Received: 28 September 2025 | Revised: 25 November 2025 | Accepted: 11 December 2025 | Published online: 25 December 2025

RESEARCH ARTICLE

Archives of Advanced Engineering Science
2025, Vol. 00(00) 1–12
DOI: 10.47852/bonviewAAES52027778

Forecasting Licensed Athletes Numbers Using Time Series and Hybrid Artificial Intelligence Models

Halil Şenol[1] and **Halil Çolak**[2,*]

[1] *Department of Energy Systems Engineering, Giresun University, Türkiye*

[2] *Department of Physical Education and Sports, Giresun University, Türkiye*

Abstract: This study comparatively applied four distinct time series and artificial intelligence-based forecasting models to predict the short-term number of licensed athletes in Türkiye: Autoregressive Integrated Moving Average (ARIMA), Convolutional Neural Network with Long Short-Term Memory (CNN-LSTM), Extremely Boosted Gradient Decision Trees with Long Short-Term Memory (XGBoost-LSTM), and gated recurrent unit with random forest (RF-GRU). Models were trained using data from 2005 to 2019; forecasts for the years 2020–2024 were subsequently compared with actual values. The ARIMA model demonstrated the highest predictive efficacy, achieving a coefficient of determination of 0.9633, a mean square error of 0.1056, and a mean absolute error of 0.1990. The CNN-LSTM model exhibited a coefficient of determination of 0.9771 and a mean absolute percentage error of 5.68%. The remaining two hybrid models (XGBoost-LSTM and RF-GRU) exhibited inferior accuracy, accompanied by comparatively elevated error levels. The results underscore the significance of data-driven decision-making in sports policy and provide scholarly contributions to strategic planning.

Keywords: number of athletes, time series, deep learning, predictive modeling, forecasting

1. Introduction

The sports sector is one of the fastest-growing industries worldwide and is progressively influencing the global economy [1]. Engagement in sports activities not only offers individuals pleasure and respite but also enhances social bonds with family and neighbors [2]. Recent years have seen a shift in the emphasis on medal rankings and international achievement, long viewed as markers of elite performance, toward strategic approaches that prioritize community-based advantages. When strategically and efficiently organized, sports become a potent instrument that can facilitate not only competitive success but also the attainment of various social and societal [3]. This shift in methodology reinterprets the societal roles of sports, placing athletes as exemplars who operate not solely for personal achievement but also with a commitment to social duty [4]. Numerous studies have identified sport as a tool for social cohesion, cultural integration, and public policy development, particularly in areas such as community engagement, youth empowerment, and public health promotion [3, 5, 6].

Currently, scholarly discourse increasingly highlights the substantial influence of the sports business on economic development, with a corresponding rise in the volume of research produced in this domain. Recognition of the strategic significance of sports as both a cultural and social domain, as well as an economic sector, is increasing [7]. Consequently, it is imperative for nations with youthful demographics, such as Türkiye, to use this potential and promote youth engagement in sports through robust sports programs. Nonetheless, the sustainable formulation of such policies necessitates data-driven strategic planning underpinned by predictive projections. In this respect, artificial intelligence (AI)-driven forecasting techniques (e.g., time series) and statistical modeling instruments can be effectively employed in both the feasibility assessment of sports investments and the long-term strategic planning of decision-makers [8]. Such foresight provides sensible foundations for infrastructure expenditures as well as essential matters such as athlete training, facility development, and regional distribution [4].

Time series analysis comprises statistical techniques used to forecast future trends and potential variations based on historical data [9]. In time series analysis, the Autoregressive Integrated Moving Average (ARIMA) model [10], Extreme Gradient Boosting algorithm (XGBoost) [11], Long Short-Term Memory (LSTM) network [12], gated recurrent unit (GRU) [13], random forest (RF) algorithm (10.1016/j.asoc.2025.113274), and Convolutional Neural Network (CNN) [14] are utilized. Traditional statistical models, as well as contemporary machine learning (ML) and deep learning methodologies, are extensively employed.

There has been a substantial growth in studies utilizing ML for sports prediction in recent years. Artificial Neural Networks (ANN), classification algorithms, k-fold cross-validation, and feature selection are often employed techniques for forecasting sports outcomes [15]. Hybrid models utilizing RF, Gradient Boosting, Support Vector Machine, Logistic Regression, and CNN in football and

***Corresponding author:** Halil Çolak, Department of Physical Education and Sports, Giresun University, Türkiye. Email: halil.colak@giresun.edu.tr

basketball betting have produced economically substantial returns [16, 17]. Research employing ANN in several sports has attained accuracy over 67% [18], while at the individual level, the error margin for elite swimming performance has been determined to be about 0.05 seconds [19]. Hybrid models incorporating Recurrent Neural Network (RNN), LSTM, and Residual Neural Networks have shown superior accuracy relative to traditional Logistic, Weibull, and Hyperbolic models in forecasting the dimensions of the sports business [20]. In the prediction of sports injury risk, image-based coding methodologies integrated with Deep Neural Network architectures attained an accuracy enhancement of over 20% [21]. Knee joint forces were modeled using wearable sensor data, demonstrating a strong correlation ($R^2 = 0.94$) [22], while fuzzy rules attained a training modeling accuracy of 68.7% [23].

Hybrid models also excel in performance forecasting. An ANN optimized using chaos theory and particle swarm optimization (PSO) attained over 90% accuracy in predicting sports performance [24]. Support Vector Machine surpassed ANN and Genetic Algorithm (GA) in forecasting material performance in sports equipment [25]. Classical ML approaches were evaluated in tennis and cricket matches; accuracy in tennis forecasts did not surpass 70% [26], whereas tree-based models showed greater efficacy in cricket. RF attained an accuracy of up to 95% in squad selection [27], whereas AdaBoost excelled in the Internet of Things-based warning system for football [28]. support vector classifier achieved the highest success rate, with an F1 score of 97.6% in the categorization of sports news [29], although Support Vector Machine and RF produced robust results in analyses including heart rate variability data [30]. Gradient boosting decision trees exhibited superior efficacy in post-concussion recovery among athletes [31], whereas the integration of CatBoost and UMAP attained 75% accuracy in athlete profiling based on coach assessments [32].

Projecting the quantity of licensed athletes is not solely a descriptive task but a strategic necessity for evidence-based sports administration. Precise forecasts underpin various policy areas:

- Funding allocation: Ministries and federations strategize staff, event subsidies, and athlete development spending according to anticipated participation levels. Infrastructure development: Long-term facility planning, encompassing sports halls, youth centers, and grassroots initiatives, necessitates dependable demand forecasts rather than historical data.
- Human resource planning: Coaching, referee training, and educator allocation are contingent upon anticipated athlete demographics across various locations and age categories. Health and social policies: Enhancing sports participation is a fundamental approach to addressing physical inactivity, obesity, and chronic diseases in Türkiye. Consequently, forecasting future involvement facilitates the alignment of sports planning with public health policy.
- Talent identification and Olympic planning: National federations require dependable projections to identify possible elite talent pools and develop long-term performance trajectories.

Thus far, hardly any research has established a quantitative framework to correlate athlete participation trends with prospective policy scenarios, especially through the utilization of AI-enhanced modeling. This study addresses a significant gap by providing forecasting methods relevant to strategic planning, resource optimization, and national sports policy.

This study employed ML modeling for the period 2005–2019, utilizing athlete data from 2005–2024, and projected the values for 2020–2024, which were subsequently compared with actual data. Consequently, the models' predictive power and accuracy were evaluated. The primary premise of the study posits that ML models utilizing historical data can yield significant and statistically valid predictions for the future. Research on predicting athlete numbers is scarce in the literature [4, 33], and no unique prediction model for Türkiye has been identified. This work provides a novel contribution to the domain of sports sciences.

2. Theoretical Method

2.1. Dataset

The research, performed on Türkiye's athletic population from 2005 to 2024, utilized data from the Ministry of Youth and Sports of the Republic of Türkiye [34] and the DrDataStats [35] data platform (Table 1). Despite the availability of annual data, the structure's inclusion of only the total annual athlete population constrains detailed analysis by discipline, age group, region, or gender; consequently, the modeling effort was executed at a general level. All raw values used in this study are publicly accessible via the Ministry of Youth and Sports and DrDataStats platforms and are fully reproducible using the values presented in Table 1.

A key limitation of the dataset is its aggregate structure, which reports only annual national totals without breakdowns by gender, sport, or region. This restricts the modeling to a univariate framework and limits the interpretability of sport-specific trends.

Table 1
Annual number of licensed athletes in Türkiye (2005–2024)

Years	Number of athletes (million)
2005	0.9178
2006	1.1230
2007	1.2629
2008	1.4694
2009	1.6213
2010	1.7648
2011	1.9512
2012	2.3312
2013	2.8178
2014	3.2193
2015	3.5343
2016	3.8416
2017	4.4288
2018	4.9080
2019	6.4726
2020	8.4541
2021	10.9139
2022	12.4500
2023	16.6783
2024	16.9277

Values are reported in millions and based on official records from the Ministry of Youth and Sports of the Republic of Türkiye [34] and DrDataStats [35] platform.

2.2. Long short-term memory

LSTM is a variant of RNN that demonstrates exceptional efficacy in learning patterns within sequential data. Due to its cell state architecture, it can encapsulate long-term dependencies [36]. The forget gate determines the deletion of previous information,

Figure 1
Long Short-Term Memory network architecture

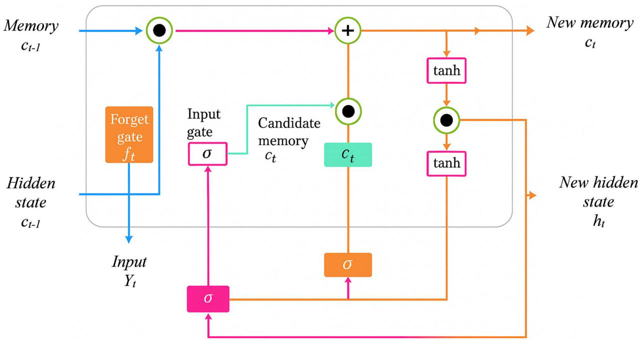
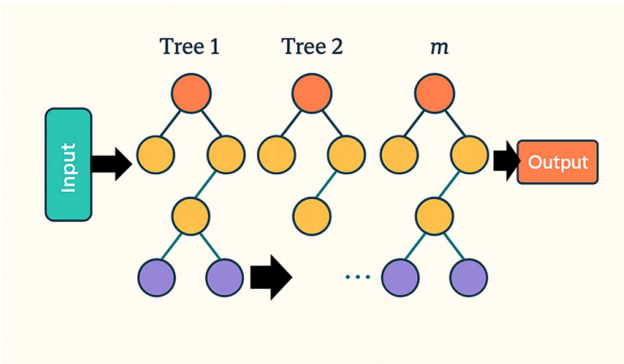


Figure 2
Extreme Gradient Boosting network architecture



the input gate governs the addition of fresh information, and the output gate selects which information from the cell to utilize [37]. These mechanisms govern the information flow, allowing LSTM to proficiently represent long-term dependencies in sequential data (Figure 1) [38].

The fundamental equations of LSTM are presented in Equations (1)–(5) [39].

$$\text{Input gate: } i_t = \sigma(W_i.[h_{t-1}, x_t] + b_i) \tag{1}$$

$$\text{Candidate cell state: } \tilde{C} = \tanh(W_c.[h_{t-1}, x_t] + b_c) \tag{2}$$

$$\text{Updated cell state: } C_t = f_t.C_{t-1} + i_t.\tilde{C}_t \tag{3}$$

$$\text{Output gate: } o_t = \sigma(W_o.[h_{t-1}, x_t] + b_o) \tag{4}$$

$$\text{Output calculation: } h_t = o_t.\tanh(C_t) \tag{5}$$

The fundamental architecture of LSTM facilitates the modeling of time-dependent interactions by regulating information flow via various gates and states. This structure comprises two sigmoid activation functions and one hyperbolic tangent (tanh) function, which are essential for governing intracellular transitions and information updates. The tanh function regulates internal memory values by constraining the cell state within the range of -1 to 1, whereas the sigmoid function generates outputs between 0 and 1, influencing the degree of information retention or obsolescence. The weight matrices (W_i, W_f, $W_{o,}$ ve W_c) utilized for the input, forget, output, and cell state gates are trainable parameters in this process, each serving a pivotal function in temporal information transfer.

2.3. Extreme gradient boosting

XGBoost is an ensemble learning technique utilizing gradient boosting, which is predicated on the sequential construction of decision trees [40]. At each iteration, the model produces new trees by allocating greater weight to instances where the preceding learner erred, hence dynamically enhancing the learning process [41]. XGBoost fundamentally integrates the predictions of numerous weak learners to formulate a robust classifier with elevated accuracy. In this procedure, each tree is trained to minimize a regularized objective (loss) function by gradient-based optimization. This mitigates the model's learning inaccuracies and enhances overall efficacy by regulating the danger of overfitting (Figure 2) [42].

The general equation of the XGBoost approach is shown in Equation (6) [43].

$$L(\theta) = \sum_{i=1}^{N}(y_i - f(x_i))^2 + \Omega(f) \tag{6}$$

$L(\theta)$: This function enhances both the predictive accuracy and generalization capacity of the model.

$\Omega(f)$: This function enhances the predictive accuracy and generalization capacity of the model.

$\sum_{i=1}^{N}(y_i - f(x_i))^2$: It is the aggregate of the squared discrepancies between the values forecasted by the model and the actual values, where y_i denotes the actual value, $f(x_i)$ signifies the model's prediction, and N indicates the number of samples.

2.4. Gated recurrent unit

GRU is a deep learning architecture that maintains long-term information via its gating mechanism. Throughout the training phase, extraneous information is systematically discarded, thereby diminishing model complexity [44]. GRU networks, due to their streamlined architecture, serve as an alternative to LSTM and are especially appropriate for resource-limited settings. They regulate information flow utilizing solely update and reset gates, rather than the intricate gate mechanism of LSTM. This architecture necessitates fewer processing resources and can surpass LSTM in certain contexts when modeling long-term interdependence [45]. The GRU design comprises two principal gates: the reset gate and the update gate. The reset gate regulates the degree of information retention from previous time steps and governs the integration of incoming information. The update gate integrates the roles of the distinct forget and input gates in the LSTM, ascertaining the extent to which prior memory is retained and how it is modified with contemporary information. This configuration enables the GRU to regulate the information flow more effectively and straightforwardly (Figure 3) [46].

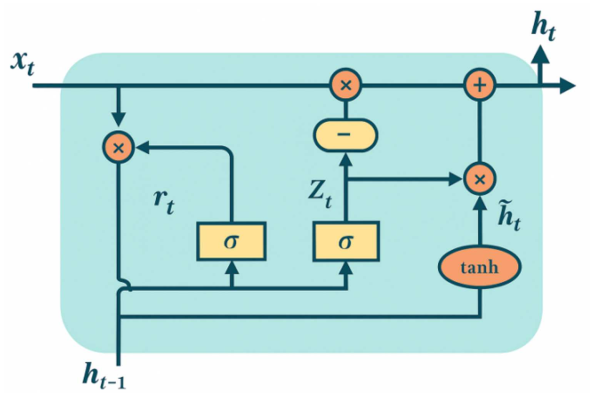
The general equations for the GRU approach are presented in Equations (7)–(9) [47].

$$z_t = \sigma(w_z x_t + u_z h_{t-1}) \tag{7}$$

$$r_t = \sigma(w_r x_t + u_r h_{t-1}) \tag{8}$$

$$h_{t=}(1 - z_t) \otimes h_{t-1} + z_t \otimes \tilde{h}_t \tag{9}$$

Figure 3
Gated recurrent unit network architecture

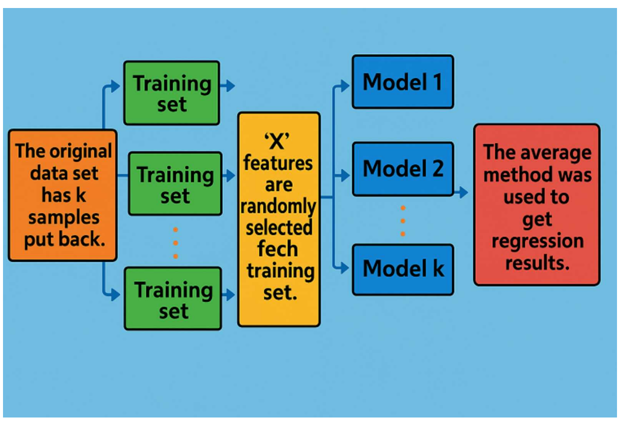


The variables W_z, U_z, W_r, and U_r in Equations (7)–(9) denote the trainable weight vectors of the model. The symbol denotes the element-wise (Hadamard) multiplication operation between two vectors; the phrase $\tilde{h}_t$ signifies the unit value of memory, and rt indicates the sigmoid activation function.

2.5. Random forest

RF is an ensemble learning system designed to enhance the restricted generalization ability of an individual decision tree utilizing the bagging method. The model's accuracy and stability are markedly enhanced by training several decision trees using random sampling and feature subsets [48]. During this procedure, numerous subsets of data are derived from the main dataset through random and iterative sampling, with each subset employed to train a distinct decision tree. During the construction of each decision tree, a randomly chosen subset of features is utilized rather than all available features in the node splitting procedure [49]. This randomization enhances model variety and mitigates the danger of overfitting. During the final prediction phase, the outputs of all decision trees are aggregated; majority voting is employed for classification tasks, whereas predictions are averaged for regression tasks. Consequently, the model's generalization capability is enhanced, yielding more balanced and stable outcomes [50] (Figure 4).

Figure 4
Random forest network architecture



Equation (10) contains the general equation for the RF method [51].

$$f(x) = \frac{\sum(T_{n(x)})}{n} \tag{10}$$

$T_{n(x)}$: It signifies the forecast of the nth decision tree for the input x.

$\sum(T_{n(x)})$: It constitutes the aggregate of the forecasts from all trees.

n: The aggregate quantity of decision trees.

2.6. Convolutional neural networks

A CNN is a neural network distinguished by convolutional operations and a profound architecture. It is typically employed to derive more abstract and complex features from raw data [52]. CNN comprises a convolution layer, an activation function, a pooling layer, a dropout layer, a fully connected layer, and an output layer (Figure 5) [53]. The convolution layer, a fundamental component of a CNN, produces feature maps by applying convolution kernels to the input data with designated stride lengths; initial layers typically extract superficial features, whereas final layers extract deep structural features. Activation functions (including ReLU, sigmoid, and tanh) provide nonlinearity to the model, facilitating the learning of more intricate correlations. A random dropout layer intermittently deactivates certain neurons or connections during training to mitigate overfitting. Consequently, the model's generalization capability enhances [54].

Equation (11) presents the universal formula for CNN [52].

$$z_{i,j}^{(k)} = \sum_{m=1}^{M} \sum_{n=1}^{N} x_{i+m-1,j+n-1} \times w_{m,n}^{(k)} + b^{(k)} \tag{11}$$

In this equation;

$z_{i,j}^{(k)}$: The output value at the (i, j) th point in the feature map generated by the kth convolution filter.

$x_{i+m-1,j+n-1}$: The pixel value at the (i+m-1, j+n-1) location of the input picture.

$w_{m,n}^{(k)}$: The weight of the kth filter at the (m, n) th location.

$b^{(k)}$: is the bias term for the kth idea.

2.7. Autoregressive integrated moving average

ARIMA is one of the most widely used models in time series analysis [9]. Developed by George Box and Gwilym Jenkins in the 1970s, this approach provided a comprehensive framework for examining stochastic time series [55]. ARIMA provides high success in short- and medium-term forecasts for series containing statistically significant patterns such as trends or seasonality,

Figure 5
Convolutional Neural Network architecture

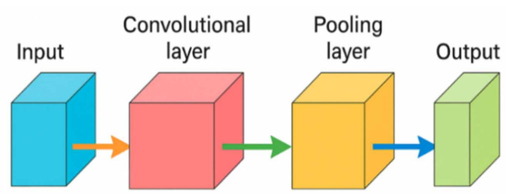
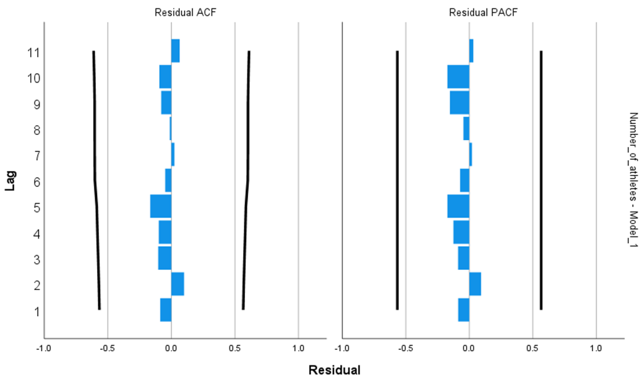


Figure 6
ACF and PACF graphs used in the refinement model



independent of exogenous factors [56]. By utilizing past values and error terms, the model produces effective forecasts, particularly in systems unaffected by exogenous shocks. ARIMA is based on three basic components: Autoregressive (AR) represents past observations, Moving Average (MA) represents past errors, and Differentiation (I) represents the series' stationarity [57]. The combination of these components forms ARMA, and when combined with differencing, forms ARIMA [58]. The ADF test is used to test stationarity in model selection, while autocorrelation function (ACF), partial autocorrelation function (PACF) (Figure 6), and Akaike information criterion (AIC) criteria are used to determine the parameters [59].

Equation (12) presents the generic formula for ARIMA [60].

$$Y_t = u + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ldots + \phi_1 \varepsilon_{t-1} + \phi_2 \varepsilon_{t-2} + \ldots + c \quad (12)$$

In time series analysis, μ denotes the mean level of the series. The φ_1 coefficient in the AR model represents the autocorrelation coefficient, signifying the influence of historical values on the present forecast. θ_1 in the MA model represents the moving average coefficient, which dictates the influence of historical error terms on the prediction. The present value is predominantly contingent upon the preceding value Y_{t-1}. The ε_t error term denotes random variations from a distribution with a zero mean and constant variance, whereas the c term signifies the stationary component of the model.

2.8. Error performance metrics

In the modeling of the number of licensed athletes in Türkiye for the years 2005–2019, root mean square error (MSE), mean square error (RMSE), sum of squared errors (SSE), mean absolute percentage error (MAPE), R^2, and mean absolute error (MAE) metrics were used (Equations 13–18) [61].

$$MSE = \sum_{i=1}^{n} \frac{(y_i - \hat{y}_i)^2}{n} \quad (13)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}} \quad (14)$$

$$SSE = \sum_{i=1}^{n} (Y_i - \hat{y}_i)^2 \quad (15)$$

$$MAPE = \frac{100}{n} \sum i = 1 - \frac{|y_i - \hat{y}_i|}{y_i} \quad (16)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(Y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad (17)$$

$$MAE = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n} \quad (18)$$

In this context, $\hat{y}i$ denotes the estimated or model-predicted value of an observation, while yi signifies the real or expected value of that observation, and n denotes the aggregate number of observations within the analyzed data collection [61].

The model selection in this study was based on theoretical diversity, interpretability, and suitability for small-sample univariate time series data. ARIMA was included as the classical benchmark for linear-trend modeling, while CNN–LSTM and XGBoost–LSTM were selected to represent deep learning and hybrid gradient boosting approaches capable of capturing nonlinear temporal dependencies. RF-GRU was employed as an alternative architecture that integrates ensemble learning with gated recurrent memory. Transformer-based models and tree-based variants such as LightGBM were excluded due to the extremely limited dataset size (15 observations) and the absence of multivariate features, which substantially increases the risk of overfitting and unstable convergence in such architectures. Future studies using larger and higher-frequency datasets may appropriately test these models.

2.9. Data preprocessing and model pipeline

Since the dataset consists of a single annual variable (total number of licensed athletes), no feature selection was required. The raw values were normalized using MinMaxScaler in the range of 0–1 prior to training and inverse transformed after prediction. For deep learning-based models, the data were converted into a supervised learning format using sliding windows. Depending on the model, the input sequence length (look-back) was set to 1 (XGBoost–LSTM), 3 (CNN–LSTM and RF–GRU), or 0 (GRU-only input from RF stage).

For ARIMA, stationarity was evaluated through differencing based on ADF test results and implemented using a (1,3,1) order.

All models used the same training/testing structure: 2005–2019 for training (70%) and 2020–2024 for testing (30%), preserving chronological order to avoid data leakage. Models were trained separately, and no cross-validation was applied due to the extremely limited dataset size (n = 15) [62, 63].

The complete modeling pipeline is shown below:

- Data extraction.
- Scaling via MinMaxScaler.
- Temporal reshaping into a supervised format.
- Model-specific training.
- Testing against 2020–2024 observations.
- Inverse transformation and evaluation using R^2, MSE, RMSE, SSE, MAE, and MAPE.

These steps ensure that the entire process can be replicated using the reported parameters and dataset. Hyperparameters for RF, XGBoost, and neural architectures were tuned using constrained grid search to reduce overfitting risk given the small sample size. The search space and selected ranges are now explicitly included to improve reproducibility.

While all model-specific hyperparameters and architecture settings are documented in Sections 3.2–3.5, this consolidated pipeline summary is intended to support transparency and reproducibility by presenting the end-to-end workflow in a single location.

The look-back windows used in this study (1–3 time steps depending on architecture) were selected based on the input dimensionality constraints of annual univariate series. Given that the dataset contains only 15 values, longer window sizes would have resulted in excessive input dimensionality relative to the available training samples, increasing the risk of overfitting. The chosen window lengths therefore represent the maximum sequence depth that can be applied without reducing the effective sample size or destabilizing model convergence.

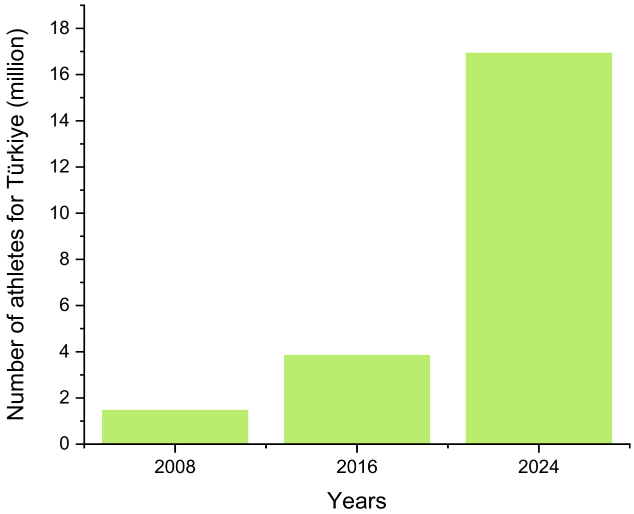
As the dataset consists of only 15 observations, alternative baseline models such as simple exponential smoothing or Holt–Winters were considered but excluded due to insufficient data points; however, these models are recommended for future comparison when additional historical data become available. Hyperparameter tuning was conducted using a constrained manual grid search. Initially, the RF model was designed to forecast the subsequent value using the previous three observations (look_back = 3). Input vectors were generated using a three-step sliding window. Grid search optimization resulted in the best performance with n_estimators = 100 and max_depth = 4. The RF model effectively captured short-term trends in the dataset while limiting sensitivity to nonlinear fluctuations and reducing overfitting risk.

3. Results and Discussion

3.1. Overview of the number of licensed athletes in Türkiye

The graph in Figure 7 indicates that the number of licensed athletes in Türkiye was 1.4694 million in 2008, rose by 161.5% to 3.8416 million in 2016, and escalated to 16.9277 million in 2024, reflecting a rise of 340.7%. The total increase between 2008 and 2024 is 1052.1%. This expansion is linked to the proliferation of grassroots sports, the advancement of school athletics, the enhancement of sports infrastructure, the improvement of facilities, and heightened individual awareness [4, 33]. Long-term planning, the promotion of sports culture, and ongoing infrastructure investments are essential for sustainability.

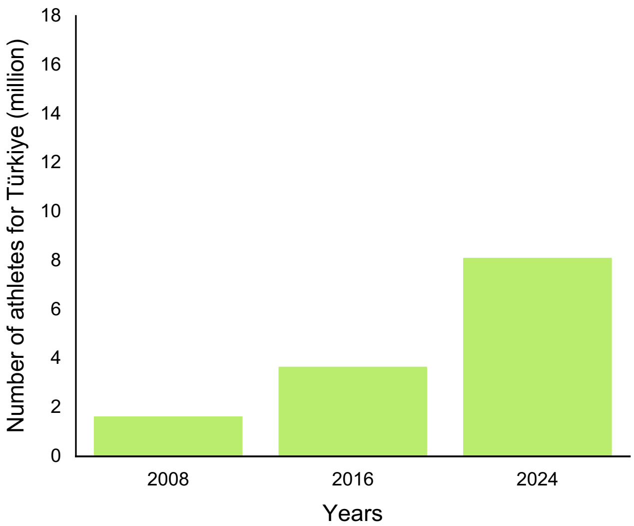
Figure 7
Change in the number of athletes in Türkiye over the years



3.2. Estimating the number of athletes in Türkiye using the XGBoost-LSTM hybrid model

This work established a hybrid methodology integrating XGBoost and LSTM models for forecasting time series data. The model was constructed using a univariate dataset including 15 time steps spanning the years 2005 to 2019. Short-term dependencies were modeled using XGBoost, and the results were input into LSTM. The dataset was partitioned into 70% for training and 30% for testing, maintaining the sequential integrity of the time series. The XGBoost model was fine-tuned with hyperparameters n_estimators=200, max_depth=5, and learning_rate=0.05, and generated predictions utilizing a historical time step (look-back=1). The LSTM architecture has 50 neurons within a solitary hidden layer and was trained for 300 epochs via a tanh activation function and an Adam optimization technique. This hybrid methodology integrates XGBoost's capacity to identify nonlinear interactions with LSTM's proficiency in learning temporal patterns, facilitating the concurrent acquisition of short- and long-term dependencies. Model estimations indicate that the quantity of licensed athletes in Türkiye is:

Figure 8
Forecasted number of licensed athletes in Türkiye using the Extreme Gradient Boosting–Long Short-Term Memory (XGBoost–LSTM) model (2005–2024)



- 2008: 1.61 million
- 2016: 3.65 million (126.71% increase)
- 2024: 8.09 million (121.64% increase)

An anticipated increase of 6.48 million is expected from 2008 to 2024, with a growth rate of 402.48% (Figure 8). The forecasts indicate a consistent and swift rise in the quantity of licensed athletes in Türkiye, with notably robust expansion persisting post-2016.

3.3. Estimating the number of athletes in Türkiye using the RF-GRU hybrid model

This study employed a hybrid methodology integrating RF and GRU models to address a time series forecasting challenge with constrained data availability (merely 15 numerical data points).

Owing to the limited dataset, the training and test data were assessed concurrently.

RF Initially, the RF model was designed to forecast the subsequent value utilizing three preceding values (look_back = 3). Input vectors were generated with a three-step sliding window technique. The optimal performance from grid search optimization was attained with hyperparameters n_estimators = 100 and max_depth = 4. The RF model effectively identified the short-term trend of the time series, minimizing sensitivity to nonlinear correlations and overfitting.

GRU step: The predictive outcomes of the RF model were inputted into the GRU during the subsequent step. The GRU model was set with a look_back of 0, indicating that only immediate RF forecasts were utilized rather than time-lagged series. The model design comprises a GRU layer featuring 64 hidden units and utilizes a "tanh" activation function. The training procedure spanned 300 epochs, with mean squared error as the loss function.

The RF-GRU hybrid model integrates the short-term forecasting abilities of RF with the nonlinear learning capabilities of GRU to generate more accurate predictions. The exclusive use of RF outputs, without the transfer of prior knowledge in the GRU layer, streamlined the model's learning architecture.

Data Preparation and Scaling: Inputs for the GRU model were normalized using MinMaxScaler and subsequently inverse transformed to the original scale post-estimation. This supplied the neural network with a suitable range of values during training and guaranteed that the outputs were translated into real-world values. According to the graph in Figure 9:

- 2008: 1.75 million licensed athletes
- 2016: 4.52 million licensed athletes (2008–2016 increase: 158.29%)
- 2024: 7.35 million licensed athletes (2016–2024 increase: 62.61%)

The cumulative gain from 2008 to 2024 surpasses 320%. The model indicates a consistent increasing trend in the number of licensed athletes in Türkiye; Nevertheless, the rate of increase from 2008 to 2016 surpasses that from 2016 to 2024.

3.4. Estimating the number of athletes in Türkiye using the CNN-LSTM hybrid model

This work built a hybrid deep learning model that integrates CNN and LSTM architectures to enhance the accuracy of predictions on time series data.
Model Structure:

- Input Layout: Future values were predicted using the past three time steps (look_back = 3).
- Data Preparation: Data were divided into three-step moving windows and normalized to 0–1 with MinMaxScaler.

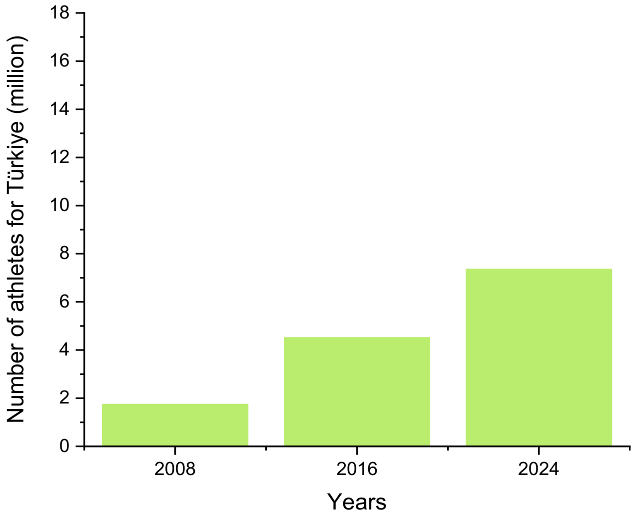
Architecture:

- Conv1D layer (64 filters, 2-unit kernel) to learn short-term patterns.
- LSTM layer (50 neurons) to model long-term dependencies.
- Dense layer to generate single-valued predictions.

Training Details:

- Optimization: Adam algorithm
- Loss function: MSE
- Training: 300 epochs

Figure 9
Forecasted number of licensed athletes in Türkiye using the random forest–gated recurrent unit (RF-GRU) model (2005–2024)



Data Split: Chronologically 70% training, 30% testing

Number of Licensed Athletes Predictions

- 2008: 1.55 million
- 2016: 4.11 million (2008–2016 increase: 165.16%)
- 2024: 9.71 million (136.27% increase from 2016 to 2024)
- Total increase (2008 to 2024): 526.45%

The model forecasts a persistent and significant upward trajectory in the quantity of licensed athletes in Türkiye, with a more pronounced absolute increase particularly following 2016 (Figure 10).

3.5. Estimating the number of athletes in Türkiye using ARIMA

The ARIMA model utilized first-order autoregressive and moving average elements, in conjunction with third-order differencing. ACF and PACF assessments of the model residuals indicated that the residuals were contained within ±95% confidence intervals and had characteristics of white noise. The AIC value was determined to be 4.80, signifying that the model demonstrated adequate accuracy and possessed an ideal parameter count. The ARIMA model forecasts indicate that the number of licensed athletes rose from 1.38 million in 2008 to 3.89 million in 2016 (an increase of 181.88%) and thereafter to 22.09 million in 2024 (a total increase of 1500%) (Figure 11).

3.6. Actual and forecast of Türkiye's athlete numbers between 2020 and 2024

This study evaluates the efficacy of the XGBoost+LSTM, RF+GRU, CNN+LSTM, and ARIMA models in forecasting the quantity of licensed athletes in Türkiye for the 2020–2024 timeframe (Figure 12). The ARIMA model generated forecasts that were most accurate for the years 2020 and 2021; nevertheless, it considerably overestimated the period of 2023–2024, demonstrating a notable limitation in its capacity to accommodate nonlinear patterns. Among deep learning methodologies, CNN+LSTM consistently yielded superior outcomes throughout all years;

Figure 10
Forecasted number of licensed athletes in Türkiye using the Convolutional Neural Network–Long Short-Term Memory (CNN–LSTM) model (2005–2024)

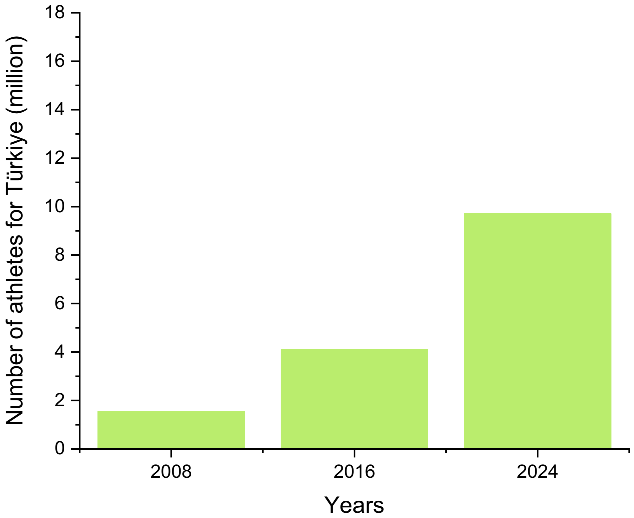


Figure 11
Estimating the number of athletes in Türkiye by year using via Autoregressive Integrated Moving Average (ARIMA) model (2005–2024).

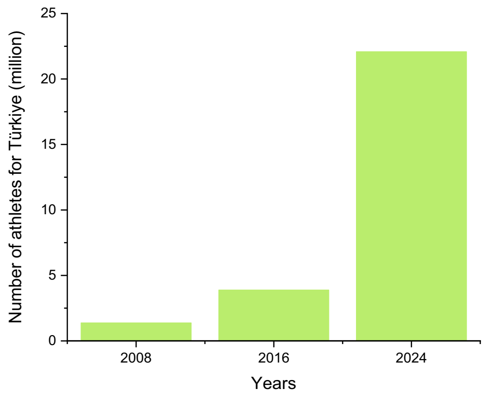
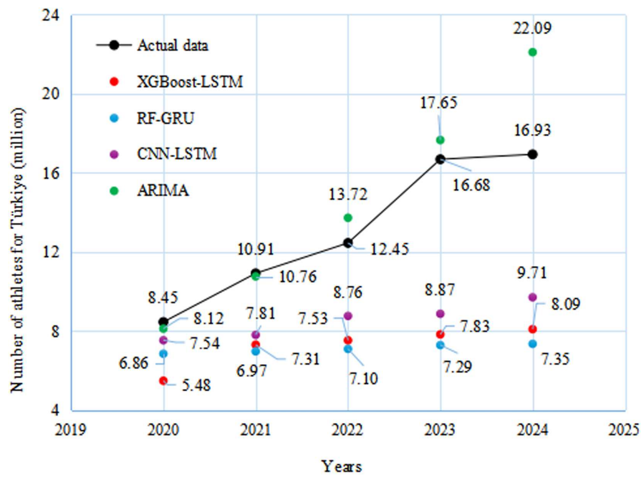


Figure 12
Comparative graph of actual and forecast methods for the number of licensed athletes in Türkiye for the years 2020–2024



nevertheless, it proved inadequate in detecting abrupt surges. The RF+GRU and XGBoost+LSTM models repeatedly undervalued the actual figures and did not accurately depict swift escalations. In the overall comparison, ARIMA showed robustness in the first phase, although CNN+LSTM yielded more consistent predictions over the long run, while the other two models displayed inferior performance.

As shown in Figure 12, ARIMA closely estimated the 2020–2021 values but substantially overpredicted the post-2022 surge, reflecting its limitations in capturing nonlinear structural breaks. This behavior is theoretically consistent with ARIMA's known tendency to extrapolate linear trends when encountering abrupt structural changes outside the learned pattern. While ARIMA provides high accuracy for stationary or near-linear series, its forecasting mechanism can intensify deterministic trend projections beyond the training interval, resulting in overestimation in rapidly increasing participation phases. Therefore, the model's strong performance metrics and its later-stage divergence are not contradictory but rather illustrate the boundary conditions under which classical models outperform deep learning in small-sample settings while still remaining vulnerable to structural regime shifts.

In contrast, CNN-LSTM maintained consistently lower error levels across all years, whereas RF-GRU and XGBoost-LSTM repeatedly underestimated the rapid growth phase, likely due to insufficient temporal depth and small-sample over-regularization. These findings highlight the importance of matching model complexity with trend behavior in rapidly expanding participation datasets.

Dalkılıç et al. [64] calculated the number of licensed wrestlers as 121,498 in 2017 and 129,890 in 2018, utilizing ANN. This study indicated that the XGBoost-LSTM model forecasted the total number of athletes to be 3.98 million and 4.30 million for the respective years.

Atasoy et al. [65] approximated the proportion of athletes in wrestling to be 3.05% of the overall athlete population in 2017 and 3.02% in 2018. Utilizing ANN, they approximated the count of licensed athletes in combat sports to be 1,157,504 in 2017 and 1,269,848 in 2018. This study indicates that the RF-GRU model forecasted a total of 5.07 million athletes in 2017 and 5.90 million in 2018, with combat athletes representing 22.83% of the total in 2017 and 25% in 2018.

Dalkılıç et al. [66] projected that 190,049 athletes engaged in underwater sports, water polo, and swimming in 2017 and 210,384 in 2018, utilizing ANN. During the same years, the ARIMA model estimated the overall number of athletes at 4.17 million and 5.11 million, respectively; these categories represented nearly 5% of the total athlete population in 2017 and 4.12% in 2018.

In the ANN study by Çolak and Şenol [33] employing the Bayesian regularization algorithm, the total number of licensed athletes in Türkiye was projected to reach 6.35 million by 2024 and was forecasted until 2030. This study observed that the XGBoost-LSTM model yielded an estimate of 1.3 times greater for the same year.

In a study using ANN, Şenol et al. [4] predicted that the total number of athletes in Türkiye would reach a maximum of 10.15 million in 2040. In this study, the lowest estimate for 2024 was found to be 7.35 million (RF-GRU), and it was determined that there would be an approximately 1.38-fold increase in the number of athletes between 2024 and 2040.

3.7. Analysis results of error performance metrics of prediction models

The ARIMA, CNN+LSTM, XGBoost+LSTM, and RF+GRU models were evaluated for their efficacy in forecasting the number of licensed athletes in Türkiye, utilizing metrics such as R^2, MSE, RMSE, SSE, MAE, MAPE, and a normalized total score (Table 2). The ARIMA model exhibited the maximum efficacy, achieving a total score of 0.9930, an R^2 of 0.9633, and minimal error values, hence demonstrating robust performance in time series fitting. The CNN+LSTM model achieved a second-place ranking (total score = 0.7870). Although proficient at identifying short-term trends, its error rates were inferior to those of ARIMA. Notwithstanding a commendable R^2 value (0.9493), the XGBoost+LSTM model exhibited subpar performance with elevated error metrics, whereas RF+GRU produced the most inferior outcomes (Total Score = 0.1488). It has been established that the standard ARIMA model yields more dependable forecasts for short-term and linear trends, but deep learning-based hybrid models offer benefits for long-term and intricate patterns.

Although AI-based hybrid models are generally expected to outperform classical approaches in nonlinear forecasting tasks, the ARIMA model yielded the highest accuracy in this study. This outcome is compatible with prior findings suggesting that deep learning architectures such as LSTM, GRU, and Transformer variants require large, high-frequency, and multivariate datasets to consistently surpass statistical models [12, 15]. In the present study, the dataset consists of only 15 annual observations and a single variable, which increases the risk of overfitting for complex models while favoring ARIMA's strengths in short-term, linear-trend prediction.

More importantly, the results contribute to an important conceptual clarification in the forecasting literature: the superiority of AI-based models is not universal but conditional on data volume, temporal granularity, and feature diversity. Recent studies have similarly shown that classical models can outperform neural architectures when sample sizes are small and exogenous predictors are absent [30]. Under such conditions, the representational capacity of hybrid models may exceed the information content of the data, leading to unstable or inferior performance. Therefore, our findings do not contradict existing theory but refine it by identifying the boundary conditions under which traditional approaches remain preferable. This highlights the importance of aligning model complexity with dataset characteristics rather than assuming algorithmic advancement guarantees improved performance.

Although formal statistical tests for forecast accuracy (e.g., Diebold–Mariano) are recommended in the literature, they could not be meaningfully applied due to the extremely small number of independent forecast-origin periods (n = 5). Future studies using higher-frequency or expanded datasets will incorporate such tests.

3.8. Future studies and recommendations

This research evaluated the efficacy of ARIMA, CNN+LSTM, XGBoost+LSTM, and RF+GRU models in forecasting the quantity of licensed athletes in Türkiye. The results indicated that ARIMA produced reliable outcomes with minimal error metrics, while CNN+LSTM exhibited exceptional performance with a high R^2 value. Nevertheless, the analysis's reliance on a singular annual series, while omitting exogenous variables, constrains its predictive efficacy. Further research, using economic indicators, demographic projections, and sports policy variables, may enhance the precision and applicability of the predictions. Moreover, the implementation of optimization-driven hybrid models (e.g., GA-ANN, PSO-LSTM) and interpretable AI tools (SHapley Additive exPlanations (SHAP), local interpretable model-agnostic explanations (LIME), partial dependence plots (PDP) would aid in minimizing forecast inaccuracies and empowering decision-makers to efficiently leverage model results. Evaluating these methodologies across many countries and locations illustrates the models' global applicability, while enhancing them with user-friendly interfaces in practice would bolster data-driven decision-making in sports strategies.

Beyond methodological extensions, the forecasts generated in this study offer direct practical value for policymakers. For example, the projected increase in athlete numbers can guide federations and the Ministry of Youth and Sports in planning annual budget allocations, coaching certification capacity, and regional facility development. A federation expecting a 40–60% increase in licensed participants within a given discipline may strategically invest in talent identification pathways, competition structures, and grassroots programs. Local governments and municipalities can also use these forecasts to prioritize infrastructure expansion in high-growth regions. Since sports participation is also linked to public health objectives, these projections provide a quantitative basis for integrating sport-based interventions into national physical activity and disease-prevention strategies. Thus, the forecasting framework presented here not only contributes methodologically but also supports multilevel evidence-based governance.

Since the dataset consists of a single univariate variable, interpretability techniques such as SHAP, LIME, or PDP would simply assign full importance to the same feature and therefore would not provide meaningful model explanations. For this reason, they were not applied in the current study. However, once multivariate datasets including policy, demographic, or economic variables become available, incorporating interpretable AI techniques will be highly valuable for improving transparency and supporting real-world adoption by nontechnical stakeholders.

Future work should expand the dataset through administrative sources that offer higher temporal resolution and disaggregated participation records, enabling multivariate modeling and explainable AI techniques. Future datasets with higher temporal resolution or multi-country structure will enable additional robustness checks,

Table 2
Analysis results of error performances of prediction methods

Methods-Metrics	R^2	MSE	RMSE	SSE	MAE	MAPE (%)	Total Score
ARIMA	0.9633	0.1056	0.3250	1.1619	0.1990	5.72	0.9930
CNN+LSTM	0.9771	4.8918	2.2117	58.7013	1.8479	5.68	0.7870
XGBoost-LSTM	0.9493	11.3652	3.3712	147.7471	2.3265	9.67	0.5332
RF+GRU	0.8929	22.9169	4.7872	275.0024	4.198	13.56	0.1488

including learning curve-based overfitting diagnostics, look-back sensitivity analyses, and transfer learning approaches. In addition to these extensions, future studies should also consider small-sample-oriented forecasting techniques such as Bayesian time series models, synthetic data augmentation, and advanced cross-validation or rolling-origin evaluation strategies. These methods were not feasible in the present work due to the limited size and structure of the dataset, but they represent promising directions for improving predictive reliability and model generalization once richer datasets are available.

Traditional overfitting diagnostics such as learning curves, rolling windows, or train–validation error monitoring were not applied because the annual univariate dataset (n = 15) is too small to produce statistically meaningful validation curves. In such conditions, well-established guidelines recommend prioritizing chronological train–test separation rather than additional splits that would further reduce the already minimal training samples [67]. However, once higher-frequency or multivariate datasets become available, such diagnostics should be incorporated.

Additionally, exponential smoothing and Holt–Winters models will be formally included as benchmark baselines in future studies once longer historical series become available.

4. Conclusion

This study examined ARIMA, CNN+LSTM, XGBoost+LSTM, and RF+GRU models for short-term forecasting of the number of licensed athletes in Türkiye. The investigation utilizing real data revealed that the ARIMA model exhibited minimal error rates (MSE = 0.1056, MAE = 0.1990) and exceptional accuracy (R^2 = 0.9633). This illustrates that classical methods remain a robust choice for predicting with stationary data. The CNN+LSTM model effectively identified patterns in the time series, achieving a high R^2 value of 0.9771 and a low MAPE of 5.68%. Conversely, the XGBoost+LSTM and particularly the RF+GRU models demonstrated subpar performance, evidenced by elevated error rates (RF+GRU MAPE = 13.56%).

Model comparisons indicated that accuracy, data structure compatibility, computational expense, and explainability are all significant factors. The ARIMA model serves as a pragmatic instrument for public institutions and policymakers, whereas deep learning and hybrid frameworks offer benefits for larger and more intricate datasets. The obtained findings offer strategic guidance for promoting grassroots sports in Türkiye and enhancing youth engagement in athletics. In this setting, data-driven sports policies are essential for infrastructure expenditures, educator recruitment, and strategic long-term planning. The research intends to furnish a definitive and empirical framework for policymakers.

Ethical Statement

This study does not contain any studies with human or animal subjects performed by any of the authors.

Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

Data Availability Statement

Data are available from the corresponding author upon reasonable request.

Author Contribution Statement

Halil Şenol: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization, Supervision, Project administration. **Halil Çolak:** Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization, Supervision, Project administration.

References

[1] Zhang, J. J., Kim, E., Mastromartino, B., Qian, T. Y., & Nauright, J. (2018). The sport industry in growing economies: Critical issues and challenges. *International Journal of Sports Marketing and Sponsorship*, *19*(2), 110–126. https://doi.org/10.1108/IJSMS-03-2018-0023

[2] Cómez-Mármol, A., Sánchez-Alcaraz Martínez, B. J., Cruz Sánchez, E. de la., Valero, A., & González-Víllora, S. (2017). Personal and social responsibility development through sport participation in youth scholars. *Journal of Physical Education and Sport*, *17*(2), 775–782. https://doi.org/10.7752/jpes.2017.02118

[3] Schulenkorf, N., Sherry, E., & Phillips, P. (2016). What is sport development? In E. Sherry, N. Schulenkorf, & P. Phillips (Eds.), *Managing sport development: An international approach* (pp. 3–11). Routledge.

[4] Şenol, H., Çolak, H., & Çolak, E. (2024). Analysis of the total number of licensed athletes using artificial neural networks for the future years in Türkiye. *Karadeniz Fen Bilimleri Dergisi*, *14*(4), 2153–2171. https://doi.org/10.31466/kfbd.1531019

[5] Coalter, F. (2013). *Sport for development: What game are we playing?* (1st ed.). Routledge. https://doi.org/10.4324/9780203861257

[6] Green, M. (2006). From 'sport for all' to not about 'sport' at all?: Interrogating sport policy interventions in the United Kingdom. *European Sport Management Quarterly*, *6*(3), 217–238. https://doi.org/10.1080/16184740601094936

[7] Li, M., Shi, Y., & Peng, B. (2022). The analysis and research on the influence of sports industry development on economic development. *Journal of Environmental and Public Health*, 3329174. https://doi.org/10.1155/2022/3329174 *2022*(1)

[8] Ratten, V., & Usmanij, P. (2020). Statistical modelling and sport business analytics. In *Statistical Modelling and Sports Business Analytics* (pp. 1–9). Routledge. https://doi.org/10.4324/9780367854454-1

[9] Şenol, H., Çolak, E., & Oda, V. (2024). Forecasting of biogas potential using artificial neural networks and time series models for Türkiye to 2035. *Energy*, *302*, 131949. https://doi.org/10.1016/j.energy.2024.131949

[10] Siddique, M. A. B., Mahalder, B., Haque, M. M., & Ahammad, A. K. S. (2025). Forecasting air temperature and rainfall in Mymensingh, Bangladesh with ARIMA: Implications for aquaculture management. *Egyptian Journal of Aquatic Research*, *51*(3), 924–303. https://doi.org/10.1016/j.ejar.2025.02.009

[11] Dehvari, M., Farzaneh, S., & Forootan, E. (2025). Forecasting rainfall events based on zenith wet delay time series utilizing eXtreme gradient boosting (XGBoost). *Advances in Space Research*, *75*(3), 2584–2598. https://doi.org/10.1016/j.asr.2024.11.013

[12] Yunita, A., Pratama, M. H. D. I., Almuzakki, M. Z., Ramadhan, H., Akhir, E. A. P., Mansur, A. B. F., & Basori, A. H. (2025). Performance analysis of neural network architectures for time series forecasting: A comparative study of RNN, LSTM, GRU, and hybrid models. *MethodsX*, *15*, 103462. https://doi.org/10.1016/j.mex.2025.103462

[13] Jin, X., Dai, C., Tao, Y., & Zhang, F. G. (2025). A comparative study of RNN, CNN, LSTM and GRU for random sea wave forecasting. *Ocean Engineering*, *338*, 122013. https://doi.org/10.1016/j.oceaneng.2025.122013

[14] F, Şener, İ., & Tuğal, İ. (2025). Optimized CNN-LSTM with hybrid metaheuristic approaches for solar radiation forecasting. *Case Studies in Thermal Engineering*, *72*, 106356. https://doi.org/10.1016/j.csite.2025.106356

[15] Horvat, T., & Job, J. (2020). The use of machine learning in sport outcome prediction: A review. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *10*(5), e1380. https://doi.org/10.1002/widm.1380

[16] Stübinger, J., Mangold, B., & Knoll, J. (2019). Machine learning in football betting: Prediction of match results based on player characteristics. *Applied Sciences*, *10*(1), 46. https://doi.org/10.3390/app10010046

[17] Hubáček, O., Šourek, G., & Železný, F. (2019). Exploiting sports-betting market using machine learning. *International Journal of Forecasting*, *35*(2), 783–796. https://doi.org/10.1016/j.ijforecast.2019.01.001

[18] McCabe, A., & Trevathan, J. (2008, April). Artificial intelligence in sports prediction. In *Fifth International Conference on Information Technology: New Generations (ITNG 2008)* (pp. 1194–1197). IEEE. https://doi.org/10.1109/ITNG.2008.203

[19] Hohmann, A., Edelmann-Nusser, J., & Henneberg, B. (2002). Modeling and prediction of Olympic performance in swimming upon neural networks. *European Journal of Sport Science*, *2*(2), 1–10. https://doi.org/10.1080/17461390200072201

[20] Liang, H. (2022). An intelligent prediction for sports industry scale based on time series algorithm and deep learning. *Computational Intelligence and Neuroscience*, *2022*(1), 9649825. https://doi.org/10.1155/2022/9649825

[21] Ye, X., Huang, Y., Bai, Z., & Wang, Y. (2023). A novel approach for sports injury risk prediction: Based on time-series image encoding and deep learning. *Frontiers in Physiology*, *14*, 1174525. https://doi.org/10.3389/fphys.2023.1174525

[22] Stetter, B. J., Ringhof, S., Krafft, F. C., Sell, S., & Stein, T. (2019). Estimation of knee joint forces in sport movements using wearable sensors and machine learning. *Sensors*, *19*(17), 3690. https://doi.org/10.3390/s19173690

[23] żyk, E. M., & Unold, O. (2011). Machine learning approach to model sport training. *Computers in Human Behavior*, *27*(5), 1499–1506. https://doi.org/10.1016/j.chb.2010.10.014

[24] Sun, W. (2022). Sports performance prediction based on chaos theory and machine learning. *Wireless Communications and Mobile Computing*, *2022*(1), 3916383. https://doi.org/10.1155/2022/3916383

[25] Zhang, L., & Li, N. (2022). Material analysis and big data monitoring of sports training equipment based on machine learning algorithm. *Neural Computing and Applications*, *34*(4), 2749–2763. https://doi.org/10.1007/s00521-021-05852-8

[26] Wilkens, S. (2021). Sports prediction and betting models in the machine learning age: The case of tennis. *Journal of Sports Analytics*, *7*(2), 99–117. http://dx.doi.org/10.2139/ssrn.3506302

[27] Shetty, M., Rane, S., Pandita, C., & Salvi, S. (2020). Machine learning-based Selection of Optimal sports Team based on the Players Performance. *2020 5th International conference on communication and electronics systems (ICCES)*, 1267–1272. IEEE,https://doi.org/10.1109/ICCES48766.2020.9137891

[28] Karakaya, A., Ulu, A., & Akleylek, S. (2022). GOALALERT: A novel real-time technical team alert approach using machine learning on an IoT-based system in sports. *Microprocessors and Microsystems*, *93*, 104606. https://doi.org/10.1016/j.micpro.2022.104606

[29] Barua, A., Sharif, O., & Hoque, M. M. (2021). Multi-class sports news categorization using machine learning techniques: Resource creation and evaluation. *Procedia Computer Science*, *193*, 112–121. https://doi.org/10.1016/j.procs.2021.11.002

[30] Estrella, T., & Capdevila, L. (2025). Identification of athleticism and sports profiles throughout machine learning applied to heart rate variability. *Sports*, *13*(2), 30. https://doi.org/10.3390/sports13020030

[31] Chu, Y., Knell, G., Brayton, R. P., Burkhart, S. O., Jiang, X., & Shams, S. (2022). Machine learning to predict sports-related concussion recovery using clinical data. *Annals of Physical and Rehabilitation Medicine*, *65*(4), 101626. https://doi.org/10.1016/j.rehab.2021.101626

[32] Teunissen, J. W., Faber, I. R., De Bock, J., Slembrouck, M., Verstockt, S., Lenoir, M., & Pion, J. (2025). A machine learning approach for the classification of sports based on a coaches' perspective of environmental, individual and task requirements: A sports profile analysis. *Journal of Sports Sciences*, *43*(1), 23–32. https://doi.org/10.1080/02640414.2023.2271706

[33] Çolak, H., & Şenol, H. (2023). *Estimating the number of licensed athletes in Türkiye until 2030 with artificial neural networks, academic evaluations in the field of sports sciences-7*. Türkiye: Duvar Publications.

[34] MYSPT. (2025). *Ministry of Youth and Sports of the Republic of Türkiye*. Retrieved from: https://gsb.gov.tr/tr

[35] DrDataStats. (2025). Retrieved from: https://www.drdatastats.com

[36] Sardarabadi, A., Ardakani, A. H., Matrone, S., Ogliari, E., & Shirazi, E. (2025). Multi-temporal PV power prediction using long short-term memory and wavelet packet decomposition. *Energy and AI*, *21*, 100540. https://doi.org/10.1016/j.egyai.2025.100540

[37] Groot, D. J., Ellerbroek, J., & Hoekstra, J. M. (2025). Comparing attention-based methods with long short-term memory for state encoding in reinforcement learning-based separation management. *Engineering Applications of Artificial Intelligence*, *159*, 111592. https://doi.org/10.1016/j.engappai.2025.111592

[38] Hong, J., Ma, S., Li, K., Liang, F., Zhang, B., Qiu, Y., & Chen, Y. (2025). Vehicle identification and battery voltage prediction using the long short-term memory neural networks for unknown real-world charging pile data oriented to vehicle-pile interaction. *Journal of Energy Storage*, *126*, 116835. https://doi.org/10.1016/j.est.2025.116835

[39] Gundu, V., & Simon, S. P. (2024). Duck shaped load curve supervision using demand response program with LSTM based load forecast. *Sādhanā*, *49*(3), 201. https://doi.org/10.1007/s12046-024-02532-w

[40] Nugroho, A., & Suhartanto, H. (2020, September). Hyper-parameter tuning based on random search for densenet optimization. In *2020 7th International conference on information technology, computer, and electrical engineering (ICITACEE)*, 96–99. IEEE.

[41] Jain, A., Singh, A., & Doherey, A. (2025). Prediction of cardiovascular disease using XGBoost with OPTUNA. *SN Computer Science*, *6*(5), 1–7. https://doi.org/10.1007/s42979-025-03954-x

[42] Vasquez-Ramos, J., Ruiz-Sandoval, M. G., Oliva, D., Ramos-Soto, O., Ramos-Frutos, J., Sharawi, M., & Pérez-Cisneros, M. A. (2025). Response surface-driven hyperparameter optimization for XGBoost. *The Journal of Supercomputing*, *81*(10), 1112. https://doi.org/10.1007/s11227-025-07600-4

[43] Zhao, T., Chen, G., Pang, C., & Busababodhin, P. (2025). Application and performance optimization of SLHS-TCN-XGBoost model in power demand forecasting. *Computer Modeling in Engineering & Sciences*, *143*(3), 2883–2917. https://doi.org/10.32604/cmes.2025.066442

[44] Liu, X., Gao, Z., Tian, J., Li, M., Pan, T., Wei, Z., . . . , & Wang, P. (2025). State of health estimation for lithium-ion batteries using attention gated recurrent unit network with partial charging curve reconstruction. *Journal of Energy Storage*, *132*, 117602. https://doi.org/10.1016/j.est.2025.117602

[45] He, J., Ma, Z., Liu, Y., Ma, C., & Gao, W. (2025). Remaining useful life prediction of lithium-ion battery based on improved gated recurrent unit-generalized Cauchy process. *Journal of Energy Storage*, *126*, 117086. https://doi.org/10.1016/j.est.2025.117086

[46] Yang, S. Y., Jhong, Y. D., Jhong, B. C., Fan-Chiang, L., & Tsai, M. C. (2025). A two-stage multi-step-ahead Gated-Recurrent-Unit-based method with the use of sewer level data for improving real-time forecasting of inundation depth. *Journal of Hydrology*, *659*, 133230. https://doi.org/10.1016/j.jhydrol.2025.133230

[47] Zhang, S., Abdel-Aty, M., Wu, Y., & Zheng, O. (2020). Modeling pedestrians' near-accident events at signalized intersections using gated recurrent unit (GRU). *Accident Analysis & Prevention*, *148*, 105844. https://doi.org/10.1016/j.aap.2020.105844

[48] Yuan, K., Huang, B., & Qin, H. (2025). Fatigue damage prediction model for broadband non-Gaussian random processes based on Bayesian optimized random forest. *Marine Structures*, *104*, 103894. https://doi.org/10.1016/j.marstruc.2025.103894

[49] Liu, M., & Hu, L. (2025). Solid particle erosion prediction in 90-degree elbows under gas and gas dominant multiphase flows based on Random Forest regression algorithm. *Powder Technology*, *464*, 121157. https://doi.org/10.1016/j.powtec.2025.121157

[50] Wei, X., Xu, Y., Li, X., Fan, G., Cheng, X., Yu, T., & Jiang, B. (2025). Study on prediction model of nitrogen oxide concentration in reprocessing plant based on random forest. *International Journal of Advanced Nuclear Reactor Design and Technology*, *7*(2), 63–69. https://doi.org/10.1016/j.jandt.2025.04.011

[51] Jimoh, K. O., Kareem, M. A., Adegoke-Elijah, A., Zevallos-Aquije, A., Ariyo, D. T., Adekunle, T. A., & Jimoh-Ademola, M. O. (2025). Concrete water cement ratio prediction system using random forest regression. *Discover Concrete and Cement*, *1*(1), 7.

[52] Chen, W. L., Deng, J., Wang, Z. Q., Liu, T. S., He, Y. J., Xiao, Y., . . . , & Bai, G. X. (2025). Combustion parameter prediction for mining conveyor belts by using convolutional neural network–long short-term memory. *Energy and AI*, *21*, 100524. https://doi.org/10.1016/j.egyai.2025.100524

[53] Sekhar, C., & Dahiya, R. (2023). Robust framework based on hybrid deep learning approach for short term load forecasting of building electricity demand. *Energy*, *268*, 126660. https://doi.org/10.1016/j.energy.2023.126660

[54] Ji, H., Zhang, X., Wang, T., Yang, K., Jiang, J., & Xing, Z. (2025). Oil spill area prediction model of submarine pipeline based on BP neural network and convolutional neural network. *Process Safety and Environmental Protection*, *199*, 107264. https://doi.org/10.1016/j.psep.2025.107264

[55] Box, G., & Jenkins, G. (1976). *Time series analysis: Forecasting and control*. USA: Holden-Day.

[56] Raza, A., Jingzhao, L., Adnan, M., & Ahmad, I. (2024). Optimal load forecasting and scheduling strategies for smart homes peer-to-peer energy networks: A comprehensive survey with critical simulation analysis. *Results in Engineering*, *22*, 102188. https://doi.org/10.1016/j.rineng.2024.102188

[57] Al-lami, A., & Török, Á. (2025). Regional forecasting of driving forces of CO2 emissions of transportation in Central Europe: An ARIMA-based approach. *Energy Reports*, *13*, 1215–1224. https://doi.org/10.1016/j.egyr.2025.01.004

[58] Sardar, I., Akbar, M. A., Leiva, V., Alsanad, A., & Mishra, P. (2023). Machine learning and automatic ARIMA/Prophet models-based forecasting of COVID-19: Methodology, evaluation, and case study in SAARC countries. *Stochastic Environmental Research and Risk Assessment*, *37*(1), 345–359. https://doi.org/10.1007/s00477-022-02307-x

[59] Sharma, V., Ghosh, S., Mishra, V. N., & Kumar, P. (2025). Spatio-temporal Variations and Forecast of PM2. 5 concentration around selected Satellite Cities of Delhi, India using ARIMA model. *Physics and Chemistry of the Earth, Parts A/B/C*, *138*, 103849. https://doi.org/10.1016/j.pce.2024.103849

[60] Alomani, G., Kayid, M., & Abd El-Aal, M. F. (2025). Global inflation forecasting and Uncertainty Assessment: Comparing ARIMA with advanced machine learning. *Journal of Radiation Research and Applied Sciences*, *18*(2), 101402. https://doi.org/10.1016/j.jrras.2025.101402

[61] Çolak, H., & Çolak, E. (2024). Estimation of prevalence distribution of pre-obesity by gender in Türkiye using artificial neural networks and time series analysis. *The Black Sea Journal of Sciences*, *14*(3), 1340–1359.

[62] Vabalas, A., Gowen, E., Poliakoff, E., & Casson, A. J. (2019). Machine learning algorithm validation with a limited sample size. *PloS One*, *14*(11), e0224365. https://doi.org/10.1371/journal.pone.0224365

[63] Krstajic, D., Buturovic, L. J., Leahy, D. E., & Thomas, S. (2014). Cross-validation pitfalls when selecting and assessing regression and classification models. *Journal of Cheminformatics*, *6*(1), 10. https://doi.org/10.1186/1758-2946-6-10

[64] Dalkılıç, M., Kargün, M., Kızar, O., & Genç, H. (2017). Estimation of licensed number of number of competitors in the wrestling of artificial neural networks. *Kilis*, *7*, 15–19.

[65] Atasoy, M., Dalkılıç, M., & Uğraş, S. (2017). Estımation of licensed sportsman-woman in area of martıal sports by artıfıcıal neural networks. *Kilis*, *7*, 33–37.

[66] Dalkılıç, M., Mamak, H., Atasoy, M., & Mihriay, M. (2017). Forecast of licensed number of living sports in water trophies. *Water Ball and Swimming Areas by Artificial Neural Networks*, *5*(50), 119–125.

[67] Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and practice*. Canada: OTexts.

How to Cite: Şenol, H., & Çolak, H. (2025). Forecasting Licensed Athletes Numbers Using Time Series and Hybrid Artificial Intelligence Models. *Archives of Advanced Engineering Science*. https://doi.org/10.47852/bonviewAAES52027778