IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received December 29, 2021, accepted January 16, 2022, date of publication January 21, 2022, date of current version February 11, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3145368

Impact of Sleep and Training on Game Performance and Injury in Division-1 Women's Basketball Amidst the Pandemic

SAMAH SENBEL[1], SRISHTI SHARMA[2], MEHUL S. RAVAL[2], (Senior Member, IEEE), CHRIS TABER[1], JULIE NOLAN[1], N. SERTAC ARTAN[3], (Member, IEEE), DIALA EZZEDDINE[4], AND TOLGA KAYA[1], (Senior Member, IEEE)

[1]Sacred Heart University, Fairfield, CT 06825, USA
[2]Ahmedabad University, Ahmedabad, Gujarat 380009, India
[3]New York Institute of Technology, New York, NY 10023, USA
[4]Tao Media, Bellevue, WA 98007, USA

Corresponding author: Tolga Kaya (kayat@sacredheart.edu)

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Institutional Review Board under Approval No. 170720A.

ABSTRACT We investigated the impact of sleep and training load of Division - 1 women's basketball players on their game performance and injury prediction using machine learning algorithms. The data was collected during a pandemic-condensed season with unpredictable interruptions to the games and athletic training schedules. We collected data from sleep monitoring devices, training data from coaches, injury reports from medical staff, and weekly survey data from athletes for 22 weeks. With proper data imputation, interpretable feature set, data balancing, and classifiers, we showed that we could predict game performance and injuries with more than 90% accuracy. More importantly, our *F1* and *F2* scores of 0.94 and 0.83 for game performance and injuries, respectively, show that we can use the prediction for informative analysis in the future for coaches to make insightful decisions. Our data analysis also showed that collegiate athletes sleep less than the recommended hours (6-7 instead of 8 hours). This coupled with a long hiatus in games and training increases the risk of injury. Varied training and higher heart rate variability (due to better quality sleep) indicated a better performance, while athletes with poor sleep patterns, were more prone to injuries.

INDEX TERMS Basketball, collegiate athletes, data mining, game performance, injury prediction, machine learning, sports analytics.

I. INTRODUCTION

Sports data analytics has been gaining significant attention through collegiate and professional leagues and E-sports. Any insights regarding athletic performance can make a difference in games in terms of athletic performance or preventing injuries, ultimately determining the overall success of a team. This transdisciplinary research has brought athletes, athletic trainers, exercise scientists, engineers, and data scientists together to investigate a Division-1 basketball team's season where COVID-19 caused unprecedented disruptions.

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

Using daily sleep patterns of the athletes combined with subjective training and survey data, this project positioned itself in a way that game performance and injuries could be predicted with machine learning (ML) methods.

A. ATHLETIC PERFORMANCE

Sleep and recovery strongly influence physical qualities such as strength, anaerobic power, flexibility, and physical performance [1]. Additionally, sleep is known to affect cognitive, emotional, and mental performance [2]. Athletes, in general, need more than eight hours of sleep to feel rested; however, their average hours of sleep fall below seven hours a night

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

due to academic, social, and athletic commitments [3]. On the other hand, sleep extension beyond 8 hours per day for several weeks was shown to improve shooting accuracy for collegiate men's varsity basketball team and reaction time for college varsity tennis players [4], [5].

B. ATHLETIC INJURIES

Collegiate athletes are reported to spend up to 41 hours per week on athletics, resulting in inadequate time for restorative sleep [6]. It has been reported that sleeping less than 8 hours per night resulted in higher injury rates in adolescent athletes [7], [8]. Furthermore, a simultaneous increase in training load, training intensity, and decreased hours of sleep resulted in a significantly higher risk of injury in this population [9]. Limited research exists addressing the association between sleep and injury rate, and data is limited to the use of subjective questionnaires to characterize sleep metrics [10].

C. MACHINE LEARNING METHODS IN BASKETBALL

Modern basketball analytics focus on action-oriented and proactive strategies [11]. ML in basketball can help roster composition, gameplay strategies, and player flow during a game [12]. Commonly, classification, regression, and clustering techniques are used in basketball analytics [13]. For example, the impact of a player's game performance on win probability was carried out by using the Bayesian linear regression model [14]. In addition, Wang *et al.* [15] used neural networks on basketball videos to classify offensive strategies such as player movement and location, in-bound plays, etc. Similarly, with the goal of improving the team performance, Piette *et al.* [16] used network analysis to find the path during the gameplay that generated maximum points and predicted the outcome.

On the other hand, injury prediction is challenging due to the multidimensional parameters coming from psychological, social, nutritional, and training workload [17]. Data scientists have used artificial neural networks (ANN), Bayesian networks, decision trees, clustering, and Support Vector Machines (SVM) for injury prediction of a player [18]. Talukder *et al.* [19] utilized game data, workload, and team schedules with random forest and an autoregressive model with logistic regression to predict player injury. Knee injury and heart defect detections have also been explored using ML [20], [21].

D. COVID IMPACT IN ATHLETICS

Covid-19 has caused the cessation of daily activities by teams and athletes in sports. Athletes missed technical, tactical, and psychological preparation that disrupted injury prevention, nutrition, and recovery programs [22] potentially raising the risk of injury, [23]. Grazioli *et al.* [24] found that after 63 days of quarantine, professional soccer players lost lean body mass and fat mass resulting in degraded sprinting ability and lower body power. It has been suggested that athletes must be gradually exposed to physical training after detraining periods [25].

E. PROPOSED WORK

This holistic study incorporated training data from strength coaches, sleep and recovery information from a wrist-worn device, and perception questionnaires from the Women's Division-I basketball team to create a platform for predicting game performance and injury risk assessment. Our research questions (RQs) were:

RQ1: Which training, sleep, and survey data features are more important for the following?

(a) Game performance
(b) Injuries

RQ2: Can we predict injury and game performance based on training, sleep, and athletes' perception for data collected during a pandemic?

Considering the above RQs the paper focuses and contributes in the following directions:

- **Selection of an interpretable feature set:** We have performed intensive feature selection to reduce the dataset complexity and created an interpretable feature set containing 12 features for game performance and 14 for injury risk prediction.
- **Game performance and injury prediction:** We applied ensemble machine learning algorithms on a reduced data set based on feature importance to predict game performance and injury with an *F1* score of 0.94 and *F2* score of 0.83 respectively.
- **Impact of pandemic on injuries:** We collected data during the pandemic, where team schedules were very different from previous seasons due to game date changes, cancellations, and quarantine protocols.

II. METHODS

A. DATASET

The Division-1 Women's Basketball team at Sacred Heart University were the subjects of this study (Institutional Review Board approval number 170720A on 9/14/2020). The methodology of the data collection is outlined in Fig. 1. Following data from 16 players were collected for 25 weeks: (1) volume load, monotony, and strain from resistance and metabolic conditioning training 3 times a week, i.e. 3/week. (2) Daily (7/week) sleep schedule and heart rates of the athletes via WHOOP strap [26]. (3) The perception of athletes' recovery and stress levels were measured with subjective surveys (twice a week, 2/week, Mondays and Thursdays). The collected data were evaluated against game statistics and injury reports for developing ML models to predict game performance and injuries. Fig. 2 shows a snapshot of the dataset. It contains a few features from sleep, recovery, training, perception (survey), game performance and injury [26].

B. QUANTIFICATION OF TRAINING LOAD

Training load was calculated as a weekly composite score by summing the total work completed in sports training, metabolic conditioning, resistance training, and gameplay. Following each session, a session rating of perceived exertion (sRPE) [27] was calculated by multiplying the time taken to

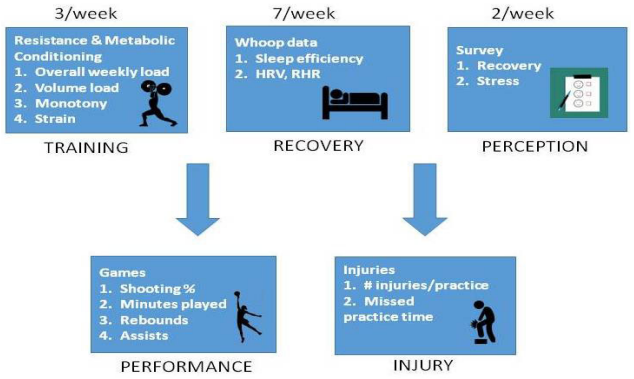


FIGURE 1. Methodology used in the study. Training (three times a week, 3/week), recovery (daily, 7/week), and perception (twice a week, 2/week) data were utilized to gain insights on the game performance and injury of the athletes.

practice by the athletes' subjective rating of perceived effort (1-10 scale) [28], [29]. Total week load and weekly standard deviation (SD) were calculated by summing sRPE for the week and calculating the SD. Weekly resistance training load was calculated by summing the total weight lifted during the week. Monotony was calculated by taking the mean daily load and normalizing it by the weekly SD of training load. Finally, training strain was calculated by taking the total weekly load and multiplying it by the monotony score. Overall, 6 data features were collected weekly [30].

C. SLEEP MONITORING

WHOOP straps were given to each athlete, and their sleep, heart, and overall recovery were monitored through WHOOP's research platform. The heart parameters include - resting heart rate, heart rate variability, sleep parameters includes - Sleep Need, Wake Periods, Sleep Disturbance. Sleep need is amount of sleep needed by athlete to recover. Wake Periods is the duration for which the athlete remains awake mid sleep. Sleep disturbances are number of times the athlete wakes up mid sleep. A total of 22 features were collected daily from the athletes using the WHOOP strap. WHOOP was validated by researchers (i.e., third party validated) via sleep clinic studies where WHOOP data was comparable to polysomnography results. More specifically, agreement, sensitivity, and specificity of WHOOP data was above 90% for sleep while the precision error for heart rate was 1.5% [31]–[33].

D. SURVEY DATA

Athletes completed the short recovery and stress questionnaire [34] via an online dashboard twice a week. This short survey collects subjective data on physical, mental, emotional, and overall recovery. The survey consists of 4 recovery- and 4 stress-related questions such as their Negative Emotional State (NES), Overall Recovery (OR), Overall Stress (OS), Mental Performance game performance Capability (MPC), Muscular Stress (MS), Physical Performance game performance Capability (PPC), Emotional Balance (EB) and Lack of Activation (LA). The athletes had to rank their values on a Likert scale from 0 to 6 [26] for each of these. This survey is both valid and reliable for athletic populations [35].

E. GAME PERFORMANCE

It is important to quantify performance of a player in a game, therefore it was calculated through statistics developed by John Hollinger [36]. The game score roughly measures player's productivity in a game [37]. As shown in Eq. 1, points scored (+), steals (+), and turnovers (-) contribute directly, whereas other game statistics are weighted.

$$\begin{aligned} &GameScore \\ &= PointsScored + (0.4 \times FieldGoals) \\ &\quad - (0.7 \times FieldGoalAttempts) \\ &\quad - (0.4 \times (FreeThrowAttempts - FreeThrows)) \\ &\quad + (0.7 \times OffensiveRebounds) \\ &\quad + (0.3 \times DefensiveRebounds) \\ &\quad + Steals \times (0.7 \times Assists) \\ &\quad + (0.7 \times Blocks) - (0.4 \times PersonalFouls) \\ &\quad - Turnover \end{aligned} \tag{1}$$

Game score holistically looks at all parameters of a box score like offensive play, defensive rebounding, steals, blocks and turnovers to measure the player's impact in that game. The parameters benefiting the team contributes positively to it while committing fouls or turnovers have negative impact. A game score of 10 is average and 40 is considered exceptional. The game score is simple to compute from easily available parameters, can be explained well and it statistically quantifies performance of a player. Therefore, we use it in the present work to measure player's performance.

F. INJURY DATA

Injury data were extracted from medical injury reports generated as injuries occurred. After the season, all original injury reports were shared with the research team, subject numbers replaced player names to ensure confidentiality, and data of interest were extracted for data analysis. This data included the date of injuries, occurrence of injuries (contact or non-contact), types of injuries (e.g., leg, shoulders, strain, tear, etc.), and how long the athlete was not available for practice or games.

III. DATA IMPUTATION

Sleep data and questionnaire responses were not available completely as reasoned below:

1) Incomplete WHOOP strap data due to improper attachment to the wrist (hence no data collection), neglecting to charge, loss of charger, and not keeping the WHOOP app on the phone running in the background to allow data to upload to the research server.

Survey Data				Training			Sleep and Recovery					Game Performance
EB	LA	NES	OS	Daily.Average	Weekly.SD	Strain	RHR	HRV	Sleep Need	Wake.Periods	Sleep Debt	Game.Score
6	0	1	0	22.50	45.00	45.00	65.00	53.00	9.71	10.00	1.47	[redacted]
6	0	1	0	22.50	45.00	45.00	54.10	98.90	9.16	11.80	0.97	[redacted]
6	0	1	0	22.50	45.00	45.00	69.00	44.00	9.28	9.00	0.73	[redacted]
2	3	3	4	100.00	86.60	346.41	60.00	91.00	9.25	5.00	0.61	[redacted]
2	3	3	4	100.00	86.60	346.41	65.00	77.00	9.03	9.00	1.03	[redacted]
2	3	3	4	100.00	86.60	346.41	56.00	79.00	8.69	8.00	0.30	[redacted]
5	2	4	2	137.50	129.71	583.03	55.00	91.00	9.47	11.94	0.66	[redacted]
5	2	4	2	137.50	129.71	583.03	56.50	103.50	9.42	11.30	1.17	[redacted]
5	2	4	2	137.50	129.71	583.03	56.00	65.00	8.63	16.00	0.00	[redacted]
6	1	0	2	90.00	60.00	540.00	72.00	69.00	9.12	4.00	1.38	[redacted]
6	1	1	2	72.50	95.00	221.32	76.00	40.00	9.08	8.00	1.52	[redacted]
4	2	2	2	72.50	95.00	221.32	74.00	24.00	9.02	13.00	1.28	[redacted]

FIGURE 2. Dataset snapshot containing features from sleep, recovery, training, perception (survey), game performance and injury. The survey parameters - EB: Emotional Balance, LA: Lack of Activation, NES: Negative Emotional State, OS: Overall Stress are covered in Section II-D. The training parameters are discussed in Section II-B. The sleep and recovery parameters like RHR: Resting Heart Rate, HRV: Heart Rate Variability are discussed in Section II-C. Game performance values were blacked out to de-identify the athlete as the game score data is publicly available.

2) Incomplete surveys as the athletes did not always complete the surveys, especially during COVID lockdown and in-season (when games were played).

Even though the game performance and injury data were only available when applicable, missing data from sleep and surveys constituted a source of bias that could negatively impact the ML algorithms. The choice of an appropriate imputation technique could subsequently ease feature selection and classification analyses. In this work, we utilized data imputation techniques to fill the missing data. Single imputation techniques such as global mean, cluster mean, conditional local mean, and K-Nearest Neighbor (KNN) fail to reflect the uncertainty about imputed values [38]. Therefore, a multiple imputation approach called Multivariate Imputation by Chained Equation (MICE) was utilized [39]. This method imputes each feature in the dataset sequentially, allowing the prior imputed data values to be used in the model to predict the following features. MICE has effectiveness for datasets with a high missingness rate (>50%). It has also been observed that datasets related to breast cancer, heart disease, 26 capital letters in English alphabets and spam emails when imputed with MICE, resulted in best feature selection when compared to a single imputation technique or no imputation [40]. Each of the datasets had large number of missing values and they include categorical and numerical features and contain data from different fields. Imputation power can be improved by increasing the number of imputations for a small dataset [41]. A key finding suggests that data missingness should not be used as a measure to use multiple imputations. Correctly specified MICE can reduce bias and improve analyses for any proportion of missingness in MAR data [42]. It means using MICE is beneficial for all proportions of missing data provided the imputation model is correctly specified and it includes all variables related to missingness. While a MICE imputer is proven to be effective, we cannot conclude that MICE is effective in all situations. Graham *et al.* [41] suggested that the number of imputations depends on the dataset size, the models used to impute, the missingness rate and computing resources.

Applying MICE reduces the bias during the subsequent feature selection technique [40]. The effectiveness of the MICE imputer is demonstrated in Table 1 which shows statistical measures of the observed and imputed values. The statistical measure of percentage bias (PB, ideal is $< 5\%$) of Resting Heart Rate (RHR) is 1.2%, Wake Periods is 0.6% and for Physical game Performance Capability (PPC) is 0.0004%, indicating effective imputation.

Table 2 shows the imputation percentages after MICE implementation. A total of 2800 data entries (rows) is possible from 16 athletes, 7 days a week of 25 weeks. It is important to mention that injury and game performance data were not imputed in order to avoid introducing bias. There were a total of 165 games and 11 injury incidents during the 25-week period. Moreover, training load data was not filled in either as non-training days were entered as 0, except monotony, which was a weekly measure and entered as the maximum value considering a worst-case scenario. In the case of monotony, defined as the differences from daily training activities and calculated as a week, no training (which happened during the pandemic-related shutdowns) weeks meant the worst-case scenario. In these weeks, there were no training, so a 0 variety, resulting a monotony of 1, the maximum value. This valuation makes the data imputation consistent and does not introduce bias.

IV. FEATURE IMPORTANCE, INTERACTION EFFECT ANALYSIS

A. FEATURE IMPORTANCE

Feature importance analysis was conducted to address our **RQ1**. With this analysis, we not only intended to reduce the dimensionality of the dataset but also aimed at understanding how significantly the features contributed to the game performance and injury respectively. The feature importance techniques explain the contribution of an individual feature of the dataset to a decision made by an ML algorithm to ultimately gain insights from a model [43], [44]. Numeric datasets such as ours primarily utilize regression models for

TABLE 1. Statistics for observed and imputed features values. N - number of samples, SD - standard deviation, min. - minimum value, max. - maximum value.

Features	Observed (Imputed)				
	N	Mean	SD	Min	Max
PPC	1280 (1520)	4.25 (4.05)	1.32 (0.81)	0.00 (0.00)	6.00 (6.00)
MPC	1280 (1520)	4.31 (3.80)	1.49 (0.94)	0.00 (0.00)	6.00 (6.00)
EB	1280 (1520)	4.03 (3.94)	1.6 (1.3)	0.00 (0.00)	6.00 (6.00)
OR	1280 (1520)	3.89 (3.99)	1.25 (0.90)	0.00 (0.00)	6.00 (6.00)
MS	1280 (1520)	2.65 (3.06)	1.43 (1.04)	0.00 (0.00)	6.00 (6.00)
LA	1280 (1520)	1.82 (2.03)	1.54 (1.04)	0.00 (0.00)	6.00 (6.00)
NES	1280 (1520)	1.68 (2.29)	1.43 (1.11)	0.00 (0.00)	6.00 (6.00)
OS	1280 (1520)	2.40 (2.88)	1.46 (1.03)	0.00 (0.00)	6.00 (6.00)
RHR	1802 (998)	58.69 (57.59)	8.38 (6.51)	43.00 (49.50)	89.00 (77.30)
HRV	1802 (998)	84.72 (88.43)	39.57 (31.57)	11.00 (11.00)	224.00 (224.00)
Recovery	1802 (998)	58.28 (60.89)	22.69 (19.15)	1.00 (1.00)	99.00 (99.00)
Sleep Score	1802 (998)	76.13 (75.71)	17.80 (13.87)	12.00 (12.00)	100.00 (100.00)
Hours in Bed	1802 (998)	7.56 (7.70)	1.74 (1.32)	1.26 (1.28)	20.64 (18.99)
Hours of Sleep	1802 (998)	6.80 (6.86)	1.53 (1.14)	1.07 (1.07)	19.27 (11.88)
Sleep Need	1802 (998)	8.90 (9.04)	1.07 (0.91)	3.35 (3.35)	11.79 (11.79)
Sleep Efficiency	1802 (998)	90.20 (89.46)	5.53 (4.12)	55.00 (58.00)	100.00 (100.00)
Wake Periods	1802 (998)	12.05 (12.08)	6.03 (4.41)	-1.58 (-0.39)	38.00 (38.00)
Sleep Disturbance	1802 (998)	9.73 (9.71)	5.06 (4.7)	0.00 (0.00)	31.00 (31.00)
Latency min.	1802 (998)	1.84 (1.64)	3.75 (3.55)	0.00 (0.00)	138.22 (138.22)
Cycles	1802 (998)	5.02 (5.27)	1.92 (1.46)	0.00 (0.00)	12.00 (12.00)
Light Sleep	1802 (998)	3.28 (3.22)	1.05 (0.75)	0.48 (0.48)	16.52 (16.52)
Awake hours	1802 (998)	0.76 (0.82)	0.52 (0.40)	0.00 (0.00)	5.53 (5.53)
Sleep Debt hours	1802 (998)	1.01 (1.07)	0.71 (0.61)	0.00 (0.00)	3.40 (3.40)
Sleep Consistency	1802 (998)	64.61 (65.70)	13.63 (10.43)	3.00 (3.32)	92.00 (92.00)
Respiratory Rate	1802 (998)	16.74 (16.98)	1.33 (1.41)	13.41 (13.41)	20.09 (19.95)
Total Cycle Nap	1802 (998)	0.29 (0.18)	0.81 (0.66)	0.00 (0.00)	8.01 (8.01)
Total Cycle Sleep	1802 (998)	7.09 (7.04)	1.66 (1.26)	1.07 (1.07)	19.27 (18.34)
REM Sleep	1802 (998)	28.49 (28.42)	8.97 (6.34)	0.00 (0.00)	56.10 (54.70)
Deep Sleep	1802 (998)	18.07 (18.28)	4.48 (3.61)	0.00 (0.00)	32.20 (28.50)
Restorative Sleep	1802 (998)	46.57 (46.72)	10.57 (9.47)	0.00 (0.00)	81.50 (79.20)

analysis where feature saliency or importance are used [45]. Not only does feature importance help explain the significance of each feature to the model, but it also reduces the curse of dimensionality [46] by suggesting features that the model may drop. The curse of dimensionality indicates that in machine learning, number of training pairs needed to estimate a function grows exponentially with respect to dimensionality of the input features [46]. It is a problem when dataset has too many input features and very few observations. It can lead to over fitting of the model as features are more than observations. The problem can be solved by reducing dimensionality of the input features by retaining only the most important ones through mechanism like feature importance.

TABLE 2. Imputations ratios of each data subset to provide a general outline of the data structure.

	Before imputation	After imputation
WHOOP	64.39% (1802 rows)	100% (2800 rows)
Perception	44.70% (1252 rows)	100% (2800 rows)

Game performance and Injury were treated as two separate target variables for feature importance analysis. With respect to each target variable, we apply correlation-driven, Extreme Gradient Boosting (XGBoost), and Random Forest (RF) [47] techniques one by one to the data and obtain ranked importance of features. The suitability of these techniques is discussed as follows: (1) **Correlation (CORR)-driven feature importance** - High correlation depicts linear dependency between attributes, helping with the feature importance and removing the redundancy and hence the bias in the dataset [48]. (2) **XGBoost** - Boosted decision trees were constructed within the model. The importance of features was given a score calculated by how much an attribute contributed to making critical decisions with the decision trees [49]. For one decision tree, importance was calculated based on how an attribute split improves the Gini Index, weighted by the node's records. In the end, the feature importance of all the trees was averaged [49]. (3) **RF** - This method is a bagging technique that uses multiple decision trees as a base model to arrive at a consensus using trees rather than relying on a single tree for decision. RF provides interpretability as it is easy to compute the contribution of features to the decision. RF uses a change in accuracy or Gini index to find the feature importance when excluded from the decision. While RF is built in parallel with aggregated results, XGBoost is built sequentially.

Fig. 3 is the correlation heat map of all sleep and recovery, training load, perception, game performance, and injury-related features. First 8 labels on the map refer to the survey results, which were explained in Section II.D. Following 6 parameters are associated with the training (Section II.B), while the next section are recovery parameters from the WHOOP data. Last two parameters are the output parameters; game performance and injury. Areas demarcated by black rectangles from left to right show the dataset groups of survey, training, and sleep, respectively. Dark red or blue colors represent high correlation (positive and negative, respectively),

and it can be seen that several correlations exist between data groups, such as training load and the survey data on stress and recovery.

For each target variable - game performance and injury, top features from the sorted set were selected by taking a weighted rank-sum (hard voting) CORR, XGB and RF for each attribute. We further dropped features with high correlation based on the correlation matrix to lower the bias. For instance, sleep disturbances are dropped while retaining wake periods as these two features are highly correlated (Fig. 3). Game statistics to calculate the game score were not included in the analysis to avoid introducing bias. In order to reduce the feature set, we tried predicting the game performance and injury using a RF regressor and classifier respectively with different sizes of feature sets. By using forward addition of features and backward elimination of features on the basis of their significance in the prediction, we tried fitting an RF regressor and classifier respectively over different sizes of feature sets and kept track of the Mean Squared Error (MSE) and R - squared (R^2) for game performance and *F2* score for injuries. We found the least MSE (0.17) and highest R^2 (0.89) values with 25 features for game prediction and the highest *F2* score was achieved with 29 features for injuries, hence these features were selected for prediction analysis. Table 3 shows the top 10 contributors to game performance and injury. Strain and Rapid Eye Movement (REM) sleep are the most significant contributors to game performance and injury, respectively. Both strain and REM sleep were assessed using daily collected values.

TABLE 3. Top 10 features ranked based on hard voting. Importance scores for these 10 features for game performance and injury are also shown.

Rank	Game Performance (CORR, XGB, RFR)	Injury (CORR, XGB, RFC)
1	Strain (0.22, 9.6%, 7.6%)	REM Sleep hours (0.11, 2.7%, 12.9%)
2	Daily Average (0.21, 7.5%, 6.5%)	Respiratory rate (0.06, 0.9%, 6.5%)
3	Heart Rate Variability (HRV) (0.16, 7.3%, 5.8%)	Monotony (0.07, 2.7%, 9.4%)
4	Emotional Balance (EB) (-0.15, 5.5%, 5.4%)	Game Score (0.12, 3.4%, 7.1%)
5	Sleep Need (0.13, 7.2%, 5.2%)	Week Trimp Total 0.12, 1.8%, 2.4%)
6	Lack of Activation (LA) (0.14, 5.4%, 4.8%)	Strain (0.12, 5.4%, 8%)
7	Overall Stress (OS) (0.11, 5.3%, 4.7%)	Weekly SD (0.06, 6%, 5.6%)
8	Sleep Debt Hours (0.09, 6.7%, 4.6%)	Repetitive Training (RT) volume load (0.02, 5%, 4.3%)
9	Weekly SD (0.18, 5.3%, 4.5%)	Awake hours (0.06, 1.06%, 1.6%)
10	Monotony (0.13, 4.2%, 4.1%)	Sleep need (0.03, 1.6%, 5.02%)

B. INTERACTION EFFECT

When multiple features have an impact or significantly contribute to the prediction of a dependent feature, their joint effect on the prediction is sometimes significantly greater than the sum of the parts. This is called the interaction effect [50]. Interaction modeling helps understand how two features work together to impact a dependent feature. It also represents the relationship between dependent and independent features better [50]. In a complex study like in the proposed paper, even the independent variables might interact with each other. It is important to find how third variable influences the relationship between an independent and dependent variable. If the real world behaves this way, it is critical to incorporate interactions in the modeling. For example, the relationship between game performance and RHR probably depends on the respiratory rate in the present work and this knowledge helps during dimensionality reduction.

Interaction modeling was performed over the dataset for game performance prediction and injury prediction individually. After removing multicollinearity from the dataset, pair wise interaction effect was checked for all the features. Forward adding and backward elimination of features was used keeping a check upon the significance level of the polynomial interaction terms on the dependent feature. Like done previously, we tried fitting an RF regressor and classifier respectively over datasets containing different combinations of significant features and interaction terms and monitored the MSE and R2 for game performance and *F2* score for injuries. We found the least MSE (0.09) and highest R2 (0.95) value with 12 features for game prediction and the highest *F2* score was achieved with 14 features for injuries. We thereby concluded that the best feature set for game performance and injuries consisted of 12 and 14 features, respectively. Table 4 shows features and interaction terms in the optimal dataset for game performance and injury.

TABLE 4. Features and interaction terms *(italics)* in optimal dataset for game performance and injury.

Optimal Dataset		
No.	**Game Performance**	**Injury**
1	Strain	REM Sleep
2	RHR	Respiratory Rate
3	Respiratory Rate	Monotony
4	Daily Average	Strain
5	*Weekly SD & REM Sleep*	Points
6	*RHR & Respiratory Rate*	*Points & REM Sleep*
7	Hours of Sleep	*Points & Respiratory Rate*
8	*REM Sleep & Deep Sleep*	Weekly Trimp Total
9	Sleep Disturbances	Sleep Disturbances
10	*Hours of Sleep and Deep Sleep*	Emotional Balance
11	Emotional Balance	RT Volume Load
12	Sleep Need	Sleep Need
13	-	*RT Volume Load & Respiratory Rate*
14	-	*REM Sleep & Deep Sleep*

There were 38 features considered in all for this analysis: 22 features from sleep monitoring using the WHOOP strap,

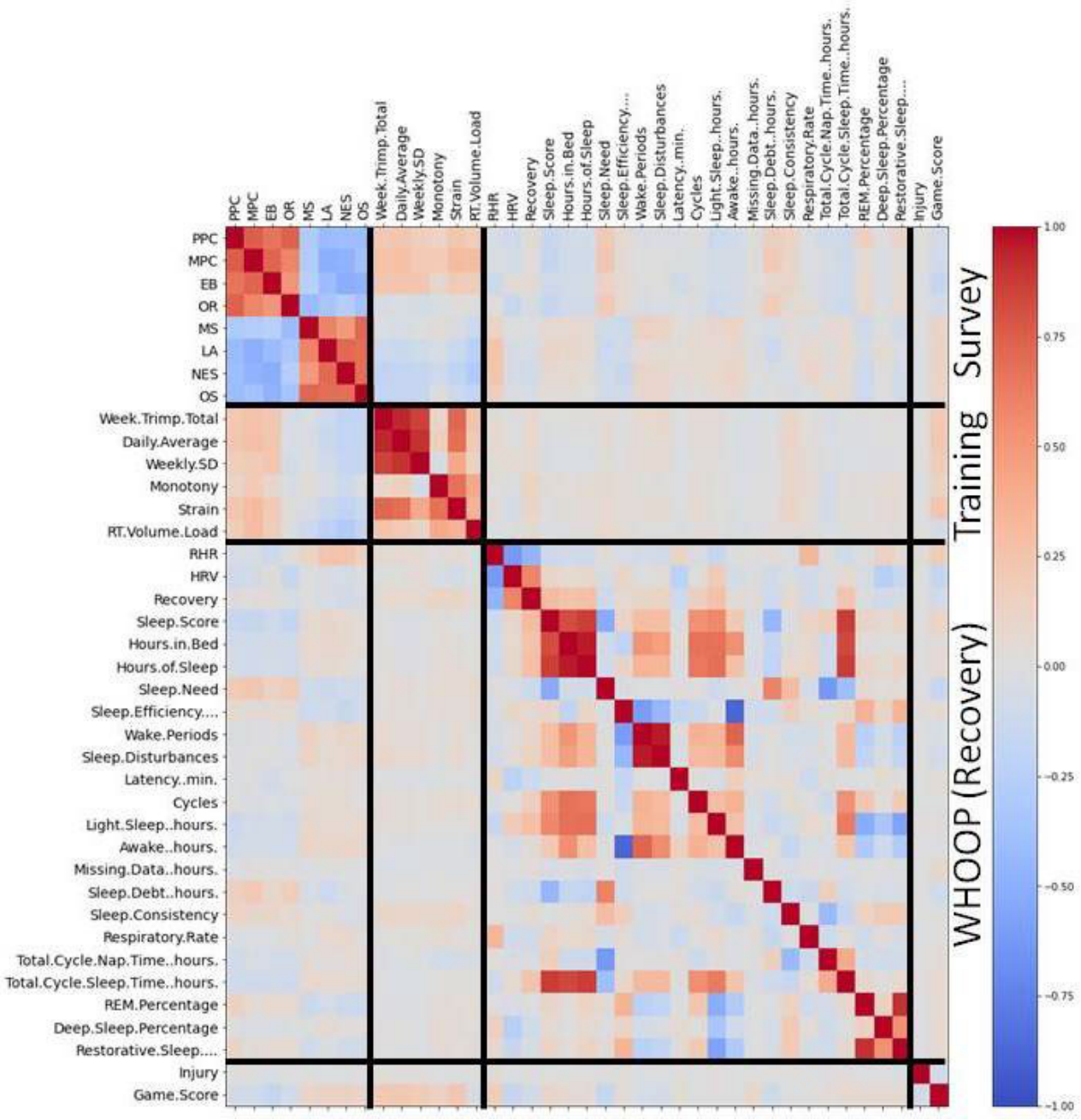


FIGURE 3. Correlation heat map for all the features.

8 features from survey data, 6 features from quantification of training load, 1 feature from game performance and 1 feature from injury data. After the feature importance and interaction analysis, we selected 12 (5 features from sleep monitoring, 1 from survey data, 2 features from quantification of training load, 3 interaction features from sleep monitoring and 1 interaction feature - a combination of sleep monitoring and training load) and 14 (4 features from sleep monitoring, 1 from survey data, 4 features from quantification of training load, 1 feature from game performance, 1 interaction feature from sleep monitoring, 2 interaction features - a combination of sleep monitoring and game performance and 1 interaction feature - a combination of sleep monitoring and training load) features for game performance and injury risk prediction respectively.

C. FEATURE INTERPRETABILITY

It is vital to interpret the impact of individual features on target variables for game performance and injury. We used Partial Dependence Plots (PDP) to show the marginal effect on the target variables and analyze each target's two most dominant features based on hard voting. As shown in Fig. 4,

training strain and daily training load average are positively correlated with game performance. The PDP of strain depicts that if an athlete trains more, she is likely to perform better. Daily average too is a training feature and its high value (>200) improves the game performance. Therefore, the weekly training strain and daily training load average are associated with improvements in performance.

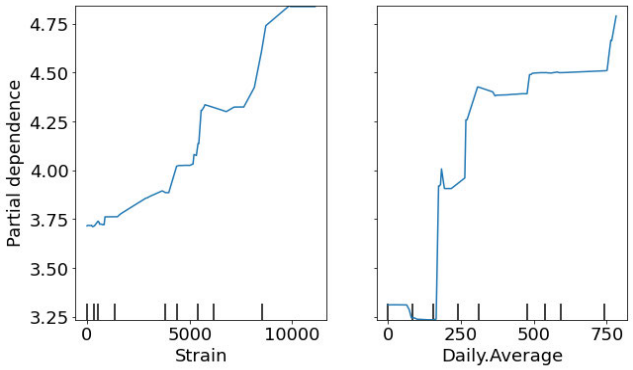


FIGURE 4. PDP for the training strain (left) and daily training load average (right) features on game performance.

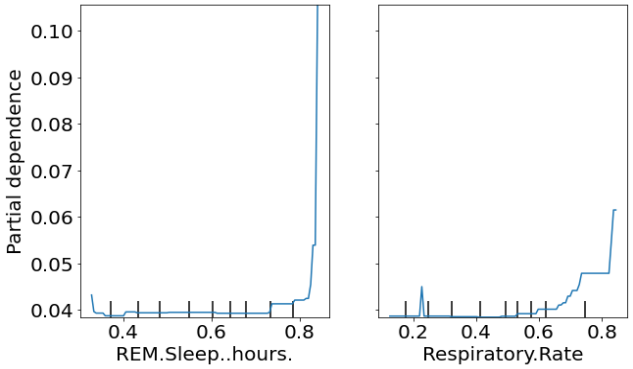


FIGURE 5. PDP for the REM% (left) and RR (right) on injury.

Similarly, we analyzed the impact of the two most dominant features for injuries (based on hard voting): REM sleep percentage (REM%) and Respiratory Rate (RR). REM sleep is a restorative sleep stage and one of the human body's four sleep cycle. In this stage human brain converts short term memories to long term ones. It is significant for a player as REM sleep helps them to retain the technical skills learnt on that day. With 22% - 28% sleep time in the REM stage i.e., 90 Minutes in approximately 7 to 8 hours of REM sleep can help athlete to gain benefit from the day's practice. [32]. The PDP shown in Fig. 5 indicates that low (<20%) and high (>30%) REM% increases the chances of injury. The RR is a reasonably stable metric and 12 - 18 respirations per minute (RPM) in a state of rest is considered normal, while any deviation indicates an unusual pattern [32]. The PDP shows that outside the typical range, sleep disturbances increase, which increases the chances of injury.

V. PREDICTION OF GAME PERFORMANCE & INJURIES

RQ2 aims at predicting the players' game performance via the game score and calculating a risk assessment analysis through injury prediction. These two features in the optimal data sets (game performance - 12 and injury 14) were set as the target variables, and ML models were applied to them individually. Data balancing techniques were used due to the small number of game performances and injury observations. The proportion of the missing data is directly related to quality of the inference and statistical analysis. However, there is no established norm on an acceptable missingness rate of the data [51]. In this work the MICE imputer has been able to neutralize the impact high data missingness during training of the ML model. Due to the limited number of games and injuries compared to the entire dataset, the data is scarce and prediction can have a significant bias, as is the case in health sciences [52]. A problem with imbalanced classification is that there are too few examples of the minority class for a model to effectively learn the decision boundary. Therefore, there is a need to increase their numbers. The best approach to overcome this problem is to synthesize new examples from the minority class using balancing techniques. These balancing techniques perform undersampling and oversampling of data which improves the performance of the decision tree based algorithms applied for prediction [53]. Preprocessing using Synthetic Minority Over Sampling with Gaussian Noise (SMOGN) was performed for data balancing on the game performance data [54]. SMOGN was not applied to injury as injury labels are binary. Pre-processed game performance data and original injury data were then evaluated on Synthetic Minority Oversampling Technique (SMOTE) [55]. Since SMOTE may introduce noise as it increases probability of class overlap, game performance data was evaluated with a combination of SMOTE and Edited Nearest Neighbor (ENN) to remove any newly generated sample that differs from two of their three nearest neighbors [56]. Borderline SMOTE creates synthetic samples along the boundary of the majority and the minority classes in the dataset [57], which is more suitable for injury prediction models.

The k-means clustering algorithm was applied over the game score to create clusters with *bad, average,* and *good* game performance labels. Similarly, the injury records were clustered to have two labels, *injury*, and *no-injury*. We used the XGBoost classifier for both predictions due to its gradient boosting framework and generalization capabilities [58].

Due to the low number of available target records in game performance and injury, accuracy percentage cannot solely evaluate the model performance. Therefore, we used *F1* and *F2* scores to evaluate the classifier's effectiveness (Fig. 6). The choice of *F1* score for predicting game score labels stems from the fact that both False Positive (FP) and False Negative (FN) are equally costly for game performance and *F1* balances them during evaluation. On the other hand, *F2* focuses more on minimizing *FN*. Also, it is crucial to predict as many injuries as possible before the game, which is

captured well by *F2*. F measure, i.e., *F(β)*, is given as follows:

$$F(\beta) = \frac{(1+\beta^2) \times P \times R}{\beta^2 \times P + R} \quad (2)$$

where, $\beta = 1$ or 2 represents *F1* or *F2*, *P* stands for precision, and *R* stands for recall. The effectiveness of a model increases as an *F1* or *F2* score approaches 1.

Stratified 10-fold cross validation was used during model training. It ensured that the fold had approximately equal number of observations for each class. This generalized the model well as estimates would have a low bias. The final dataset, which was balanced using the SMOGN followed by SMOTE with ENN, contained 72 records for game performance prediction after the oversampling and undersampling. By using the traditional 70: 30 training: prediction ratio, 51 records were used for training while 21 (totaling to 72 with 70: 30 ratio) were used for testing. While there were 72 records for the game performance, there were only 11 injury recordings, which was only balanced with Borderline SMOTE without oversampling. The entire dataset of 352 records was divided into 289 and 63 where 9 and 2 injuries were present in the data, respectively. The XGB classifier predictions resulted in an *F1* score of 0.94 for the game performance and *F2* score of 0.83 for injury. Values on the diagonal of each matrix in Fig. 6 represent the correct prediction while the cells labeled with red (1 for game performance and 2 for injury) are false negatives (errors in prediction).

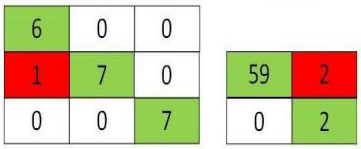


FIGURE 6. Confusion matrix on test set for game performance (left) and injury (right).

While stratified 10-fold cross validation partitions the dataset into training and testing sets randomly, to take the subject-to-subject variations and the tendency of autocorrelations in the time series data belonging to a single subject into account, a leave-one-subject-out analysis was also performed [59]. In this approach, for each athlete, the data belonging to the rest of the athletes was used for training and that belonging to her was used for testing. Repeating the exercise for all athletes, an average F1 score of 0.94 was achieved using the XGB classifier for prediction of game performance and average F2 score of 0.80 for injury.

VI. DISCUSSION

A. GAME PERFORMANCE

A multifaceted approach was taken to the monitoring of game performance to quantify various pressures placed on the athletes. Several features were highlighted for importance (Table 3) including *training strain, daily training load average*, and *HRV*. In our study the athletes averaged lower (6.8 hours per day) than the recommended sleep hours which likely increased the sleep need and sleep debt during the data collection period. Sleep, both in amount and quality, is important for recovery and subsequent adaptations from training. This fact is also corroborated as *sleep need* and *sleep debt hours* appear in the top 10 features for game performance. Team average *Resting Heart Rates (RHR)* (58.70 bpm) and *HRV* fluctuations (11-224) during this period may be reflective of the overall stress placed on the athletes from training, competition, and academic stress.

The second game performance monitoring approach was the measurement of subjective stress as reported by the athletes. While overall higher scores for recovery and lower scores for stress are desired, our analysis revealed that *emotional balance* (recovery metric) and *lack of activation* (stress metric) reflected game performance as they both are highly ranked at 4^{th} and 6^{th} place in feature importance, respectively. As the student athletes in this study were training and competing during times that overlapped with academic pressures, emotional stress may have reflected overall pressure placed on them by sports-related and academic domains. The observations are valid as *EB*, and *overall stress* features also appear in the top 10 features impacting game performance.

The final approach was the monitoring of the entire training and competition workload experienced by the athletes. The primary stresses placed on the athletes are both physiological and biomechanical. From the feature selection, the most predictive features were *training strain* and *daily average training load*. These measures reflect total work completed through the week and the variability of that work. Athletes should attempt to avoid large spikes in training volume and periods of no or little training. During this study, two periods of reduced or absent training were present due to COVID-19-related protocols. The lack of training for 7 to 14 days (potential reduction in physical capabilities) followed by training or competition resulted in decrease in *LA*, increased *monotony*, and decreased *weekly SD*, which all showed up in the feature importance analysis. Based on these results, it can be suggested that through consistent monitoring of workload, return to play protocols should be developed to reintegrate athletes back into training and competition slowly.

B. INJURY

Consistent, varied, and structured training is essential in the physical development of athletes. In this study, a total of 11 injuries occurred during or immediately following four separate shutdown periods, ranging from 10-16 days in length, in which no games were played. These frequent periods of unexpected shutdowns resulted in the athlete's acute (1-week) training load far exceeding that which they had done over a several week (chronic) period, similar to [60]. The majority of injuries in the present study were chronic/overuse injuries such as tendonitis, which aligned with the findings of a study on youth basketball players [61].

In addition to erratic and unstructured training, another factor that has been shown to increase injury prevalence is

sleep dysfunction, which was also correlated to the respiratory rate through recovery. Consistent with previous research, we found that our athletes were sleeping too little (Mean: 6.8 hours, Standard deviation: 1.5 hours). As observed, sleep indeed impacts injury, and it was validated as *REM, awake hours,* and *sleep need* appearing in the top 10 features.

C. PREDICTION

Accuracies of the XGBoost classifier for game performance and injury prediction were found at 95.3% and 96.8%, respectively. However, these could be misleading due to the very low number of records for game score (165) and injuries (11) compared to the size of the sleep and training load data. Therefore, F measures were used to evaluate the performance of the prediction models. With F1 of 0.94 and F2 of 0.83, we believe our models can correctly predict the game score and injury with reasonable confidence. The proportion of imputed data may impact the prediction and the bias introduced due to imputation. In the present case, we assumed that the data is Missing at Random (MAR), which ensured the production of unbiased data after imputation and was evident from the low PB values. However, due to the large proportions of missing data (Table 2), the predictions must be considered as insights or hypothesis-generating results rather than final decisions for coaches.

The 3×3 confusion matrix on the test set for game performance has only one FN in a three cluster classification. The Hollinger game score [36] used to predict game performance is only a rough measure of player's productivity. Thus the predicted game performance provides only partial information to the coaches, and other parameters such as game contribution and practice efforts can be considered to have a more complete assessment of the true impact of a player during the season.

As seen in the 2×2 confusion matrix, the binary injury model predicted four injuries (2 True Negative (TN) and 2 FN) and 59 no-injury (all True Positive (TP), 0 FP). In the injury case, FN is better than FP (injured player predicted as not injured). The coaches prefer over prediction because unknowingly making an injured player play (FP) can prove far more costly to the team compared to a non-injured player predicted as injured (FN).

VII. CONCLUSION

Due to the inherently chaotic nature of sports, coaches and scientists are constantly searching for insights related to an athlete's game performance and injury potential. This study attempted to quantify both; the game performance aspects associated with basketball while simultaneously understanding potential injury risks. Through a multi-disciplinary approach, the authors tried to quantify inputs (training, stressors, and perceptions) and outputs (game performance and injury) in a collegiate basketball setting. The game performance analysis revealed **(RQ1 a)**: 1) questionnaire data, training load, HRV, sleep need, and debt are associated with game performance; 2) Collectively, a combination of questionnaire data, training load metrics, and whoop band metrics can provide insights into game performance.

Similarly, the analysis revealed that significant outcomes associated with injury are **(RQ1 b)**: 1) Restorative sleep, respiratory rate, training monotony, training strain, and total minutes played; 2) A compressed competition schedule and training irregularities likely increased injury risk; 3) The athletes measured in this study slept less than the recommended amount for athletes.

Given the high injury rates previously reported in basketball, specifically collegiate basketball, it is imperative to further explore this relationship within the context of collegiate basketball to quantify training loads and allow clinicians to quantify the risk of subsequent injury. Collectively, this data highlights the need for gradual reintroduction to activities for the athletes' following layoffs and a consistent return to play protocol.

Our prediction models were reliable with high *F1/F2* scores that are very encouraging to be used during a basketball season by working with the coaches to close the loop on insightful decisions **(RQ2)**. Regardless of whether or not we will face another pandemic in the near future, our results will pave the way for understanding the impact of lengthy breaks in a season or unforeseen player absences. As collegiate athletes struggle with lack of sleep and meeting academic, athletic, and social responsibilities, we believe our data collection methodology with a validated machine learning framework will help improve the performance of athletes while reducing the risk of injuries.

ACKNOWLEDGMENT

The research team wants to thank the players and coaching staff for their cooperation. They are also grateful for the support from Welch College of Business & Technology for obtaining the WHOOP straps and the research platform.

(Samah Senbel and Srishti Sharma contributed equally to this work.)

REFERENCES

[1] T. Reilly and B. Edwards, "Altered sleep–wake cycles and physical performance in athletes," *Physiol. Behav.*, vol. 90, nos. 2–3, pp. 274–284, Feb. 2007.

[2] D. van H. van der Kloet, T. Giesbrecht, and H. Merckelbach, "Sleep loss increases dissociation and affects memory for emotional stimuli," *J. Behav. Therapy Experim. Psychiatry*, vol. 47, pp. 9–17, Jun. 2015.

[3] C. Sargent, M. Lastella, S. L. Halson, and G. D. Roach, "How much sleep does an elite athlete need?" *Int. J. Sports Physiol. Perform.*, vol. 1, no. aop, pp. 1–12, 2021.

[4] C. D. Mah, K. E. Mah, E. J. Kezirian, and W. C. Dement, "The effects of sleep extension on the athletic performance of collegiate basketball players," *Sleep*, vol. 34, no. 7, pp. 943–950, Jul. 2011.

[5] J. Schwartz and R. D. Simon, "Sleep extension improves serving accuracy: A study with college varsity tennis players," *Physiol. Behav.*, vol. 151, pp. 541–544, Nov. 2015.

[6] E. Kroshus, J. Wagner, D. Wyrick, A. Athey, L. Bell, H. J. Benjamin, M. A. Grandner, C. E. Kline, J. M. Mohler, J. R. Prichard, N. F. Watson, and B. Hainline, "Wake up call for collegiate athlete sleep: Narrative review and consensus recommendations from the NCAA interassociation task force on sleep and wellness," *Brit. J. Sports Med.*, vol. 53, no. 12, pp. 731–736, Jun. 2019.

[7] M. D. Milewski, D. L. Skaggs, G. A. Bishop, J. L. Pace, D. A. Ibrahim, T. A. L. Wren, and A. Barzdukas, "Chronic lack of sleep is associated with increased sports injuries in adolescent athletes," *J. Pediatric Orthopaedics*, vol. 34, no. 2, pp. 129–133, 2014.
[8] P. von Rosen, A. Frohm, A. Kottorp, C. Fridén, and A. Heijne, "Too little sleep and an unhealthy diet could increase the risk of sustaining a new injury in adolescent elite athletes," *Scandin. J. Med. Sci. Sports*, vol. 27, no. 11, pp. 1364–1371, Nov. 2017.
[9] P. von Rosen, A. Frohm, A. Kottorp, C. Fridén, and A. Heijne, "Multiple factors explain injury risk in adolescent elite athletes: Applying a biopsychosocial perspective," *Scandin. J. Med. Sci. Sports*, vol. 27, no. 12, pp. 2059–2069, Dec. 2017.
[10] T. M. Burke, P. J. Lisman, K. Maguire, L. Skeiky, J. J. Choynowski, V. F. Capaldi, J. N. Wilder, A. J. Brager, and D. A. Dobrosielski, "Examination of sleep and injury among college football athletes," *J. Strength Conditioning Res.*, vol. 34, no. 3, pp. 609–616, 2020.
[11] E. Morgulev, O. H. Azar, and R. Lidor, "Sports analytics and the big-data era," *Int. J. Data Sci. Analytics*, vol. 5, no. 4, pp. 213–222, Jun. 2018.
[12] O. Kostakis, N. Tatti, and A. Gionis, "Discovering recurring activity in temporal networks," *Data Mining Knowl. Discovery*, vol. 31, no. 6, pp. 1840–1871, Nov. 2017.
[13] V. Sarlis and C. Tjortjis, "Sports analytics—Evaluation of basketball players and team performance," *Inf. Syst.*, vol. 93, Nov. 2020, Art. no. 101562.
[14] S. K. Deshpande and S. T. Jensen, "Estimating an NBA player's impact on his team's chances of winning," *J. Quant. Anal. Sports*, vol. 12, no. 2, pp. 51–72, Jan. 2016.
[15] K.-C. Wang and R. Zemel, "Classifying nba offensive plays using neural networks," in *Proc. MIT Sloan Sports Analytics Conf.*, vol. 4, 2016, pp. 1–9.
[16] J. Piette, L. Pham, and S. Anand, "Evaluating basketball player performance via statistical network modeling," in *Proc. 5th MIT Sloan Sports Analytics Conf.*, 2011, pp. 1–10.
[17] V. Sarlis, V. Chatziilias, C. Tjortjis, and D. Mandalidis, "A data science approach analysing the impact of injuries on basketball player and team performance," *Inf. Syst.*, vol. 99, Jul. 2021, Art. no. 101750.
[18] J. G. Claudino, D. D. O. Capanema, T. V. de Souza, J. C. Serrão, A. C. Machado Pereira, and G. P. Nassis, "Current approaches to the use of artificial intelligence for injury risk assessment and performance prediction in team sports: A systematic review," *Sports Med.*, vol. 5, no. 1, pp. 1–12, Dec. 2019.
[19] H. Talukder, T. Vincent, G. Foster, C. Hu, J. Huerta, A. Kumar, M. Malazarte, D. Saldana, and S. Simpson, "Preventing in-game injuries for nba players," in *Proc. MIT Sloan Sports Analytics Conf.*, Boston, MA, USA, 2016, pp. 11–12.
[20] E. Adetiba, V. C. Iweanya, S. I. Popoola, J. N. Adetiba, and C. Menon, "Automated detection of heart defects in athletes based on electrocardiography and artificial neural network," *Cogent Eng.*, vol. 4, no. 1, Jan. 2017, Art. no. 1411220.
[21] S. Ge, "Research on the factors of basketball injury in physical teaching based on artificial neural network," *Rev. Fac. Ing*, vol. 32, pp. 415–422, 2017.
[22] L. V. Andreato, D. R. Coimbra, and A. Andrade, "Challenges to athletes during the home confinement caused by the COVID-19 pandemic," *Strength Conditioning J.*, vol. 42, no. 3, pp. 1–5, 2020.
[23] F. Sarto, F. M. Impellizzeri, J. Spörri, S. Porcelli, J. Olmo, B. Requena, L. Suarez-Arrones, A. Arundale, J. Bilsborough, M. Buchheit, J. Clubb, A. Coutts, D. Nabhan, L. Torres-Ronda, A. Mendez-Villanueva, I. Mujika, N. A. Maffiuletti, and M. V. Franchi, "Impact of potential physiological changes due to COVID-19 home confinement on athlete health protection in elite sports: A call for awareness in sports programming," *Sports Med.*, vol. 50, no. 8, pp. 1417–1419, Aug. 2020.
[24] R. Grazioli, I. Loturco, B. M. Baroni, G. S. Oliveira, V. Saciura, E. Vanoni, R. Dias, F. Veeck, R. S. Pinto, and E. L. Cadore, "Coronavirus disease-19 quarantine is more detrimental than traditional off-season on physical conditioning of professional soccer players," *J. Strength Conditioning Res.*, vol. 34, no. 12, pp. 3316–3320, 2020.
[25] R. Ross, L. Irvin, R. Severin, and B. Ellis, "Return-to-play considerations after COVID-19 infection in elite athletes," *J. Athletic Training*, vol. 56, no. 10, pp. 1061–1063, Oct. 2021.
[26] C. Taber, S. Senbel, D. Ezzeddine, J. Nolan, A. Ocel, N. Artan, and T. Kaya, "Sleep and physical performance: A case study of collegiate women's division 1 basketball players," in *Proc. 43rd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Oct. 2021, pp. 6787–6790.
[27] C. Foster, J. A. Florhaug, J. Franklin, L. Gottschall, L. A. Hrovatin, S. Parker, P. Doleshal, and C. Dodge, "A new approach to monitoring exercise training," *J. Strength Conditioning Res.*, vol. 15, no. 1, pp. 109–115, Feb. 2001.
[28] M. Haddad, G. Stylianides, L. Djaoui, A. Dellal, and K. Chamari, "Session-RPE method for training load monitoring: Validity, ecological usefulness, and influencing factors," *Frontiers Neurosci.*, vol. 11, pp. 1–14, Nov. 2017.
[29] G. Borg, *Borg's Perceived Exertion and Pain Scales*. Champaign, IL, USA: Human Kinetics, 1998.
[30] M. McGuigan, *Monitoring Training and Performance in Athletes*. Champaign, IL, USA: Human Kinetics, 2017.
[31] S. Berryhill, C. J. Morton, A. Dean, A. Berryhill, N. Provencio-Dean, S. I. Patel, L. Estep, D. Combs, S. Mashaqi, L. B. Gerald, J. A. Krishnan, and S. Parthasarathy, "Effect of wearables on sleep in healthy individuals: A randomized crossover trial and validation study," *J. Clin. Sleep Med.*, vol. 16, no. 5, pp. 775–783, May 2020.
[32] D. J. Miller, M. Lastella, A. T. Scanlan, C. Bellenger, S. L. Halson, G. D. Roach, and C. Sargent, "A validation study of the WHOOP strap against polysomnography to assess sleep," *J. Sports Sci.*, vol. 38, no. 22, pp. 2631–2636, Nov. 2020.
[33] Y. Sekiguchi, W. M. Adams, C. L. Benjamin, R. M. Curtis, G. E. W. Giersch, and D. J. Casa, "Relationships between resting heart rate, heart rate variability and sleep characteristics among female collegiate cross-country athletes," *J. Sleep Res.*, vol. 28, no. 6, Dec. 2019.
[34] M. Kellmann and K. W. Kallus, *Recovery-stress questionnaire for athletes: User manual*. Champaign, IL, USA: Human Kinetics, 2001.
[35] M. Kellmann and S. Kölling, *Recovery and Stress in Sport: A Manual for Testing and Assessment*. Evanston, IL, USA: Routledge, 2019.
[36] J. Hollinger, *Pro Basketball Forecast*. Lincoln, NE, USA: Potomac Books, Univ. of Nebraska Press, 2005.
[37] J. Kubatko, D. Oliver, K. Pelton, and D. T. Rosenbaum, "A starting point for analyzing basketball statistics," *J. Quant. Anal. Sports*, vol. 3, no. 3, pp. 1–22, Jan. 2007.
[38] R. Arnab, *Survey Sampling Theory and Applications*. New York, NY, USA: Academic, 2017.
[39] B. Sv and K. Groothuis-Oudshoorn, "Mice: Multivariate imputation by chained equations in R," *J. Stat. Softw.*, vol. 45, pp. 1–67, Dec. 2010.
[40] M. Mera-Gaona, U. Neumann, R. Vargas-Canas, and D. M. López, "Evaluating the impact of multivariate imputation by MICE in feature selection," *PLoS ONE*, vol. 16, no. 7, Jul. 2021, Art. no. e0254720.
[41] J. W. Graham, A. E. Olchowski, and T. D. Gilreath, "How many imputations are really needed? Some practical clarifications of multiple imputation theory," *Prevention Sci.*, vol. 8, no. 3, pp. 206–213, Aug. 2007.
[42] P. Madley-Dowd, R. Hughes, K. Tilling, and J. Heron, "The proportion of missing data should not be used to guide decisions on multiple imputation," *J. Clin. Epidemiol.*, vol. 110, pp. 63–73, Jun. 2019.
[43] *Interpretable Machine Learning*, Molnar, Lulu, Abu Dhabi, United Arab Emirates, 2020.
[44] E. Tjoa and C. Guan, "A survey on explainable artificial intelligence (XAI): Toward medical XAI," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 11, pp. 4793–4813, Nov. 2021.
[45] A. Fisher, C. Rudin, and F. Dominici, "All models are wrong, but many are useful: Learning a variable's importance by studying an entire class of prediction models simultaneously," *J. Mach. Learn. Res.*, vol. 20, no. 177, pp. 1–81, 2019.
[46] R. Bellman, "Dynamic programming," *Science*, vol. 153, nos. 37–31, pp. 34–37, 1966.
[47] J. Tang, S. Alelyani, and H. Liu, "Feature selection for classification: A review," *Data Classification: Algorithms and Applications*. U.K.: CRC Press, 2014, p. 37.
[48] A. M. Kowshalya, R. Madhumathi, and N. Gopika, "Correlation based feature selection algorithms for varying datasets of different dimensionality," *Wireless Pers. Commun.*, vol. 108, no. 3, pp. 1977–1993, Oct. 2019.
[49] T. Hastie, R. Tibshirani, and J. Friedman, "Unsupervised learning," in *The Elements of Statistical Learning*. New York, NY, USA: Springer, 2009, pp. 485–585.
[50] J. Jaccard, R. Turrisi, and J. Jaccard, *Interaction Effects in Multiple Regression*, no. 72. Newbury Park, CA, USA: Sage, 2003.
[51] Y. Dong and C.-Y.-J. Peng, "Principled missing data methods for researchers," *SpringerPlus*, vol. 2, no. 1, pp. 1–17, Dec. 2013.
[52] B. Krawczyk, "Learning from imbalanced data: Open challenges and future directions," *Prog. Artif. Intell.*, vol. 5, no. 4, pp. 221–232, 2016.
[53] S. Sreejith, H. Khanna Nehemiah, and A. Kannan, "Clinical data classification using an enhanced SMOTE and chaotic evolutionary feature selection," *Comput. Biol. Med.*, vol. 126, Nov. 2020, Art. no. 103991.

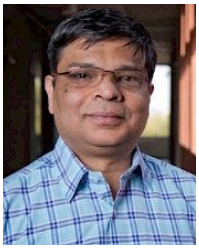
[54] P. Branco, L. Torgo, and R. P. Ribeiro, "Smogn: A pre-processing approach for imbalanced regression," in *Proc. 1st Int. Workshop Learn. Imbalanced Domains, Theory Appl.*, 2017, pp. 36–50.
[55] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," *J. Artif. Intell. Res.*, vol. 16, pp. 321–357, Jun. 2002.
[56] L. Shao, D. Zhang, H. Du, and D. Fu, "Deep forest in ADHD data classification," *IEEE Access*, vol. 7, pp. 137913–137919, 2019.
[57] H. Han, W.-Y. Wang, and B.-H. Mao, "Borderline-smote: A new over-sampling method in imbalanced data sets learning," in *Proc. Int. Conf. Intell. Comput.* New York, NY, USA: Springer, 2005, pp. 878–887.
[58] N. S. Rajliwall, R. Davey, and G. Chetty, "Cardiovascular risk prediction based on XGBoost," in *Proc. 5th Asia–Pacific World Congr. Comput. Sci. Eng. (APWC CSE)*, Dec. 2018, pp. 246–252.
[59] Y. Gao, X. Zhang, S. Wang, and G. Zou, "Model averaging based on leave-subject-out cross-validation," *J. Econometrics*, vol. 192, no. 1, pp. 139–151, May 2016.
[60] L. Bowen, A. S. Gross, M. Gimpel, S. Bruce-Low, and F.-X. Li, "Spikes in acute: Chronic workload ratio (ACWR) associated with a 5–7 times greater injury rate in English premier league football players: A comprehensive 3-year study," *Brit. J. Sports Med.*, vol. 54, no. 12, pp. 731–738, Jun. 2020.
[61] L. C. Benson, O. B. A. Owoeye, A. M. Räisänen, C. Stilling, W. B. Edwards, and C. A. Emery, "Magnitude, frequency, and accumulation: Workload among injured and uninjured youth basketball players," *Frontiers Sports Act. Living*, vol. 3, p. 49, Apr. 2021.



SAMAH SENBEL joined the Faculty, Sacred Heart University, in 2017. She is currently an Assistant Professor with the School of Computer Science and Engineering, Sacred Heart University. Her teaching focus includes programming, data structures, distributed systems, web programming, and operating systems. Her research interests include data analytics, distributed systems, cloud computing, machine learning, and big data analytics.



SRISHTI SHARMA is currently pursuing the Ph.D. degree in computer science with Ahmedabad University. Her research interests include machine learning and sports data analytics.



MEHUL S. RAVAL (Senior Member, IEEE) received the bachelor's degree in ECE, the master's degree in EC, and the Ph.D. degree in ECE from the College of Engineering Pune, University of Pune, India, in 1996, 2002, and 2008, respectively. He has more than 24 years of experience as an academic with visits to Okayama University, Japan, under Sakura Science Fellowship, an Argosy Visiting Associate Professor at the Olin College of Engineering, USA, during fall 2016, and a Visiting Professor at Sacred Heart University, Fairfield, CT, USA, in 2019. He is currently an Associate Dean of experiential learning and a Professor with Ahmedabad University. His research interests include computer vision, image processing, machine learning, and engineering education. He served the IEEE Gujarat Section, during 2008–2015, the IEEE Signal Processing Society (SPS) Chapter, the IEEE Gujarat Section, as the Vice Chair and an Exe-Com Member, in 2014. He is also serving as the Chair for the IEEE Computational Intelligence Society Chapter, IEEE Gujarat Section.



CHRIS TABER received the Ph.D. degree in sport and exercise physiology from East Tennessee State University. He is currently the Program Director of exercise and sport science M.S. program and an Assistant Professor at the College of Health Professions, Sacred Heart University. His research interests include strength and power development, periodization, muscle physiology, and monitoring and testing of athletic performance.



JULIE NOLAN received the Ph.D. degree in kinesiology from the University of Connecticut, in 2013. She is currently an Assistant Professor with the Athletic Training Department, Sacred Heart University. Her research interests include exertional heat stroke, hydration and maximizing performance in the heat, thermoregulation and thermal physiology, and preventing sudden death in sports.



N. SERTAC ARTAN (Member, IEEE) received the Ph.D. degree in electrical engineering from New York University (formerly Polytechnic University). Before joining New York Tech, he was on the Faculty, New York University School of Engineering. He also worked as an Application Specific Integrated Circuit (ASIC) Design Engineer and designed integrated circuits for commercial, academic, and military applications. He is currently an Assistant Professor of electrical and computer engineering at the NYIT College of Engineering and Computer Sciences. His current research interests include low-power embedded systems and very large scale integration (VLSI) circuits for implantable medical devices, and efficient algorithms for biomedical signal processing for applications, such as epileptic seizure detection.



DIALA EZZEDDINE received the bachelor's degree in pure mathematics with a strong teaching experience in mathematics and statistics and the master's degree in statistics, and the Ph.D. degree in computer science from Lumiere University, France, where she specialized in machine learning and data mining. She is currently an Independent Data Scientist. She taught mathematics and statistics for more than ten years.



TOLGA KAYA (Senior Member, IEEE) is currently a Professor and the Director of Engineering programs under the School of Computer Science and Engineering. He oversees the growth of engineering programs and the IDEA Laboratory (makerspace) at Sacred Heart University (SHU). Before joining to SHU, he was an Associate Professor at the School of Engineering and Technology, Central Michigan University, from 2010 to 2017. From 2007 to 2010, he worked as a Research Scientist at the School of Engineering and Applied Sciences, Yale University. He worked in industry as a Technology Consultant with Brightwell Corporation and a Senior Analog Circuit Designer with Microelectronics, Inc., for several years. His research interest includes wearable devices to monitor physiology of athletes. He is well known in the field of hydration research and sports data analytics. He also focuses on the Internet of Things (IoT) devices and 3D printing technologies.