently of kicking form. We also constructed CNN, convLSTM, and LSTM deep learning models to test their shot and pass detection capabilities in real-world circumstances. They collected IMU data from over 800 football players in lab and real-world settings to achieve this. Three scenarios illustrate different model evaluation difficulty levels. They tested the models using segmented laboratory ball contact data. For the second and third techniques, they modified a peak detection technique from the literature and added a sliding window approach to identify candidates before classifying them. The strategy [13] suggests can be applied to any sport, not only soccer. Thus, managers and coaches' player selection must be improved and supported. Soft computing methods, such as artificial intelligence and neural networks, are designed to make logical conclusions in complex situations. Data mining and machine learning in sports are explored in this paper [14]. Machine learning is used to predict the badminton tournament's outcome. We will forecast the Australian, Malaysian, German, and Singapore Open badminton championships using three classifiers: Hyper Pipes, CHIRP, and NB-CBFW. SVMST was proposed in [15] to evaluate student sports efficiency. Opponent data, participant data, traditional game statistics, and person quality criteria inform sports training prototypes. Large and moderate are student success grades. Supervised learning and categorization are used to create a student sports training efficiency template. Learning, data collection, model evaluation, and sports performance prediction difficulties are addressed by SVM. In [16], authors used subjective wellbeing metrics from a commercial PmSys digital health monitoring system to predict soccer players' performance. We analyze two Norwegian female soccer teams' two-year data. Daily, players reported their readiness to play, mood, tension, muscle discomfort, exhaustion, sleep duration, and quality. We use univariate and multivariate time series models to predict preparation. We empirically evaluate the prediction performance of numerous time series models, including exclusively recurrent models, mixed recursive convolutional models, an ensemble of deep CNN models, and multivariate recurrent models, with an emphasis on peak detection. They examine next-day and next-week forecast accuracy using different input and prediction windows. We also investigate forecasting individual players using team data instead of player data in individual player models. We concatenate arrays, fill repeated values, and replace all gaps with zeros to fix missing data. In [17], authors examined the relationship between digital technology utilization and corporate success, as well as organizational innovation and digital transformation strategy's mediating roles. A Taiwanese financial supervisor survey was used to perform an empirical study. 227 companies responded to questionnaires. The results demonstrated that digital technology boosts organizational innovation, digital transformation strategy, and business performance. Additionally, organizational innovation and digital transformation strategy totally mediated the association between digital technology utilization and corporate success. In [18], authors examine secondary data and modern sports literature to discuss sports with technology, focusing on sports management. Using cutting-edge technologies to benefit all sports ecosystem players is the goal. The chapter examines player or team performance enhancement, audience involvement, and alternative sports consumption, where technology may be crucial in the future.

International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 13s, 2025
https://www.theaspd.com/ijes.php

A. *Problem Statement*

Even with major advances in technology, deep learning and artificial intelligence applications in sports are still in their infancy. For performance monitoring and strategy planning, many teams and athletes depend on antiquated techniques that are frequently laborious and imprecise. To give real-time, actionable information and promote innovation in sports performance and strategy, a complete methodology integrating AI and deep learning is required.

III. PROPOSED METHODOLOGY

Our study suggests utilizing AI and deep learning by using complex model construction with Conv-LSTM and IQPSO, substantial data collecting and preprocessing, and enhanced feature selection and extraction techniques. These techniques allow for thorough performance analysis and strategic planning, and they are always improving thanks to data updates and iterative feedback.

A. Dataset collection and Preprocessing

Here the Sports dataset was collected from below mentioned link:
https://www.kaggle.com/datasets/sujaykapadnis/comprehensive-sports-database
And the dataset was preprocessed using the normalization is a feature scaling technique that aims to ensure that all attribute values are on an equal scale. Standardized moment, z-score normalization, and min-max normalization are among many different normalization methods. The min-max normalization technique was implemented in this work. At the final stage, the Min-Max scaling of the data placed each characteristic inside the same range of [0, 1].

$$Z_{Norm} = Z - \frac{\min(a)}{\max(b)} - \min(b) \qquad (1)$$

Here the attribute Z's minimum and maximum values are denoted by min (b) and max (b). Z and Z-score Normalization represent the original and normalized values of the characteristic, Z, respectively. This is an essential step in the data analysis process since it improves algorithm performance and facilitates the selection of pertinent features from the data, as well as helping the model prepare the data for the next stage. In this proposed study, we have preprocessed using Z-Score Standardization. Particularly, one of the most frequently utilized normalization techniques is called Z-score normalization and involves converting and scaling the data according to its mean and variance. For many machine learning methods, an important preprocessing step is Z-score normalization. It is scaling every feature to have a mean of 0 and a standard deviation of 1. We remember that the Z-score Z of a collection of data X can be calculated as, for example, have already demonstrated the effect of representing each domain via z-score on the classification of the signals. Each value in a data set is z-score normalized using the algorithm below:

$$\mathrm{Z}(\mu, \sigma, \mathrm{A}) = \frac{\mathrm{a}-\mu}{\sigma} \qquad (2)$$

Where:

A: Original Amount
μ: Mean of data
σ: Data Variance.
The authors highlighted the way traditional Z-scores are presented.

B. Feature Selection and Extraction

After preprocess the data the features were selected using the Random Forest and Genetic Algorithm (RF-GA). Features extracted from this dataset include player statistics, game results, team performance indicators, and historical match data. The step of choosing suitable characteristics to construct a better model comes next. By removing unnecessary, empty, or noisy data, the model can be made simpler, its performance enhanced, and its accuracy and security maximized.

a) Random Forest

Random Forest (RF) classification is an ensemble technique that continually trains many decision trees using bootstrapping, averaging, and bagging. Utilizing different subsets of available attributes, several

International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 13s, 2025
https://www.theaspd.com/ijes.php

independent decision trees can be built simultaneously on different portions of the training samples. Bootstrapping ensures that each decision tree within the RF is distinct, reducing RF variance. RF classification combines several tree judgments for the final judgment, resulting in a robust generalization. The RF model builds and aggregates a set of learning trees depending on the decision tree-based model's properties.Fig 1 explains the RF. The principle is to integrate multiple decision trees, with the dataset selected at random every time, while randomly only a few of the features as input; hence, the RF Algorithm can also be considered a decision tree used as the predictor in a collecting algorithm. Each decision tree is trained on a randomly selected subset of training data. Finally, the selection of classification is reached using a system of voting among the decision trees.

- *Forest*

The construction method is as follows:

i. Sample selection: Bootstrap sampling is used to randomly choose different subgroups from the initial set of data

$$\{V_M \; m = 1,2,3 \ldots, M\} \qquad (3)$$

ii. Feature Selection: M features will be selected randomly within M features.

iii. Decision tree construction: Create a decision tree based on the samples from the data collection. At every node, the best partition feature is selected based on Gini impurity.

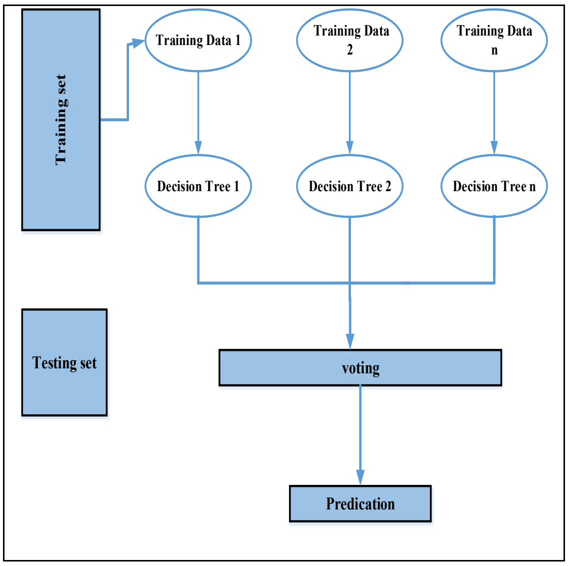

Fig 1: The Structure of Random

Integrated prediction involves a majority vote based on the predicted results for each decision tree to establish a sample's classification. Implementing the RF method might be difficult, because of the dataset's high-dimensional data and the class difference resulting in unsatisfactory Achievement of requirements like the precision of classification. In addition, the algorithm's classification performance is highly influenced by two factors: the accuracy of individual decision trees and the bias in voting results across numerous decision trees. To improve RF overall performance in IDS, we are investigating combining the dimensionality of data reduction methods with economical methods of learning. Essential steps and concepts for improvement are suggested.

Algorithm: 1 RANDOM FOREST
Require: Sample features f_s

Ensure: b=1 to B
Create a bootstrap sample Z of size S from f_s
Create a Random Forest tree T_n to bootstrapped data

For every node, repeat the process recursively.
in T_n
Stop when the minimum size b_{min} is reached
While $f_s \neq$ **0 do**
if $S \neq$ **0 then**
Select variable m at random from the p variable
Pick the best-split point among the m variable
Splitting the node into two daughter nodes
Output the ensemble of trees $T\frac{P}{p}$
else if S > 0 **then**
Predict at the new point of x
Regression: $f\frac{1}{B}(orex) \;= \frac{1}{B}\sum_{b=1}^{B} T_b(x)$
Classification: Let $C_b(x)$ be the class prediction of the RF

$$C_{rf}^{B}(x) = Majority\ vote[C_{rf}^{B}]_1^B$$

b) **Genetic Algorithm**The biological optimization approach is a biologically inspired optimization methodology that utilizes evolutionary optimization approaches like transmission, crossings, and variation. The proposed system includes a Genetic Algorithm (GA)- based component for feature selection, which serves as the main learning module. This work uses a GA to select optimal features in an unsupervised manner. The GA is evaluated with various parameters as specified in the experimental setup. Below fig 2 explains the GA. The GA consists of the following steps:

i. Initial population and structure of chromosomes
ii. Function of Fitness
iii. Approach for the Selection
iv. Mutation and crossing across

The fundamental learning technique for the proposed system is the component for selecting features based on GA. To tackle the present problem, this investigation uses a GA for optimum feature selection in an unsupervised manner

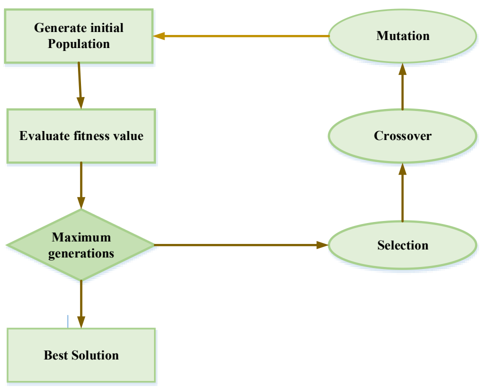

Fig 2: Structure of Genetic Algorithm

International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 13s, 2025
https://www.theaspd.com/ijes.php

Here the setup of the experiment lists the different parameters that are examined with the GA. The GA consists of the following phases: Cross-over, mutation, final next-generation production, selection technique for choosing parents who generate children for the next generation, and initial population generation are all covered in this paper.

i.Initial Population and Structure of Chromosomes

Generation of the First group is the beginning stage of GA deployment. Here the first population is created at random. Interacting the distinct genes chosen at random results in chromosomes. To prevent duplicate genes inside a chromosome, the genes are selected at random as features from the initial data characteristics, taking into account the total number of genes present on every chromosome. In the population, every chromosome stands for a single solution to the selection problem. An array of N cells in one dimension is the chromosome. This work has N set to 10. Nevertheless, several studies have been carried out later on with different values of N. A numerical value V is stored in each chromosome cell, where feature set length ≥V ≥0. The feature number is shown by the numerical value in every gene. A value of 8 for the gene at index 0 indicates the eighth feature from the original data, for instance. This implies that a part that is later optimized to be the perfect solution is created by each chromosome holding ten features from the feature set.

ii.Fitness function

In the absence of class labels, the correlation between the features selected is computed, and this research introduces a new fitness function. A tool to identify those features with little similarity among them is provided by the fitness function. This makes it feasible to gather features with a greater diversity and the ability to represent the majority of the dataset's information. By comparing a solution with different options and assigning a fitness score to each, it determines how well a solution fits. The fitness score is used to determine the possibility that an alternative will be chosen. Let us assume that α and σ remain constant:

$$A[h] = \alpha + \sigma h \qquad (4)$$

The correlation average is determined by the proposed function of fitness. Once the correlation of the provided features has been determined. The values of the average correlation achieved have come to be optimized by raising them over GA generations after the correlation average is obtained. There has to be a solution where the chosen characteristics have to be lowly correlated and diversified. Transformation of the correlation of average value to non-correlation of average is necessary for this. Taking the difference of the average correlation value from 1 (the higher value is achievable) It has succeeded at each in age. In this method, the values of correlation and accuracy averaged together might be shown in an increasing or decreasing order. After that, one can compute the fitness values by utilizing the accuracy of that specific chromosome and the resulting average correlation value. Here is the average of the changed values and the precision. A given fitness of the chromosome is the outcome value. The main objective of the suggested approach is to improve the correlation of average values and simultaneously accuracy to display the suggested function of the fitness.

$$corr_{avg=\frac{Sum(g) of\ values\ above\ the\ diagonal}{number\ of\ Values}} \qquad (5)$$

$$Corr_{avg}^{t} = (1 - Corr_{avg}) \qquad (6)$$

$$F_I = \frac{A_{i+(1-M_i)}}{2} \qquad (7)$$

$corr_a$ls the correlation the accuracy value; $Coor_{avg}^{t}$is the transformed un-correlation the average; g is the accuracy; and M_i is the computed correlations matrix. Returns the chromosome's fit value. Here the fitness function employs two approaches: the scaling function and the objective function. The objective function this work optimizes is the acquired accuracy. The function of un-correlation, which scales the accuracy performance, is how this is achieved.

iii.Selection method

The most effective approaches are chosen in this stage, and their genes are allowed to continue to the following iteration. This is comparable to Darwinism's theory that the fittest people pass on their genes to the following generation. The fitness scores are used to choose each person's parents. A significant amount of fitness is required for an individual to be chosen for reproduction. It makes use of Selection probability distribution, which is the selection probability for individual fitness of string to the proportional. The probability of the process of the selection (Prob.) is expressed mathematically as follows:

$$S = \frac{\text{Prob}*[\text{selection of fittest string}]}{\text{Prob}*[\text{string selection of average}]} \quad (8)$$

For this contribution, the client's choice method to produce a future generation is applied using the wheel of fortune selecting technique. The wheel of fortune selection method has the advantages of being faster to execute when used in parallel and not requiring sorting or scaling like other selection methods. The selection of parents has a huge impact on the fitness values that the function of fitness produces. It suggests that selection chances will rise with rising fitness value. The fundamental technique used to weight the slots in roulette wheel selection is linear search, which is based on the fitness values of the individual chromosomes. Because the chromosome covers more area on the Wheel of Fortune, higher fitness ratings suggest a higher probability of its selection.

iv. Crossover and mutation

A significant role of diversity in EAs is played by reproduction operators. Both mutations and crossovers are applied in this experiment. After every chromosome in the population is fit, the crossover and mutation operations are carried out. We test three different crossover rates: 0.25, 0.5, and 0.75. We choose the best one. Where better result was predicted by the simulations with a cross-over probability of 0.5. Crossover has a particular design:

For crossover, two parents are chosen by use of the roulette wheel procedure.

- The chromosome gets direct transfer of the second half of the first parent.
- The second parent duplicates the remaining genes in the same order as the child's chromosome.
- Until all necessary people in the population are found, the aforementioned two procedures are repeated.
- The chosen parent couple produces just one child.

The field of search dimensions is determined by the mutation rate that is utilized in the GA to provide techniques for exploration. The convergence of GA is impacted by the rate of mutation. As GA typically loses important components of the solution before convergence, a high mutation rate will almost always lead to poor convergence. The supplied mutation procedure has been suggested.

A. Feature Extraction

The selected features were extracted using the Kernel PCA.

i.Kernel PCA

Several variables with correlation can be converted into Principle Components Analysis (PCAs), which are made up of several independent variables, using this technique. Reducing the number of dimensions of the input variables while preserving as much variation as feasible in the derived variables is the primary objective of PCA. By extending PCA to nonlinear dimensionality reduction, kernel PCA improves the scope of PCA. Assume that a set of m matrices is an input $y_1, y_2 \dots y_m$ there are n characteristics in every vector. Using a nonlinear mapping function, kernel PCA converts a training set into a high-dimensional feature space. The connection between the feature vector and the mapping equation $\mathcal{E}$are is expressed formally as follows:

$$\mathcal{E}: y \in R^n \to \mathcal{E}(y) \in S \quad (9)$$

Where

$$\sum_{i=1}^{n} \mathcal{E}(y_i) = 0 \quad (10)$$

We can find the Covariance of matrix $\mathcal{E}(y_i)$ can be calculated by

International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 13s, 2025
https://www.theaspd.com/ijes.php

$$\mathcal{D} = \frac{1}{n}\sum_{i=1}^{n}(\mathcal{E}(y_i - mean)(\mathcal{E}(y_i) - mean)^T \quad (11)$$

Where mean = $\frac{1}{n}\sum_{i=1}^{n}\mathcal{E}(y_i)$ (12)

$$\mathcal{D}y = \lambda_i y \quad (13)$$

A linear equation makes it simple to expand each of $\mathcal{D}$ eigenvectors, y. The linearly expanded may be explained as follows.

$$y = \sum_{i=1}^{m}(\alpha_i \mathcal{E}(y_i)) \quad (14)$$

To determine the quotiety α_i , a kernel matrix is defined by us L of dimension$n \times n$, whose components are determined using the equation that follows

$$L_{ij} = (\mathcal{E}(y_i)^T \mathcal{E}(y_i) = \mathcal{E}(y_i) \cdot \mathcal{E}(y_j) = m(y_i y_j) \quad (15)$$

Where

$$m(y_i y_j) =< \mathcal{E}(y_i), (y_j) > \quad (16)$$

If the predicted dataset $\mathcal{E}(y_i)$ is missing a Zero mean, we replace the matrix of the kernel L with the Gram matrixL'. Here is how the Gram matrix L' is constructed.

$$L' = L - \mathcal{N}L - L\mathcal{N} + \mathcal{N}L\mathcal{N} \quad (17)$$

In this case, the matrix $\mathcal{N}$ is $n \times n$ and each element has a size are$\frac{1}{n}$. The Eigenvalue problem in the equ (4) can be resolved by rewriting the preceding equation as follows.

$$L'\alpha = n\lambda\alpha \quad (18)$$

The orthonormal characterization of $\mathcal{L}$, represented by the vectors α associated with the p biggest positive the eigenvalue namely $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p$. Thus, the orthonormal eigenvectors y_i of $\mathcal{D}$ may be expressed as follows.

$$y_i = \frac{1}{\sqrt{\lambda_i}}\mathcal{E}(y_i)\alpha_i \quad (19)$$

The translation of y_{new} onto eigenvectors y_i may be stated as follows; $\mathcal{E}(y_{new})$ is the function that maps for the newly generated vector samples y_{new} to the spaces of features.

$$t = (y_1, y_2, y)^T \mathcal{E}(y_{new}) \quad (20)$$

The i^{th} transformed feature t_i of kernel PCA can be calculated by

$$t_i = y_i^T f(y_{new}) = \frac{1}{\sqrt{\lambda_i}}\alpha_i^T M(y_i, y_{new}) \quad (21)$$

Significantly, the kernel matrix may be immediately constructed using the data set used for training. Numerous kernel functions have been suggested by investigators; the most often used function in kernels is the Gaussian kernel. The Gaussian kernel, which may be stated as follows, is the kernel function that we have chosen for this investigation:

$$M(a, b) = e^{(\frac{-||x-y||^2}{2 \times sigma^2})} \quad (22)$$

Using the kernel PCA feature extraction approach reduces the dimensionality and noise in the data being processed, which speeds up the processing of data and boosts the system's efficiency.

C. Train the Model using the Novel Conv-LSTM-IQPSO

The suggested method that is the **CONV-LSTM** (Long Short-Term Memory and Convolutional Neural Networks) saves a great lot of time and effort by working directly with the raw data and automatically extracting significant features.

a. Conv-LSTM

The input matrix of Conv-LSTM is given below, where f indicates a specific spot inside a region and t indicates time.

$$S = \begin{bmatrix} S_1^1 & S_1^2 & \cdots & S_1^f \\ S_2^1 & S_2^2 & \cdots & S_2^f \\ \vdots & \vdots & \ddots & \vdots \\ S_t^1 & S_t^2 & \cdots & S_t^f \end{bmatrix} \quad (24)$$

A time series vector $S_t^f = (S_1^f,\ S_2^f,\ S_3^f, \dots ,\ S_t^f)$ can be used to represent the matrix. Every element in the vector reflects the traffic flow information for every place in the area around the spot that has to be predicted simultaneously. The elements in vectors is represented as $S_t = (S_1,\ S_2,\ S_3, \dots ,\ S_t)$. To process each element inS_t^f , 1D-Conv is used. Convolution information of the local perceptual domain is obtained by sliding filtering via a one-dimensional convolution kernel filter. It is advantageous to extract the spatial features between nearby places using such a procedure. The global feature is then created by combining the local features. Every step creates a unit node that moves over the vector. The node of the unit is shown as:

$$H(i) = S(A\omega + B) \quad (25)$$

Where A is the input node's value, B is bias, S is the activation function, and ω is the node filtering weight. The data vector is subjected to the pooling filter. The pooling filter differs in that it doesn't carry out complicated convolutional processes. Rather, the numbers are only averaged. Pooling efficiently decreases vector sizes. In order to acquire more abstract data, some irrelevant data is filtered away during the pooling process. Through pooling, the feature sequence M that is the resulting value is reduced to 1/2 of its original dimension. When handling traffic flow data, the deep neural network has a better distortion tolerance because to these two feature extractions.

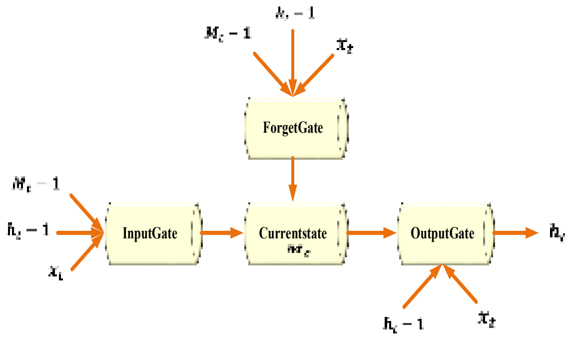

Fig. 3: LSTM Structure

The output of the results come out as a time series vector like $M_t = (M_1,\ M_2,\ M_3, \dots ,\ M_t)$ after every element present in the time series vector has undergone the corresponding convolution and pooling operations. Fig. 3 displays the LSTM structure. The LSTM block is consisting of three gates which will regulate any of the updates to the output part of the LSTM block and also to the cell which is recording the state of block cell. To identify which portion of the data is lost from the previous state of the cell, the LSTM cell first performs the Forget Gate. The cell must then choose whether information from the input data should be retained in the second step. The input gate will decide the set of values which should be updated in the LSTM module. Thirdly, it will use the acceleration data in this phase and then the output of LSTM module at the end; second, an intermediate state vector R_t is created using a $\tan h$ activation function. After that, the cell value is changed. Next, the value of the cell is updated. To retrieve the values used at the last stage; multiply the value of the Forget Gate with the cell value in the moment *t – 1 to* get it. The value that will be added in the cell is the outcome of the value after multiplying of the input gate by the state vector which is produced by $\tan h$function. The newly updated rate of the cell is increased by the sum of the two values. Ultimately, the output gate yields the LSTM output. Only the current cell value is relevant to the output, and the current cell value is completely dependent on the value of previous cell. This process can also be represented as:

ForgetGate $H_f = \sigma(W_f.[K_{t-1}, I_t] + d_f)$ (26)

InputGate $H_i = \sigma(W_i.[K_{t-1}, I_t] + d_i)$ (27)

Intermediatestate $R_t = \tan h\,(W_c.[K_{t-1}, I_t] + d_c)$ (28)

UpdateCell $M_t = H_f * M_t - 1 + H_i * R_t$ (29)

OutputGate $H_o = \sigma(W_o.[K_{t-1}, I_t] + d_o)$ (30)

Output $K_t = H_o * \tan h\,(M_t)$ (31)

Where R_t is the intermediate state of the *t* moment, and W_c and d_c are the weight and bias of the intermediate state. σ is the activation function, and H_f, H_i and H_o are the ForgetGate, InputGate, and OutputGate. W_f, W_i and W_o are their weights and d_f, d_i and d_o are their biases. Overfitting may be successfully prevented by using the Dropout structures found in each layer of the LSTM structure. Fig. 4 displays the Conv-LSTM module construction.

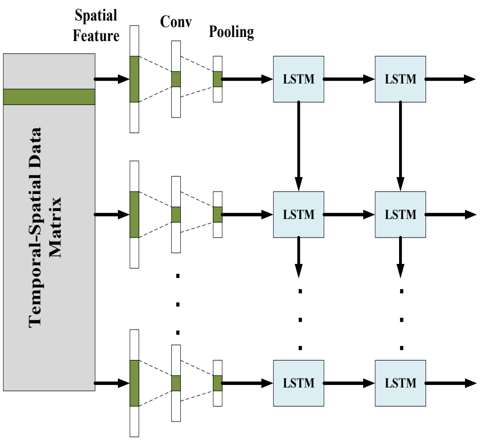

Fig.4: Conv-LSTM structure

a. Improved Quantum Particle Swarm Optimization

A popular and well known population-based metaheuristic algorithm, the PSO has been effectively used to optimize a number of issues, including function optimization. The location and velocity of bird flocks are the two primary factors that inform the development of the PSO. Taking into account the optimization process, the PSO equation for the nth iteration is expressed as follows:

$P_{n+1} = P_n + V_{n+1}$ (32)

Where P_n and V_n are the particle position and velocity. The updated velocity in Equation 18 may be obtained using the following relation based on the best swarm location (P_s) and the personal best value (P_w):

$V_{n+1} = b.V_n + a_1.t_1(P_n - P_w) + a_2.t_2(P_n - P_s)$ (33)

In this case, t_1 and t_2 stand for two random coefficients that are created at each iteration, *w* is the inertia weight, and a_1 and a_2 are the acceleration coefficients. Until the desired criterion is met, the repeating process is carried out. The fundamental idea of PSO is that each potential solution to an optimization issue is shown as a particle moving across the search space. Each particle in the swarm is optimized based on the associated fitness value, and the other particles follow the optimal particle to find the best solution in the search space. The search model uses a velocity vector which is made up of flying and distance direction. However, monitoring movement is how the best procedure is carried out, because the particles can fly in a limited search space than flying in the complete search space due to the velocity which includes magnitude and direction, certain PSO algorithms are unable to find the global optimum solution. In order to prevent such issues, QPSO has been suggested.

b. Quantum Particle Swarm Optimization (QPSO)

"Quantum theory" is been utilized in QPSO throughout the search procedure. QPSO asserts that bound states derived from attractive potentials have an impact on a particle's motion. The particle which represents a solution should reach the optimal point instead of changing to infinity, in the feasible search space, because the particle impacted by a particular bound state may emerge at any position in the search space with a defined probability. Particle mobility in the traditional QPSO model is established by common knowledge and personal experience. The particle variable and the fitness function determine knowledge of an individual. The organization of variable in particle that is the structure of particle is shown in Fig. 5. In QPSO, there are three import positions: "$pbest$, $gbest$, and $mbest$". $pbest$ may be thought of as the best for knowledge of individual as it reflects the finest solution the particle has ever found in its history of movement. It is considered as one of the best social knowledge since it reflects the better result identified by all the particles from the complete history of the movement.As compared to, all the average level of particles from the recent iteration; $mbest$ can be considered as the accepted standard. The movement of the particle is influenced by both the mainstream ideas and best individual knowledge, but the best social knowledge in the learning process is the optimal answer for the current iteration.

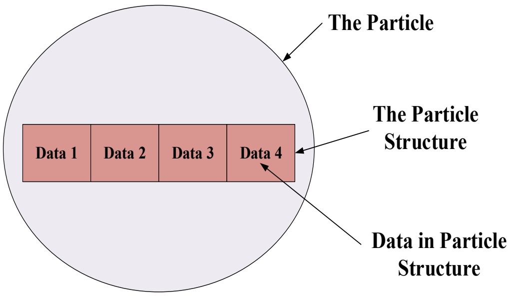

Fig.5: Particle Structure

Assume that $v_i(t)$ denotes the particle's ith dimension at time 't', $pbest_i$ denotes pbest's ith dimension, and $mbest_i$ denotes mbest's ith dimension. The iterative equation 18 governs the particle motion:

$$\begin{cases} v_i(t+1) = pbest_i + \beta \times |mbest_i - v_i(t)| \times \ln\left(\frac{1}{x}\right), y \geq 0.5 \\ v_i(t+1) = pbest_i - \beta \times |mbest_i - v_i(t)| \times \ln\left(\frac{1}{x}\right), \ y < 0.5 \end{cases} \tag{34}$$

Where x and y in the interval [0, 1] are the values produced by the uniform probability distribution functions, and β stands for the contraction–expansion coefficient. Equation 19 can be used to get the $mbest$ when n is the swarm's capacity.

$$mbest_i = \frac{1}{n_i}\sum_{i=1}^{n} pbest_i(t) \tag{35}$$

c. Improved QPSO

In this study an improved version of QPSO that is "Opinion Leader-Based Quantum Particle Swarm Optimization (OLB-QPSO)" is used. In this context, an opinion leader is a person who often shares knowledge, viewpoints, and counsel with others, so exerting influence on their choices. Opinion leaders gather information first, evaluate it in light of their expertise, and then distribute it to others before it reaches the general public. As a result, the opinion leader has the power to affect how others interpret the information. The phenomenon mentioned earlier in QPSO demonstrates how the particle with the best fitness in relation to the others is related to the creation of mainstream thinking. The OLB-QPSO best may be attained by

International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 13s, 2025
https://www.theaspd.com/ijes.php

$$mbest_i = \alpha \times pbest_i^l(t) + (1-\alpha)/(n-1) \times \left(\sum_{e=1}^{n} pbest_i^e(t) - pbest_i^l(t)\right)$$
(36)

Where $pbest_i^l(t)$ is the pbest serving as an opinion leader, $pbest_i^e(t)$ is the pbest of the eth particle, α is the opinion leader's weight, and n is the swarm capacity. Experience has shown that setting α to a value between [0.3, 0.5] yields better results. Assuming that the swarm has a capacity of three, Fig.6 illustrates the evolution of the mbest from equation 21 and 22. As a result, the mbest may get close to the appropriate pbest.

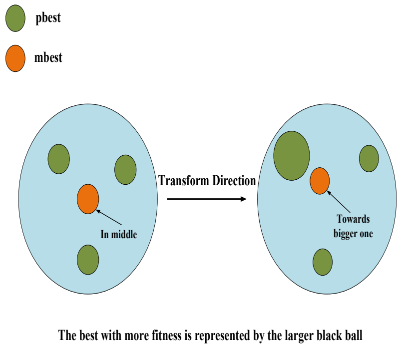

Fig.6: Evolution of mbest

By following these steps, we systematically explored and selected the most suitable model for our prediction problem, ultimately improving the accuracy and effectiveness of our system. These models enabled AI-driven performance analysis to identify players' strengths and shortcomings, guiding training programmes and strategic recommendations. We also used simulations and predictive analytics to help coaches and management predict opponent tactics, optimize game strategies, and make data-driven judgments during matches. Our iterative process improved strategy and performance analysis by incorporating fresh data and feedback.

IV. RESULT AND DISCUSSION

We analysis the performance of proposed and existing methods such as Random Forest [19], Recurrent Neural Network [20] using the metrices such as accuracy, precision, Recall and F-measure.

a. Accuracy

An organization model's accuracy is determined by how many of its predictions turn out to be accurate. Equ (37) the following is used to compute it:

$$\mathfrak{S} = \frac{\mathfrak{P}+\mathfrak{N}}{\mathfrak{P}+\mathfrak{N}+\mathfrak{A}+\mathfrak{F}} \quad (37)$$

True Positive ($\mathfrak{P}$) represents the number of Insider threat cases that have been classified as such and appropriately determined. True Negative ($\mathfrak{N}$) indicates the number of non-insider threats that were accurately classified as such. False positive ($\mathfrak{F}$) indicates the number of non-insider threats that were incorrectly assigned to the threats classification. False negative ($\mathfrak{A}$) signifies the number of insider threats that were incorrectly classified as insider threats. Fig. 3 indicates the Accuracy and Table 3 represents the numerical outcomes of accuracy.

TABLE 1 Numerical Outcomes of Accuracy

No of Epochs (x-axis)	Accuracy (%)- (y-axis)		
	RF	RNN	Proposed
10	55	65	75
20	60	70	80
30	70	80	85
40	80	85	90
50	82	85	96

International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 13s, 2025
https://www.theaspd.com/ijes.php

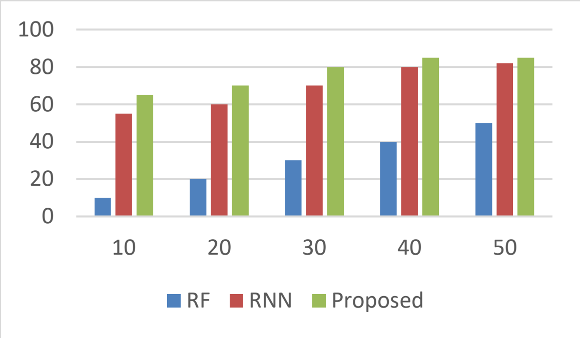

Fig. 7 Number of Epochs vs. Accuracy (%)

The Number of epochs vs. Accuracy (%) and the results of Accuracy (%) are shown in Fig 7 and Table 1. The graph displays the accuracy % for each of the three approaches to the user base. The accuracy starts at 10 epochs and goes up to 10% for RF, 55%, for RNN 65%, and 75% accuracy for the suggested. The suggested method achieves 85% accuracy at 30 epochs for RNN 80% and RF 70% respectively. Ultimately, the suggested method achieves a high accuracy of 96% at 50 epochs, which is better than the 85% accuracy for RNN and the 82% accuracy for RF.

b. Precision

This metric indicates the percentage of effectively-identified positive predictions. The following equ (38) is used to determine the precision.

$$Precision = \frac{\mathcal{J}}{\mathcal{L}+\mathcal{U}} \tag{38}$$

TABLE 2
Numerical results of precision comparison

Number of epochs (x-axis)	Classification Precision (%) (Y-axis)		
	RF	RNN	Proposed
10	40	50	60
20	50	60	70
30	60	70	80
40	70	80	85
50	80	83	95

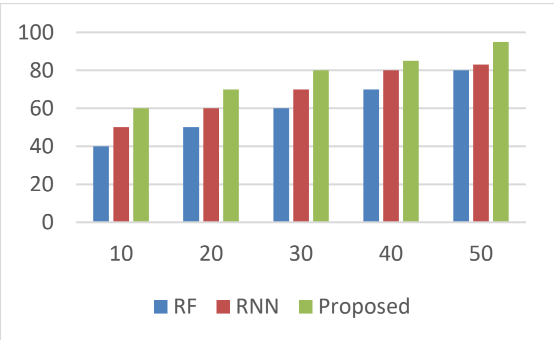

Fig. 8 Number of Epochs vs. Precision (%)

A comparison of the precision analysis is shown in Table 4 and Fig 4. The previously mentioned figure makes it transparent that the suggested approach achieves 95% classification precision at 50% repetitions.

International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 13s, 2025
https://www.theaspd.com/ijes.php

In contrast, the current method RNN 85%achieves and RF 80% precision iterations; while RNN achieves 70% precision for 30 epochs and RF for 60% of the 30 epochs and improves the proposed method for 80%. whereas the existing approach the10 epochs for RF 40% of accuracy and 50% of RNN of the precision it improves the 60% of the proposed method. The results for the suggested approach achieve 95% classification precision at 50% repetitions. In contrast, the current method RNN 83%achieves and RF 80% of the precision.

c. F-measure

An equation that compares the System F-measure with the number of epochs is given by

$$F(n) = \frac{F_{max}}{1+n^{-k(n-n_0)}} \tag{39}$$

where,

- $F(n)$ is the F-measure after n epochs.
- F_{max} is the maximum possible F-measure.
- n is the number of epochs.
- k is a positive constant that controls the growth rate.
- n_0 is the epoch at which the F-measure reaches half of its maximum value.

The above equation demonstrates that the number of epochs and the maximum F-measure that each epoch is capable of achieving independently are correlated with the system's overall performance. If all other conditions stay the same, increasing the number of epochs or increasing the F-measure will result in a greater total performance of the system.

TABLE 3
Numerical analysis of F-Measure

(x-axis) – Number of Epochs	F-Measure (%) - (y-axis)		
	RF	RNN	Proposed
10	73.6	75.3	93.2
20	74.5	79.2	94.7
30	75.8	82.5	95.4
40	76.6	83.2	96.8
50	83	84	98

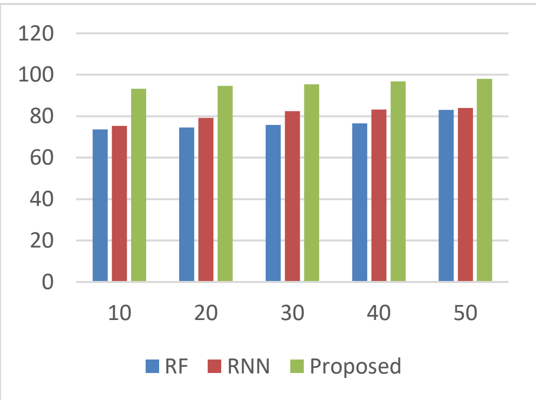

Fig. 9: Number of Epochs vs. F-Measure

Fig. 9 and Table 3 represent the numerical analysis of F-Measure. The F-Measure of three distinct approaches like GM, IOHMM and the proposed method — across a range of Epochs counts is compared numerically. The Proposed technique consistently beats RF and RNN as the number of Epochs rises from 0 to 50. The F-Measure in the proposed technique is more, for instance, 98% at 50 epochs, while RF and RNN only have 77% and 84% respectively. These findings show that the proposed method have increased F-Measure for increase in Epochs number.

d. Recall

To calculate the Recall, the following formula can be used:

$$R(n) = R_{max} \cdot n^{-n^{-k(n-n_0)}} \quad (40)$$

where,

- $R(n)$ is the Recall after n epochs.
- R_{max} is the maximum possible Recall.
- n is the number of epochs.
- k is a positive constant that controls the growth rate.
- n_0 is the epoch at which the growth rate changes most rapidly.

TABLE 4 Numerical analysis of Recall

(x-axis) – Number of Epochs	Recall (%) - (y-axis)		
	RF	RNN	Proposed
10	80.5	83.7	95.3
20	81.5	85.2	95.9
30	81.9	86.4	96.4
40	82	87.3	96.9
50	82.47	88	97

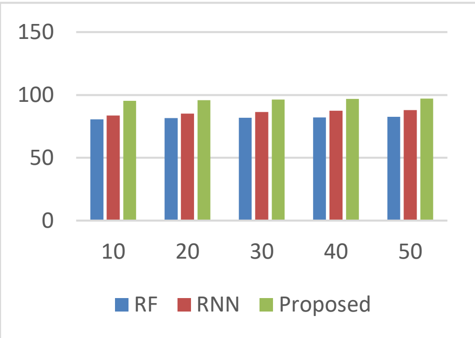

Fig. 10: Number of Epochs vs. Recall

Fig. 10 and Table 4 represent the numerical analysis of Recall. The Recall of three distinct approaches such as RF, RNN and a proposed method — across a range of epochs count is compared numerically. The Proposed technique consistently beats RF and RNN as the number of epochs rises from 0 to 50. The Recall in the proposed technique is more, for instance, 97 at 50 epochs, while RF and RNN only have 82.47 and 88 respectively. These findings show that the proposed method has increased Recall and the function of the system is enhanced.

V. CONCLUSION

Artificial intelligence (AI) and deep learning have transformed sports performance enhancement and strategy planning. Methodically applying modern data processing techniques creates robust models, as shown in this study. These methods include K-PCA for feature extraction, RF-GA for feature selection, and Z-score normalization for feature normalization. The Novel Convolutional Long Short-Term Memory with Improved Quantum Particle Swarm Optimization (Conv-LSTM-IQPSO) models are outstanding for identifying subtle patterns and correlations in performance data that standard methods would miss. These models' AI-driven performance analysis allows for extensive study of a player's strengths and weaknesses, making it easier to design personalised training plans. This tailored approach boosts team performance and player development. AI's strategic planning allows coaches and managers to mimic situations, foresee opponents' strategies, and make informed decisions that can change game outcomes. This study shows how an iterative feedback loop from data collection and model refinement maintains strategies agile and adaptable. The smooth integration of AI and deep learning has given teams and players enhanced strategic planning and performance optimization tools, a major sports science accomplishment. Technology has ushered in a new era in sports, when data-driven insights and predictive analytics are crucial to success on and off the field.

REFERENCES

[1] Yadav, J., Misra, M., Rana, N. P., Singh, K., & Goundar, S. (2023). Blockchain's potential to rescue sports: A social media perspective. In *Distributed computing to Blockchain* (pp. 405-414). Academic Press.
[2] Wang, F., Ma, M., & Zhang, X. (2024). Study on a portable electrode used to detect the fatigue of tower crane drivers in real construction environment. *IEEE Transactions on Instrumentation and Measurement*.
[3] Zhao, S., Liang, W., Wang, K., Ren, L., Qian, Z., Chen, G., ... & Ren, L. (2023). A multiaxial bionic ankle based on series elastic actuation with a parallel spring. *IEEE Transactions on Industrial Electronics*.
[4] Westerbeek, H., & Karg, A. (2022). International sport business: current issues, future directions. Routledge.
[5] Yadav, J., Misra, M., Rana, N. P., Singh, K., & Goundar, S. (2023). Blockchain's potential to rescue sports: A social media perspective. In *Distributed computing to Blockchain* (pp. 405-414). Academic Press.
[6] Miao, Y., Wang, X., Wang, S., & Li, R. (2022). Adaptive switching control based on dynamic zero-moment point for versatile hip exoskeleton under hybrid locomotion. *IEEE Transactions on Industrial Electronics*, *70*(11), 11443-11452.
[7] Cheng, K., Guo, Q., He, Y., Lu, Y., Xie, R., Li, C., & Wu, H. (2023). Artificial intelligence in sports medicine: could GPT-4 make human doctors obsolete?. *Annals of Biomedical Engineering*, *51*(8), 1658-1662.
[8] Song, L., & Guo, Y. (2023). Design of large-scale sports event management system under the internet of things CAD technology. *Computer-Aided Design and Applications*, *20*, 78-88.
[9] Bai, Z., & Bai, X. (2021). Sports big data: management, analysis, applications, and challenges. *Complexity*, *2021*(1), 6676297.
[10] Powell, D., Stuart, S., & Godfrey, A. (2021). Sports related concussion: an emerging era in digital sports technology. NPJ digital medicine, 4(1), 164.
[11] Ma, S. (2022). College sports intelligence using human–computer interaction system for education. *International Journal of Human–Computer Interaction*, 1-9.
[12] Stoeve, M., Schuldhaus, D., Gamp, A., Zwick, C., & Eskofier, B. M. (2021). From the laboratory to the field: IMU-based shot and pass detection in football training and game scenarios using deep learning. *Sensors*, *21*(9), 3071.
[13] Vijay Fidelis, J., & Karthikeyan, E. (2022). Optimization of artificial neural network parameters in selection of players for soccer match. In *Sustainable Advanced Computing: Select Proceedings of ICSAC 2021* (pp. 275-288). Singapore: Springer Singapore.
[14] Sharma, M., Monika, Kumar, N., & Kumar, P. (2021). Badminton match outcome prediction model using Naïve Bayes and Feature Weighting technique. *Journal of Ambient Intelligence and Humanized Computing*, *12*, 8441-8455.
[15] Kewei, S., Díaz, V. G., & Kadry, S. N. (2022). Evaluating the Efficiency of Student Sports Training Based on Supervised Learning. *International Journal of Technology and Human Interaction (IJTHI)*, *18*(2), 1-17.
[16] Kulakou, S., Ragab, N., Midoglu, C., Boeker, M., Johansen, D., Riegler, M. A., & Halvorsen, P. (2022). Exploration of different time series models for soccer athlete performance prediction. *Engineering Proceedings*, *18*(1), 37.
[17] Tsou, H. T., & Chen, J. S. (2023). How does digital technology usage benefit firm performance? Digital transformation strategy and organisational innovation as mediators. *Technology Analysis & Strategic Management*, *35*(9), 1114-1127.
[18] Basu, B. (2023). Perspectives on the intersection between sports and technology. In *Sports Management in an Uncertain Environment* (pp. 143-168). Singapore: Springer Nature Singapore.
[19] Ćwiklinski, B., Giełczyk, A., & Choraś, M. (2021). Who will score? a machine learning approach to supporting football team building and transfers. Entropy, 23(1), 90.
[20] Alghamdi, W. Y. (2023). A novel deep learning method for predicting athletes' health using wearable sensors and recurrent neural networks. Decision Analytics Journal, 7, 100213.